Microsoft

Microsoft Learn
Spark possibility

Microsoft

Microsoft Purview Workshop

Date | 9 & 10 Januari 2023

Few words about me

Tobias Koprowski

Bachelor Degree in Banking
Higher national diplomas in: European Law & Corporate Governance
Three years in personal and home insurance
Five years in consumer & corporate banking
Ten years in physical Data Center
Microsoft Certified Trainer (MCT) & Educator (MCE)
CertNexus Authorized Instructor (CAI)
ISO 27001 Lead Auditor (PCA/TÜV Nord)

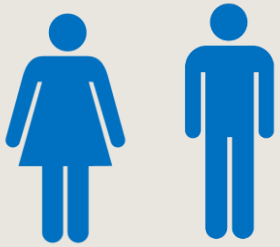
Member of:
| **BCS** (The Chartered Institute of IT)
| **IAPP** (International Association of Privacy Professionals)
| **ISSA** (Information Security System Association)
| **ISACA** (Information Systems Auditing & Control Association)
| **ISC²** (International Information System Security Certification Consortium)
| **CSA** (Cloud Security Alliance) – AI Usage Policy Working Group

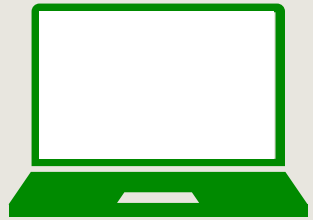
STEM Ambassador | Royal Voluntary Service

Social Media: KoprowskiT @ [TW|LI|BS|FB]

Facilities

- Agenda (next slide)
- Building hours
- Parking
- Restrooms
- Meals
- Phones
- Messages
- Internet access
- Recycling
- Emergency procedures



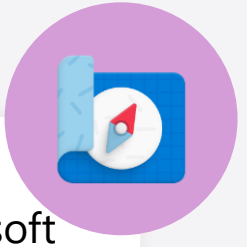


Agenda | Dag Ett | 9 Januari | Sales & Executive Track

Under denna dag kommer vi att fokusera på Microsoft Clouds: Microsoft 365 (på morgonen) och Azure (på eftermiddagen) ur perspektivet av chefer, säljare och personer inom icke-tekniska områden.

Dagen kommer att vara fylld med information som kan hjälpa organisationer att förstå, planera och implementera Data Governance och Information Protection.

Vi kommer att ha slides, korta demos, frågor, svar och diskussioner.



During this day, we will focus on Microsoft Clouds: Microsoft 365 (in the morning) and Azure (in the afternoon) from the perspective of executives, salespeople, and people in non-technical fields.

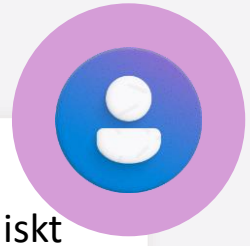
The day will be filled with information that can help organizations understand, plan, and implement Data Governance and Information Protection.

We will have slides, short demos, questions, answers and discussions.

Agenda | Dag Ett | 9 Januari | Sales & Executive Track

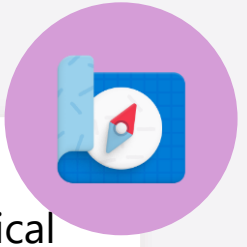
08:15 Välkomstfika & morgonbullar
08:30 **Keynote Microsoft** (30 min)
09:00 **Pass ett** | Overview, features, and requirements for Microsoft Purview [M365] *(60 min)*
10:00 Fika (rast)
10:15 **Pass två** | The Importance of Planning and Implementation of Information Governance *(90 min)*
11:45 Lunch
12:45 **Pass tre** | Zero Trust Approach – not only externally *(60 min)*
13:45 Fika (paus)
14:00 **Pass fyra** | Overview, features, and requirements for Microsoft Purview [Azure] *(60 min)*
15:00 Fika (rast)
15:15 **Pass fem** | Governance of the Digital Data Estate in the organization *(60 min)*
16:15 Sammanfattning, länkar, läxa efter eventet, Q&A och avslutning dag ett
16:30 AW på Microsofts kontor

Agenda | Dag Två | 10 Januari | Technical Dive



Under denna dag kommer vi att ha ett mer tekniskt tillvägagångssätt för Data Governance och Information Protection i Microsoft 365 (på morgonen) och Azure (på eftermiddagen).

Det blir lite färre slides och längre demos, med möjlighet att ställa frågor, få svar och diskutera alla aspekter av Microsoft Purview.



During this day, we will have a more technical approach to Data Governance and Information Protection in Microsoft 365 (in the morning) and Azure (in the afternoon).

There will be fewer slides and longer demos, with the opportunity to ask questions, get answers, and discuss all aspects of Microsoft Purview.

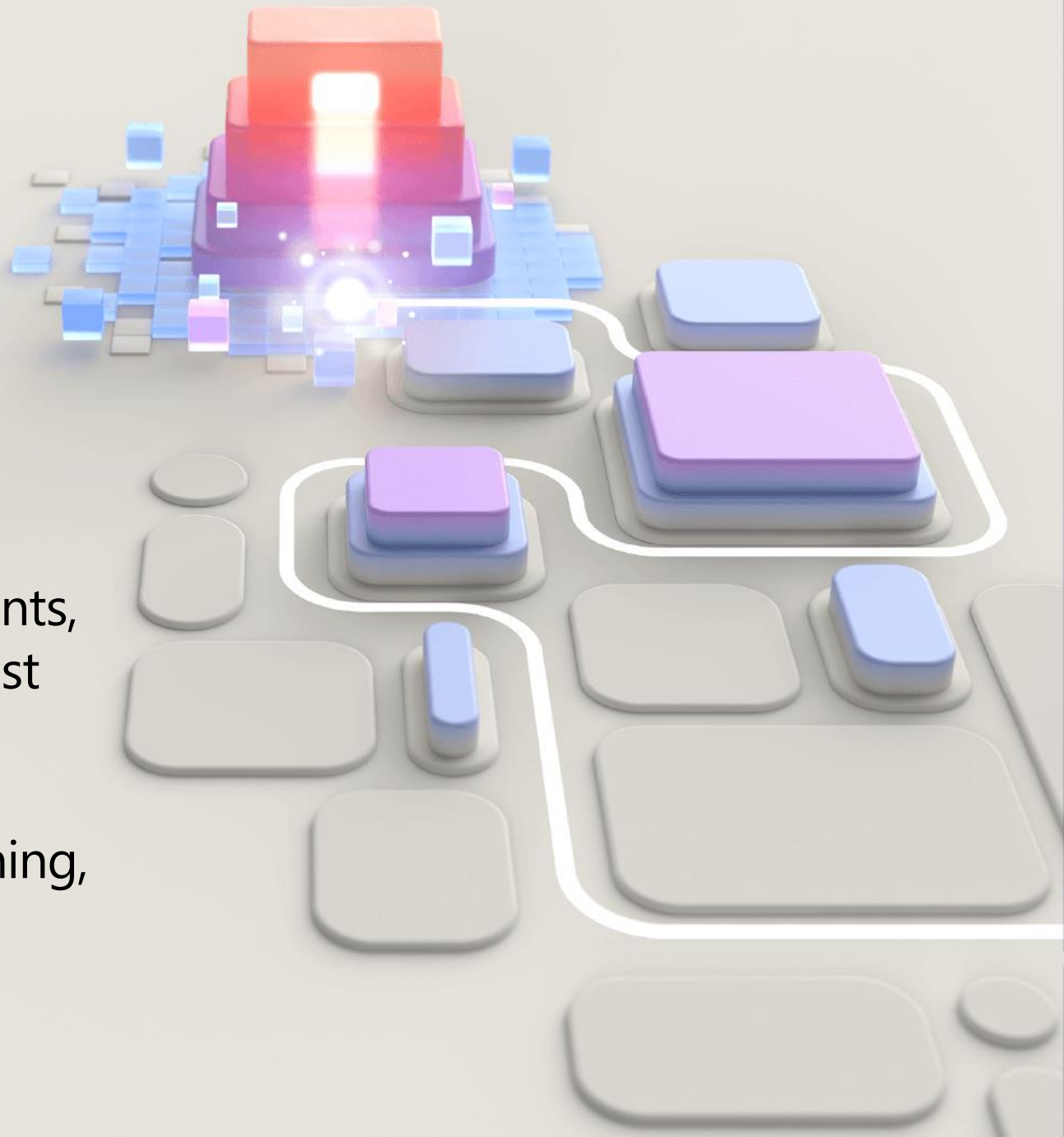
Agenda | Dag Två | 10 Januari | Technical Dive

08:15 Välkomstfika & morgonbullar
08:30 **Pass ett** | Microsoft Information Protection in Practice *(90 min)*
10:00 Fika (rast)
10:15 **Pass två** | Search and analysis of sensitive content *(90 min)*
11:45 Lunch
12:45 **Pass tre** | Diving into Microsoft Purview in Azure portal *(60 min)*
13:45 Fika (paus)
14:00 **Pass fyra** | Terminology, design, and execution of Microsoft Purview Part One *(60 min)*
15:00 Paus
15:15 **Pass fem** | Terminology, design, and execution of Microsoft Purview Part Two *(45 min)*
16:00 **Pass sex** | Getting hands dirty at home – do it yourself after the event *(15 min)*
16:15 Sammanfattning, länkar, läxa efter eventet, Q&A och avslutning dag två

Microsoft Security

Microsoft

Dag Ett | Day One | 09 Januari 2023



Förmiddagspass | Morning Sessions

Microsoft 365 Purview: features, requirements, planning, information governance, zero trust

Eftermiddagspass | Afternoon Sessions

Azure Purview: features, requirements, planing, digital data estate

JUST IN CASE, JUST IN CASE...

Data Privacy and Governance in The Kingdom of Sweden

Sensitive Information in Sweden

National Legislation

General data protection laws

The General Data Protection Regulation (EU) (2016/679) ("GDPR").

In Sweden, the *GDPR* is supplemented by the National Implementation Law (Lag 2018:218 of 19th of April 2018, **Lag (2018:218) med kompletterande bestämmelser till EU:s dataskyddsförordning (dataskyddslagen)** | not legaly binding in English)

Lag (2018:218) med kompletterande bestämmelser till EU:s dataskyddsförordning | Sveriges riksdag (riksdagen.se)

government.se/contentassets/467ef1335aac404c8840c29f9d02305a/act-containing-supplementary-provisions-to-the-eu-general-data-protection-regulation-sfs-2018218/

Entry into force

The *GDPR* has applied since 25 May 2018.

The Act containing supplementary provisions to the EU entered into force at the same time as the *GDPR*, on 25 May 2018.

Scope of the Swedish Data Protection Authority’s work

General Data Protection Regulations

The Camera Surveillance Act
The Credit Information Act
The Patient Data Act
The Debt Recovery Act

Sensitive Information in Sweden

National Supervisory Authority

Details of the competent national supervisory authority

Datainspektionen founded in 1973 with the uintroduction of *Datalagen* (Data Act) in same year.

On 1 January 2021, the name was changed to ***Integritetsskyddsmyndigheten ("IMY")***.

Swedish Data Protection Authority

Visiting address: Fleminggatan 14, 7th Floor, Stockholm

Postal address: Integritetsskyddsmyndigheten, Box 8114, 104 20 Stockholm, Sweden

E-mail: imy@imy.se

Switchboard: +46 (0)8 657 61 00| | https://www.imy.se/

Types of legal compliance requirements

There are a range of different laws that you must be aware of in relation to the various different roles that you may undertake.

We will take a look at the following today:

- HIPAA
- SOX
- GDPR
- NISD
- E-Privacy Directive

hipaa

Health Insurance Portability and Accountability Act is a series of American Federal Regulatory standards that outlines the use and disclosure of protected health information. This is a living culture that health care organisations must implement to protect the privacy, security and integrity of protected health information.

There are a range of difference consequences for violating HIPAA, starting at $100 per violation.

hipaa

There are two different possible options for a violation of HIPAA.

Civil penalties

- Starting at **$100** per violation
- Where there have been multiple violations of the same type, the fine can rise to **$25,000**
- Civil penalties will only apply when the individual is aware that they are breaching the rules. When there is no wilful intent, and the violation was corrected within 30 days, civil penalties will not apply

hipaa

- Criminal Penalties
  - These will be much more severe
  - The minimum fine for wilful breach of HIPAA is **$50,000**, but the maximum fine for an individual is **$250,000**
  - However, in the case of criminal penalties, there is also likely to be a prison sentence involved. These can start from **one** year where a person is deemed negligent. Obtaining information under false pretences can lead to up to **five** years. Knowingly violating the rules with malicious intent can lead to up to **ten** years
  - There is also a mandatory two year term for aggravated identity theft

SOX

SOX - Sarbanes Oxley is a US law meant to protect investors from fraudulent activities by corporations. It is designed to ensure that financial disclosures are made by organisations to prevent accounting fraud.

Whilst this law does not explicitly require anything of the IT industry, it does have a significant impact on the system that stores information, and the security of the information.

SOX

The consequences of non-compliance with SOX can be severe and the SEC (Securities and Exchange Commission) who enforce this law can impose the following:

- Fines
- Freezing back accounts
- Permanent bans for people being directors
- Removing companies from the stock exchange
- Invalidating insurance policies

Fines can be very high:

Knowingly certifying a report that doesn't meet guidelines can lead to a fine of **$1 million**, **10 years** in prison – or both

Willingly certifying a report that doesn't meet guidelines can lead to a fine of **$5 million**, **20 years** in prison – or both

Companies that discriminate against whistle blowers under this law are also subject to civil penalties.

gdpr

- GDPR (**General Data Protection Regulations**) - whilst there are two different GDPR laws (Europe and UK) we will focus on the European law here.

- GDPR is a European directive which covers both data privacy and data security around companies all over the world. It affects any organisations who hold data regarding European residents, whether the company is in Europe or not.

gdpr

GDPR sets out a number of principles:

- Lawfulness, Fairness and Transparency
- Purpose Limitation
- Data Minimisation
- Accuracy
- Storage Limitation
- Integrity and Confidentiality (Security)
- Accountability

gdpr

- The consequences of breaching GDPR can be significant and can include some of the following:
  - Issuing warnings and reprimands
  - Ordering rectification or destruction of data
  - Suspending their ability to transfer data to third countries
  - Imposing bans on data processing

- However, they also can lead to fines. The maximum could be **€20 million** (about **£18 million**) or **4%** of annual global turnover – whichever is greater.

nisd

The **National Information Systems Directive** is a European wide directive that identifies the baseline level of security requirements for network and information systems to ensure the continuity of essential services.

The NIS Directive has three main parts:

National Capabilities – EU Member states must have certain national cyber security capabilities

Cross-border Collaboration – collaboration between EU countries

National supervision of critical sectors – EU Member states must supervise the cybersecurity of critical market operators

nisd

There are a range of penalties that are set by each individual EU member state. The UK has adopted the following:

Penalty	Type of Contravention
Up to £1,000,000	Any material contravention which we determine could not cause an incident, such as a failure to comply with an Information Notice or lack of co-operation with an inspection.
Up to £3,400,000	Any material contravention which we determine has caused, or could cause, an incident leading to a reduction in the provision of your service.
Up to £8,500,000	Any material contravention which we determine has caused, or could cause, an incident leading to the disruption of your service.
Up to £17,000,000	Any material contravention which we determine has caused, or could cause, an incident that results in a threat to life or in significant adverse impact on the UK economy.

E-privacy directive

This 2002 ePrivacy Directive is an important legal instrument for privacy in the digital age, and more specifically, the confidentiality of communications and the rules regarding tracking and monitoring. Being an EU directive, it is up to each individual EU member state to bring it into law within their country.

The UK has implemented this as the Privacy and Electronic Communications Regulations, also known as PECR.

Examples of industry compliance requirements

There are a number of different compliance requirements that have been introduced for a variety of reasons, three of which we will understand are below:

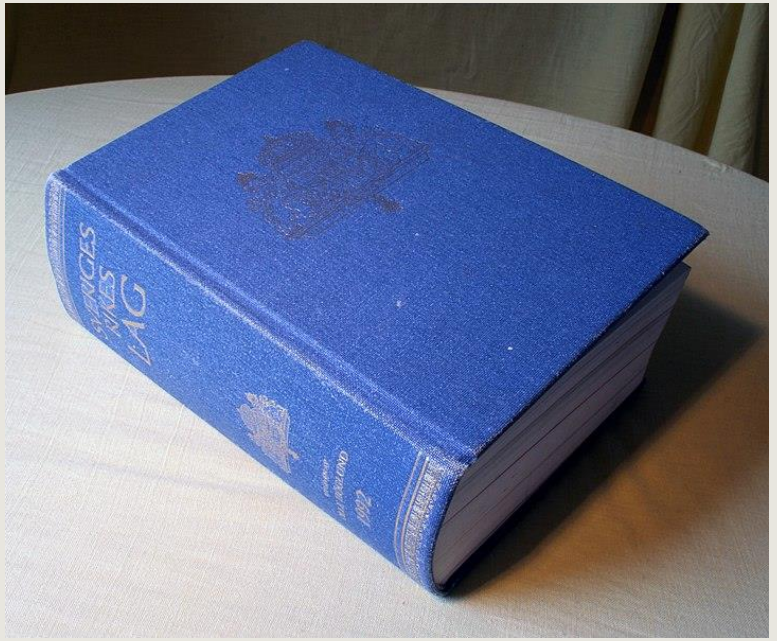
PCI DSS (Payment Card Industry Data Security Standard) – this specifies the expectations for any organisation that is holding cardholder data

ISO 27001 – this is an international standard that sets expectation levels for how an organisation manages their information security

NIST (National Institution of Standards and Technology) – this is a non-regulatory US government agency who develop standards that aim to push organisations to achieve the best possible standards

EXTRAS: Few examples about legal acts around cloud computing in Sweden

The **Swedish Code of Statutes** (Swedish: *Svensk författningssamling*, Swedish law collection; **SFS**)
contains the chronological session laws of the Riksdag,
regulations of the Government,
and ordinances, collectively called *författning*.



