The Two Sides of the Data Coin: *Data Protection vs. Data Retention in Practice*

Anna Bordioug
Protiviti M365 Consultant

Joanne Klein
NexNovus M365 Consultant, Microsoft MVP

Microsoft 365 Security & Compliance User Group | October 29, 2025

Joanne C Klein

NexNovus M365 Consultant

joannecklein.com

https://www.linkedin.com/in/joannecklein/

Anna Bordioug

Protiviti M365 Consultant

www.annabordioug.com

https://www.linkedin.com/in/anna-bordioug/

Agenda

- Data Classification: An Important Underpinning to Purview
- Data Protection and Data Retention: What They Each Look Like in Purview
- How Does AI Affect Both Data Protection and Data Retention
- Practical Lessons We've Learned from Customers
- Common Questions We Hear from Customers
- Q&A (time permitting)

Leave with our exclusive Top 20 Label Behavior Comparison Chart (at the end of this deck!)

Data Classification

An Important Underpinning to Purview

What is data classification and why is it important?

Core Role of the Data Classification service

To store classification details of your content to reflect the current Purview classifier definitions (match, confidence level, occurrence count). This helps to **identify information** your org finds sensitive in a **scalable, automated** way.

Sensitive information types	Trainable classifiers	Exact Data Match classifier	Document Fingerprint
• Pre-built • Custom	• Pre-trained • Custom	• Custom	• Custom

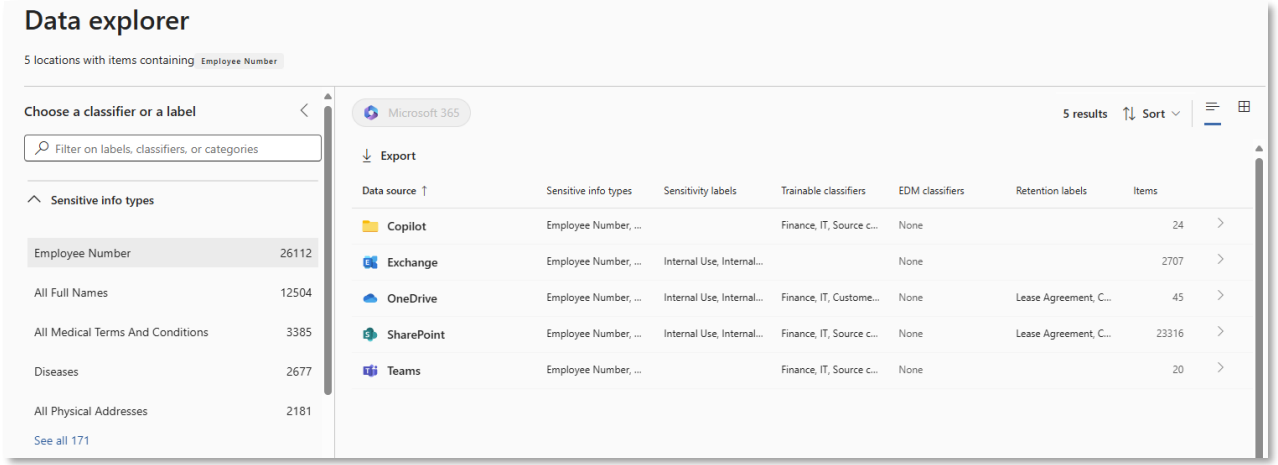
Actionable Compliance Workflows

Classified data enables **policy enforcement**, **automated labeling**, data loss prevention, and insider risk management **risky activity detection**.

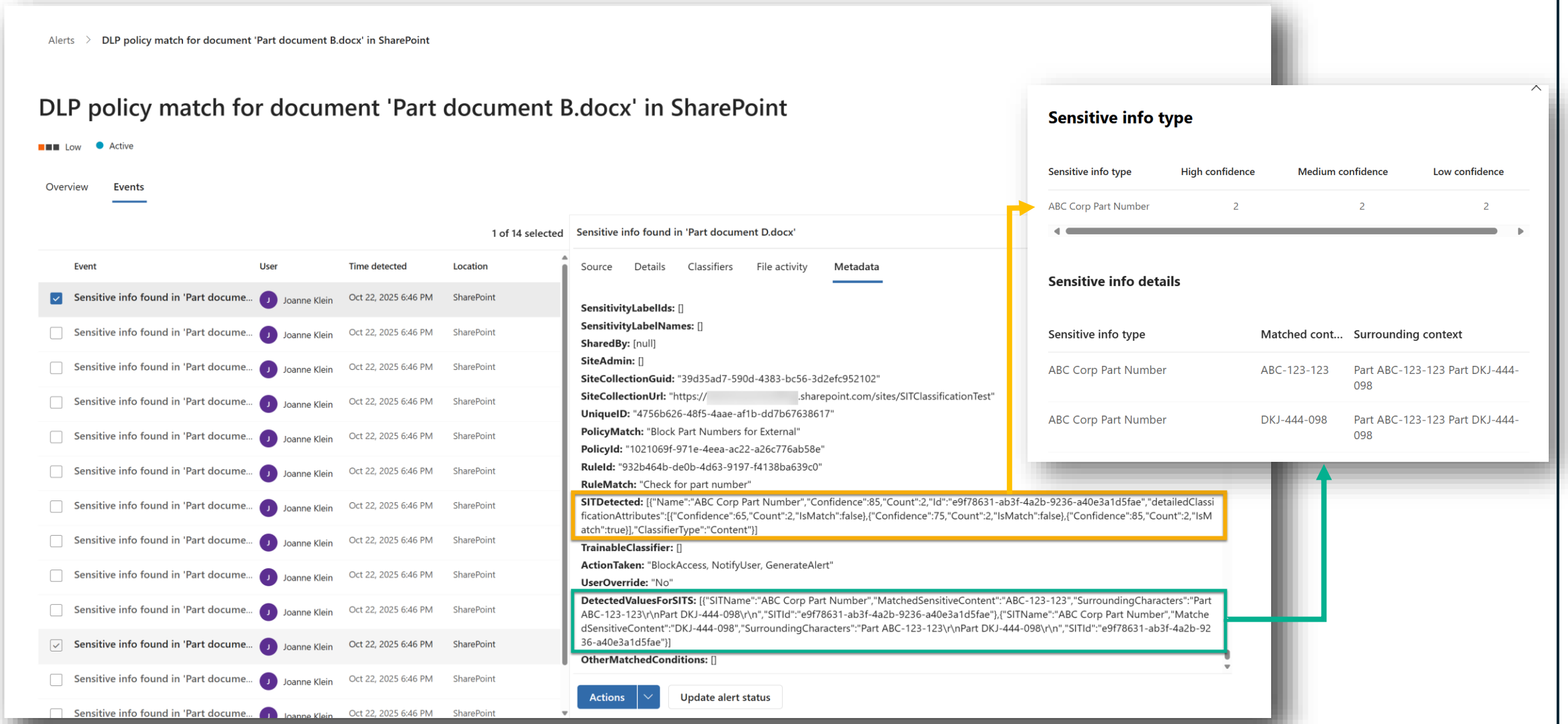
Where do you see classification details for your content?

A lot of places! Some of it is summarized for you in a nice UI format; some of it is raw classification details.

(e.g., Purview Data Explorer)



What is data classification and why is it important?



What is data classification and why is it important?

What triggers data to be classified?

Client-side classification

- Users create/update content
- Exchange/SharePoint/OneDrive

“data in motion”

Re-index SharePoint site

- Initiated via UI or PowerShell
- SharePoint/OneDrive

“data at rest”*

On-demand classification

- Initiated via Purview… UI or PowerShell
- SharePoint/OneDrive/Endpoint

“data at rest”

Downstream effects of having your data classified…

Once data is classified, many other Purview solutions can use the classification details as input. This can reduce the risk of AI tools surfacing unlabeled or unprotected data.

PURVIEW SOLUTION	CLASSIFICATION IMPACT
Information Protection	Auto-labeling based on current classification of content
Data Loss Prevention (DLP)	Policy actions can be triggered by classified content
Insider Risk Management	Flags risky behavior involving sensitive data
Records Management	Applies retention labels based on current classification of content
eDiscovery & Audit	Enables targeted search and review of sensitive content in classified content
Communication Compliance	Detects sensitive or inappropriate communications for classified content

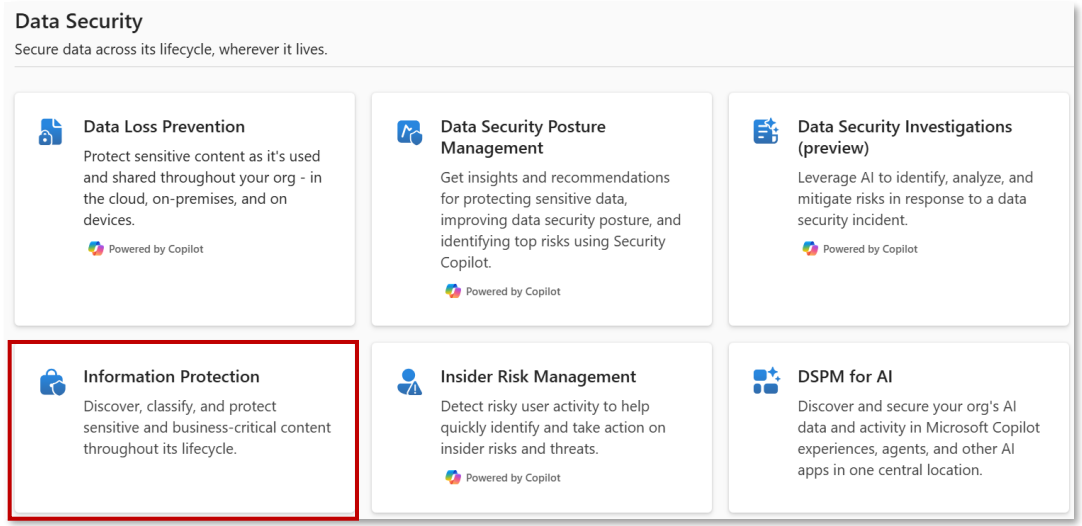
Technical Best Practices

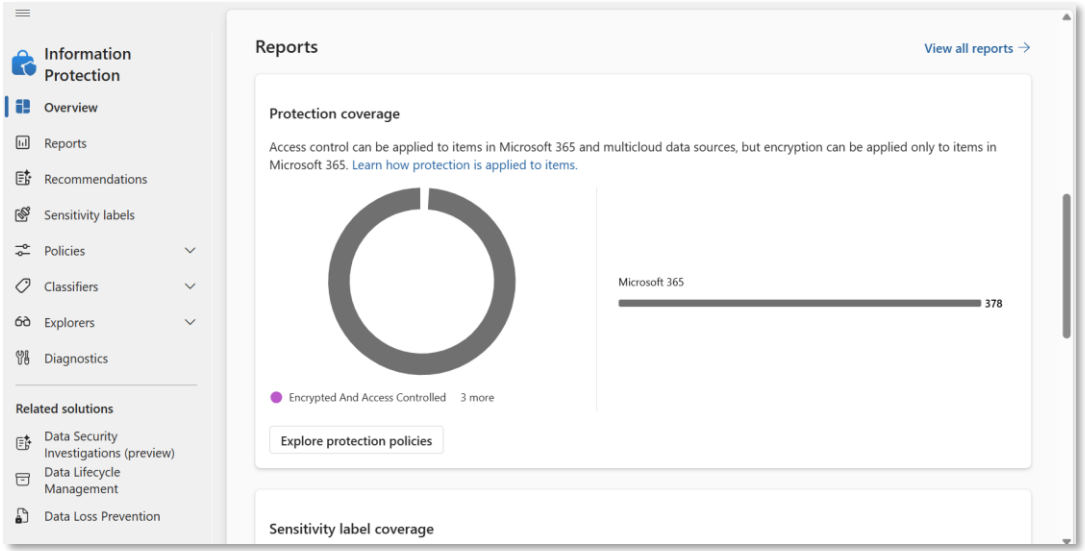
BEST PRACTICE	DESCRIPTION
Start with Built-in Types	Start with predefined sensitive info types for validation for accuracy in your tenant. Don't underestimate the time/effort in this step!
Expand to Custom Classifiers as early as possible	Create organization-specific classifiers for tailored detection. The earlier you can define these in your tenant lifecycle the better.
Testing classifiers	Test via Exchange PowerShell: *$r = Test-DataClassification -ClassificationNames "Canada Physical Addresses" -TextToClassify "value"* *$r.ClassificationResults* Testing via UI: Purview Classifiers… SIT classifiers and EDM classifiers (single file only)
Validating post-classification	Use Data Explorer, Content Explorer, eDiscovery search, Activity Explorer to verify classification coverage. Syntax to search for a SIT: *SensitiveType:"SIT Name",1..99,86..100* SIT name or GUID · Count range · Confidence range
Cross-Team Collaboration	Align security, compliance, and governance teams for a unified strategy