Sveriges rikes lag, the *de facto* statute book, containing a selection of current laws from the SFS

EXTRAS: Few examples about legal acts around cloud computing in Sweden

I. Data Protection Act || **SFS 2018:1218**
II. NIS Directive 2016 adapted in 2018 || **SFS 2018:1174**
III. Electronic Communication Act || **SFS 2003:38** & Electronic Commerce Act || **SFS 2002:562**
IV. Public Access To Information And Secrecy Act || **SFS 2009:400**
V. Archives Act || **SFS 1990:782** + Administrative Procedure Act || **SFS 1971:291** + Public Procurement Act || **SFS 2016:1149**+ Patient Data Act || **SFS 2008:395**
VI. Security Protection Act || **SFS 1996:627** replaced by **SFS 2018:589**
VII. Banking And Finance Business Act || **SFS 2008:297**
VIII. Bancrupcy Act || **SFS 1987:672** + Enforcement Act || **SFS 1981:174**
IX. Distance and Off-Premises Contracts Act || **SFS 2005:59** + Consumer Contracts Act || **SFS 1994:1512** + Electronic Commerce Act || **SFS 2002:562** + Consumenr Sales Act || **SFS 1990:932** + Consumer Services Act || **SFS 1985:716**

PASS ETT | SESSION ONE

Overview, features, and requirements for Microsoft Purview [M365]

Here Should be embedded video...

https://go.microsoft.com/fwlink/p/?linkid=2236827&clcid=0x809&culture=en-gb&country=gb

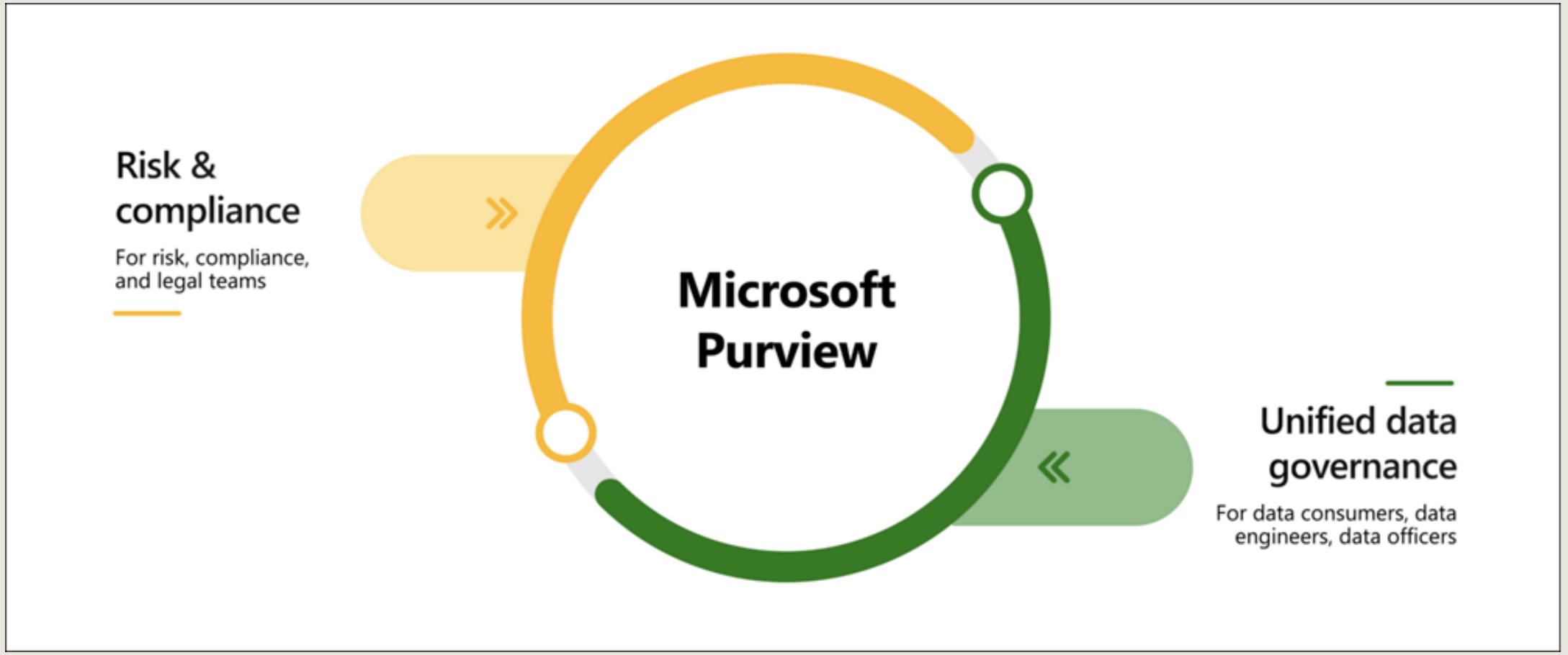
What is Microsoft Purview

Microsoft Purview is a family of data governance, risk, and compliance solutions that can help your organization govern, protect, and manage your entire data estate. Microsoft Purview solutions provide integrated coverage and help address the recent increases in remote user connectivity, the fragmentation of data across organizations, and the blurring of traditional IT management roles.

Microsoft Purview combines the former **Azure Purview** and **Microsoft 365 compliance** solutions and services together into a single brand. Together, these solutions help your organization to:

- Gain visibility into data assets across your organization
- Enable access to your data, security, and risk solutions
- Safeguard and manage sensitive data across clouds, apps, and endpoints
- Manage end-to-end data risks and regulatory compliance
- Empower your organization to govern, protect, and manage data in new, comprehensive ways

What is Microsoft Purview



Microsoft Purview risk and compliance solutions

Microsoft Purview includes **risk and compliance solutions** that support services included in Microsoft 365. These services include Microsoft Teams, SharePoint, OneDrive, Exchange, and others.

These compliance and risk solutions help your organization to:

Protect sensitive data across clouds, apps, and devices

Identify data risks and manage regulatory compliance requirements

Get started with regulatory compliance

Microsoft Purview unified data governance solutions

Microsoft Purview includes **unified data governance solutions** that help you manage data services across your on-premises, multi-cloud, and software-as-a-service (SaaS) estate. That includes Azure storage services, Power BI, databases like SQL or Hive, file services like Amazon S3, and many more.

These governance solutions are accessible through the Microsoft Purview governance portal, which provides tools to enable your organization to:

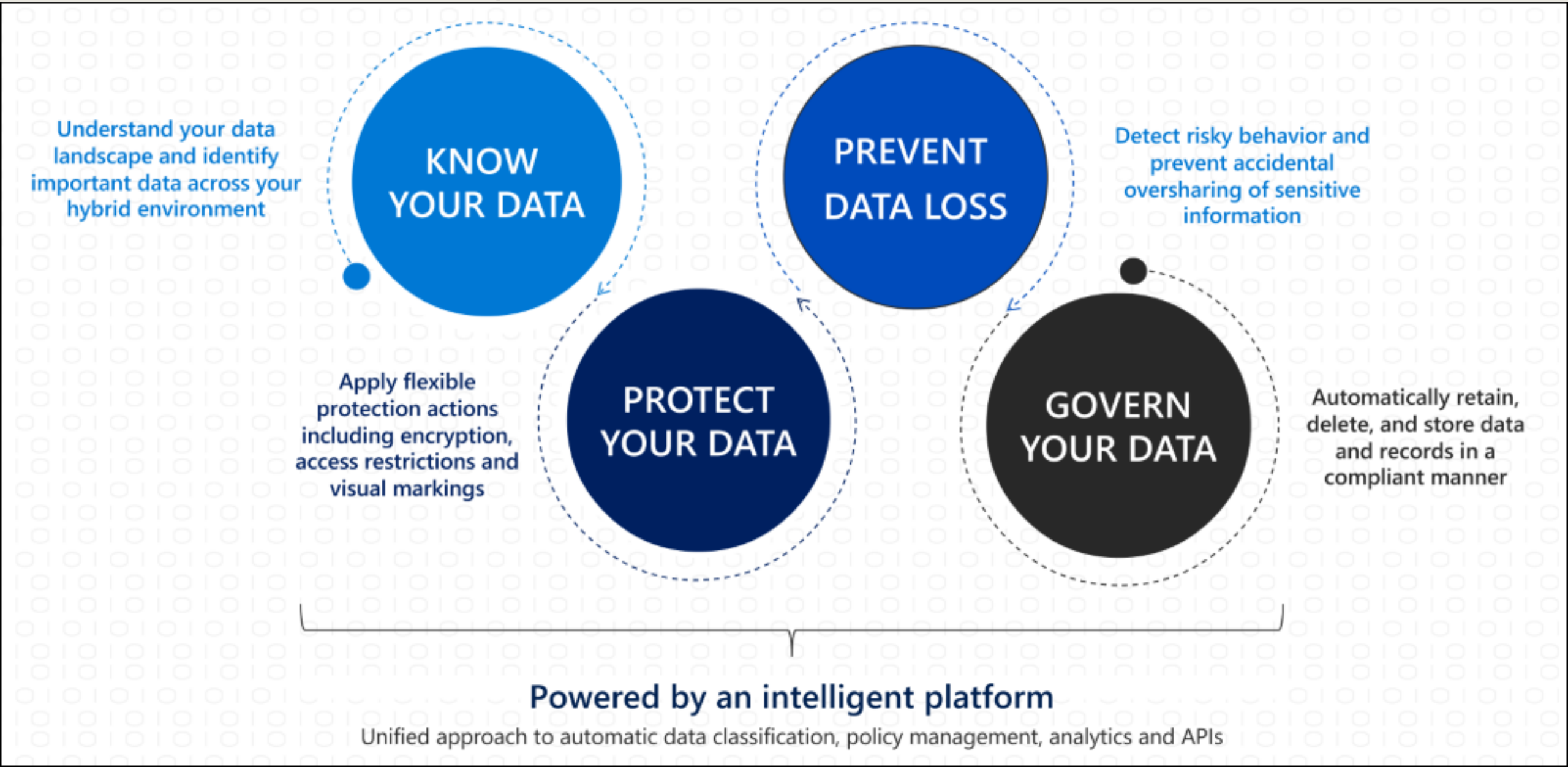
Create an up-to-date map of your entire data estate that includes data classification and end-to-end lineage

Identify where sensitive data is stored in your estate

Create a secure environment for data consumers to find valuable data

Generate insights about how your data is stored and used

Introduction to information protection



Know your data

Data classification concepts, apply one or more of the following to your data:

Sensitive information types

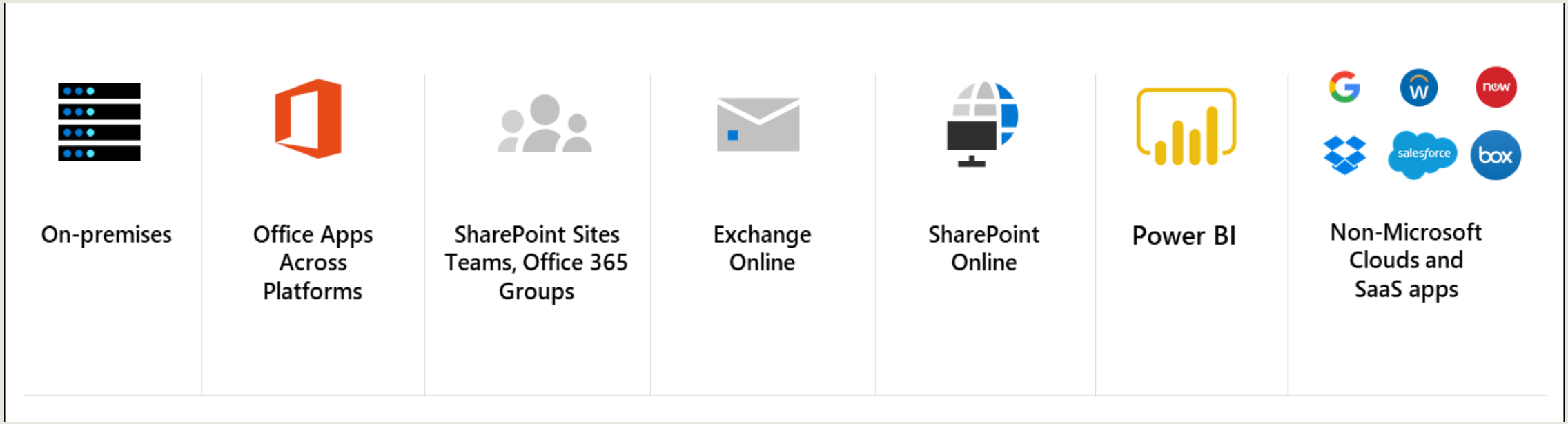
Trainable classifiers

Labels

Policies

Protect your data

Information protection is integrated into Microsoft 365 apps and services like:



Prevent data loss

Secure Data

- Enforce conditional access to sensitive data
- DLP action to block sharing
- Encrypt files and emails based on sensitivity label
- Prevent data leakage through DLP policies based on sensitivity label
- Business data separation on devices
- Secure email with encryption & permissions

Enable productivity

- Manually apply sensitivity label consistently across applications and endpoints
- Show recommendations and tooltips for sensitivity labels with auto-labeling and DLP
- Visual markings to indicate sensitive documents across apps and services (like watermarks, lock icons, sensitivity column in SharePoint Online)
- Co-author and collaborate with sensitive documents
- Enable searching of encrypted files in SharePoint
- Allow users to open and share encrypted PDF files in Edge in addition to Adobe Acrobat Reader

Govern your data

Themes of the Microsoft solutions for governing your data:

Streamlined administration

Automation at scale

Tailored workflows

Data classification overview

Microsoft 365 data classification components:

Overview

Provides snapshots of how sensitive information types and labels are being used.

Content explorer

Explore the email and documents in your organization that contain sensitive information or have labels applied.

Activity explorer

Review activity related to content containing sensitive info or has labels applied, such as what labels were changed, files were modified, and more.

Sensitive info types

Manage the built-in and custom sensitive information types available to classify data.

Trainable classifiers

Manage the classifiers used to identify content based on what the item is, not by the elements in the item.

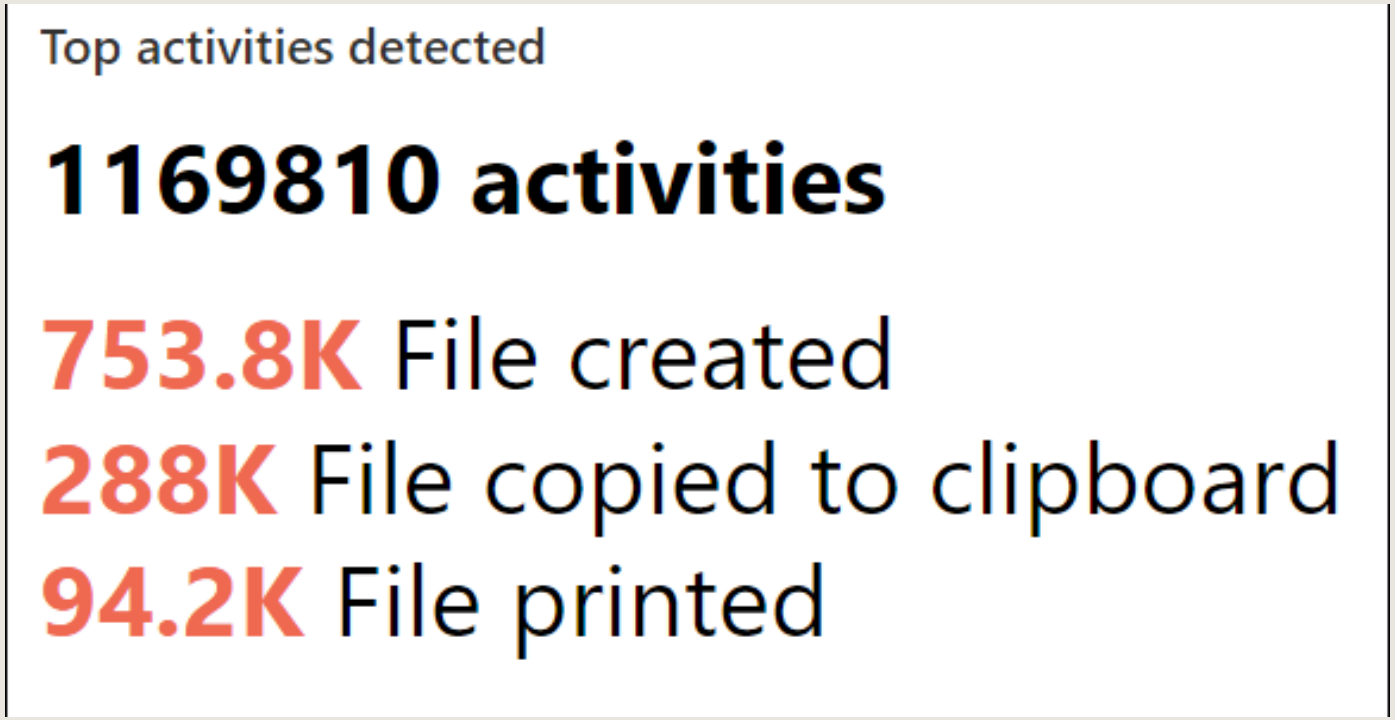
Review sensitive information and label usage

The Overview page can answer questions like:

What sensitive data is out there?

What labels are being used the most?

Is sensitive data being copied or shared outside the organization?



Classify data using sensitive information types

Information protection and governance solutions and where in those solutions you can use sensitive information types:

Information protection: Sensitivity label auto-labeling policies

Data loss prevention (DLP): DLP policies

Data Lifecycle Management: Retention policies and retention label auto-apply policies

Records management: Retention label auto-apply policies

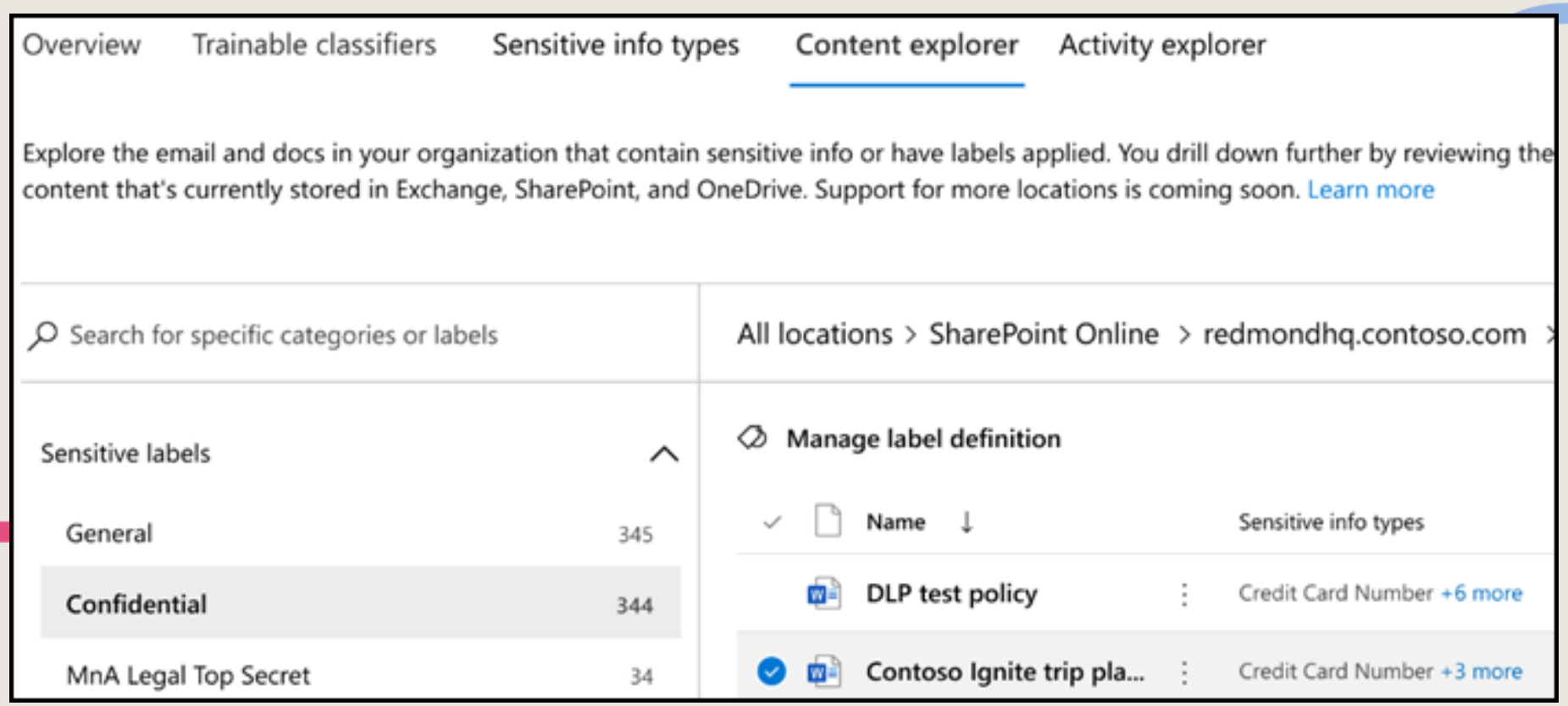
Explore labeled and sensitive content

Here is a summary of what the content explorer provides:

Visibility into the amount of sensitive data in a document that triggered the classification to be applied.

Ability to filter by label or sensitive information type to get a detailed view of the locations where the data is stored.

Integrated viewer to display documents, providing context for the circumstances in which sensitive information is being detected.



Activities related to your data

Activity explorer provides the following:

Visibility into document-level activities like label changes and label downgrades.

Ability to filter to see all the details for a specific label including file types, users, and activities.

Understand a broad-spectrum of sensitivity label activities across Microsoft 365.

Demo Time

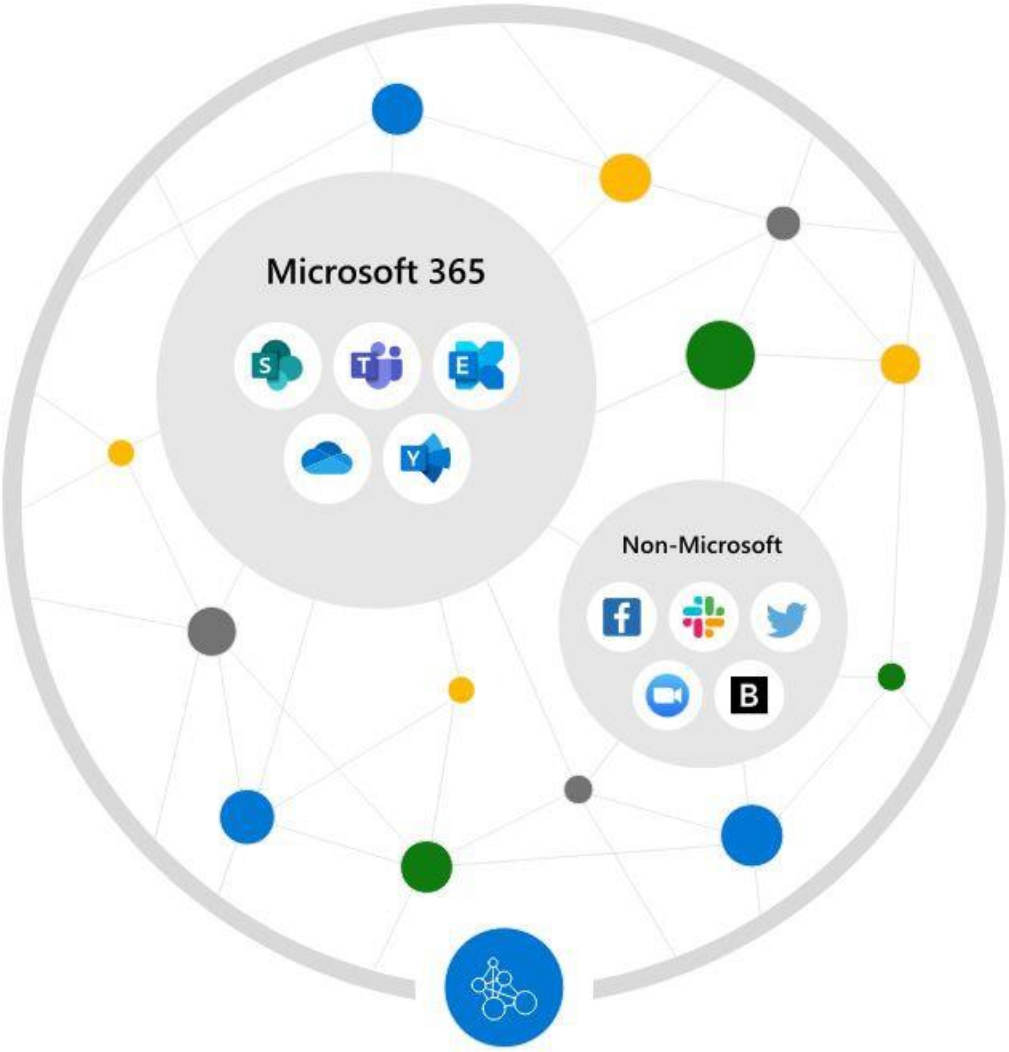
Fika (rast) [15 min]

Create image of Swedish Fika with Kaffe Te and Cinnamon bun

PASS TVÅ | SESSION TWO

The Importance of Planning and Implementation of Information Governance

Core aspects of Information Governance



Core aspects of Information Governance

Accountability

Across the organization, you'll need team members to take control of your data—if no one takes that responsibility, then there's no data governance. You and your IT team must implement ownership and responsibility. Establish a data governance team with representatives from other departments to ensure cross-organization accountability.

Rules and regulations

You'll need standardized rules and regulations for everyone to follow—developed by your data governance team to implement and create criteria for all data usage.

Data administration

Selecting a dedicated data administrator, also known as a data steward, is key to enacting and ensuring the proper protection of your data governance. As a data steward, this person's responsibility is to report to the data governance team and enforce data rules and regulations, ensuring they're followed regularly.

Data quality

You'll need high-quality, clean, and reliable data to make informed business decisions. To do this, your data steward will create a shared set of standards to improve data quality.

Transparency

All data governance processes need to be transparent as possible. Maintaining permanent records of all functions and steps ensures any future audits can determine data usage, what data was used, how you handled the data, and why your team used it.

Govern YOUR Data

Built-in information governance | Seamlessly classify, retain, review, dispose, and manage content in Microsoft 365.

Intelligent capabilities | Use intelligent machine learning capabilities to classify content and automatically apply appropriate policies.

Defensible policies | Get disposition reviews, proof of disposal, and documented audit trails with information governance.

Data governance definition

Data governance is the collection of processes, policies, roles, metrics, and standards that ensures an effective and efficient use of information. This also helps establish data management processes that keep your data secured, private, accurate, and usable throughout the data life cycle.

A robust data governance strategy is crucial for any organization that uses data to drive business growth, make improved decision-making, and ensure successful outcomes in a competitive market. When collecting vast amounts of internal and external data, you'll need to have a strategy that manages risks, reduces costs, and executes business objectives effectively.

7 data governance key foundations | by Gartner

No. 1: Align data and analytics governance with business outcomes
No. 2: Maintain a model of accountability and decision rights
No. 3: Implement trust-based governance
No. 4: Value digital ethics and transparency
No. 5: Consider risk management and information security
No. 6: Deploy governance training and education
No. 7: Encourage cultural change and collaboration

https://gtnr.it/3S6wTwh

The benefits of data governance

A big part of data governance is building a program that breaks down data silos through a collaborative process with stakeholders from disconnected business units. Your data governance program will need to do the heavy lifting to ensure that organized data is appropriately used and accurately entered into systems.

Implementing a robust data governance strategy helps ensure that your information is:

- Cleanly audited
- Evaluated
- Documented
- Managed
- Protected
- Trustworthy

Reliable Data…

✓Having a single source of truth.
✓Improved data quality.
✓Improved data management.
✓Faster, consistent compliance.
✓Reduced costs and a better profit margin.
✓A stellar organizational reputation.

The challenges of data governance

- ✓Company-wide acceptance.
- ✓Poor data management.
- ✓Standardization.
- ✓Aligning stakeholders.
- ✓Assignment of responsibilities.

The best practices of data governance

- ✓Accountability
- ✓Rules and regulations
- ✓Data administration
- ✓Data quality
- ✓Transparency

The best “PRACTICAL” practices of data governance

1. **Think big but start small. |** Document your high-level goals but keep in mind your project objectives and milestones.
2. **Appoint an executive sponsor. |** This person will advocate your data governance strategy to your high-level executives, as well as the broader organization.
3. **Build your case. |** Create the business case you'll need to justify why you need to implement a successful data governance plan as soon as possible.
4. **Develop the right metrics. |** Too many or too few metrics will make it difficult to understand if you're reaching your goals. The users, operators, and teams will need to quickly determine which metrics are and aren't necessary as you meet their objectives.
5. **Keep communicating with all levels. |** Stay open to this new process, especially encouraging those adverse to change. You'll need to provide context and transparency to many who might not understand your process and its importance

What is MIP?

Microsoft Information Protection (currently Microsoft Purview Information Protection) features help you identify, classify, and protect sensitive information while ensuring that your organization's productivity and collaboration are not impacted.

Contrary to common belief, MIP (Microsoft Information Protection) is not a single product but a collection of technologies that are integrated into several components of the Microsoft 365 environment. Microsoft 365 Compliance includes MIP capabilities, which provide you with the tools to know your data, protect it, and prevent data loss.

Data loss prevention overview

Each DLP policy contains:

Where to protect the content

Content is protected in locations like SharePoint Online, Exchange Online, OneDrive accounts, Microsoft Teams chat and channel messages, and Windows 10 or higher devices.

When and **how** to protect the content

When and how to protect the content is defined by enforcing rules. A policy contains one or more rules, and each rule consists of conditions and actions at a minimum.

Identify content to protect

Use the following to identify content:

Content explorer

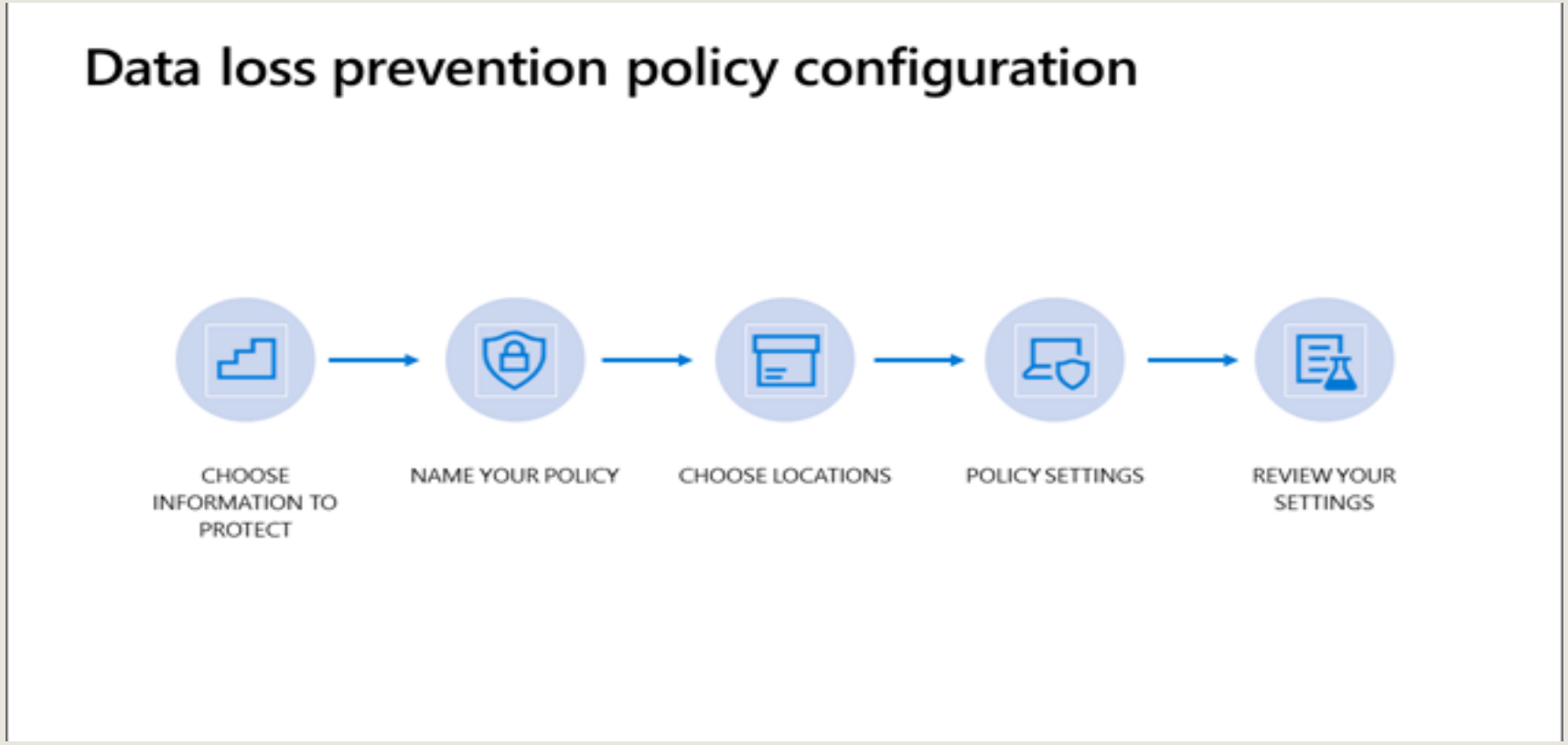
Content explorer identifies the email and documents in your organization that contain sensitive information.

Activity explorer

Activity explorer includes information on activity related to content that contains sensitive information, which can also inform what should be protected by DLP policies.

Define policy settings for your DLP policy

To create a DLP policy go to the **Microsoft Purview Portal**



Choose the information to protect

DLP policy templates consist of one or more sensitive info types grouped into categories:

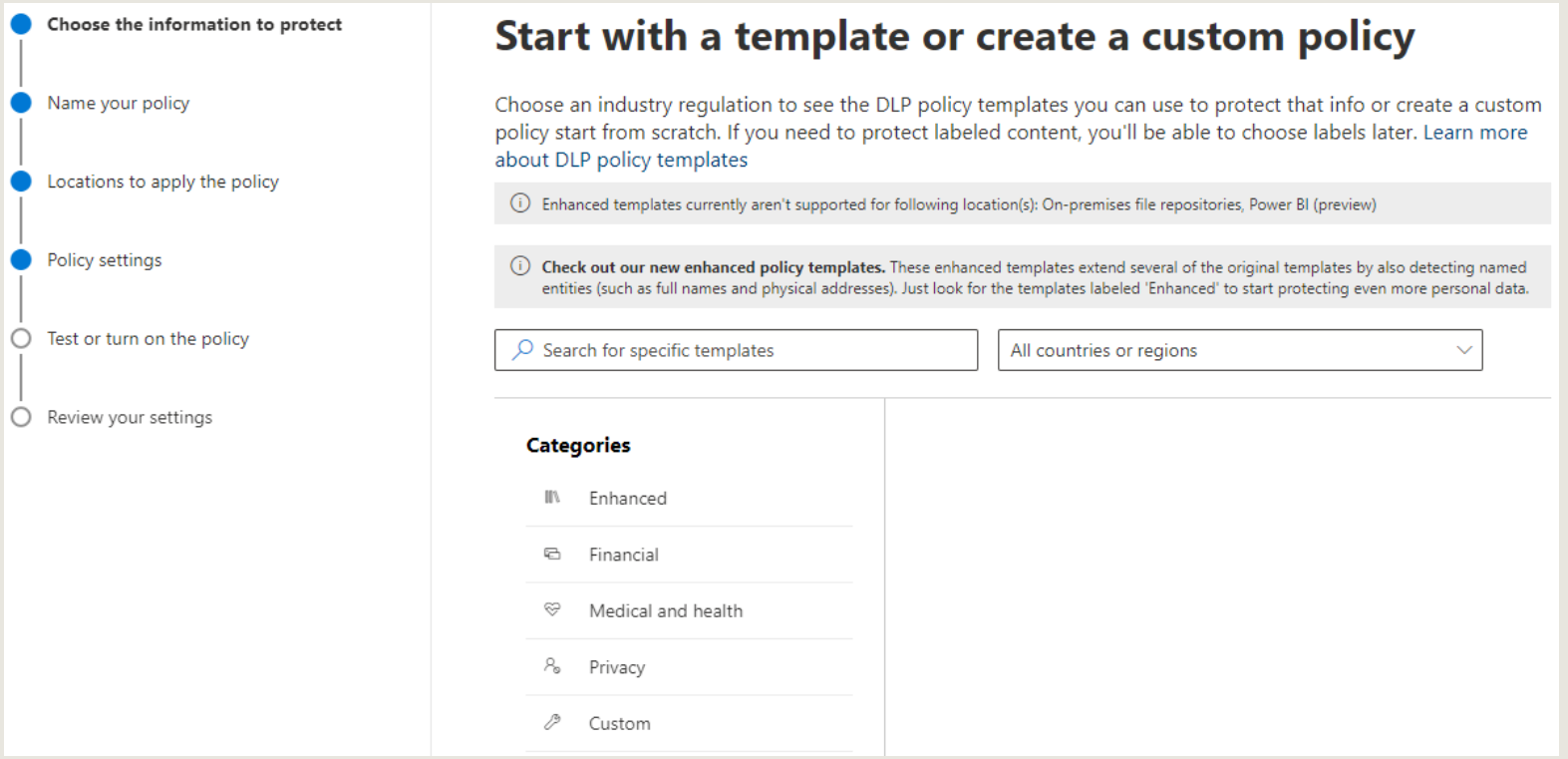
Enhanced

Financial

Medical and health

Privacy

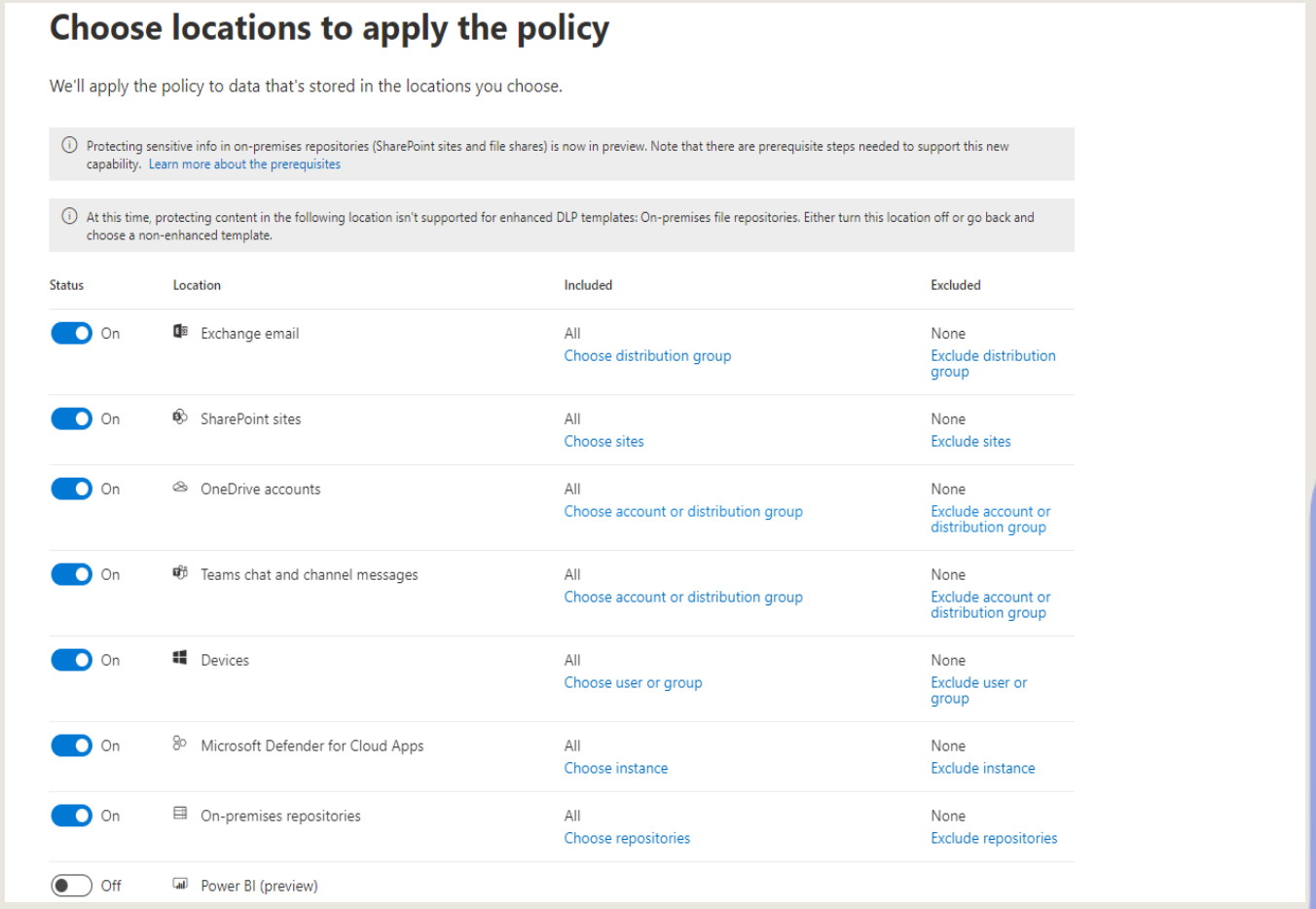
Custom



Choose locations to apply the policy

Locations are places or service the DLP policy will apply to:

- *Exchange Online email*
- *SharePoint Online sites*
- *OneDrive accounts*
- *Microsoft Teams chat and channel messages*
- *Devices*
- *Microsoft Defender for Cloud Apps*
- *On-premises repositories*
- *Power BI (preview)*



Define policy settings

DLP policy rules include:

Conditions

Determine what types of information you are looking for, and when to take an action.

Exceptions

Prevents the application of a rule for content matching the exceptions.

Actions

When content matches a condition in a rule, you can apply actions to automatically protect the content.

User notifications

Use notifications to educate your users about DLP policies and help them remain compliant without blocking their work.

User overrides

Allows the user to override the policy and share the content.

Incident reports

With a matched rule, you can send an incident report to your compliance officer with details of the event.

Test or turn on your DLP policy

Use test mode to gauge impact before policy activation.
Policy matches will be reported to you in emails or through DLP reports.
Test mode allows you to activate policy tips without enforcing protective actions.
Policy tips allow users to flag false positives.
Configure exceptions to the policy to reduce false positives.

Demo Time

Time for Lunch Break [60 min]

PASS TRE | SESSION THREE

Zero trust Approach – not only externally

What is Insider Risk Management

Microsoft Purview Insider Risk Management is a compliance solution that helps minimize internal risks by enabling you to detect, investigate, and act on malicious and inadvertent activities in your organization.

Insider risk policies allow you to define the types of risks to identify and detect in your organization, including acting on cases and escalating cases to Microsoft eDiscovery (Premium) if needed.

Risk analysts in your organization can quickly take appropriate actions to make sure users are compliant with your organization's compliance standards.

Insider Risk Management – Pain Points

Managing and minimizing risk in your organization starts with understanding the types of risks found in the modern workplace. Some risks are driven by external events and factors that are outside of direct control. Other risks are driven by internal events and user activities that can be minimized and avoided. Some examples are risks from illegal, inappropriate, unauthorized, or unethical behavior and actions by users in your organization.

These behaviors include a broad range of internal risks from users:

- Leaks of sensitive data and data spillage
- Confidentiality violations
- Intellectual property (IP) theft
- Fraud
- Insider trading
- Regulatory compliance violations

Insider Risk Management - Principles

Insider risk management is centered around the following principles:

Transparency: Balance user privacy versus organization risk with privacy-by-design architecture.
Configurability: Configurable policies based on industry, geographical, and business groups.
Integration: Integrated workflow across Microsoft Purview solutions.
Actionability: Provides insights to enable reviewer notifications, data investigations, and user investigations.

Insider Risk Management – Risk Analysis

This evaluation can help your organization identify potential areas of higher user risk and help determine the type and scope of insider risk management policies you may consider configuring. Analytics scans offer the following advantages for your organization:

Easy to configure:

To get started with analytics scans, you can select Run scan when prompted by the analytics recommendation or go to Insider risk settings > Analytics and enable analytics.

Privacy by design:

Scan results and insights are returned as aggregated and anonymized user activity, individual user names aren't identifiable by reviewers.

Understand potential risks through consolidated insights:

Scan results can help you quickly identify potential risk areas for your users and which policy would be best to help mitigate these risks.

Insider Risk Management – Risk Activities

Analytics scans for risk activity events from several sources to help identify insights into potential areas of risk. Depending on your current configuration, analytics looks for qualifying risk activities in the following areas:

Microsoft 365 audit logs:

Included in all scans, this is the primary source for identifying most of the potentially risky activities.

Exchange Online:

Included in all scans, Exchange Online activity helps identify activities where data in attachments are emailed to external contacts or services.

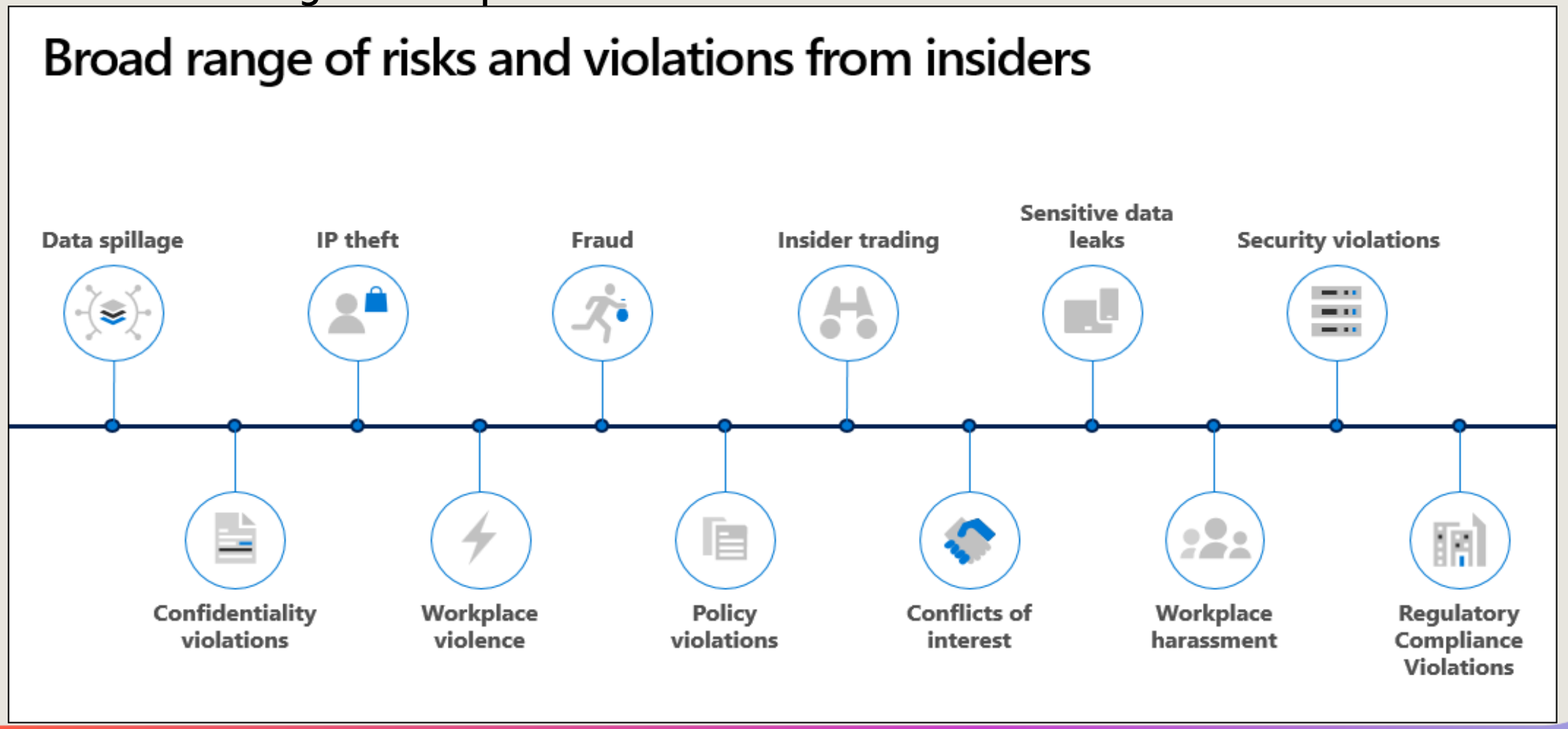
Azure Active Directory:

Included in all scans, Azure Active Directory history helps identify risky activities associated with users with deleted user accounts.

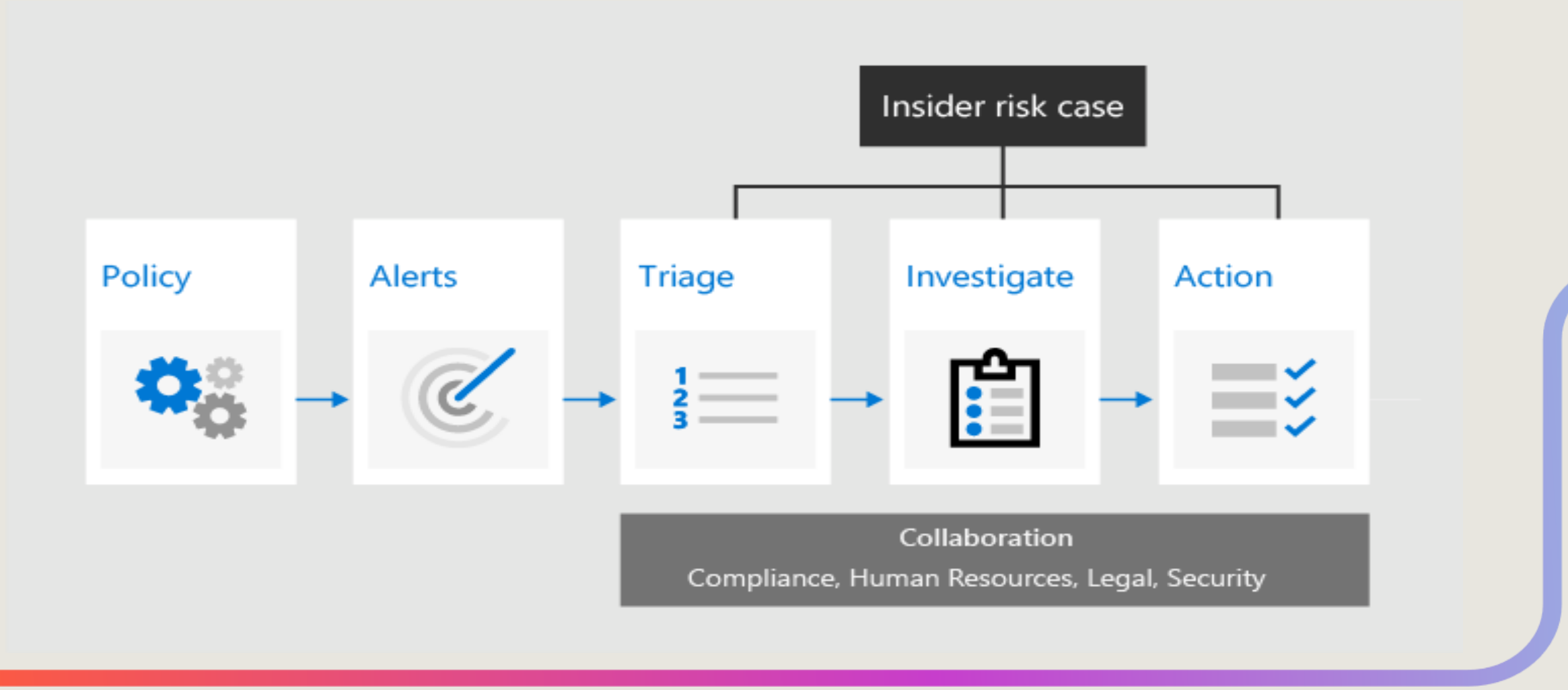
Microsoft 365 HR data connector:

If configured, HR connector events help identify risky activities associated with users that have resignation or upcoming termination dates.

Insider risk management explained



Insider risk management workflow



Manage insider risk policies

Policy templates

Departing employee data theft

Data leaks

Offensive language in email

Policy settings

Privacy and Indicators

Policy timeframes

Intelligent detections

Demo Time

Fika (paus) [15 min]

Create image of Swedish Fika with Kaffe Te and Cinnamon bun

PASS FYRA | SESSION FOUR

Overview, Features, and Requirement for Microsoft Purview [Azure]

[Tala! Det är så mörkt]

Microsoft

Microsoft Purview

Unified Data Governance to Maximize the Business Value of Data

Data-driven transformations yield significant benefits

54%
increase in revenue performance

44%
faster time to market

62%
improvement in customer satisfaction

54%
increased profit results

Source: How to lead a data-driven digital transformation by Harvard Business Reveiw

Today's data realities

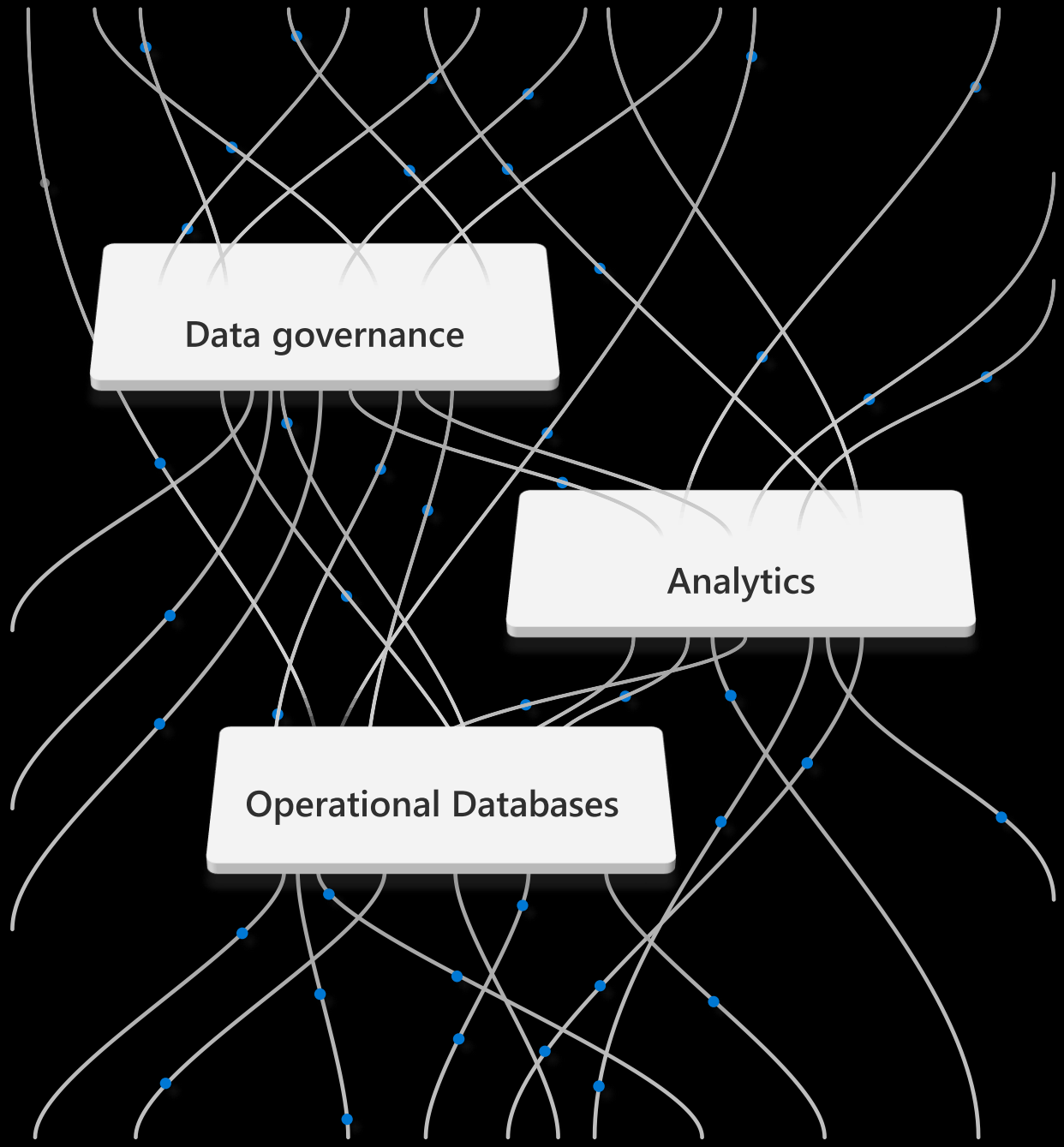
What data do I have?

Is it trustworthy?