Data Protection and Data Retention

What they each look like in Purview

Data Protection in Microsoft Purview

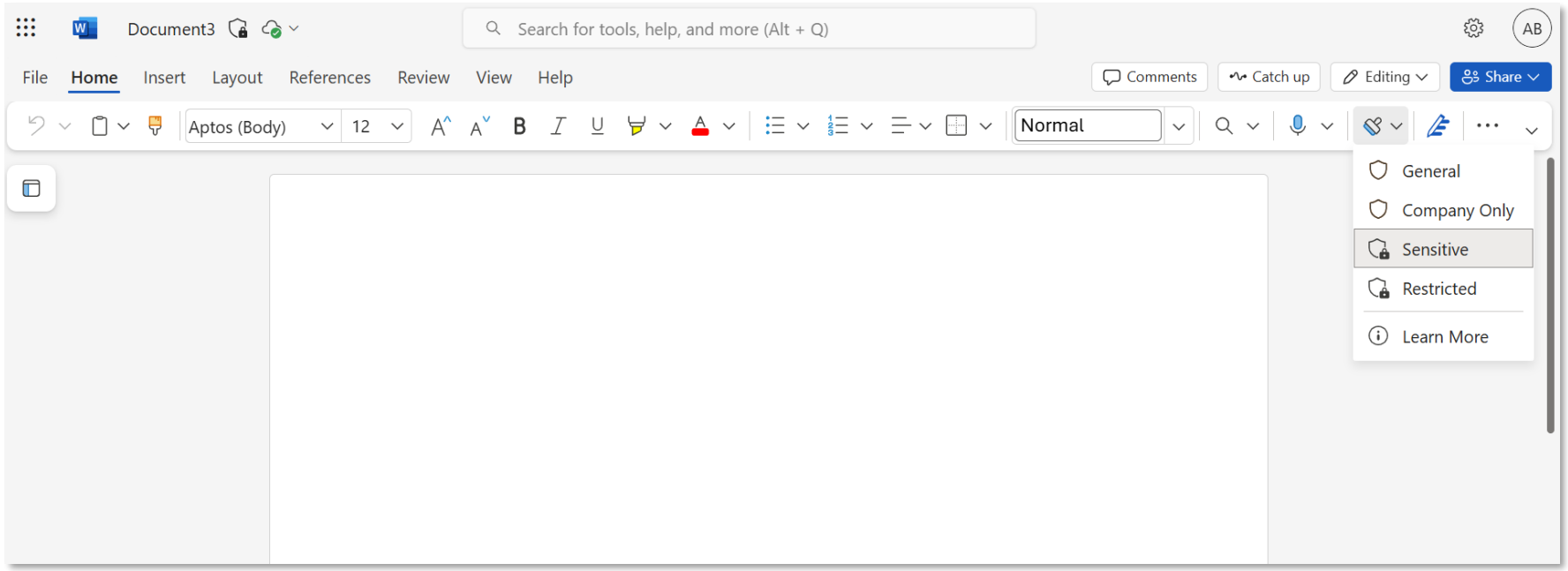


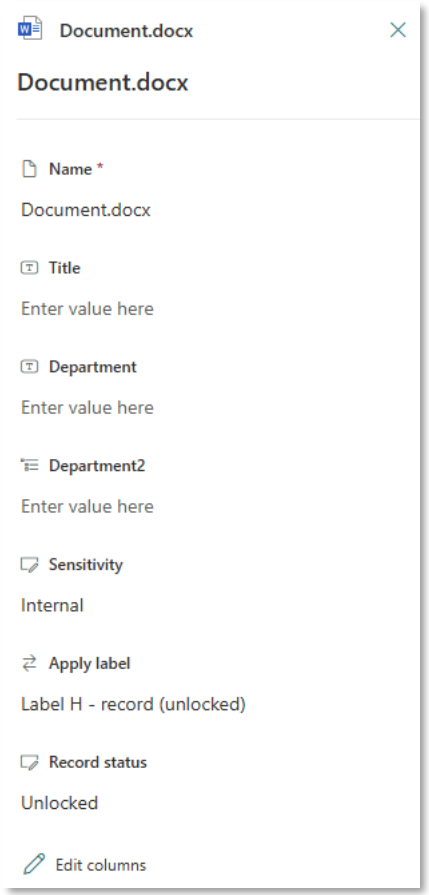


Information Protection

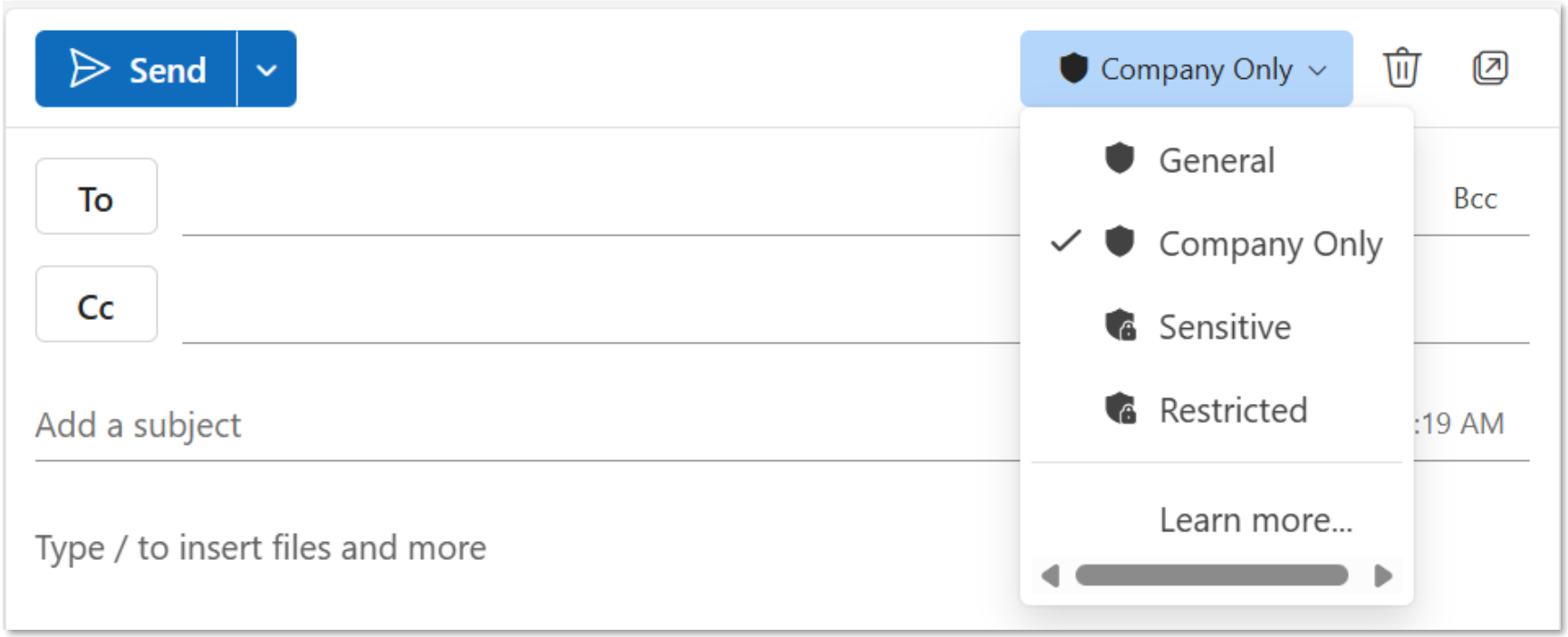
- Classify, label, and protect sensitive data
- Encrypt files and emails to enforce access controls
- Automatically apply sensitivity labels if sensitive information exists in content
- Apply default sensitivity labels to increase labeling coverage
- Protect SharePoint sites, Microsoft 365 groups, and Teams meetings

Microsoft Purview Information Protection Demo

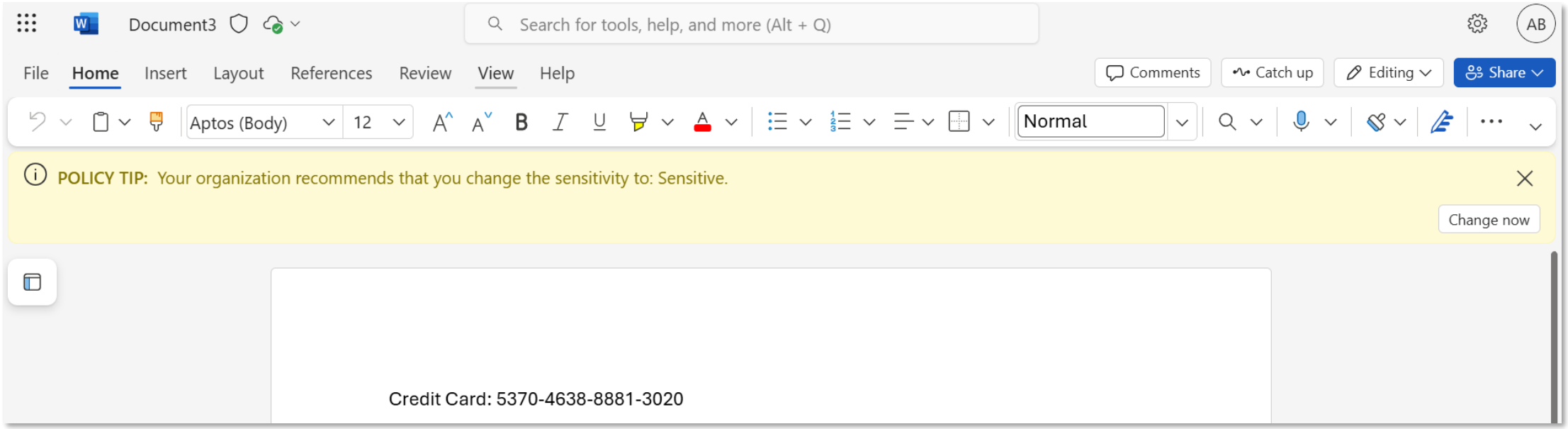




Microsoft Purview Information Protection Demo

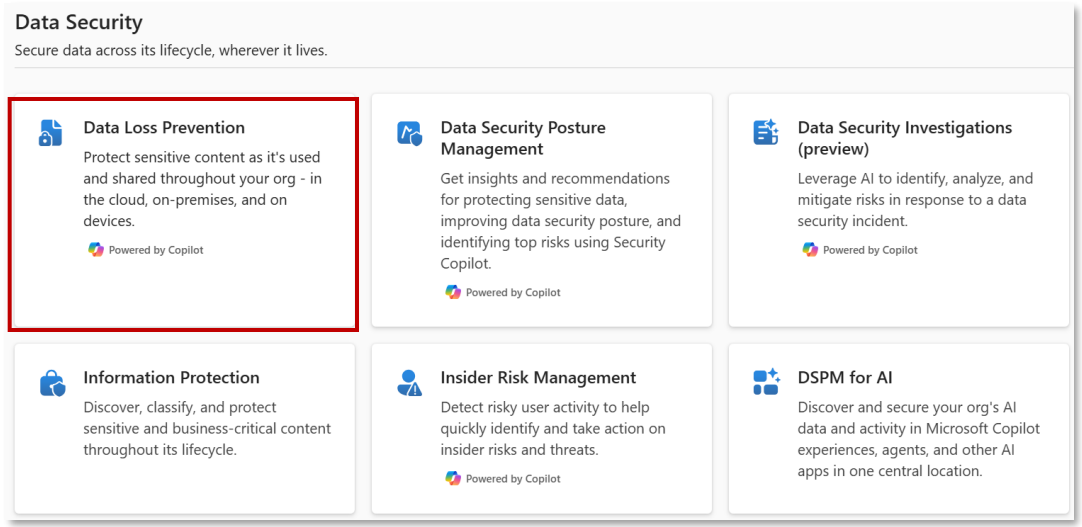


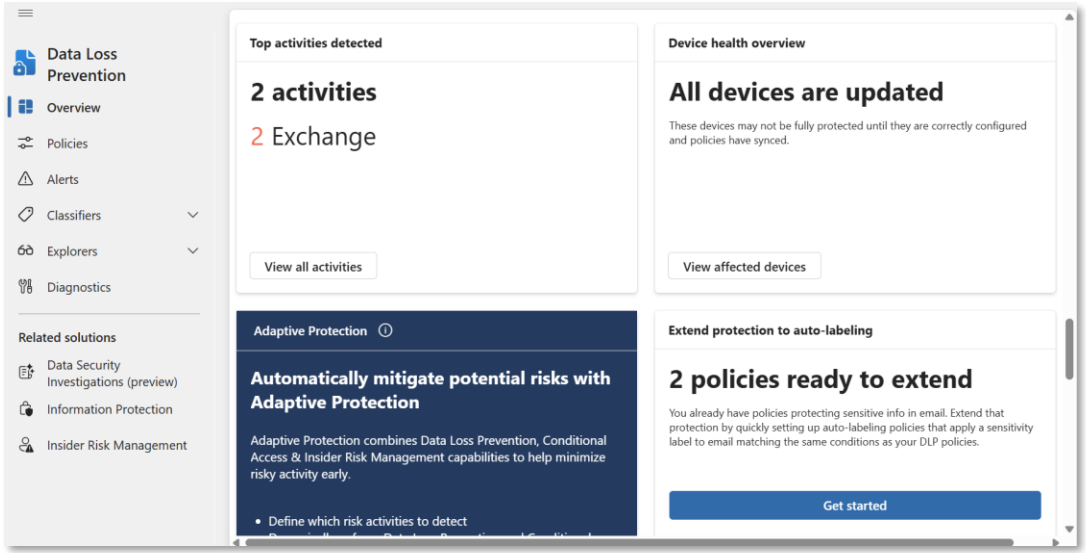
Microsoft Purview Information Protection Demo



Microsoft Purview Information Protection Demo

Data Protection in Microsoft Purview

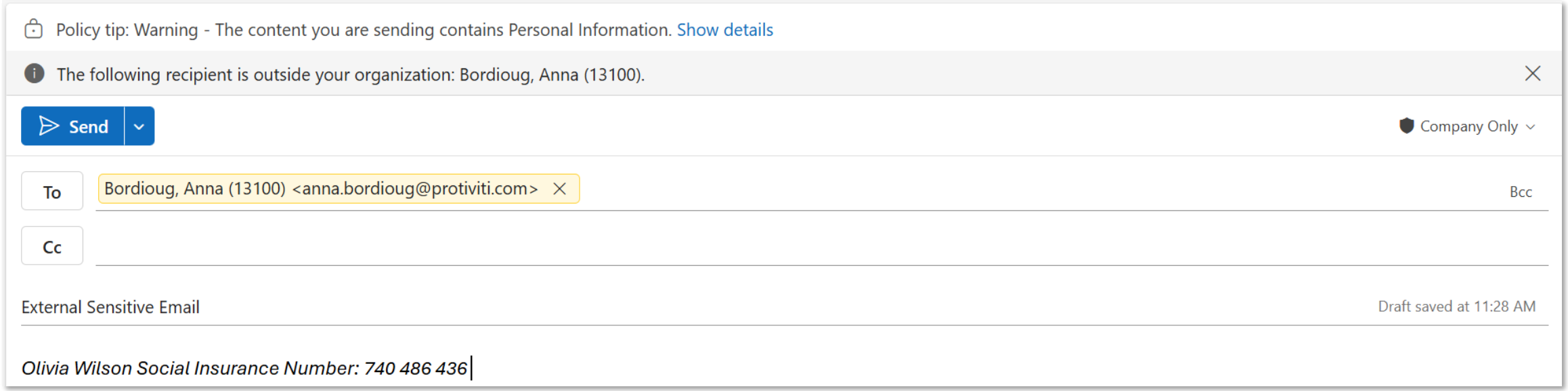


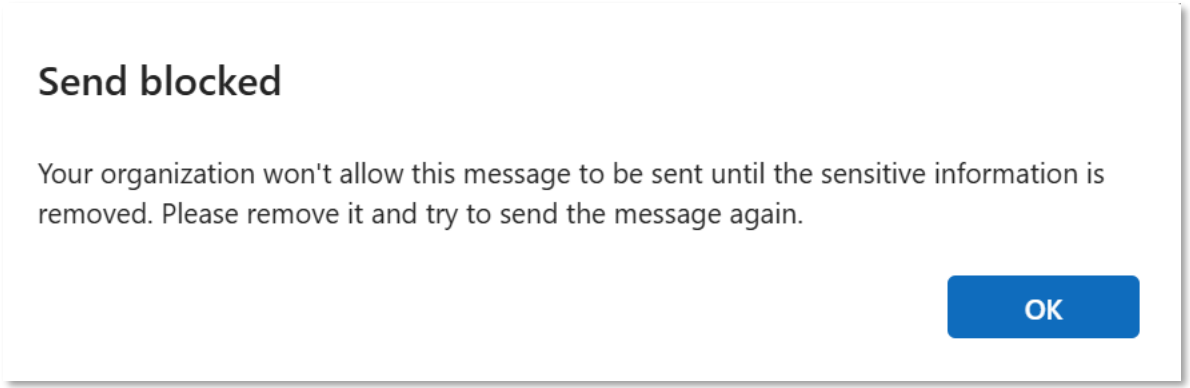


Data Loss Prevention

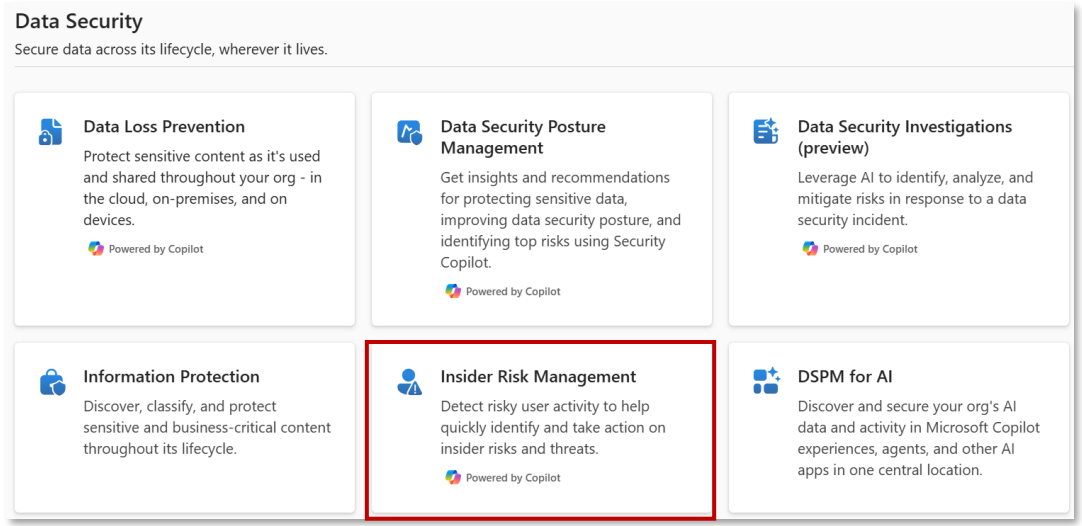
- Implement policies to audit and block data loss across Microsoft 365 and other cloud locations
- Monitor user activity and protect data on devices
- Educate users on unapproved information flows
- Protect data on-premises
- Review and escalate realized data security risks tracked by policies