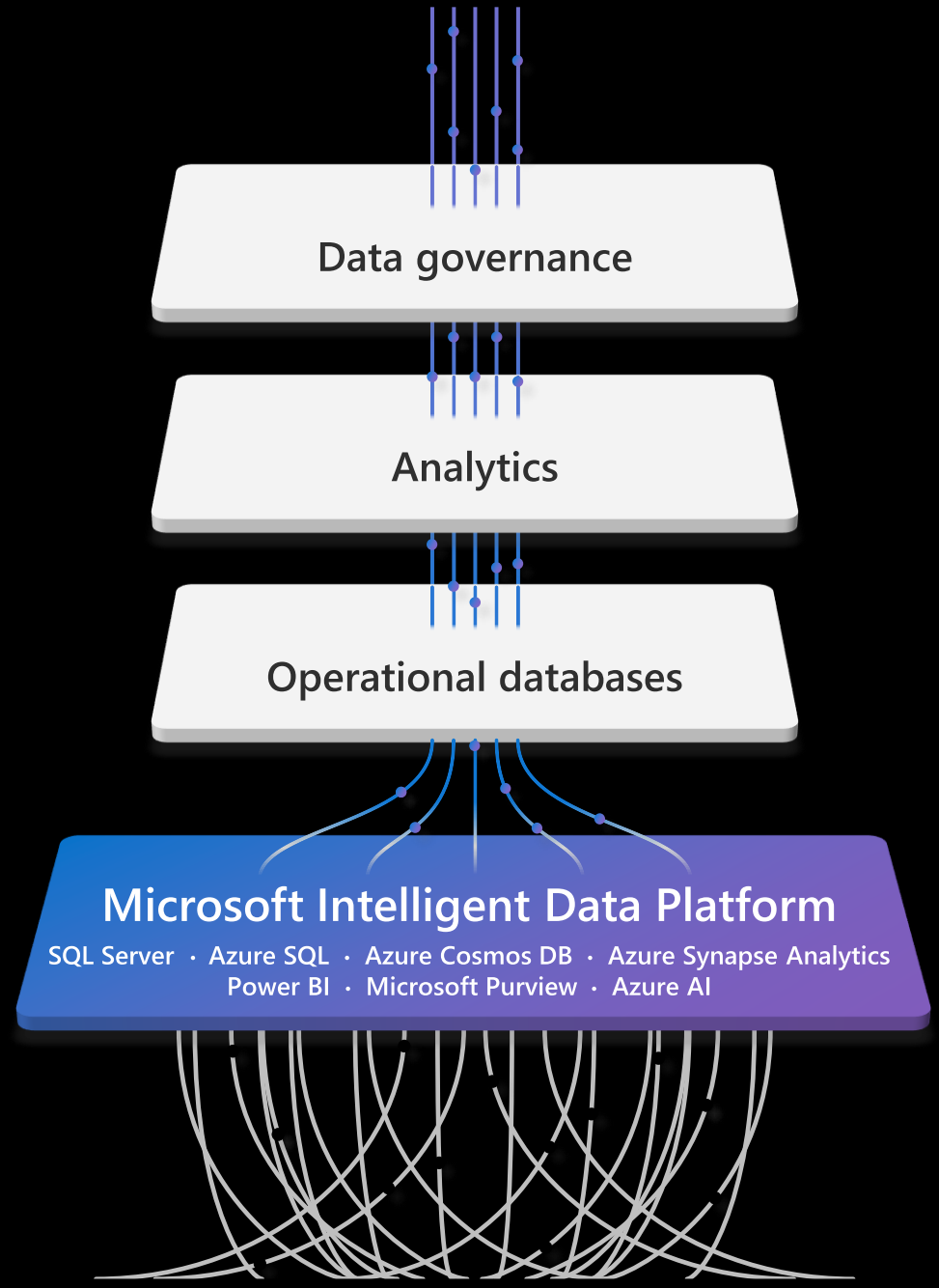
Can people access the data needed to make the right decisions?

How can I enable faster business insights?

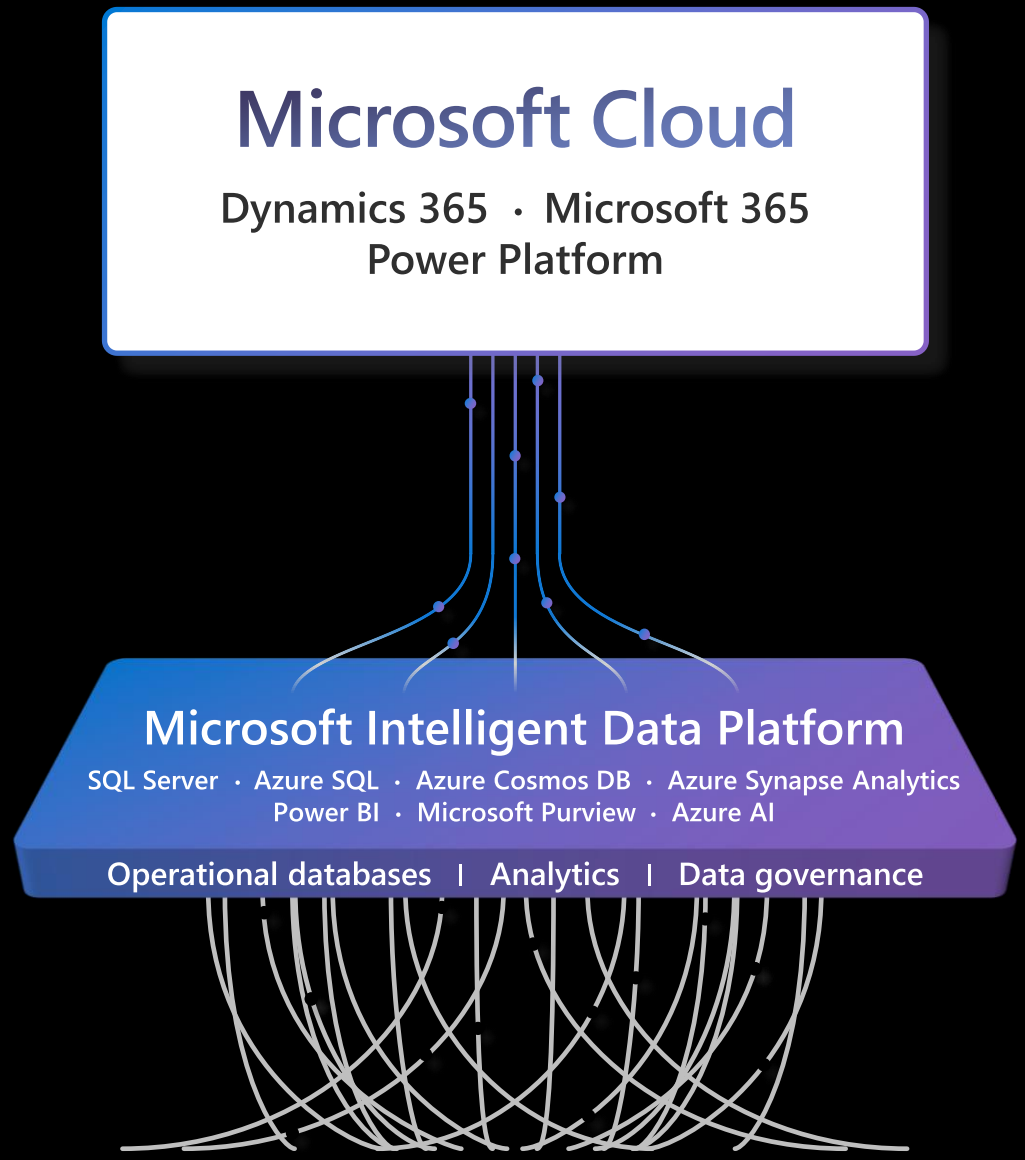
What's my compliance exposure?



Introducing Microsoft Intelligent Data Platform



Transform with the Microsoft Cloud



Data is your most strategic asset

90%

Of corporate strategies will cite information as a critical enterprise asset by 2022

GARTNER

Why Data and Analytics are key to digital transformation. Christy Pettey. Mar, 2019

175 ZB

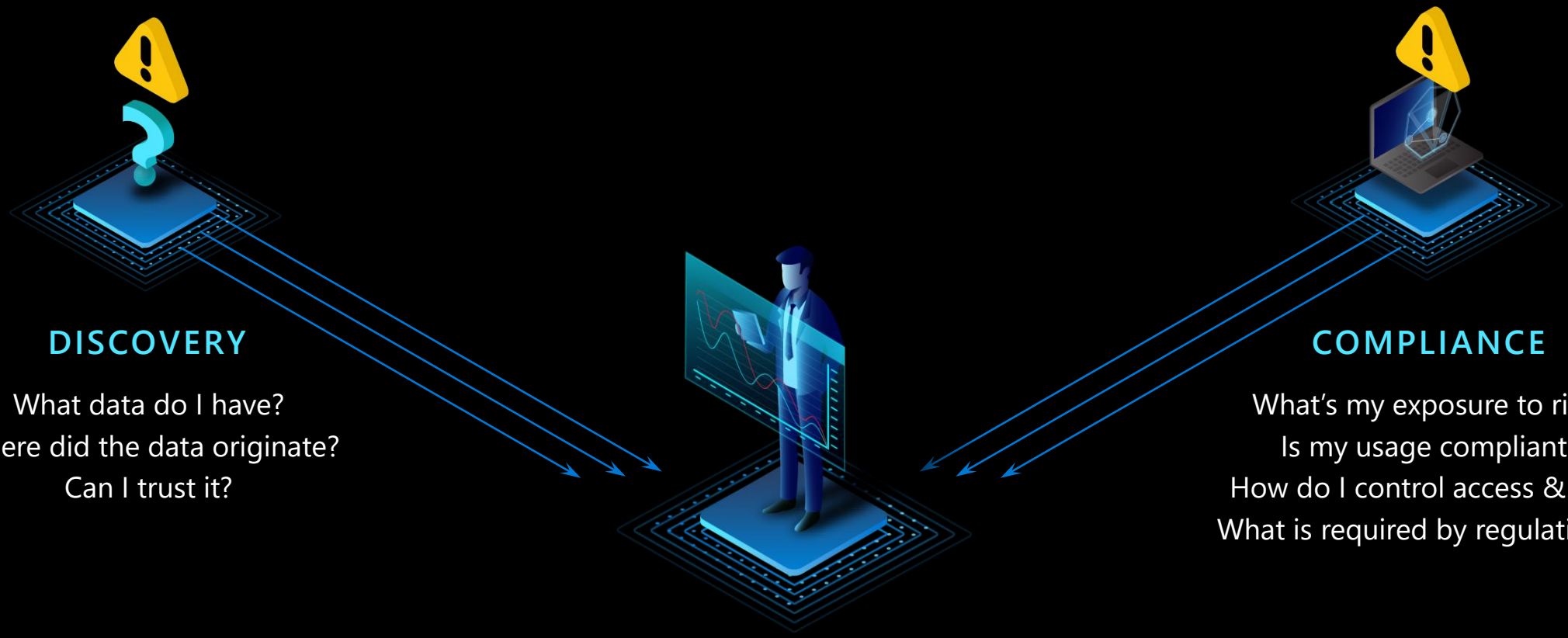
Expected global volume of data generated annually by 2025

IDC

IDC Data Age 2025, Dave Reinsel,

Microsoft

Data governance
is becoming increasingly interdisciplinary



DISCOVERY

What data do I have?
Where did the data originate?
Can I trust it?

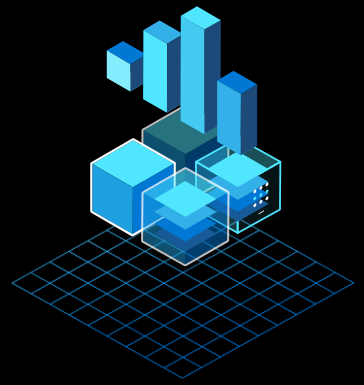
COMPLIANCE

What's my exposure to risk?
Is my usage compliant?
How do I control access & use?
What is required by regulation X?

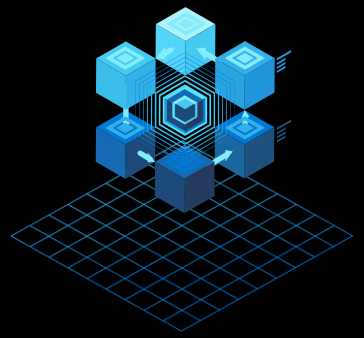
Chief Data Officer

Microsoft

Elements of successful data governance



Manage growing data landscape



Overcome operational silos

Increase data agility

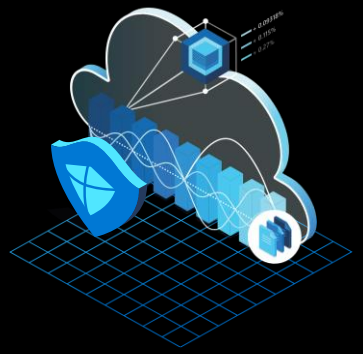
Comply with industry regulations

Reimagine data governance with Microsoft Purview

Unified | Hybrid | Open

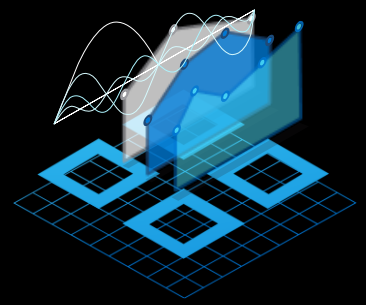
Microsoft

Microsoft Purview enables unified data governance



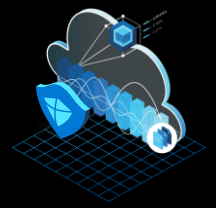
Reimagine data governance in the cloud

Set the foundation for effective data governance



Maximize business value of data for data consumers

Gain strategic insight into data use across the estate



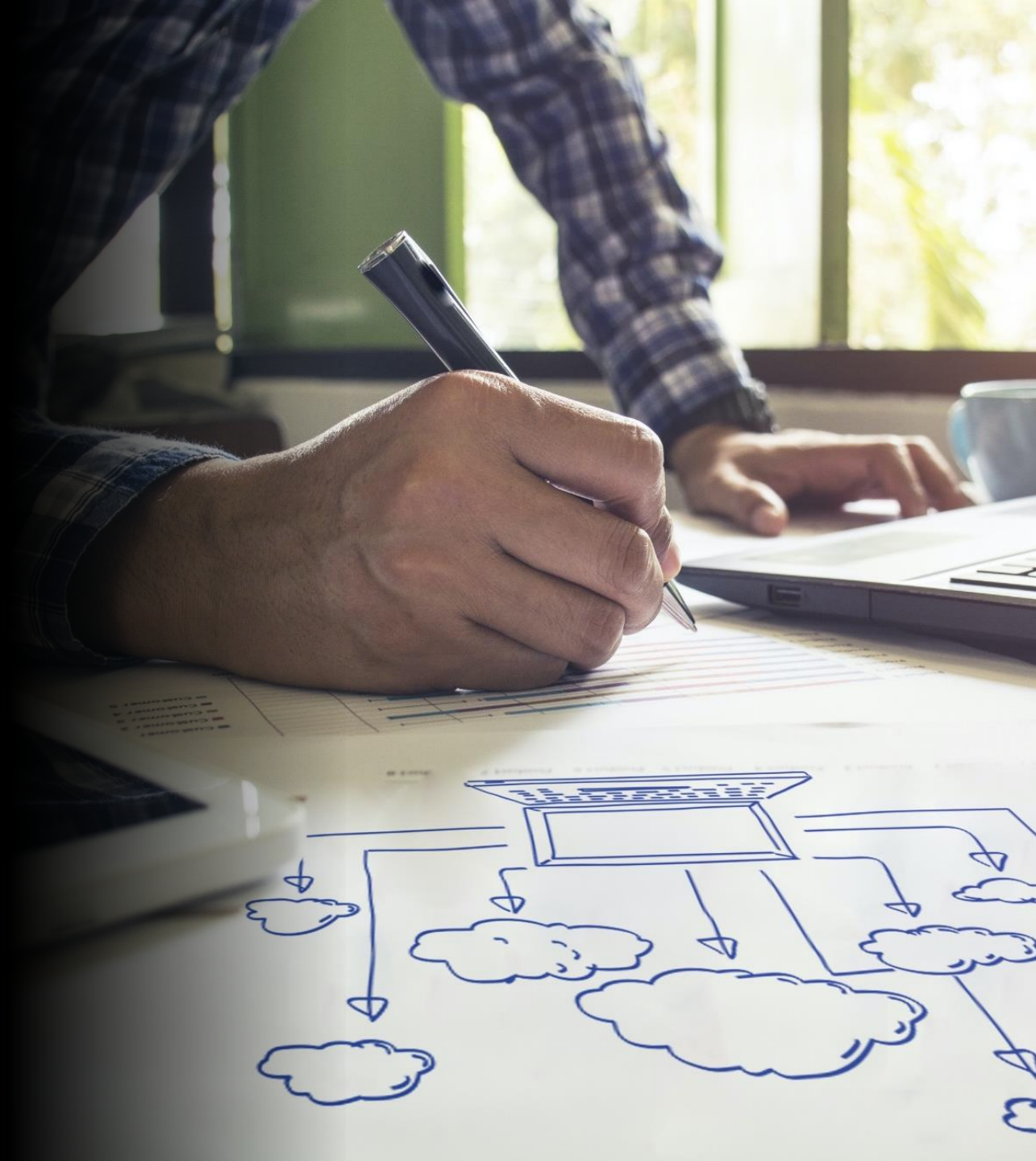
Reimagine data governance in the cloud

- Manage and govern operational, transactional and analytical data
- Cloud-native, purpose-built service to address discovery and compliance needs
- Fully managed, serverless, PaaS service
- Eliminate manual, ad-hoc and homegrown solutions

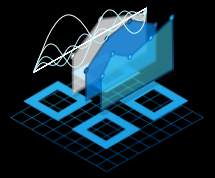
Microsoft

Establish the foundation for effective data governance

- Automate discovery of data in on-premises, multi-cloud and SaaS sources
- Classify data at scale to specify sensitivity, compliance, industry, business and company-specific value
- Know where data came from and what was derived from it with data lineage



Microsoft



Maximize business value of data

- Connect business and technical data analysts, data scientists, and data engineers to a trusted data catalog
- Enable users to quickly find data and view its lineage and sensitivity
- Deliver a curated and consistent glossary of business terms and definitions

Microsoft

Gain insight into data use across the estate

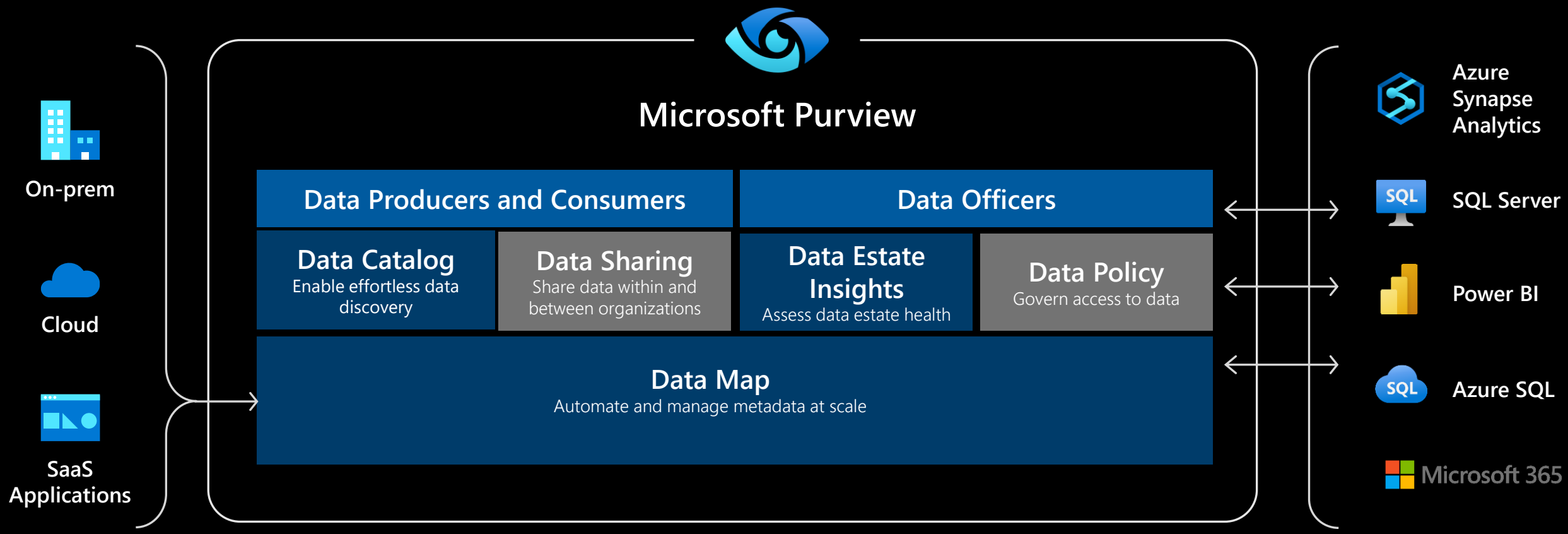
- Understand at a glance how data is being created and used across your data estate
- Visually assess the state of data assets, scans, business glossary and sensitive data

Microsoft

Unified Data Governance with Microsoft Purview

Generally Available

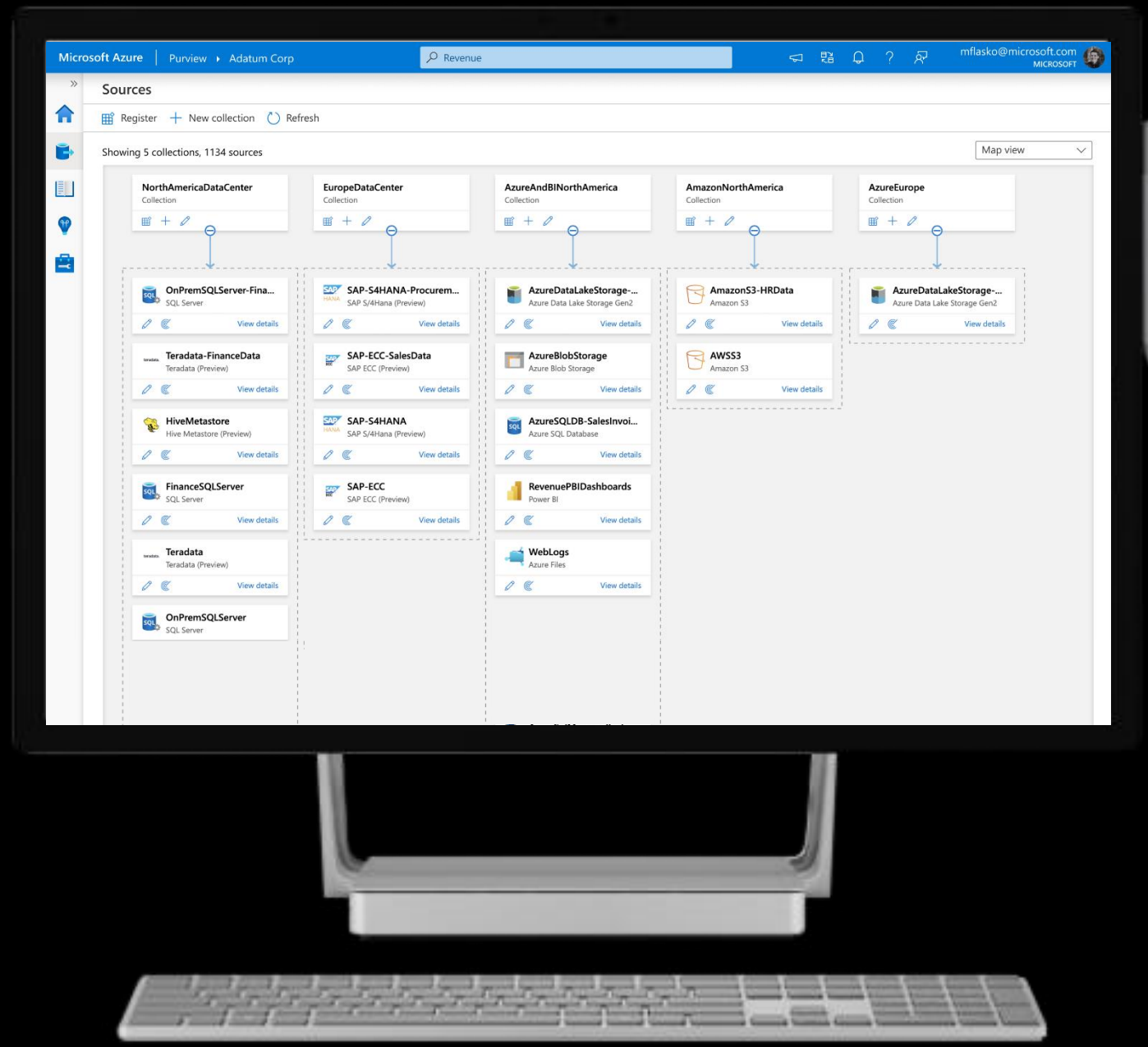
Preview



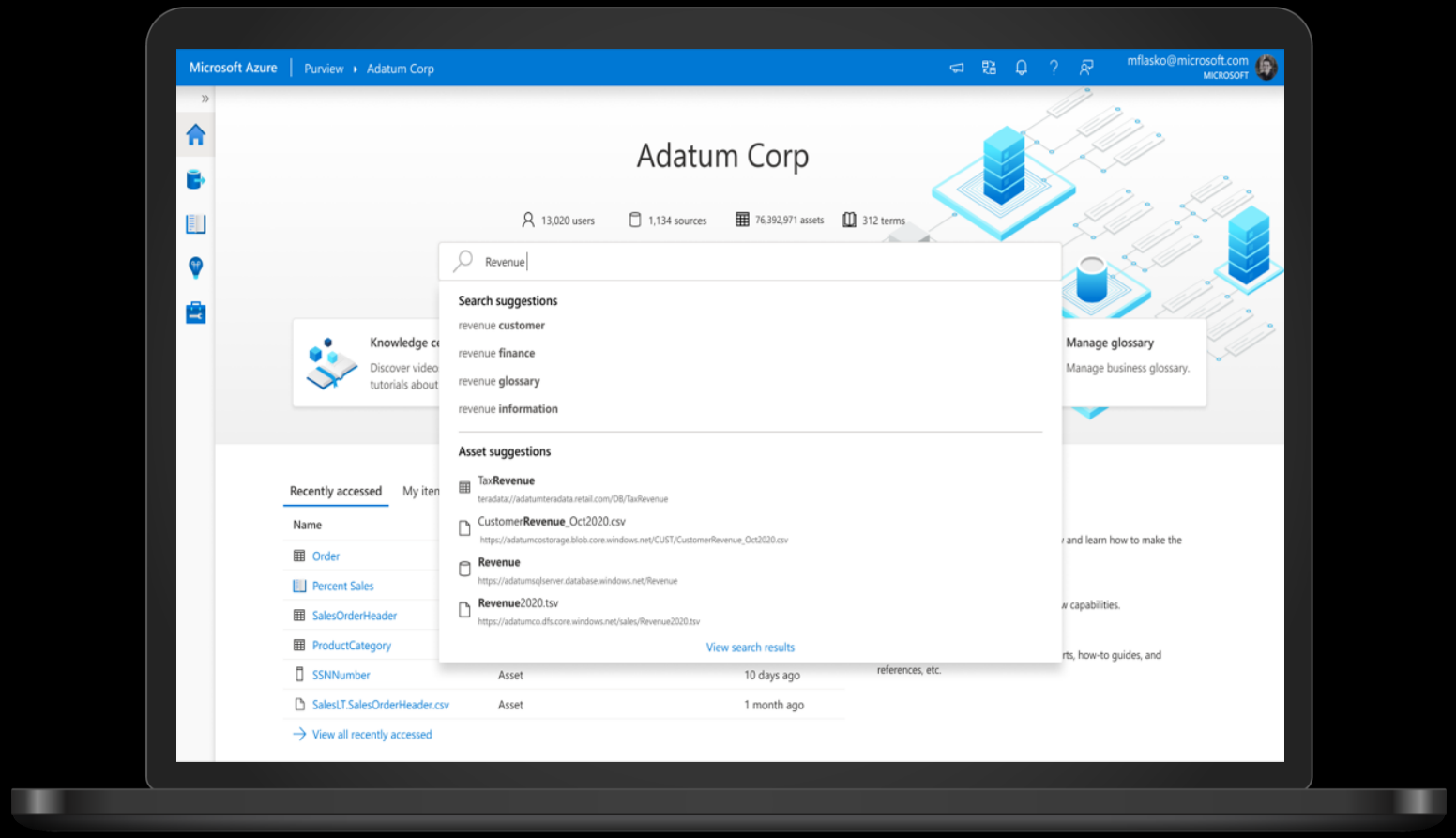
Microsoft

Unify and make data meaningful using Microsoft Purview Data Map

- A graph describing the data assets and their relationships across your data estate
- Automated data scanning, classification and lineage extraction of hybrid data stores
- 200+ built-in data classifiers
- Support for 35 data sources and growing



Microsoft



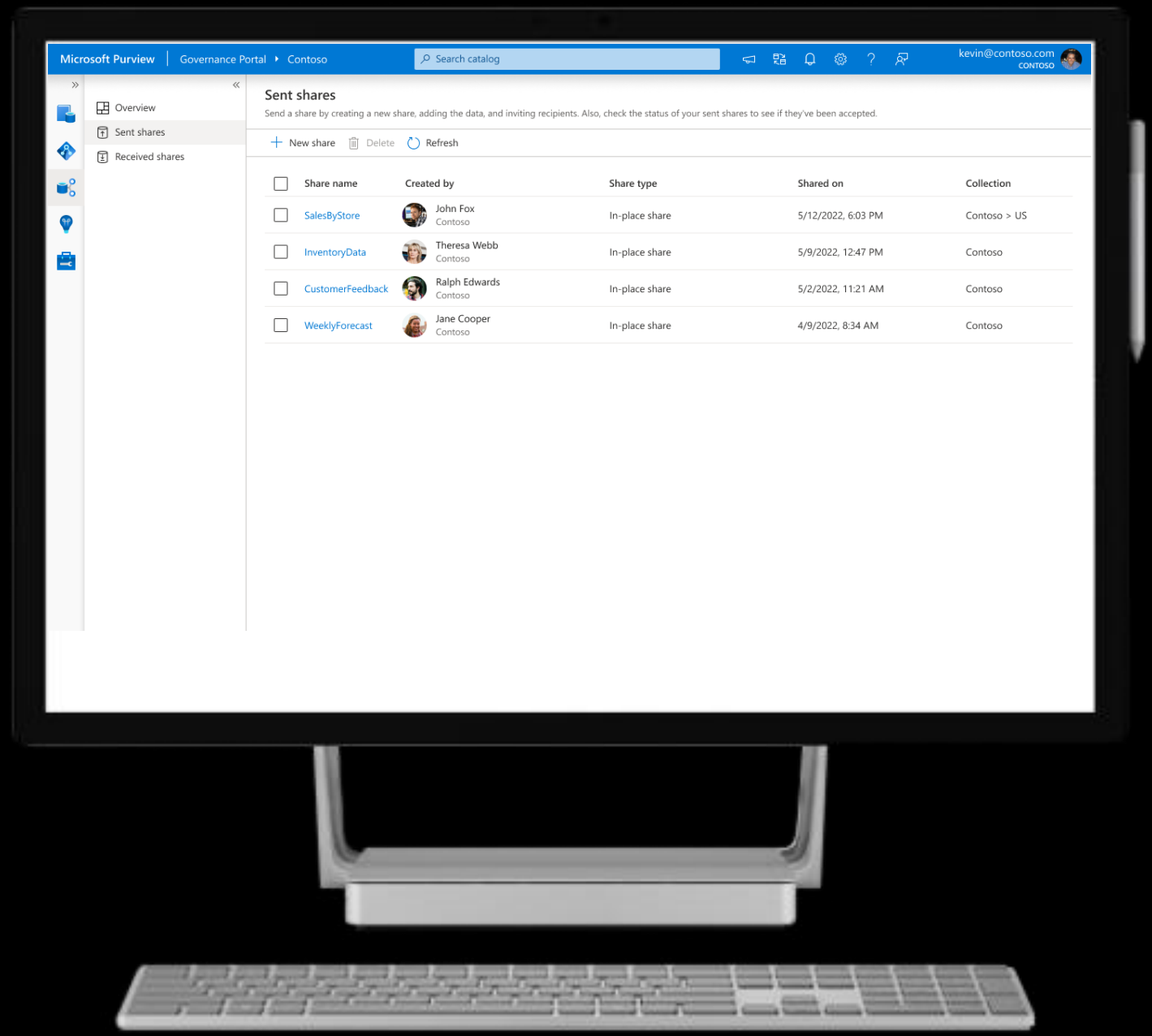
Enable effortless discovery of trusted data with Microsoft Purview Data Catalog

- Search, browse & curate
- Intelligent recommendations
- Lineage visualization

Microsoft

Share data within or between organizations using Microsoft Purview Data Sharing

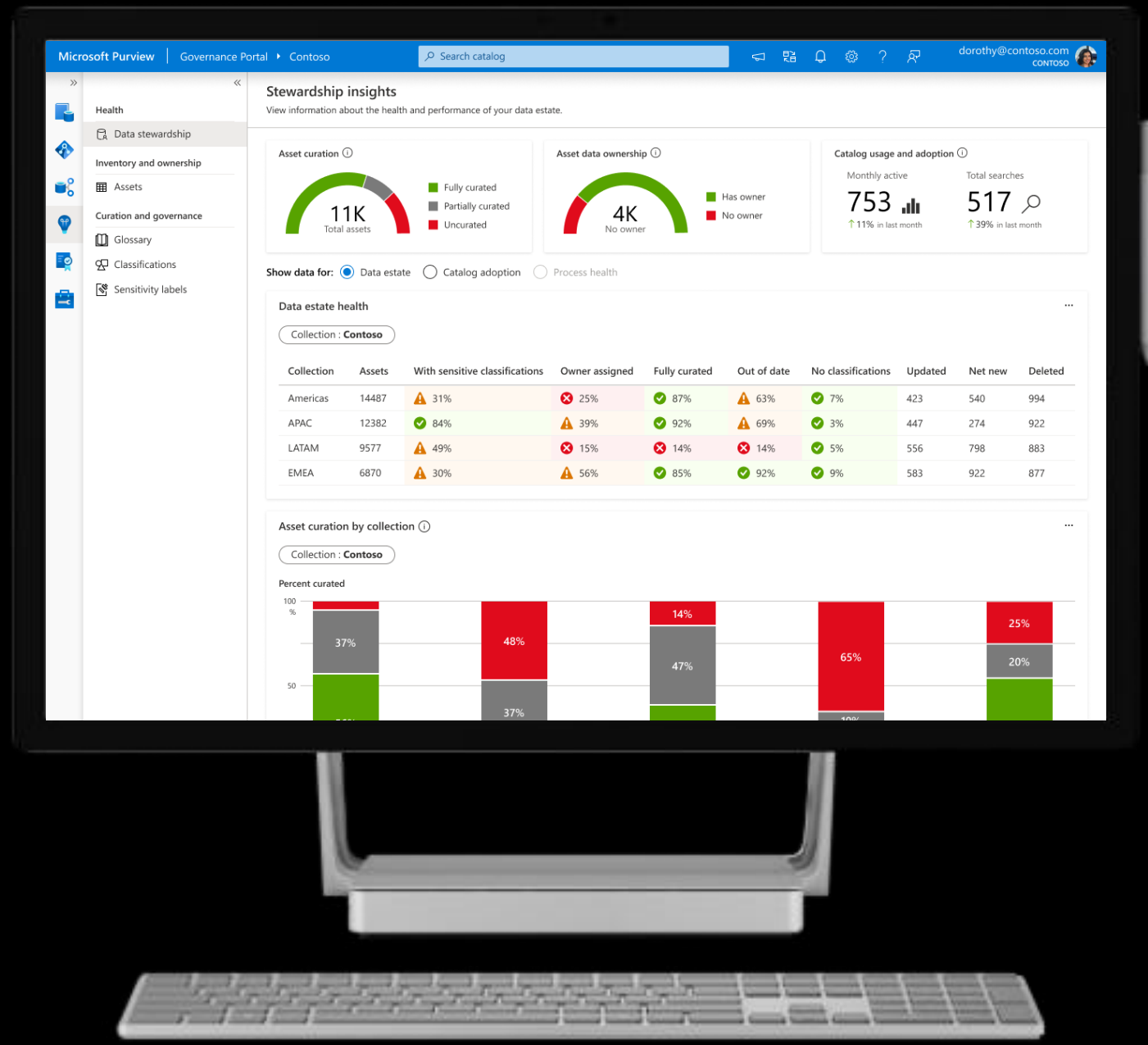
- Easily share data in-place with no data duplication
- Near real time access to shared data
- Centrally manage sharing relationships
- Supports Azure Data Lake Storage (ADLS Gen2) and Blob Storage



Microsoft

Get a bird's-eye view of your data landscape with Microsoft Purview Data Estate Insights