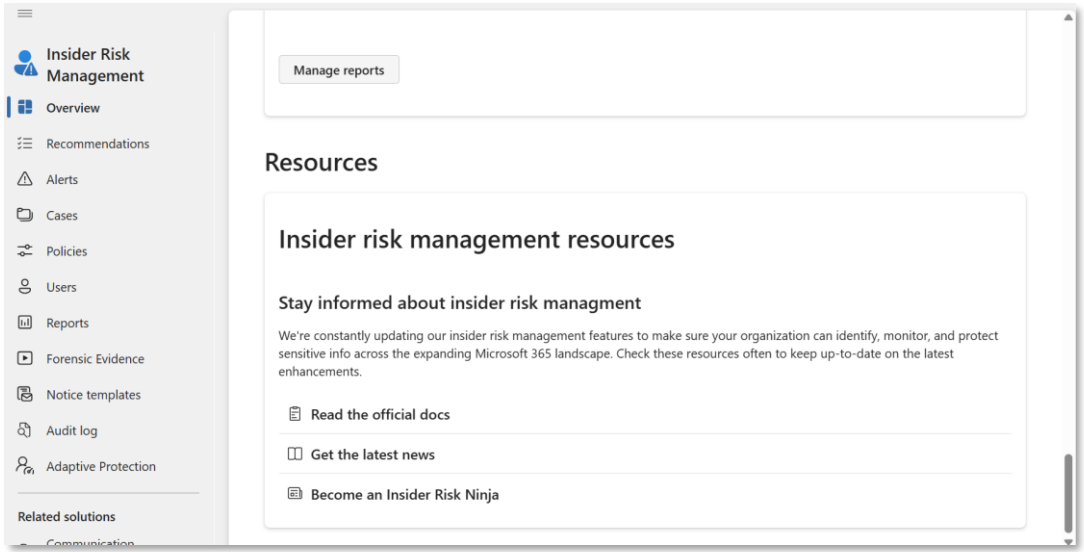
Microsoft Purview Data Loss Prevention Demo





Data Protection in Microsoft Purview

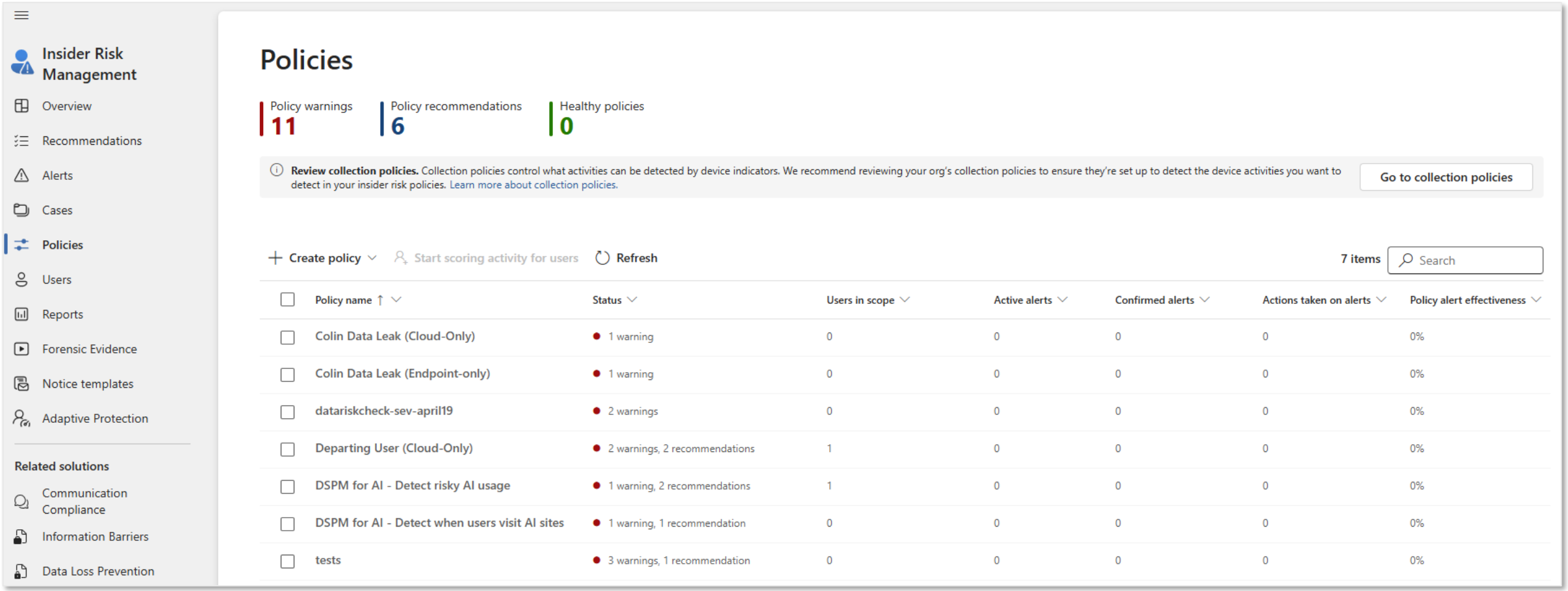




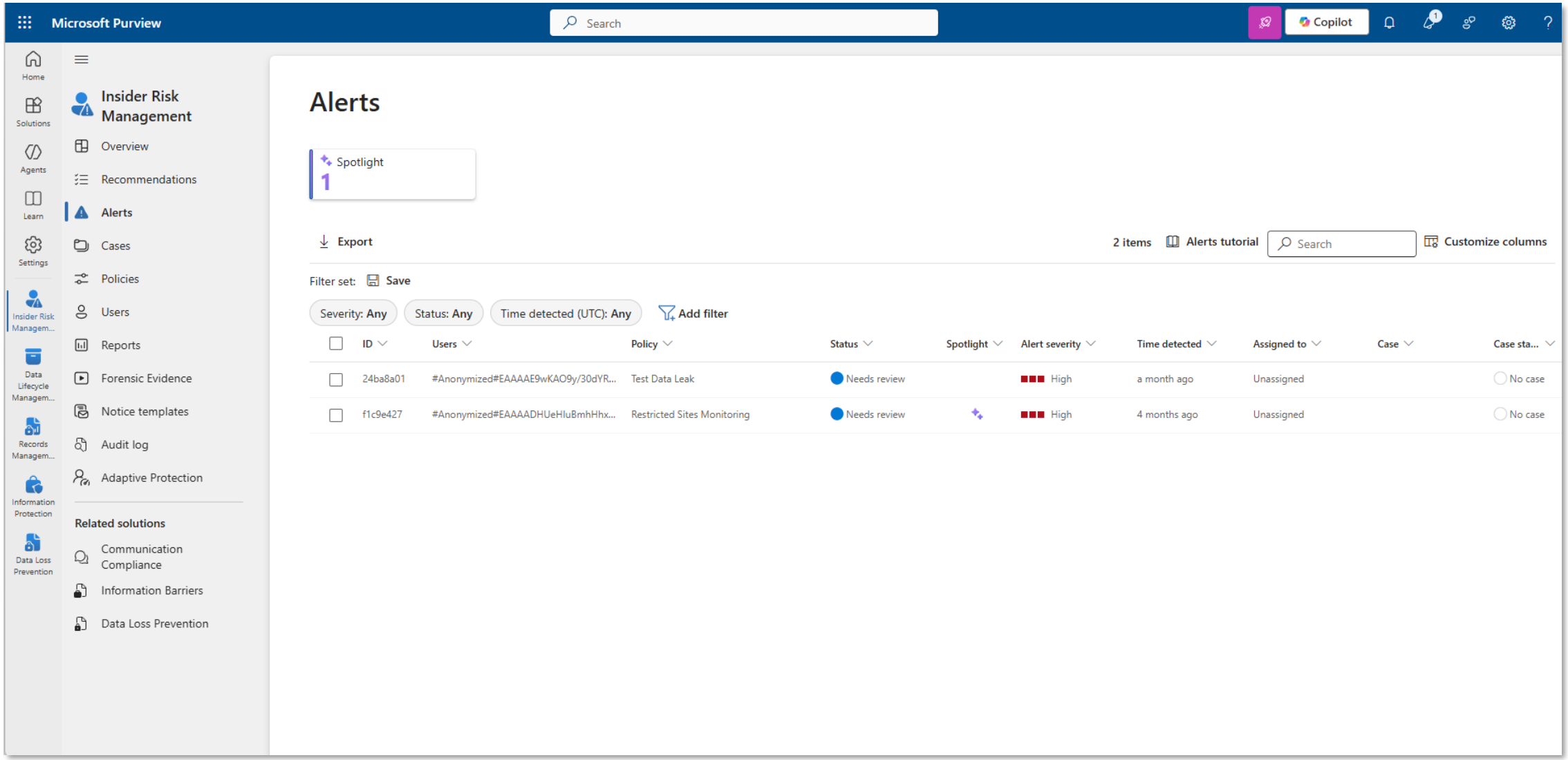
Insider Risk Management

- Implement policies to proactively identify insider risks (e.g., exfiltration of data by departing users)
- Tailor policies to industry-specific insights such as misuse of health records
- Dynamically apply security controls based on a user's risk level
- Review and escalate data security risks realized by insiders and tracked by policies

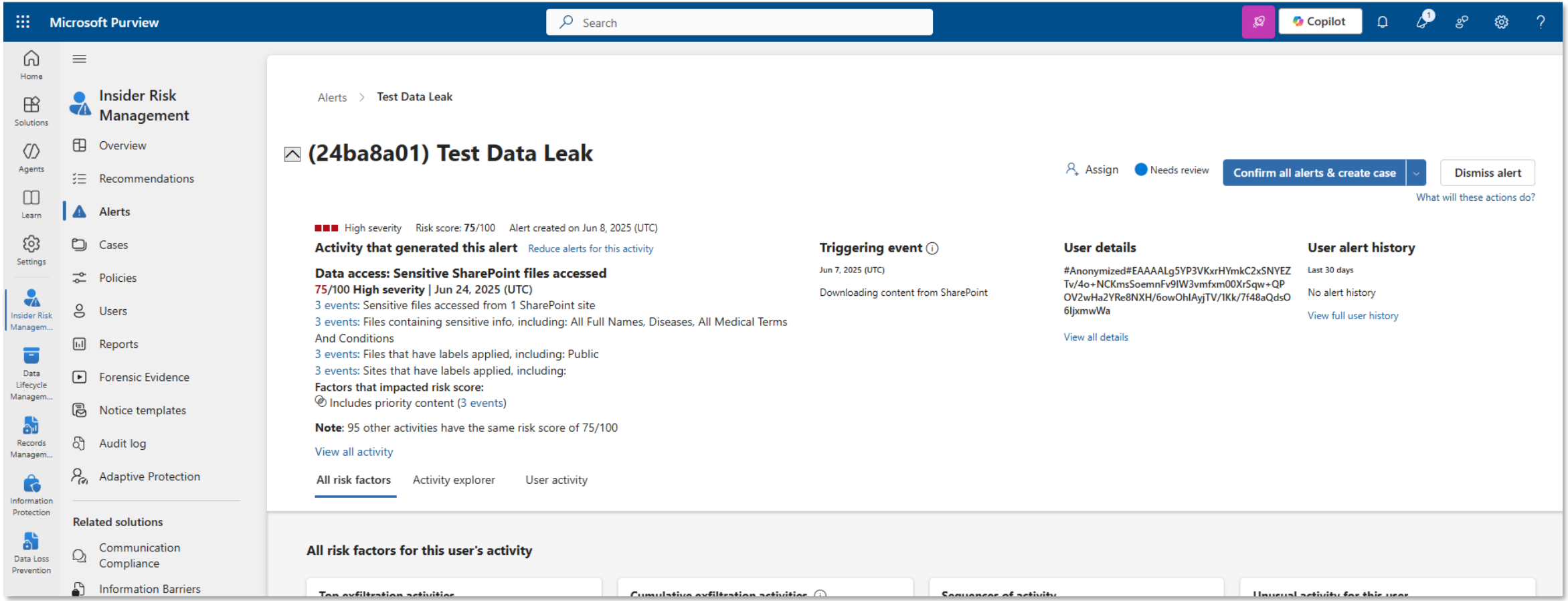
Microsoft Purview Insider Risk Management Demo



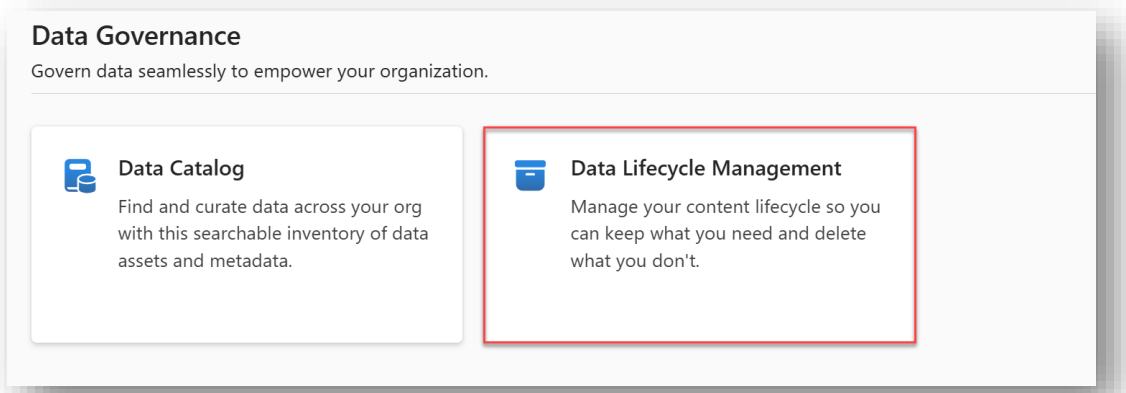
Microsoft Purview Insider Risk Management Demo