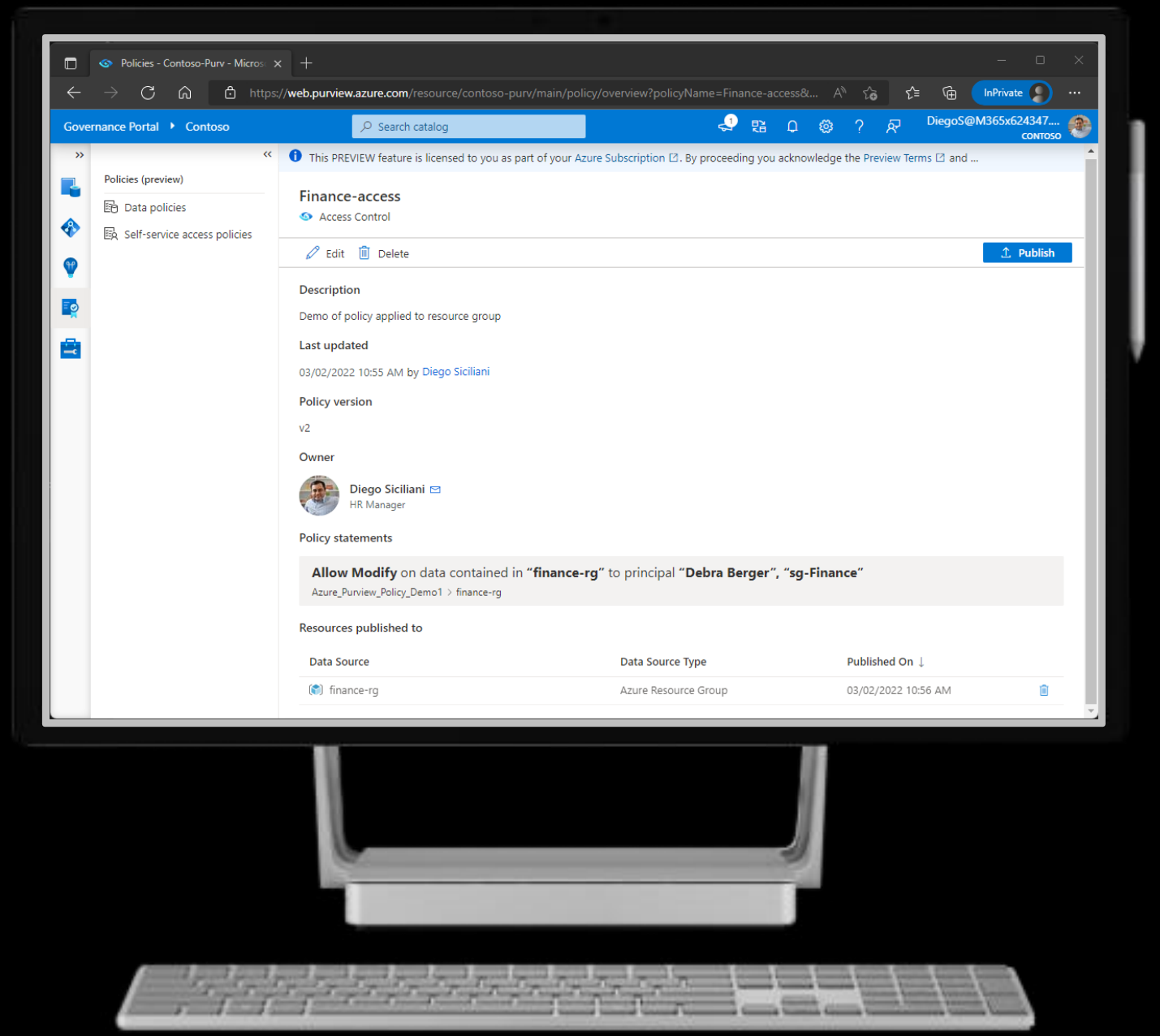
- Data Stewardship
- Data Asset distribution
- Sensitive Data
- Business Glossary Utilization



Microsoft

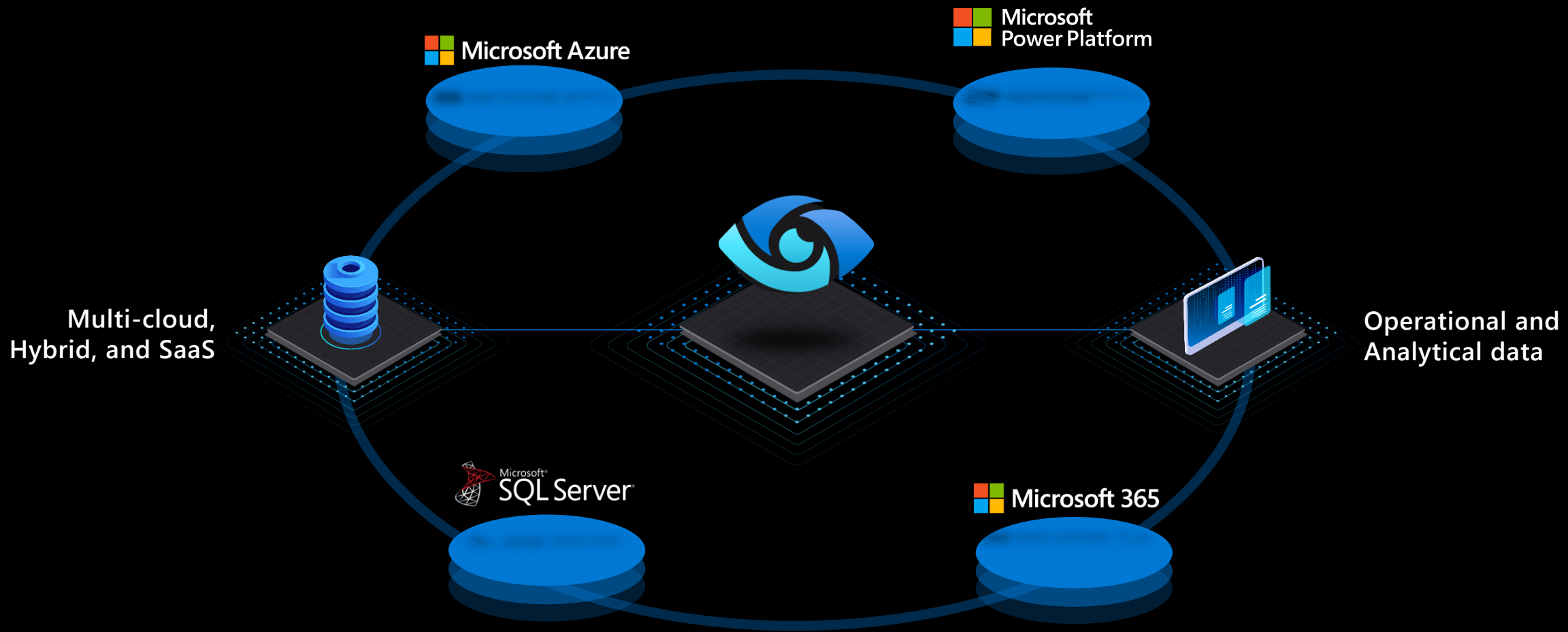
Provision access to data through an intuitive authoring experience with Microsoft Purview Data Policy

- Enable data engineers and owners to provision access to data assets
- Supports subscriptions, resource groups, Azure Blob Storage, Azure Data Lake (Gen2), Azure SQL DB, and SQL Server 2022



Microsoft

Make the most of your Microsoft investments



Microsoft

Microsoft Purview Features

Microsoft Purview Data Map	Generally Available	In Preview
Automated scanning of hybrid sources	●	
Multi Cloud Scanning for AWS S3	●	
Data Classification	●	
Apache Atlas API support	●	
Microsoft Purview Data Catalog		
Search and Browse	●	
Business Glossary	●	
Data Lineage	●	
Microsoft Purview Data Estate Insights		
Data Stewardship report	●	
Assets report	●	
Glossary report	●	
Classification and Labelling Reports	●	
Asset-level drill down by sensitivity Trends, drill downs and ability to take action	●	
Export asset list into CSV file for offline tracking with data owners	●	
Microsoft Purview Data Policy		
Author and enforce data access policies for subscriptions, resource groups, Azure Blob Storage, Azure Data Lake (Gen2) and SQL DevOps roles		●
Microsoft Purview Data Sharing		
In-place sharing for Azure Blob Storage and Azure Data Lake (Gen2)		●

Microsoft

Microsoft Purview Customer Spotlight

"Purview not only addresses the governance needs of our ambitious data estate, but the latest discovery features and simple UI will empower our analyst community to search and quickly understand the data across Heathrow. The ability to easily discover assets and browse lineage will ensure we follow trusted data across the enterprise based on common data standards, driving us to be more data-driven and efficient."

ANDREW ISENMAN | Head of Technology: Cloud and Data

Heathrow

Microsoft

Microsoft Purview Trusted Partners

PASS FEM | SESSION FIVE

Governance of the Digital Data Estate in the organisation

7 data governance key foundations | by Gartner

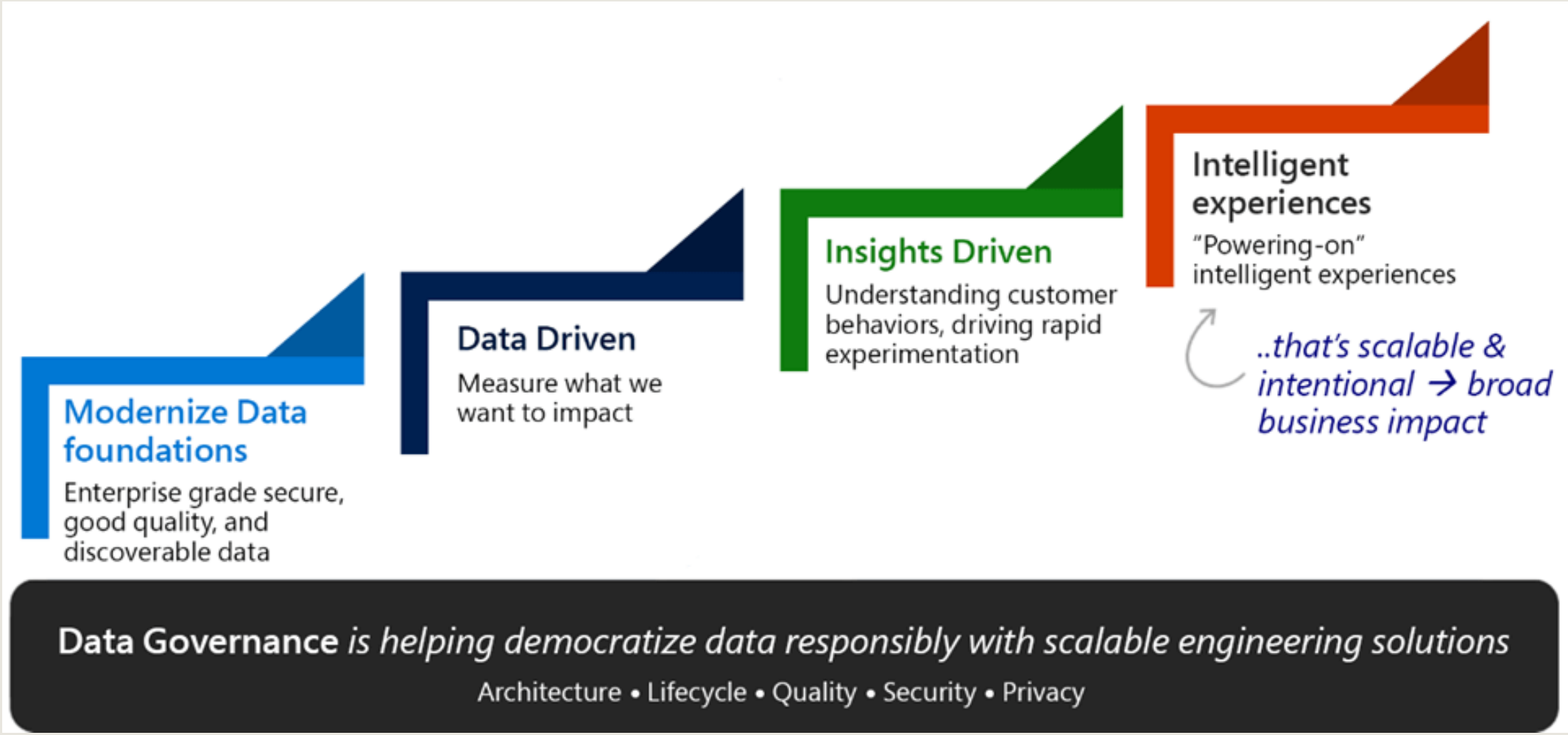
No. 1: Align data and analytics governance with business outcomes
No. 2: Maintain a model of accountability and decision rights
No. 3: Implement trust-based governance
No. 4: Value digital ethics and transparency
No. 5: Consider risk management and information security
No. 6: Deploy governance training and education
No. 7: Encourage cultural change and collaboration

https://gtnr.it/3S6wTwh

Modern Data Governance | by Microsoft

In the simplest terms, data governance is about managing data as a strategic asset.
It involves ensuring that there are controls in place around data, its content, structure, use, and safety. To provide effective data governance, we need to know
- what data exists,
- whether the data is of good quality,
- whether the data is usable,
- who's accessing it,
- who's using it,
- what are they using it for, and
- whether the use cases are secure, compliant, and governed.

Microsoft's approach to Data Governance | part one

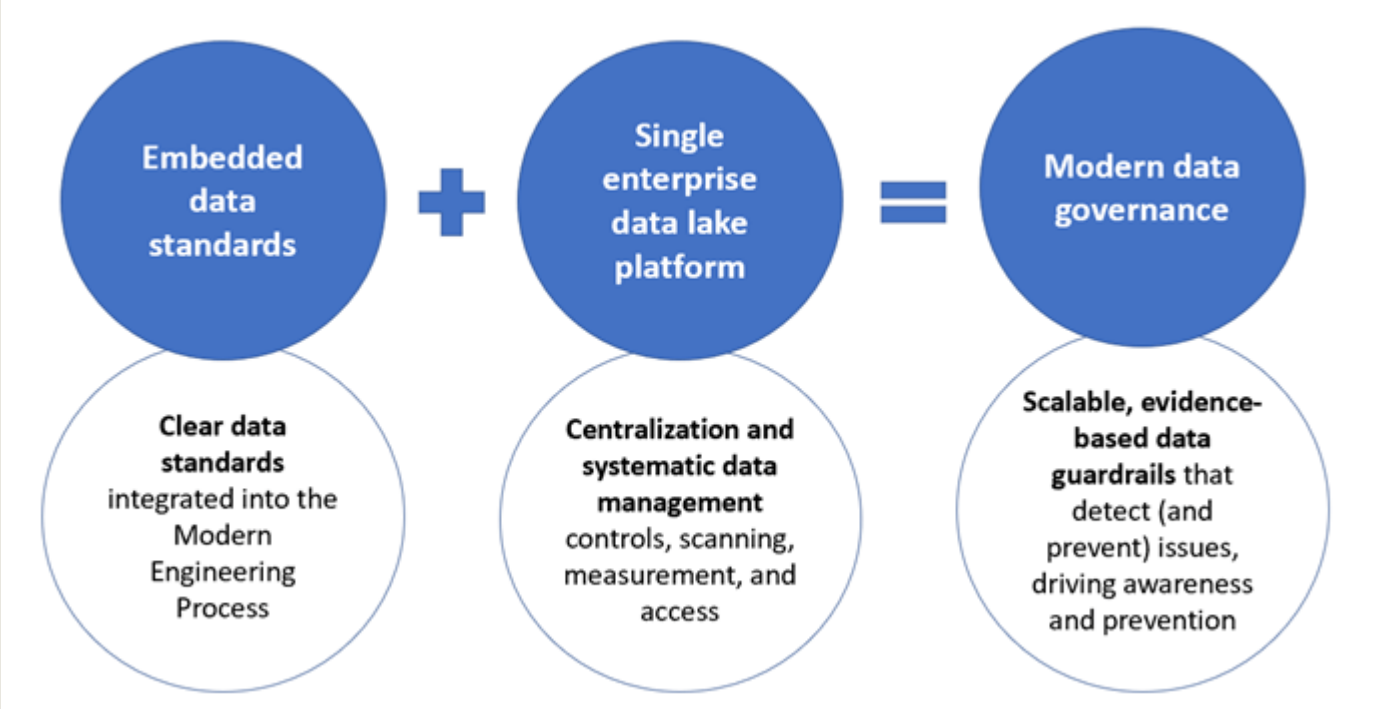


Microsoft's approach to Data Governance | part two

We developed [Microsoft Digital Employee Experience / MDEE] a more modern data governance strategy with five goals in mind:

1. *Reduce data duplication* and sprawl by building a single Enterprise Data Lake (EDL) for high-quality, secure, and trusted data.
2. *Connect data* from disparate silos in a way that creates opportunities to use that data in ways not possible in a siloed approach.
3. Power responsible *data democratization* across Microsoft.
4. *Drive efficiency* gains in the processes Microsoft employs to gather, manage, access, and use data.
5. Meet *or exceed compliance and regulatory requirements* without compromising Microsoft's ability to create exceptional products.

Microsoft's approach to Data Governance | part two



Our approach to modern data governance has two key components. First, we embed clear data standards and build them into our application development process. This move helps us automate and proactively manage data governance issues and data policy compliance. Second, we leverage the EDL platform, to centralize and systemically scan and monitor the data.

Building modern foundations for trusted and connected data

Trusted data services to ensure data quality, security, compliance, and governance

A single source of truth where connected enterprise data is collected, shaped into trusted forms, secured, made accessible, and conformed to applicable governance controls

Connected data products, including unified master data, data from disparate sources conformed to common enterprise data models, and entity hierarchies

Modern systems and tools to build and operate data products with sufficient guardrails to prevent improper data proliferation to edge systems and applications

A unified data catalog for democratized access to the data and data products that teams require to power their own digital transformation

Measuring what matters with metrics and scorecards

Define metrics that matter with scorecards consisting of:

Standardized, governed metric definitions for consistent calculations and reporting

Automated data pipelines for data collection and measurement

Reporting capabilities that are generated and refreshed automatically, and that include dashboards with trends, dimensional pivots, and self-service capabilities

Generating actionable insights with analytics

These insights uncover distinct data states and trends to spur timely action, so our robust analytics capabilities include:

Trusted and connected enterprise data that directly relate to metrics that matter (defined with scorecards)

Self-service analytic tools that enable data analysts and domain experts to generate actionable insights by querying and/or visualizing data, creating analytics modes, and exploring deeper data correlations

Tools for data democratization so data analysts and domain experts can publish and share their insights for benefit elsewhere in the company

Converting actionable insights to intelligent experiences with machine learning and AI

Investments in machine learning and AI support those goals through predictive, prescriptive, and cognitive intelligence that bolster products and internal systems.

This includes:

Infrastructure to enable data scientists to build and operate ML/AI models, with tools and services to cleanse and prep data when they need to integrate model-specific data with enterprise data

DevOps services and tools ensuring that when data scientists build, test, deploy, and operate ML/AI models, they do so in a secure, compliant, and scalable way

A **repository of reusable ML/AI models** and services that are available to non-data scientists for use in their own products and systems

Democratizing data responsibly with modern data governance

Data breaches and compliance violations could not only damage Microsoft's reputation as a trusted brand, but also create obstacles to achieving a responsible, data-driven culture. Accordingly, our approach to modern data governance includes:

Forming and staffing a data governance team to define and operationalize modern data governance across the enterprise. We're using a hub and spoke model, with the data governance team forming the hub and data stewards in each team publishing or using data to scale governance practices.

Governance processes that utilize either automation or human workflows (or some combination of the two), depending on the goals and context.