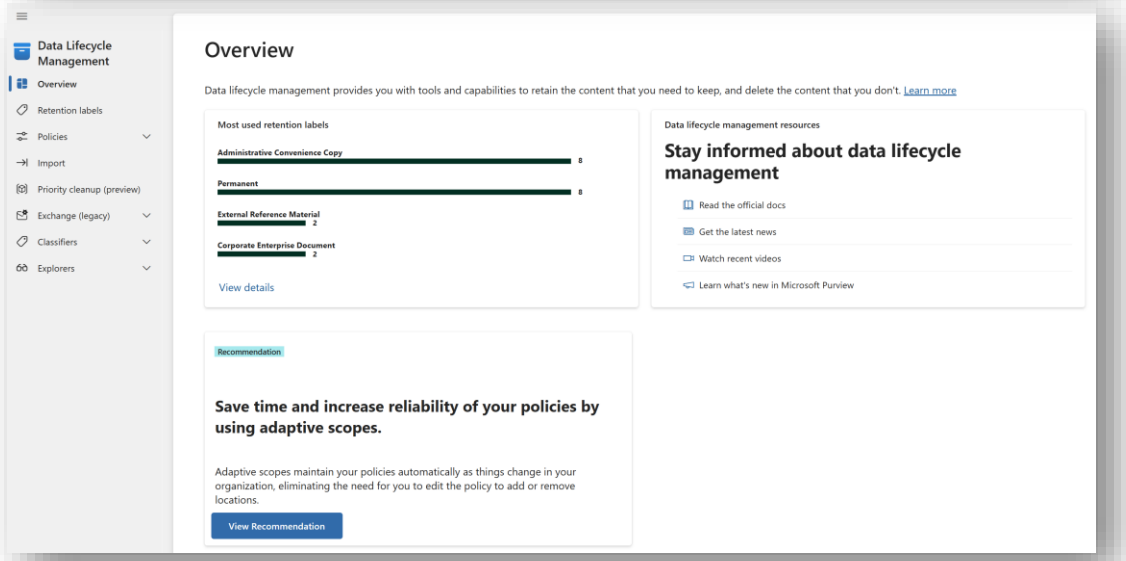


Microsoft Purview Insider Risk Management Demo



Data Lifecycle Management in Microsoft Purview

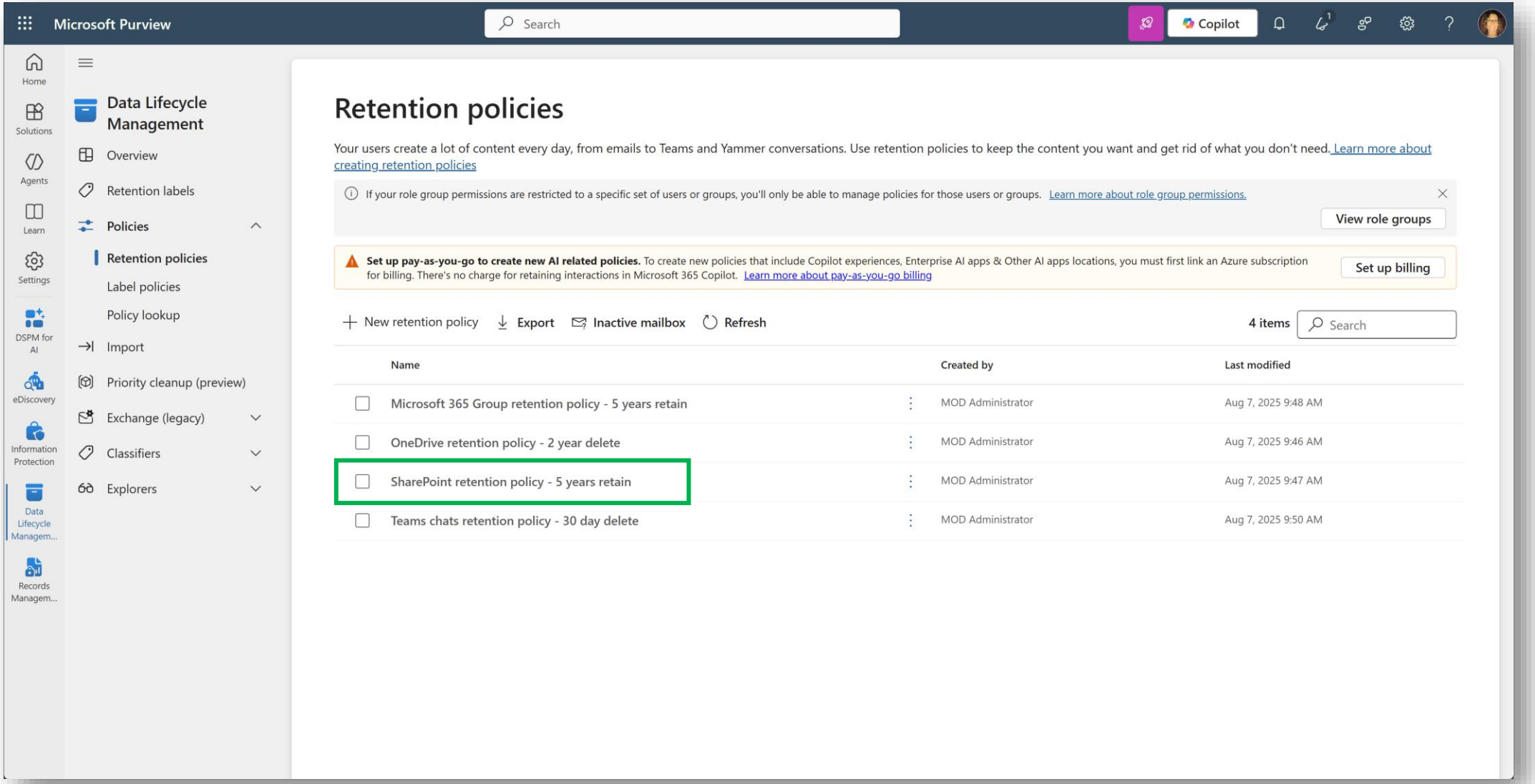




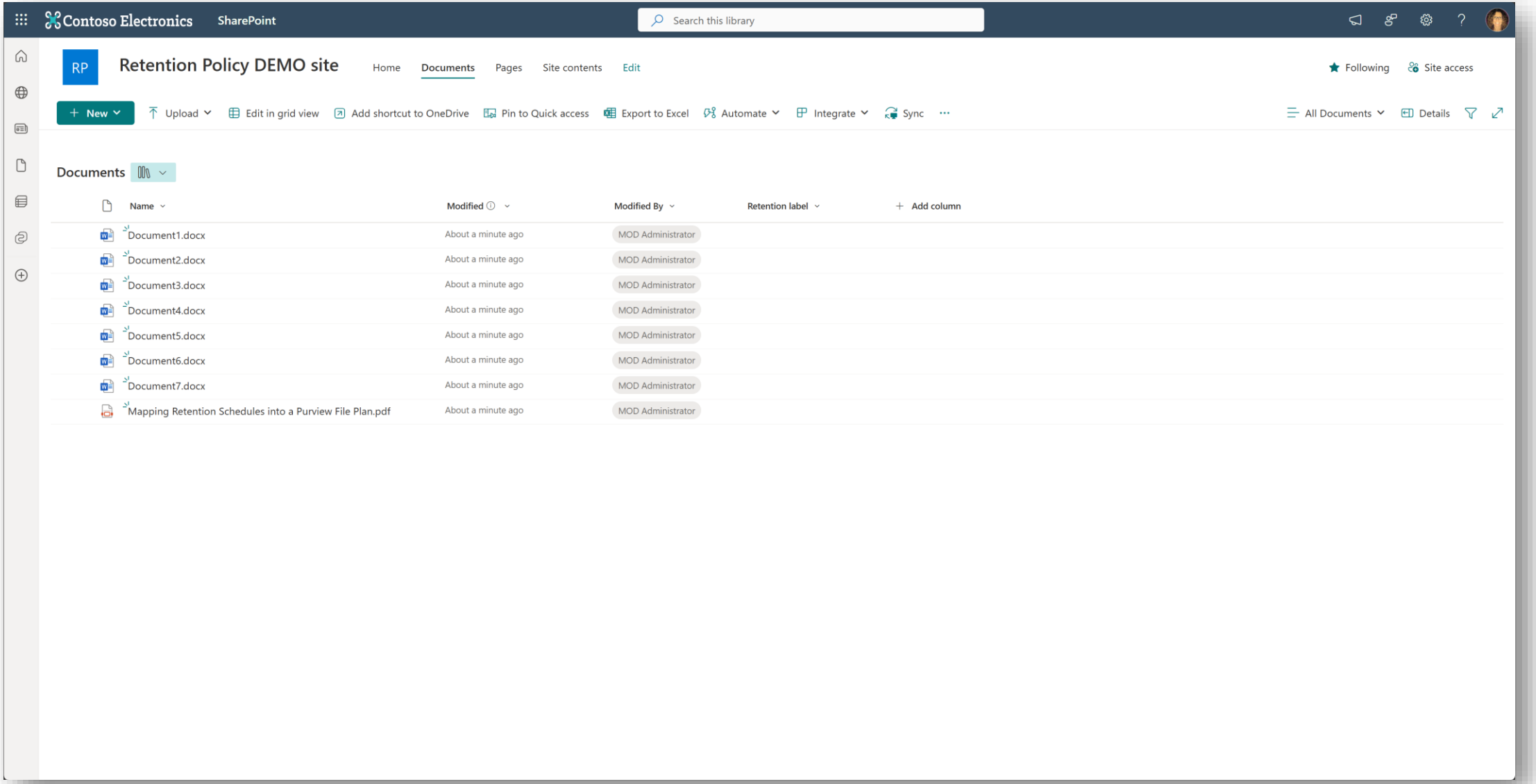
Data Lifecycle Management

- Implement retention policies to automatically retain and/or delete content (e.g., retain all executive's emails for 10 years, delete all Teams chats after 30 days)
- Tailor controls to government and industry-specific regulations such as CCPA, GDPR
- Dynamically apply retention policies based on a user's role
- Preserve content due to inadvertent or malicious deletes based on elevated user risk level

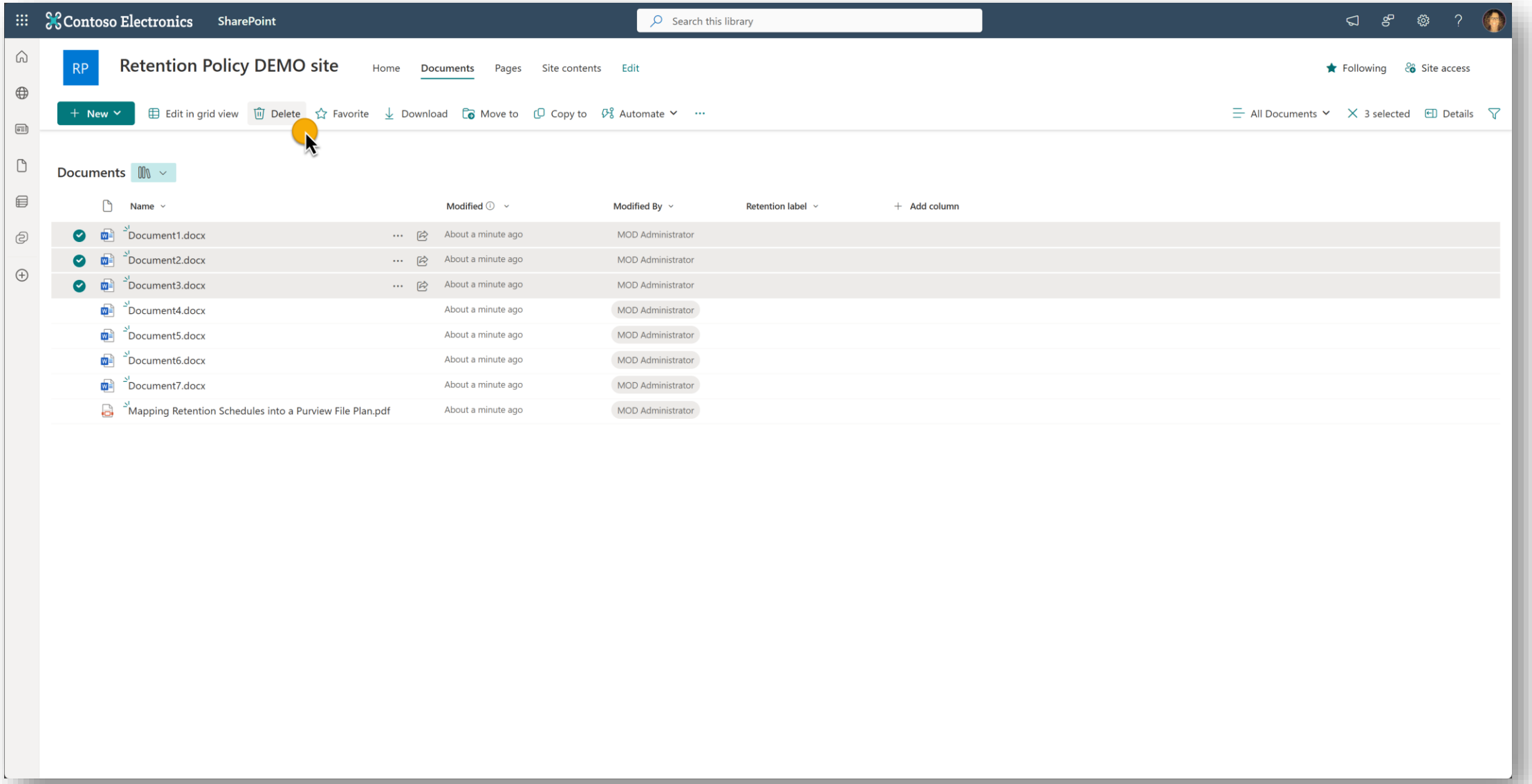
Microsoft Purview Data Lifecycle Management



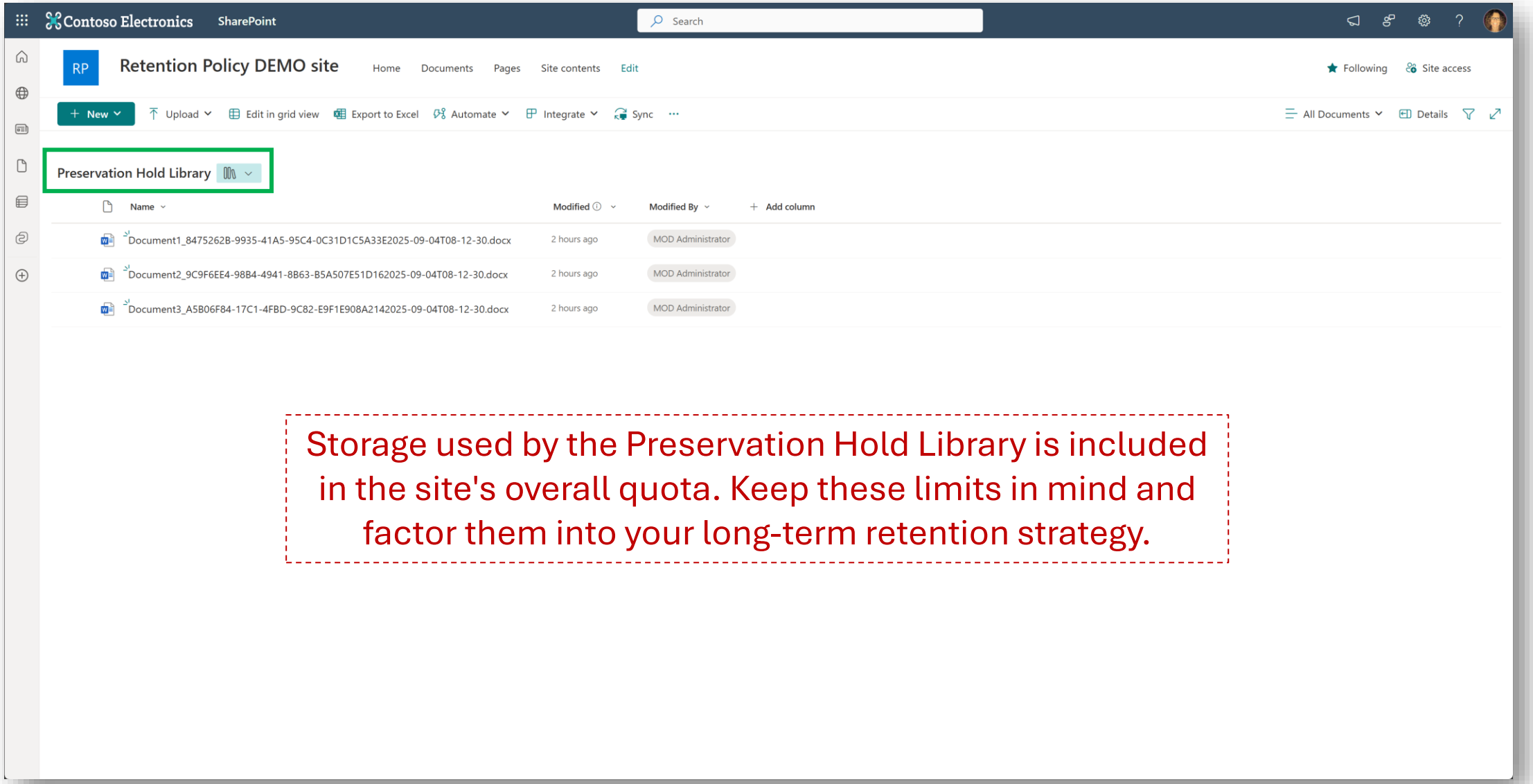
Microsoft Purview Data Lifecycle Management



Microsoft Purview Data Lifecycle Management

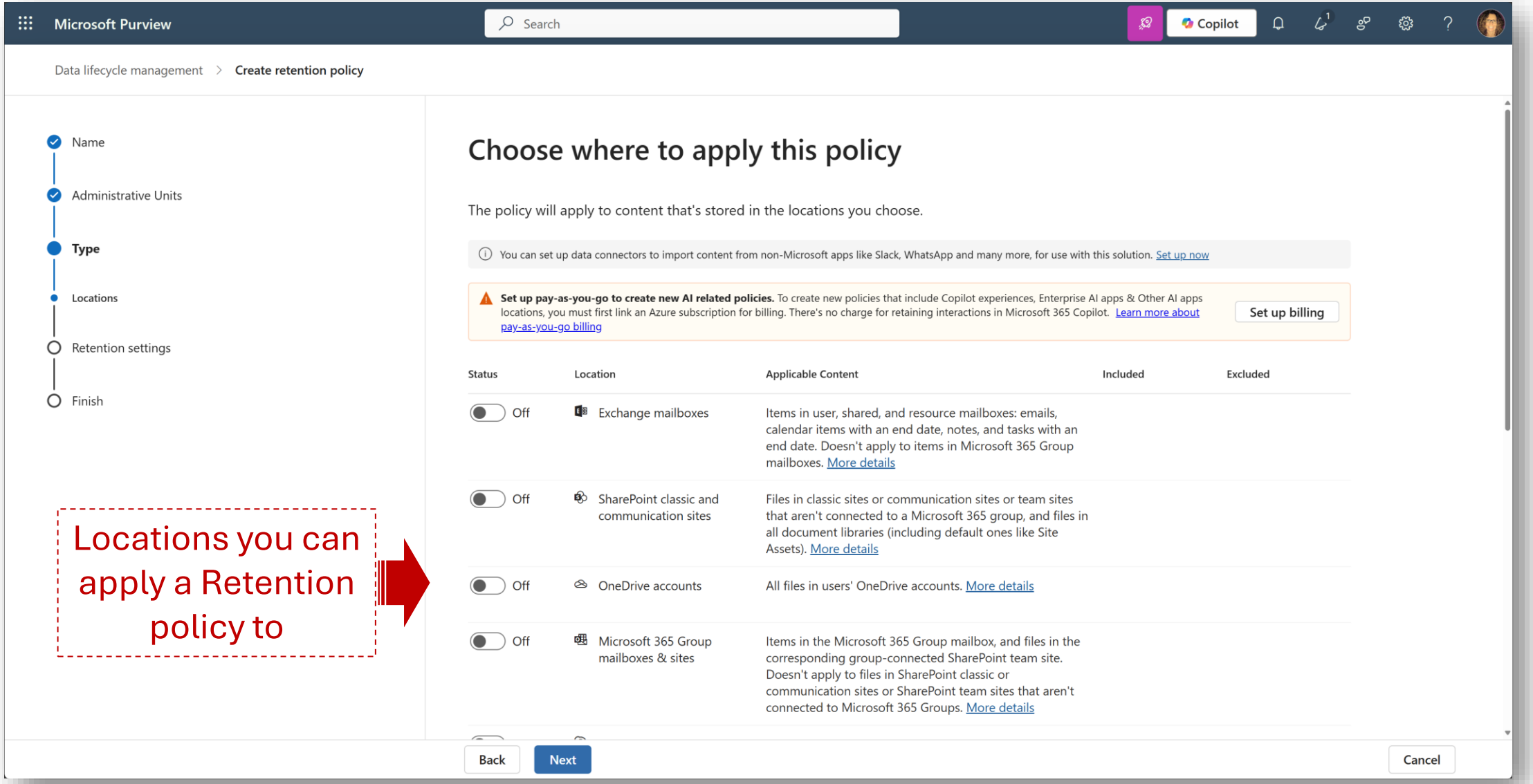


Microsoft Purview Data Lifecycle Management



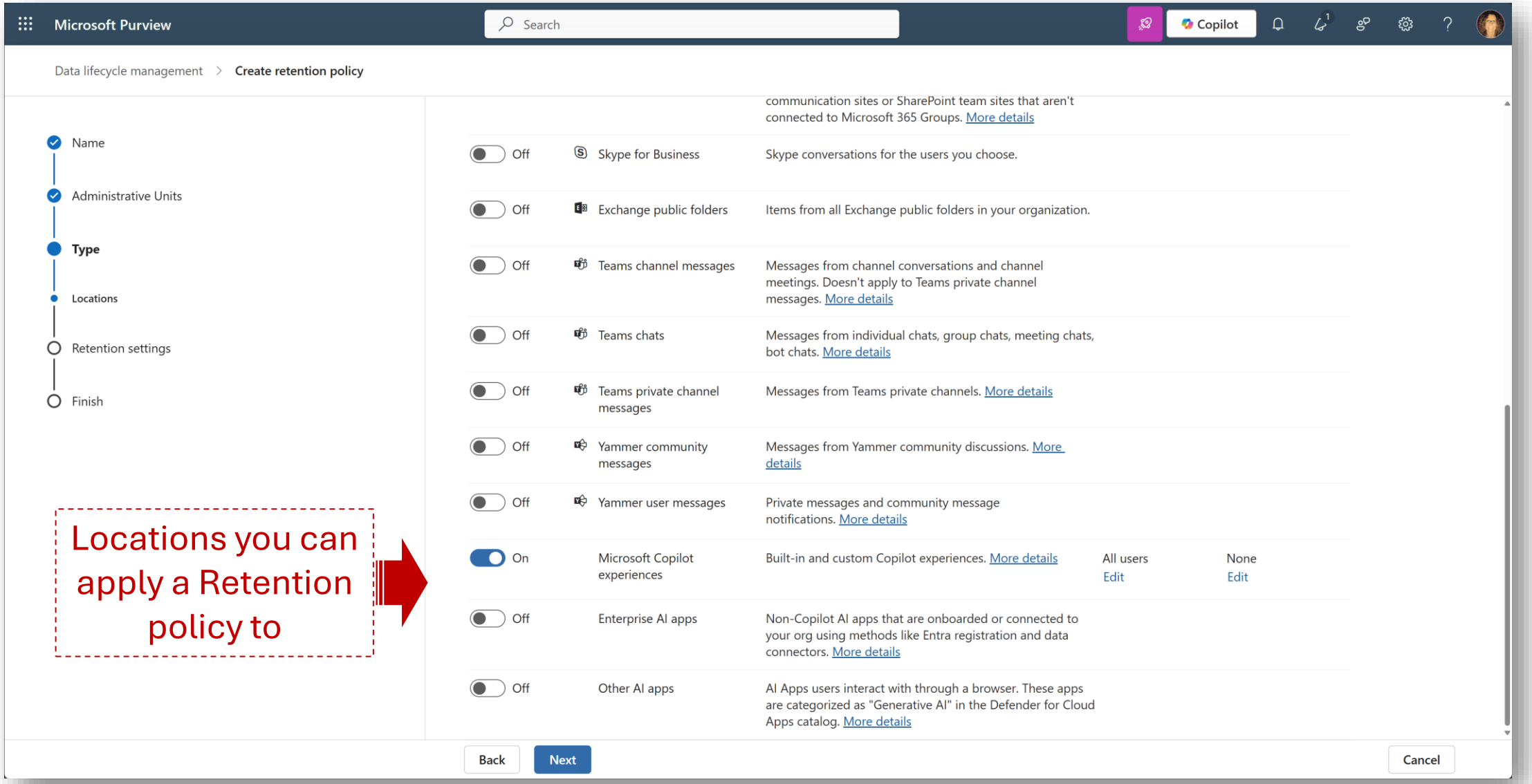
Storage used by the Preservation Hold Library is included in the site's overall quota. Keep these limits in mind and factor them into your long-term retention strategy.

Microsoft Purview Data Lifecycle Management

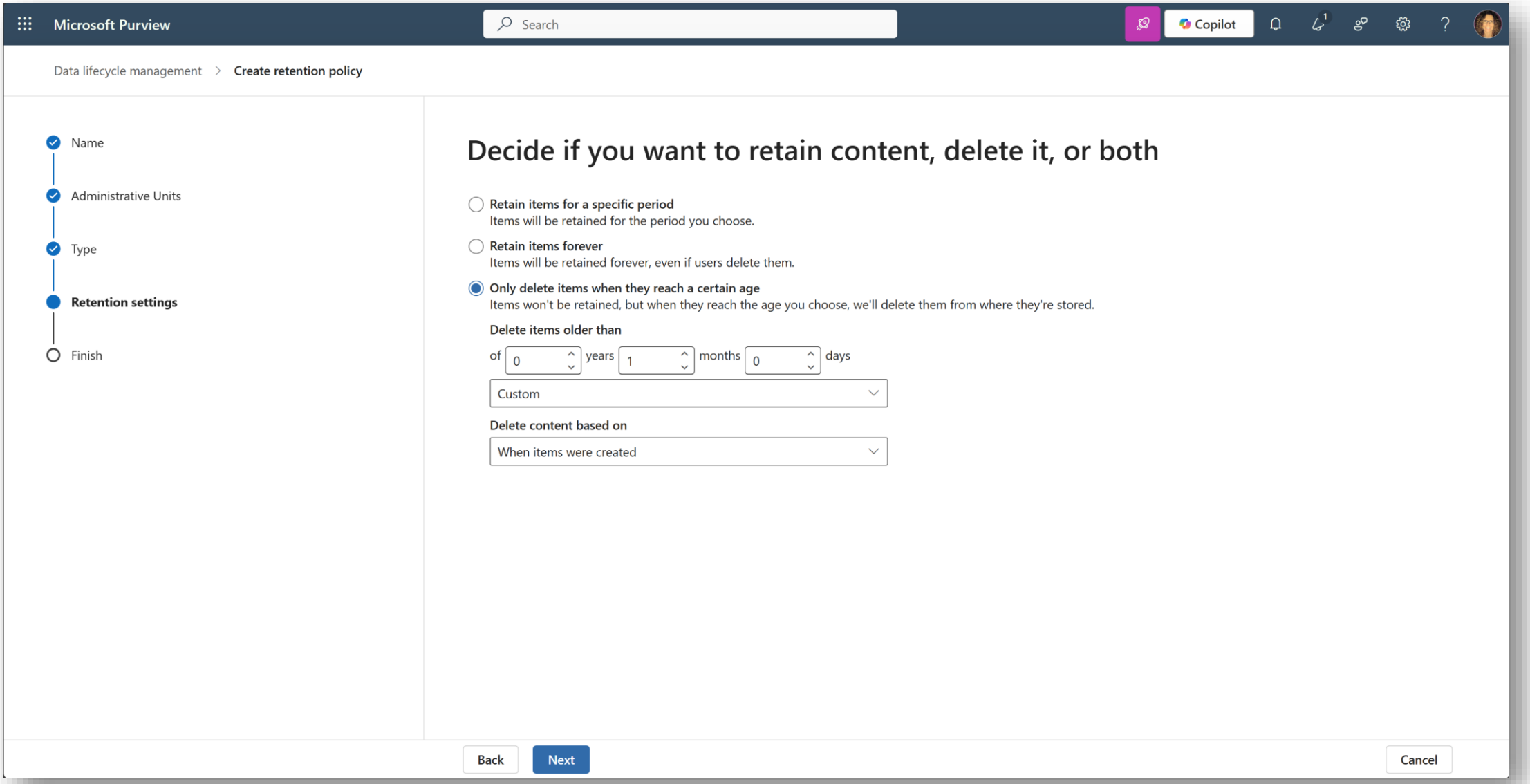


Locations you can apply a Retention policy to

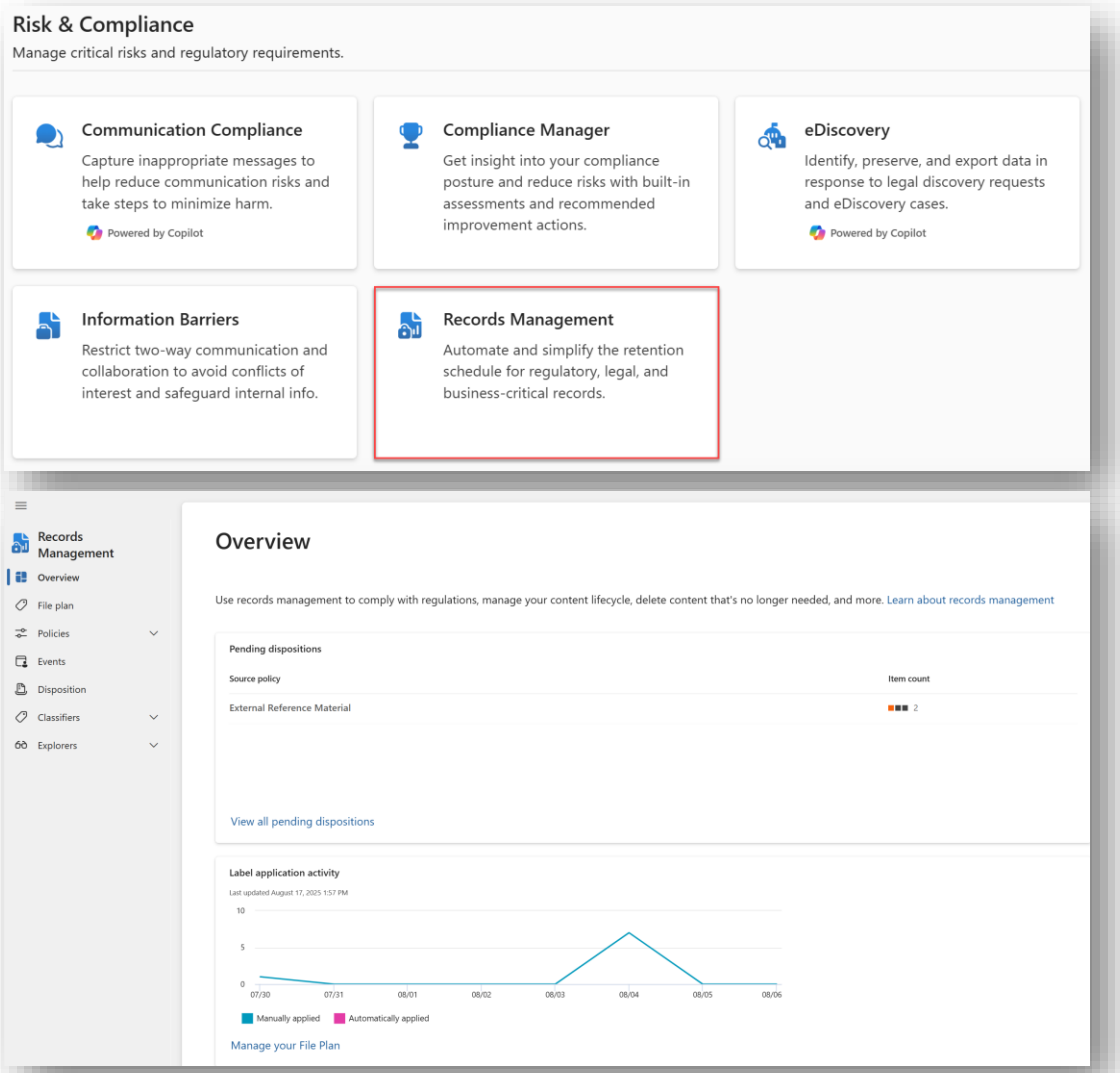
Microsoft Purview Data Lifecycle Management



Microsoft Purview Data Lifecycle Management



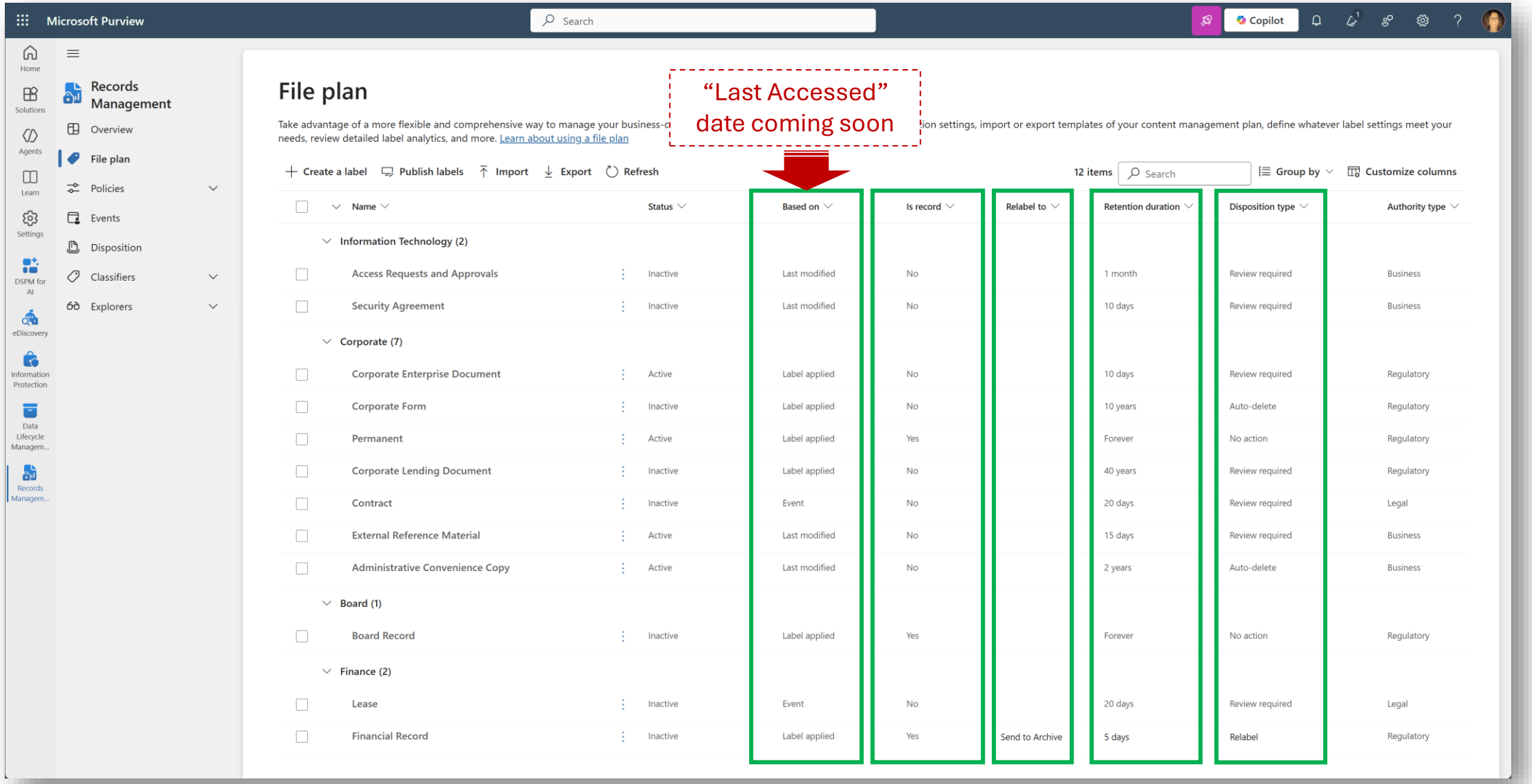
Records Management in Microsoft Purview



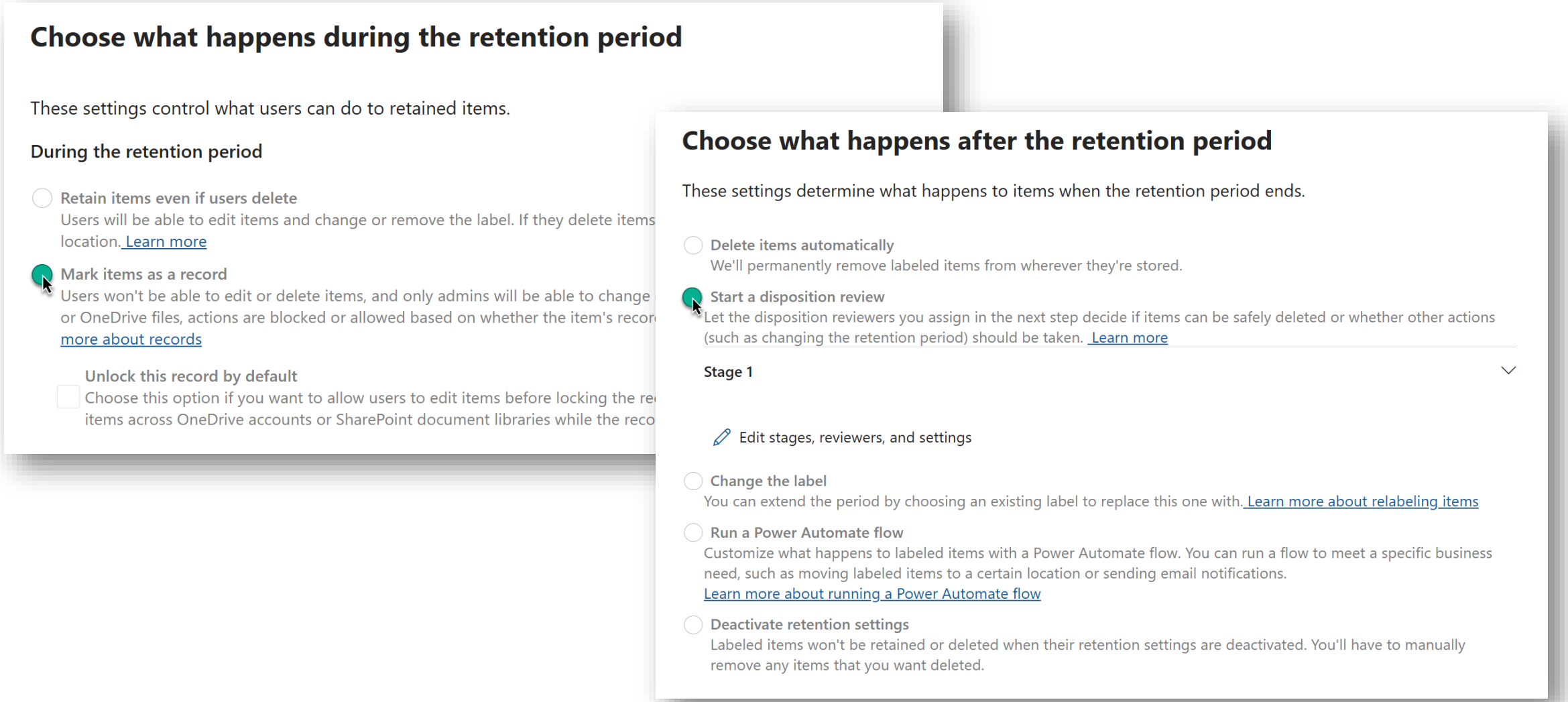
Records Management

- Implement retention labels in a file plan to identify business records (e.g., contracts, budgets, maintenance records, board meeting minutes)
- Labels provide regulatory-grade control such as immutability and defensible disposition
- Automatically apply labels based on location, metadata, or Purview classifiers

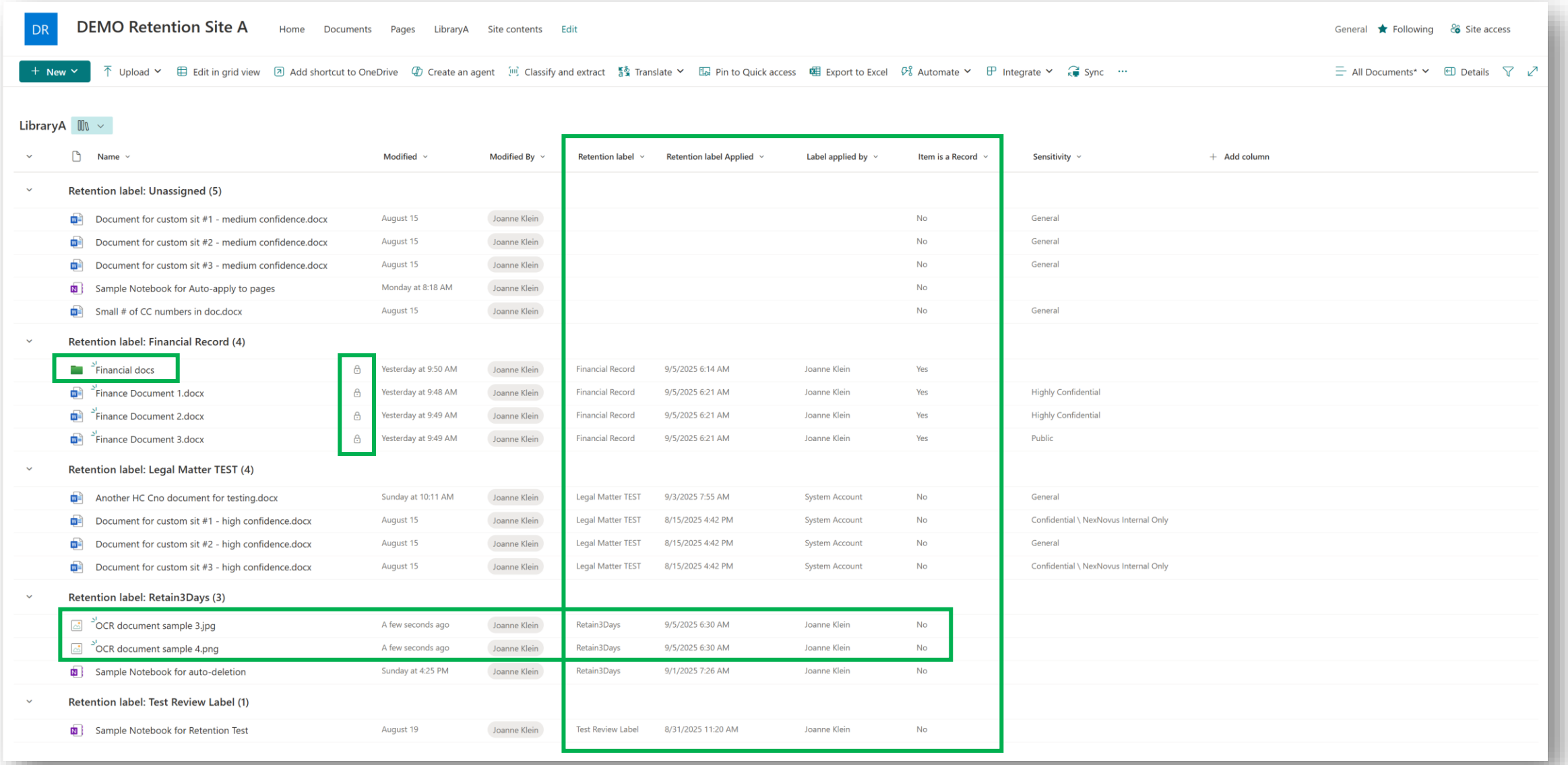
Microsoft Purview Records Management



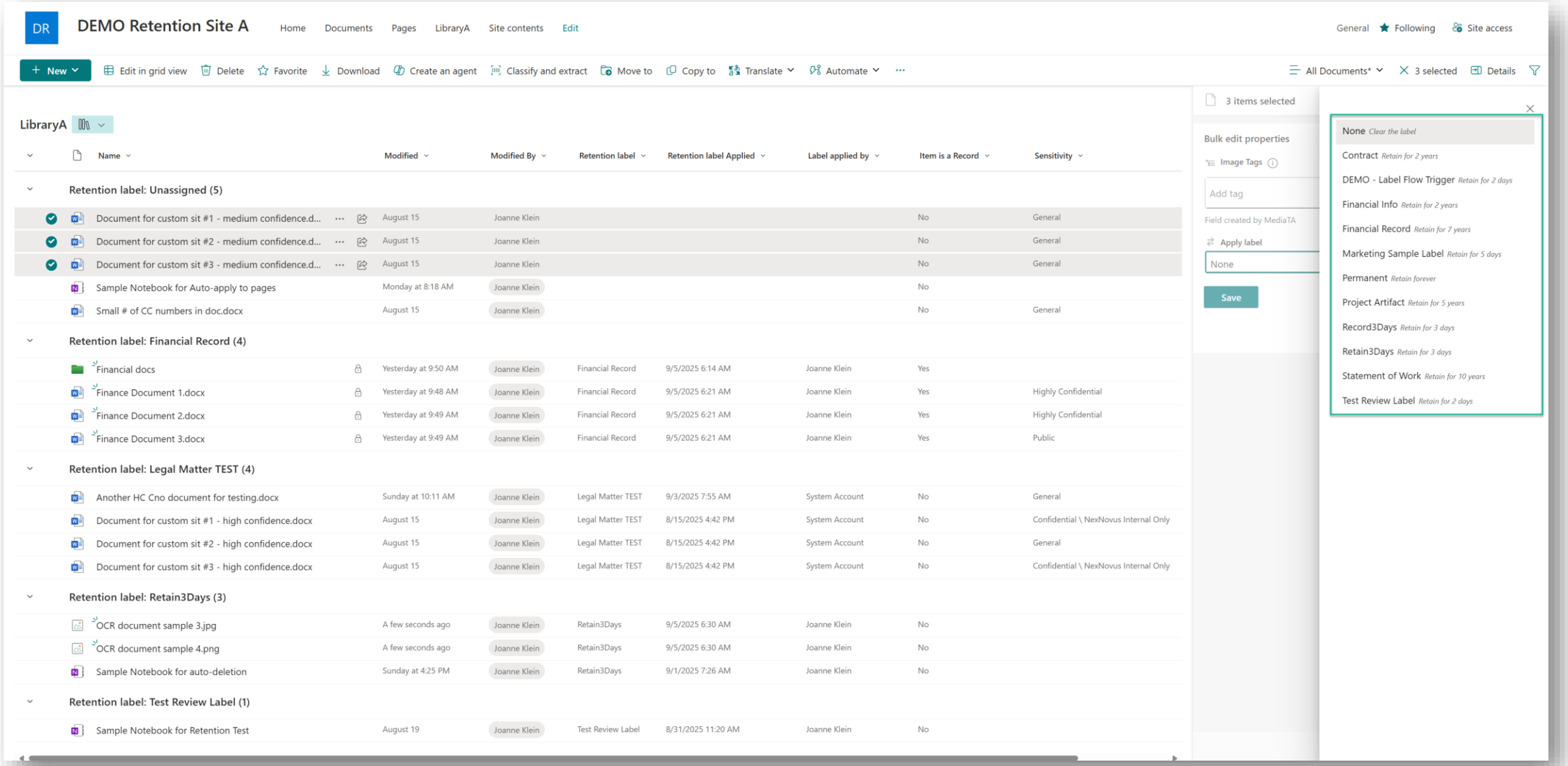
Microsoft Purview Records Management



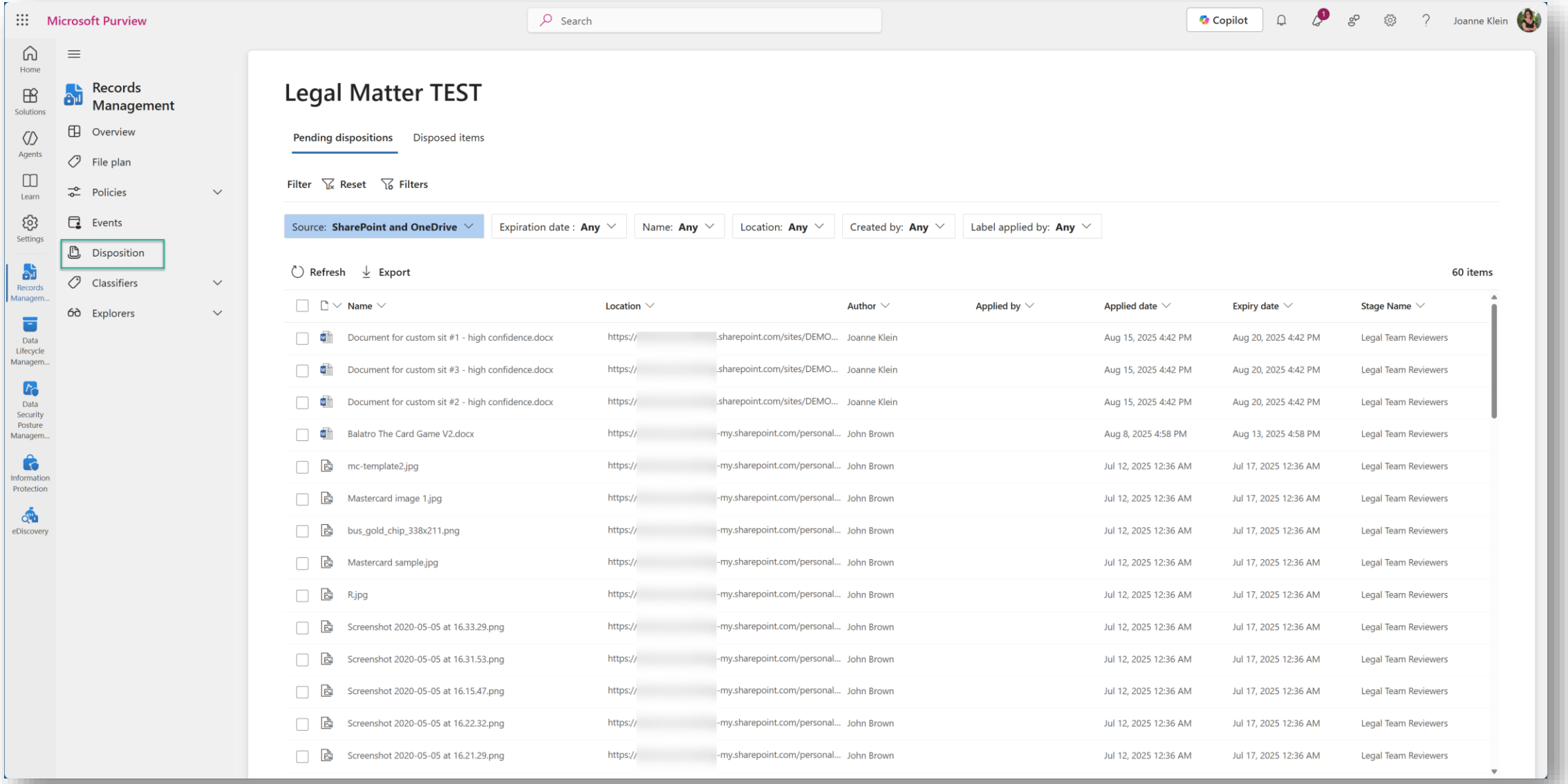
Microsoft Purview Records Management



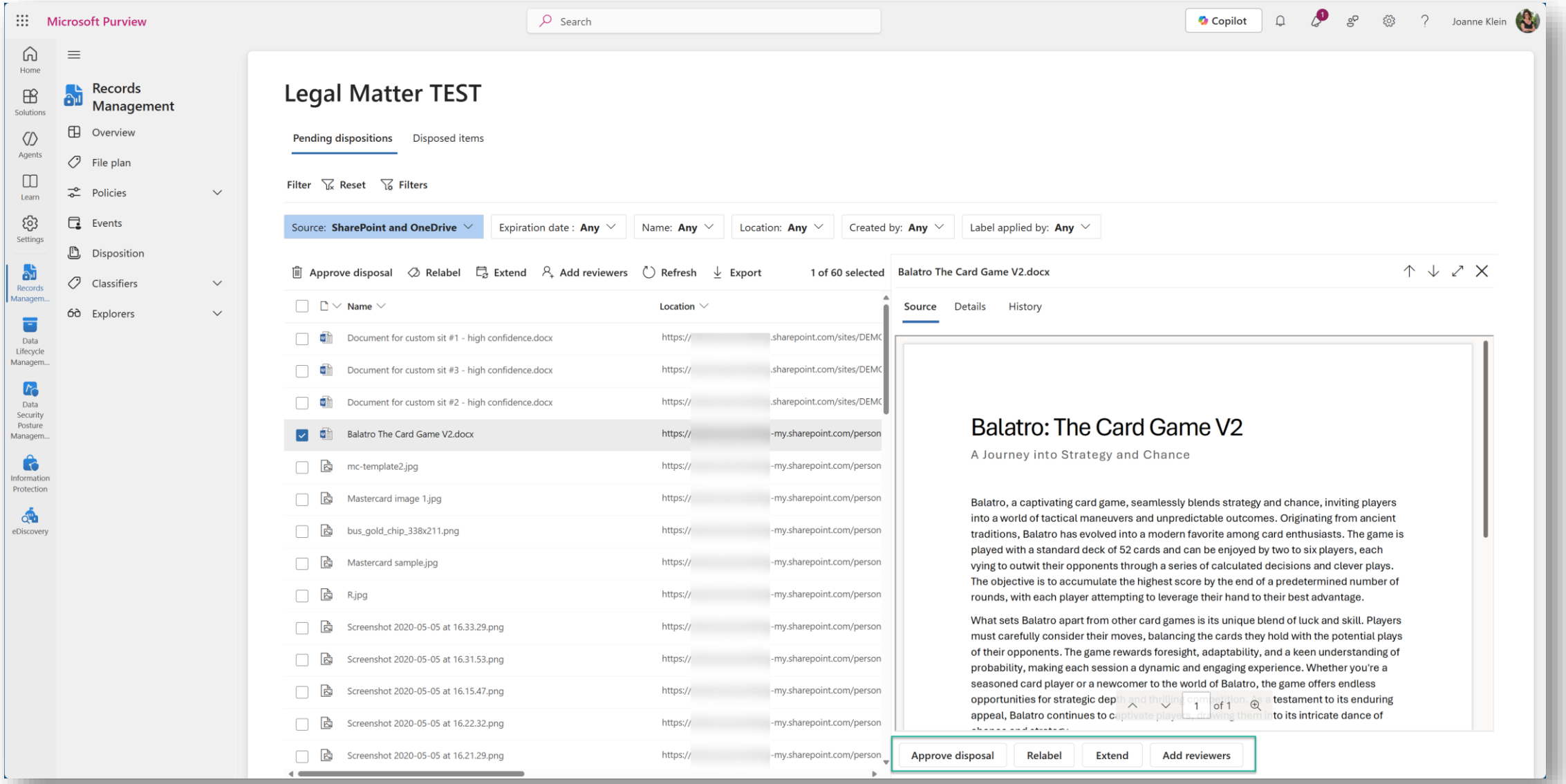
Microsoft Purview Records Management



Microsoft Purview Records Management



Microsoft Purview Records Management



Data Protection and Data Retention

How Does AI Affect Each of these?

AI and Data Protection … why it matters

Reason #1: Protect Sensitive Data from Unauthorized Access

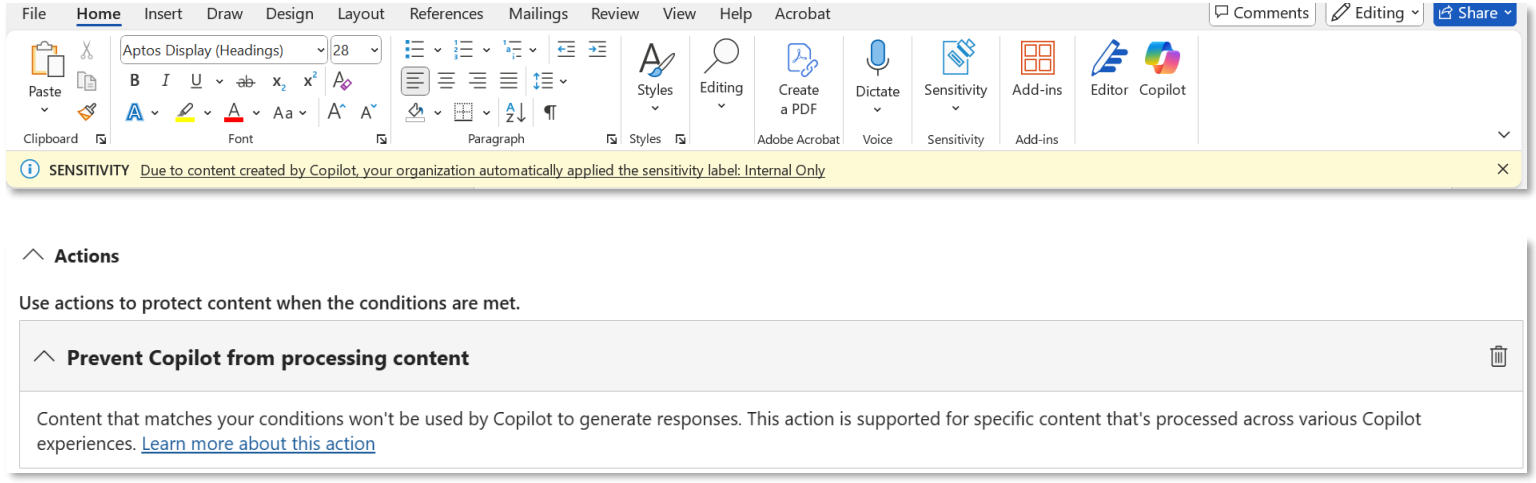
Data Protection controls can prevent unauthorized users from accessing, rapidly replicating, and oversharing sensitive data using AI, whether maliciously or inadvertently.

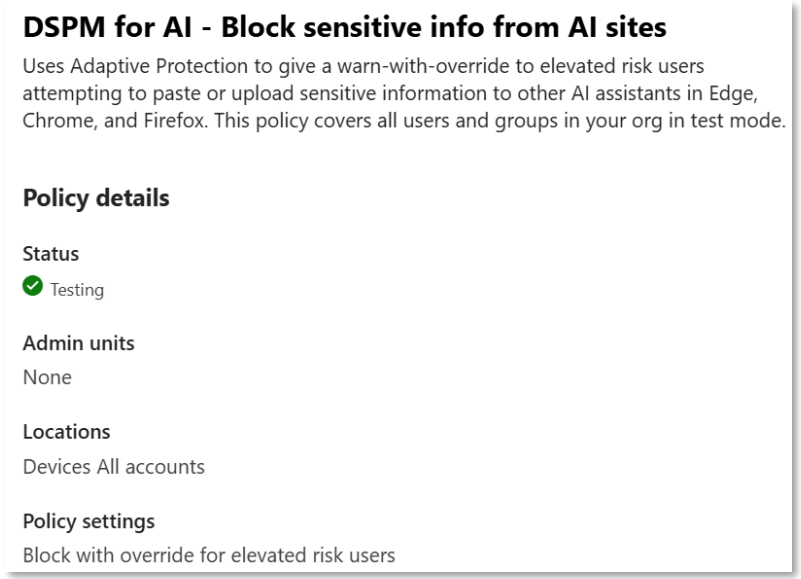
- Classify, label, and protect sensitive data

  E.g., “Prevent Microsoft 365 Copilot from accessing data labeled **Restricted**.”

- Implement policies to audit and block data loss

  E.g., “Block pasting of sensitive information into third-party AI applications.”





AI and Data Protection … why it matters

Reason #2: Monitor and Audit AI interactions to proactively identify data security risks

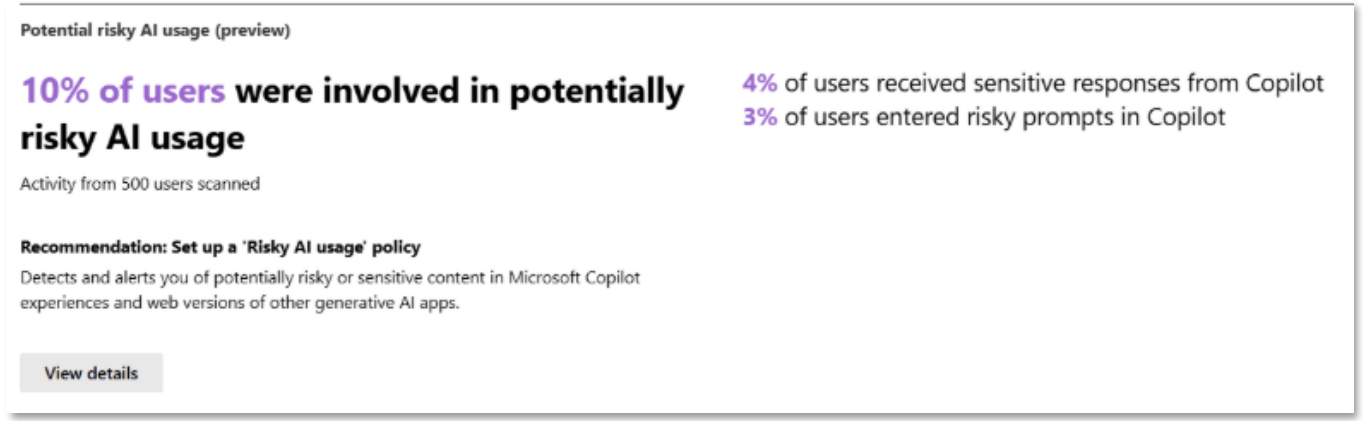
Monitoring and auditing AI interactions will help security administrators to proactively identify data protection policy gaps and remediate data security risks before a breach or leak takes place.

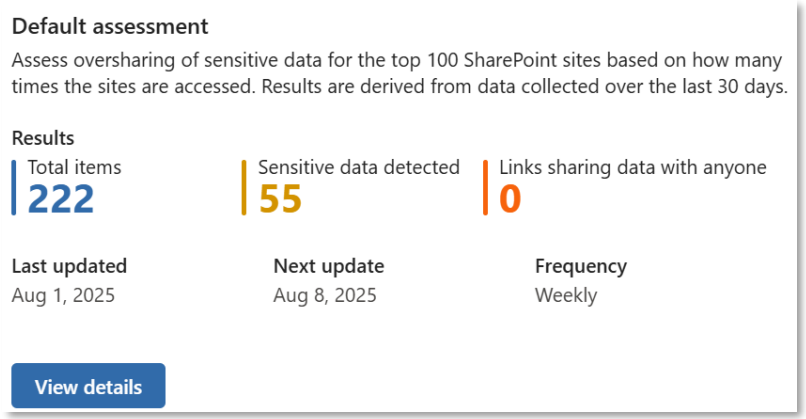
- Identify, investigate, and remediate risky AI interactions

  E.g., Identify a departing user entering prompts containing sensitive information in Copilot.

- Scan the environment for risks of data oversharing

  E.g., Leverage DSPM for AI data risk assessments, SharePoint Advanced Management data governance access reports, SharePoint site permissions and sharing report, or Microsoft Graph Data Connect for SharePoint.



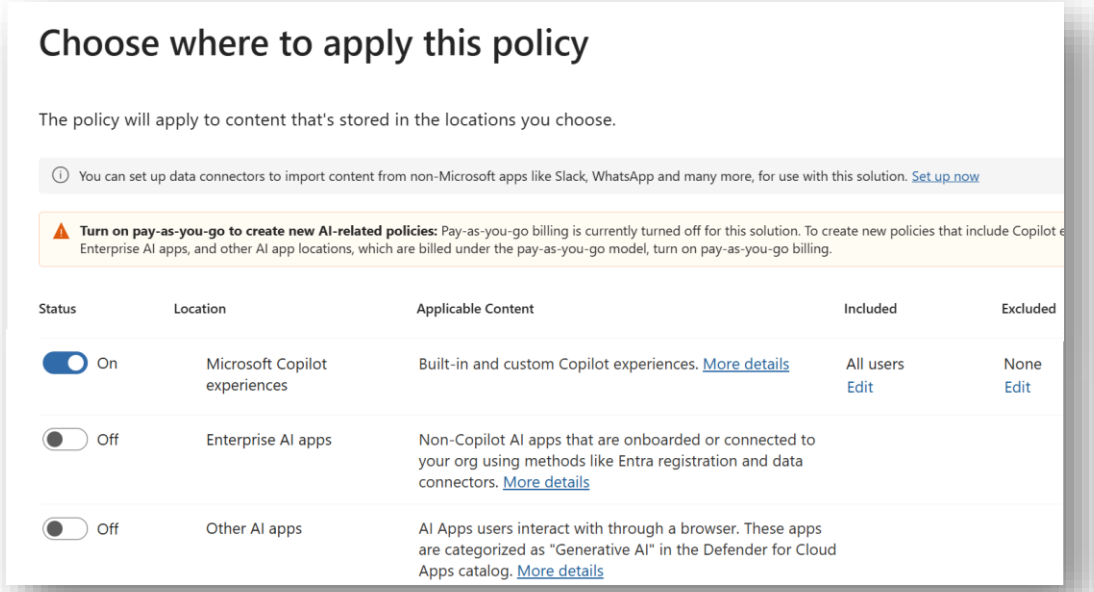


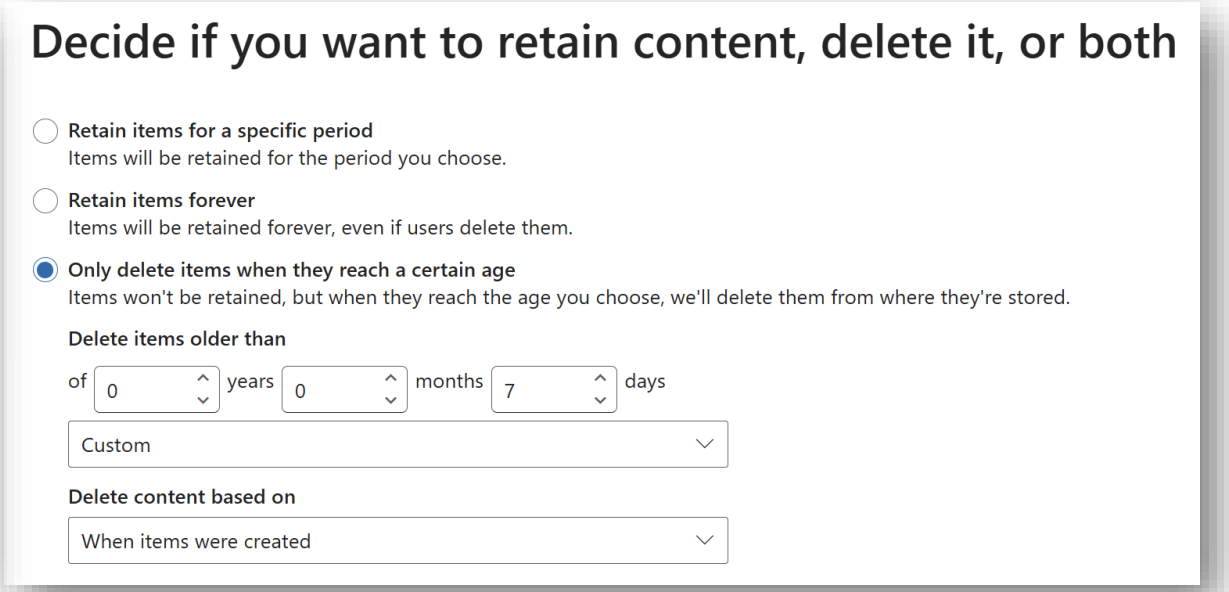
AI and Retention… why it matters

Reason #1: AI Prompts and Responses can be a compliance and discovery risk

You may want to apply retention controls for compliance or risk reasons and to remain compliant with AI regulations. If they exist, they're discoverable!

E.g., "Automatically **delete** Copilot prompts and responses after 7 days with a **retention policy**"





AI and Retention… why it matters

Reason #2: Retention and Deletion controls can improve Data Quality

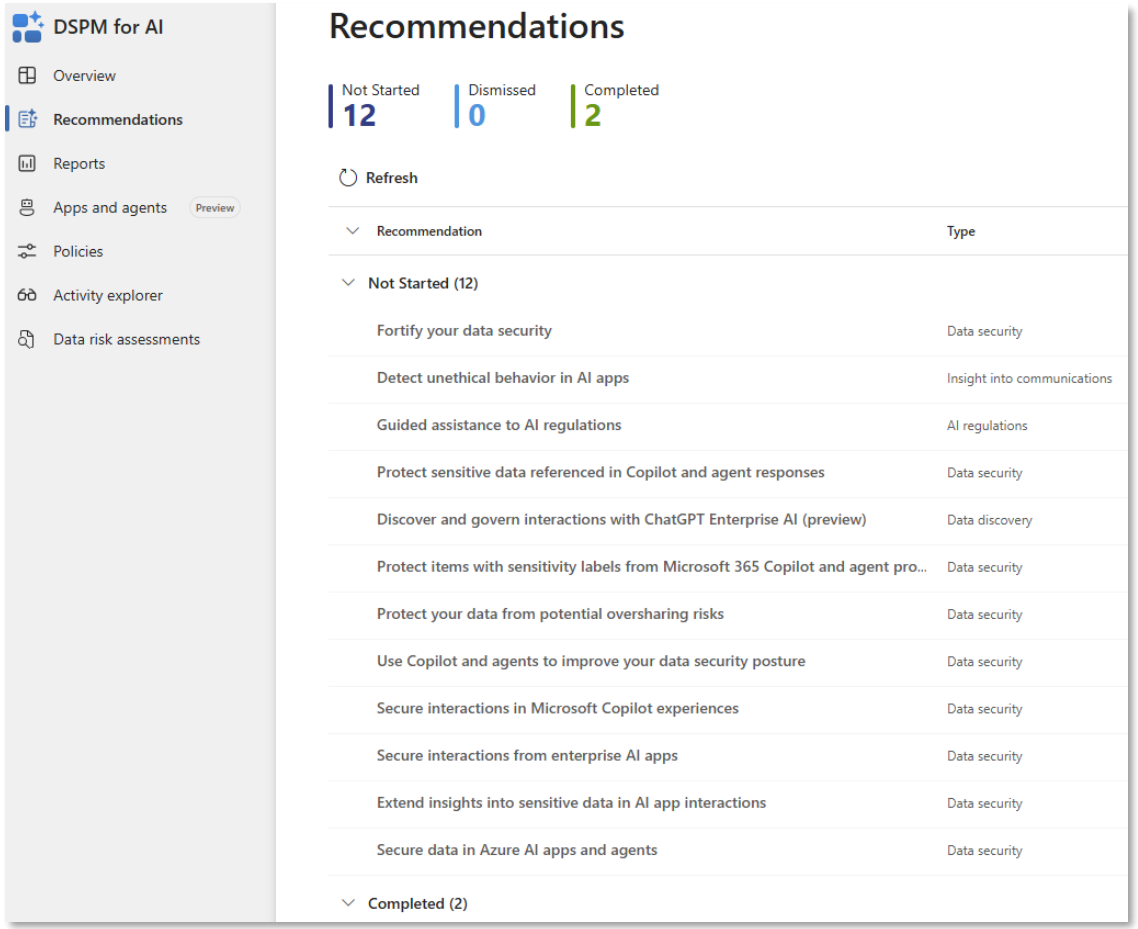
Retention controls can improve the quality of your data, so AI tools have better information to work with.

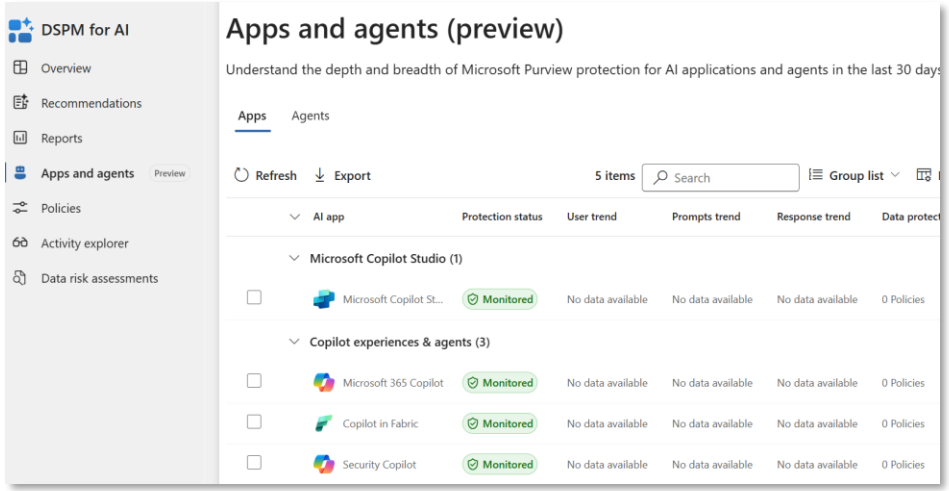
Combining these 2 retention controls can help with this:

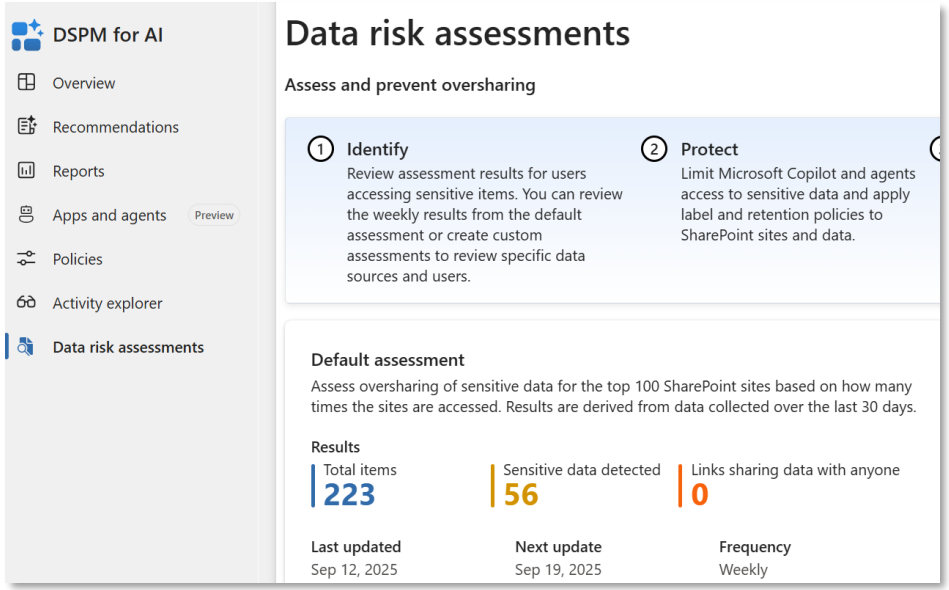
- **Retention policies**
  - Deleting ROT (**R**edundant, **O**bsolete, **T**rivial) content across your Microsoft 365 digital landscape with **retention policies**
  - Purview's Data Security Posture Management for AI (DSPM for AI) solution will suggest you create a retention policy for this very reason

- **Retention labels**
  - Ensuring you are retaining your records of legal, historical and business value with **retention labels** … "*Keep the gold, delete the rest*"

Data Security Posture Management for AI (DSPM for AI)







Data Protection and Data Retention

Practical Lessons We've Learned with Customers

Practical Lessons We've Learned About Data Protection

1. Make it easy for your users to label content
   - Minimize the number of sensitivity labels available for users to select
   - Follow a clear naming convention that aligns with the sensitivity of the data
   - Couple end-user facing changes with strong change management efforts

2. Follow a phased rollout approach, especially with blocking policies
   - Do not block actions until a detailed understanding of business case exceptions is gained
   - Allow for users to request required business case exceptions

3. Ensure alert volume remains manageable
   - A higher volume of alerts does not mean a more effective data protection program
   - Fine tune policies and automate alert response where possible to support analysts in focusing on high priority alerts

Practical Lessons We've Learned About Data Retention

1. Keep your Purview file plan as simple as your regulation will allow
   - Everything in your retention schedule doesn't need to become a Purview retention label!!
   - There are both technical and practical limits to the number of retention labels – be careful here

2. Have a well-governed SharePoint/Teams setup to reduce complexity
   - Good governance around provisioning, site ownership, content organization will make this easier as it gives Purview something to "automate against"
   - Most retention projects **also** end up being a "SharePoint cleanup" project because of this!

3. Perfection is the enemy of good (enough)
   - The (messy) reality of the Microsoft 365 collaborative environment is often at odds with records managers' expectations. Have mitigating controls in place (DLP, Audit, Adaptive protection) to alert/prevent premature deletion of data in case you don't have the right retention controls in place.
   - Defaulting retention labels as much as possible removes the burden on the end-user to apply a label or set a piece of metadata to trigger retention (the new **SharePoint Knowledge Agent** can help with automatic extraction and generation of metadata… and then apply a retention label based on it)

Data Protection and Data Retention

Common Questions we hear from Customers

Common questions we hear from customers

1. Is it best to do data protection or data retention first?

2. Can you do them at the same time?

3. What are the most challenging parts of each?

Closing Thoughts and Q&A

Thank you for joining us today!

Joanne C Klein

NexNovus M365 Consultant

joannecklein.com

https://www.linkedin.com/in/joannecklein/

Anna Bordioug

Protiviti M365 Consultant

www.annabordioug.com

https://www.linkedin.com/in/anna-bordioug/

Data Protection and Data Retention

Sensitivity Labels vs Retention Labels

Labels, Policies, and Label Policies

Data Protection

Sensitivity Labels

- Label files, emails, meetings, groups, and sites according to sensitivity.
- Enforce granular protection controls.
- Automatically label data in use.

Label Publishing Policies

- Publish sensitivity labels to users and groups.
- Require labeling, justification, and specify default label.

Auto-Labeling Policies

- Automatically label data at rest and data in transit based on the content of the item.

Data Retention

Retention Labels

- Apply retention controls to **files and emails**.
- Manage item-level retention, deletion, and disposition actions.

Retention Policies

- Apply retention controls to **containers**.
- Manage retention and deletion automatically.

Retention Label Policies

- Publish retention labels OR auto-apply a retention label to locations across Microsoft 365 (e.g., SharePoint, OneDrive, Exchange).

Contrast and Comparison of Label Behaviours (1 thru 10)

		PURVIEW LABEL TYPE COMPARISON	
	FUNCTIONALITY	SENSITIVITY LABEL[1] **Purpose**: applies data handling and protection controls	RETENTION LABEL **Purpose**: enforces retention and disposition controls to meet regulatory/legal/ business requirements
1	**TYPICAL VOLUME OF LABELS**	Sensitivity labels are created based on your data classification schema and are typically under ~10.	Retention labels are created from your retention schedule and can number in the tens and hundreds.
2	**CAN CHANGE THE LABEL NAME**	Yes. The display name can be changed after creation.	No
3	**CAN HAVE MULTLINGUAL LABEL NAMES**	Yes. Requires PowerShell.	No
4	**CAN MANUALLY APPLY THE LABEL**	Yes (only 1 file at a time)	Yes (1 a time and bulk)
5	**CAN DEFAULT A LABEL FOR A USER**	Yes (Different defaults can be set for files, emails, meetings, fabric, power bi)	No
6	**CAN DEFAULT A LABEL ON A DOCUMENT LIBRARY**	Yes (Applies only to new/edited **office files and pdfs** based on label priority) ***change 485732 coming soon*	Yes (Applies to **all** unlabeled items **regardless of filetype.** Can apply to either all items or only new/edited based on a setting)
7	**CAN DEFAULT A LABEL ON A SPECIFIC FOLDER**	No	Yes
8	**CAN RECOMMEND A LABEL TO A USER**	Yes	No
9	**CAN AUTO-APPLY A LABEL ON CLIENT-SIDE**	Yes	No
10	**CAN AUTO-APPLY A LABEL ON SERVICE-SIDE**	Yes (via auto-labeling policy; does NOT apply to Exchange)	Yes (via auto-apply label policy)

Contrast and Comparison of Label Behaviours (11 thru 20)

		PURVIEW LABEL TYPE COMPARISON	
	FUNCTIONALITY	SENSITIVITY LABEL[1] **Purpose**: applies data handling and protection controls	RETENTION LABEL **Purpose**: enforces retention and disposition controls to meet regulatory/legal/ business requirements
11	**CAN MAKE THE LABEL MANDATORY**	Yes	No
12	**CAN REQUIRE JUSTIFICATION TO CHANGE THE LABEL**	Yes	No
13	**CAN AUTOMATICALLY OVERWRITE A LABEL THAT'S ALREADY APPLIED**	Yes (There are rules based on taxonomy priority)	No with 1 exception (Only by changing 1 default label to another default label)
14	**CAN ASSIGN A LABEL PRIORITY**	Yes (Priority defined in label taxonomy; manual always wins)	No (First label to be applied to an item wins; manual always wins)
15	**LABEL REMAINS WITH CONTENT EVEN OUTSIDE OF TENANT**	Yes	No
16	**CAN SUPPORT A LABEL HEIRARCHY END-USERS SEE**	Yes (label groups and sublabels)	No
17	**CAN DELETE A LABELED ITEM**	Yes (assuming you have the usage rights in the label)	Yes (depends on label type (standard/record), user permission and tenant setting to allow for deletion
18	**WHERE LABELS CAN BE PUBLISHED TO**	Users, groups (distribution groups, mail-enabled security groups, Microsoft 365 groups)	Locations (exchange mailboxes, SharePoint sites, OneDrive accounts, Microsoft 365 group mailboxes and sites)
19	**COPILOT INTEGRATION**	Encrypting label without the EXTRACT and VIEW usage rights will prevent Copilot from summarizing the data.	Copilot interactions cannot be individually targeted for retention/deletion with a retention label – a retention policy is needed. Use labels to keep your records of value for Copilot to use.
20	**CAN LEVERAGE ADMINISTRATIVE UNITS**	YES	YES