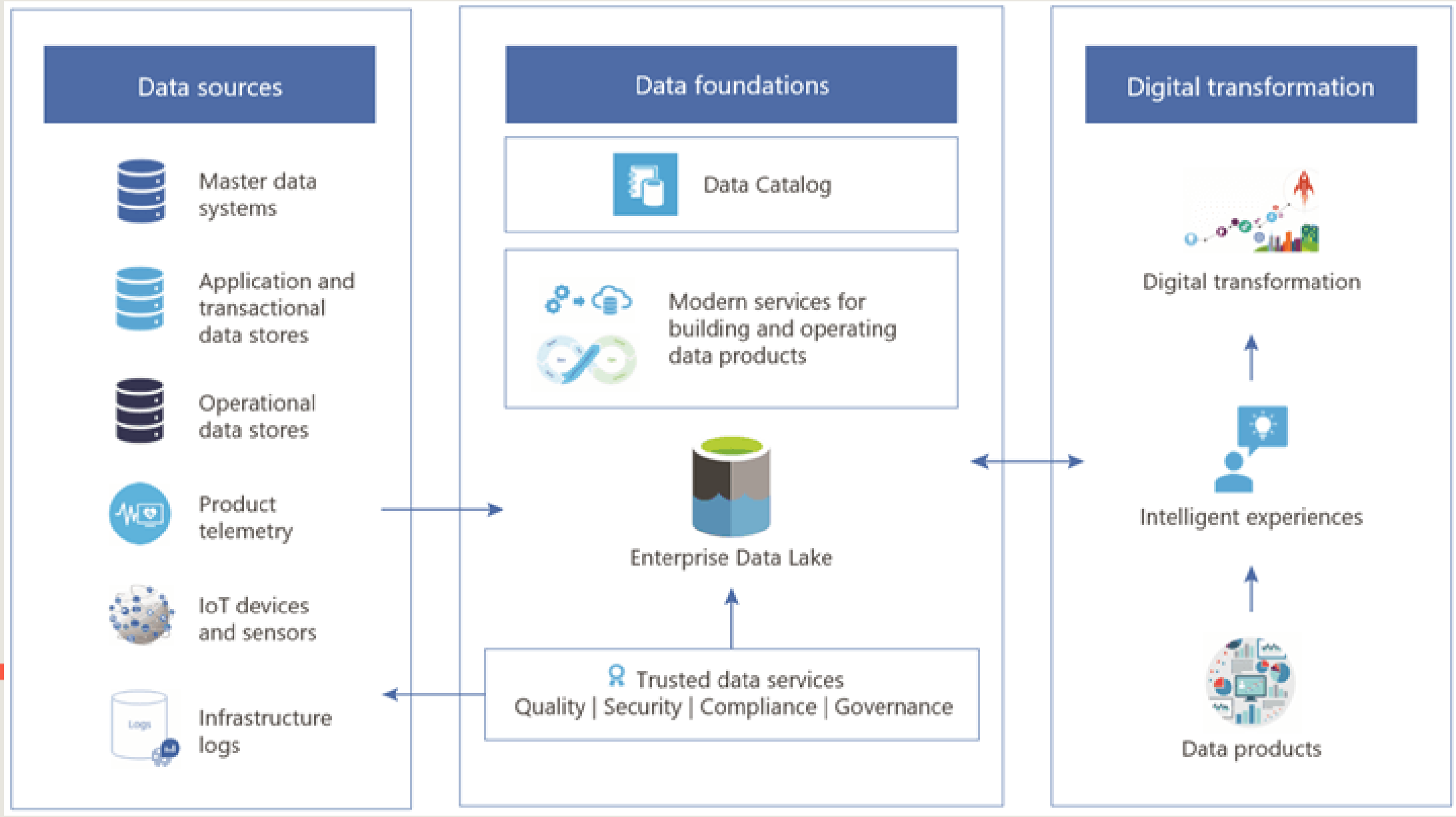
Strong, scalable technological foundations to embed governance practices in data management, data quality management, data security, data access management, compliance, and governance process automation.

Growing and scaling the data community

Since democratized access to data is a foundational goal of the digital transformation, non-technical users and teams will require support and training to evolve their use of data. We're fostering a community within Microsoft to train teams and apply shared learnings. We've created working groups for data topics, training sessions, consultation forums, and shared sources for this purpose.

More broadly, we've also made data literacy a core tenet of our software engineering, product management, and design teams and processes. Within our product and service development teams, dedicated data professionals now work to foster a data-driven mindset and healthy data practices. Such structural change requires strong leadership, buy-in, and momentum.

Implementing the Modern Data Foundations



The Enterprise Data Lake & Trusted data services

Consolidating and standardizing each data generation and publishing system created over the past several decades simply isn't feasible. The Enterprise Data Lake (EDL)—built on Azure Data Lake, Azure Data Factory, and Azure Synapse Analytics—addresses this challenge by serving as the enterprise's system of intelligence. There, data from across the enterprise is ingested, conformed, standardized, connected, democratized, and served for enterprise-wide applications in analytics as well as ML/AI.

Data is valuable only insofar as it's trusted. We create trusted data by making investments in data quality, security, compliance, and governance services, created by using and extending related Azure services. Data quality services include probabilistic, rules-based data quality scanning, as well as closed-loop data publisher workflows. Standardized schemas and shared data and analytics models conform data for shared entities, and data sources from multiple systems (and that don't include common connector attributes) are unified to generate golden records.

Security and management services are built using Azure Active Directory (AAD), Azure Entitlements Lifecycle Management (ELM), and Azure Key Vault. These services are secure by design, an achievement made possible by abstracting the complexities of creating, managing, and operating security and access management capabilities.

Is there any CEO out there?

6 best practices for good data governance

Identify critical data elements and treat data as a strategic resource

Set policies and procedures for the entire data lifecycle

Involve business users in the governance process

Don’t neglect master data management

Understand the value of information

Don’t over-restrict data use

SUMMARY

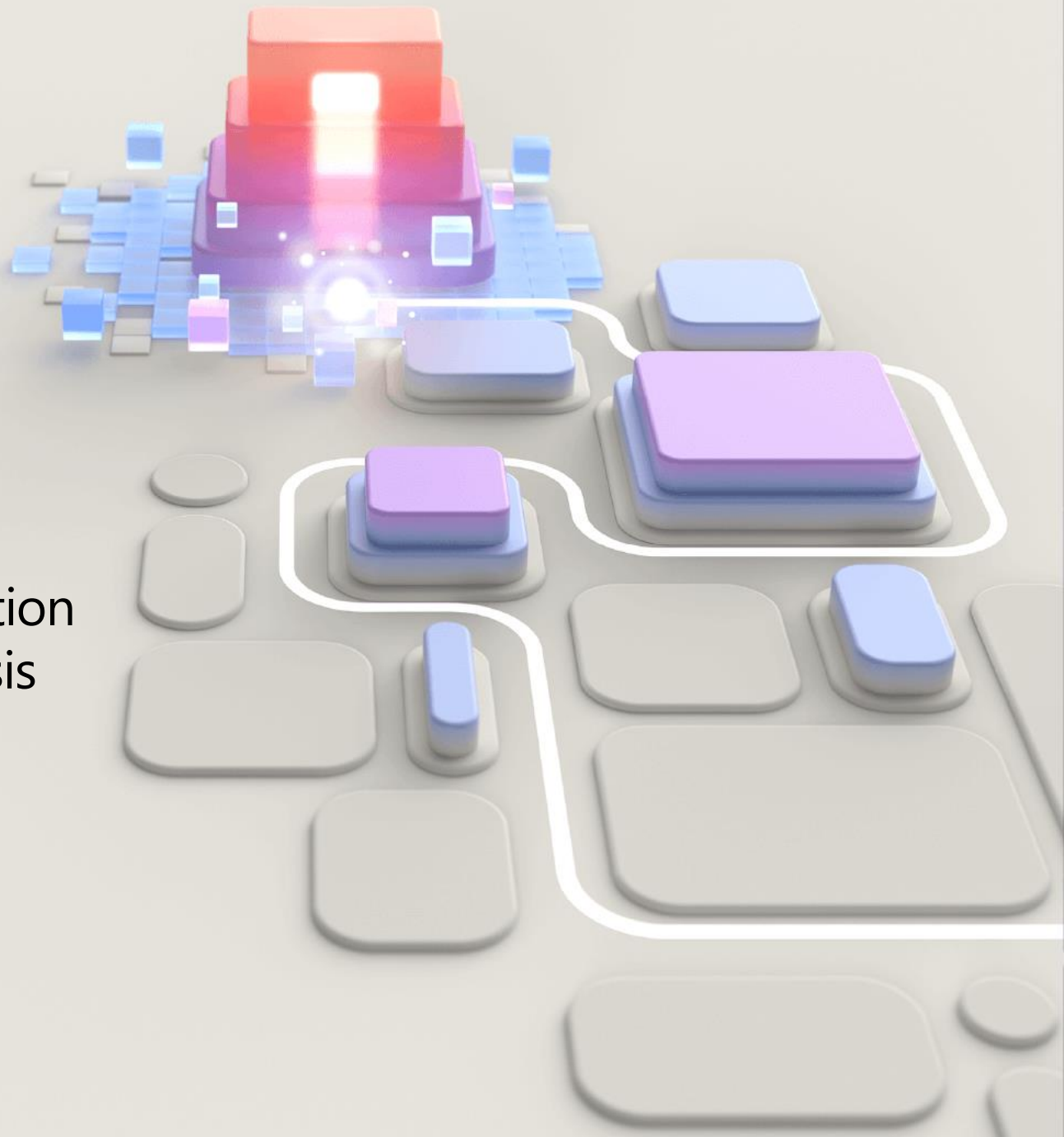
Q&A | Discussion

- Questions?

- Answers?

Microsoft

Dag Två | Day | Two | 10 Januari 2023



Förmiddagspass | Morning Sessions

Microsoft 365 Purview: information protection in practice, sensitive content, search, analysis

Eftermiddagspass | Afternoon Sessions

Azure Purview: diveing into purview, terminology, design, execution

PASS ETT | SESSION ONE [A]

Compliance and Sensitivity Labels with MIP

What is A Sensitive Information

Sensitive information refers to privileged or proprietary information that only certain people are allowed to see and that is therefore not accessible to everyone. If sensitive information is lost or used in any way other than intended, the result can be severe damage to the people or organization to which that information belongs.

Some examples of sensitive information are as follows:

Personal information, including Social Security Number and bank credentials

Trade secrets

System vulnerability reports

Pre-solicitation procurement documentation, including work statements

Computer security deficiency reports

Taken from Techopedia [What is Sensitive Information? - Definition from Techopedia]

What is CIA Triad

According the Computer Security Act of 1987 (United States, first implemented globally), organizations must be held responsible for protecting their own sensitive information by providing one or more of the following:

Confidentiality: Sensitive information should only be accessible to those who are allowed to see it, not just those who wish to see it.

Integrity: Unauthorized users should not be able to make changes to the information, thus compromising its integrity.

Availability: Information must be accessible during a certain time and may not be destroyed during that time frame. People with permission to view the data must be able to view it.

Taken from Techopedia [What is Sensitive Information? - Definition from Techopedia]

Compare built-in versus custom sensitive information types

- Built-in Sensitive Information Types
  - More than ~~260~~ over 300 built-in sensitive information types.
  - Includes default patterns managed by Microsoft.
  - Fulfills basic protection of common information types.
  - Starting point for implementations.

- Custom Sensitive Information Types
  - Allows to detect business individual sensitive information.
  - Consists of customized patterns.
  - Provides several special features:
    - Exact Data Match (EDM)-based classification
    - Document Fingerprinting
    - Keyword dictionaries

Create and manage custom sensitive information types

Sensitive Information Type parts:

Primary pattern

Search pattern for detection, consisting of keywords or regular expressions.

Additional evidence

Second search pattern for higher matching accuracy of the primary pattern, consists of keywords.

Character proximity

Detection window in characters of primary patterns and additional evidence.

Confidence level

Supporting level of pattern and evidence matching accuracy.

Create a keyword dictionary

Efficient way to manage large lists of words that are regularly subject to changes

Supports up to 100KB of terms after compression.

Source can be cleartext files such as .txt and .csv files.

Keyword Dictionary Creation Best Practices:

- For a school, get together with a class of students to find words and phrases you don't want in an education environment.
- For companies, use various options to collect keywords:
  - Search typical words from some departments e.g., using Microsoft Forms.
  - Information from employees e.g., from HR or legal to create a list of typical words.
  - Create an employee audit and then create the list out of the outcome.

Describe custom sensitive information types with exact data match

Phase	Requirements
Step 1: Set up EDM-based classification	• Read access to the sensitive data • Database schema in XML format • Rule package in XML format • Admin permissions to the Security & Compliance Center
Step 2: Hash and upload the sensitive data	• Custom security group and user account • Local admin access to machine with EDM Upload Agent • Read access to the sensitive data • Process and schedule for refreshing the data
Step 3: Use EDM-based classification in policies	• Microsoft 365 subscription with DLP • EDM-based classification feature enabled

Sensitive information type based on database matching

Supports up to 100 million rows of sensitive data, 32 columns (fields) per data source and up to 5 columns (fields) marked as searchable.

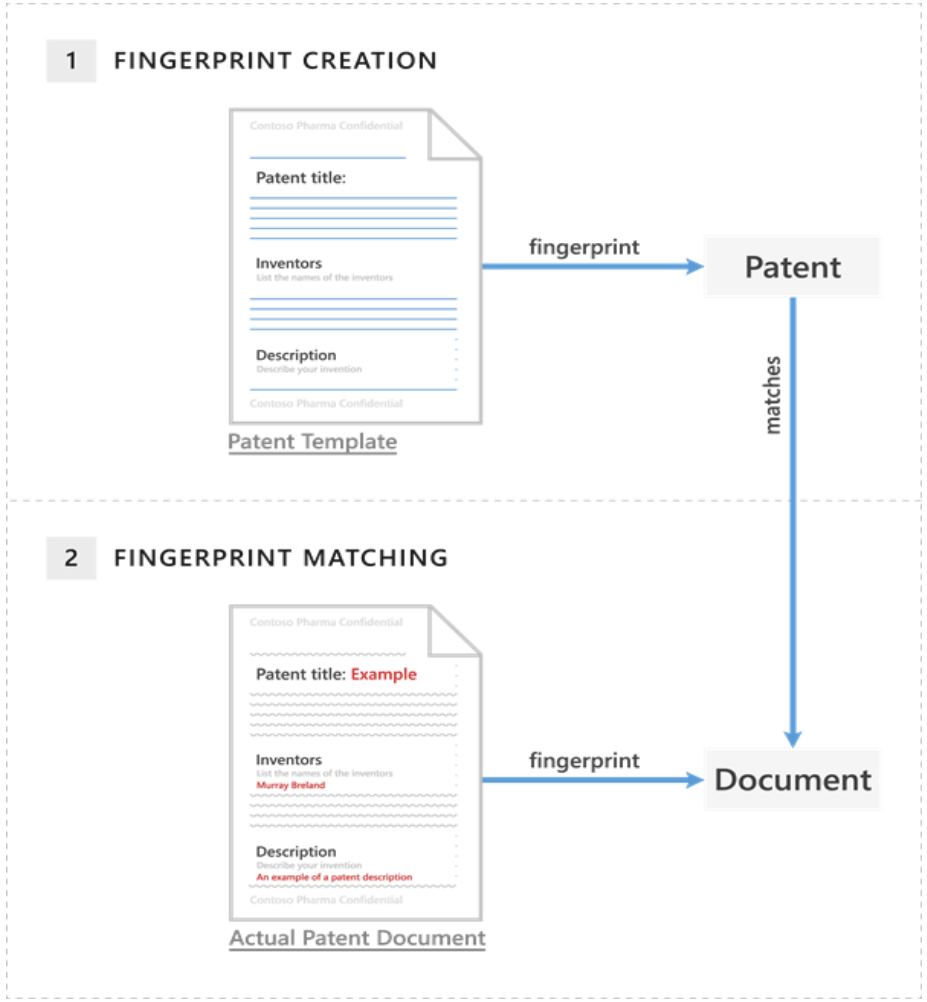
Additional license requirements

Configuration done in three steps:

- Step 1: Set up EDM-based classification
- Step 2: Hash and upload the sensitive data
- Step 3: Use EDM-based classification in policies

EDM – Exact Data Match

Implement document fingerprinting



Digital fingerprint from empty template document.

Detection of documents created from a template.

Requires all fields from the original template document.

Exchange Online only.

Password protected files are not supported, as well as documents containing images.

Implement document fingerprinting

Examples of forms that you can upload include:

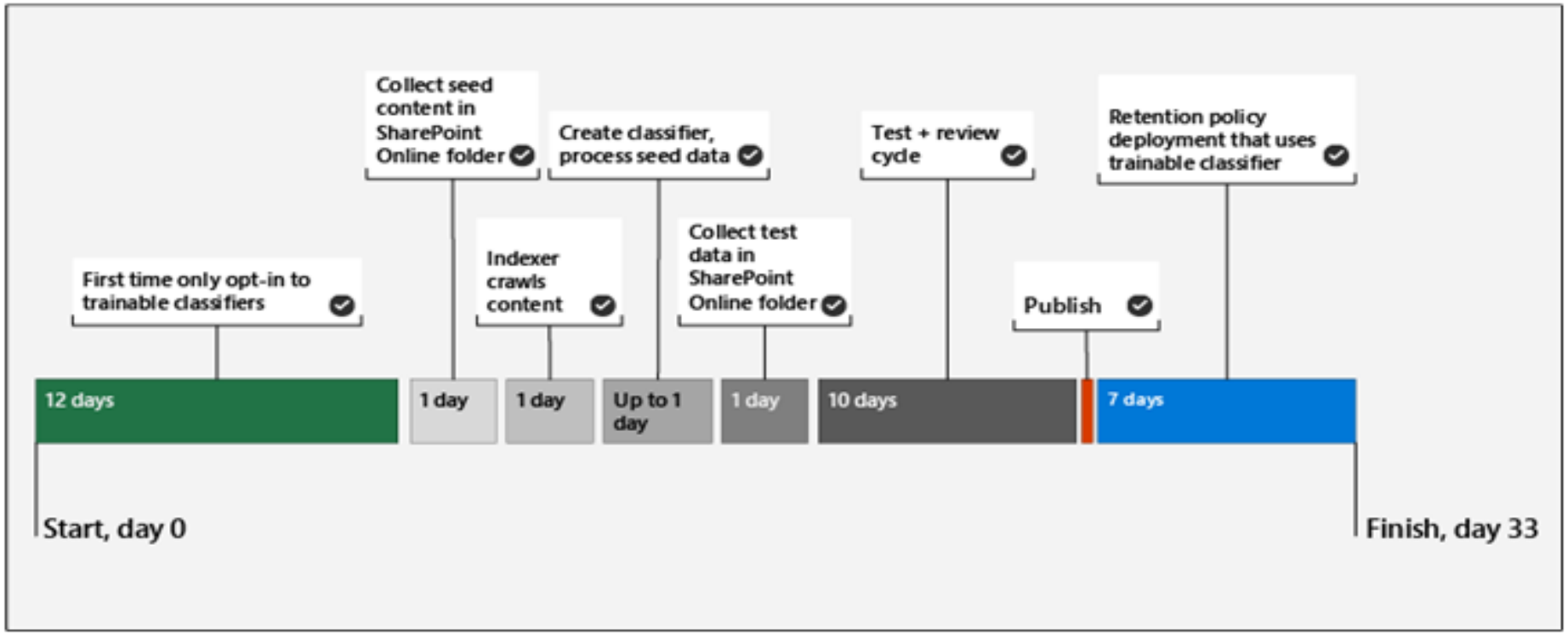
- Government forms
- Health Insurance Portability and Accountability Act (HIPAA) compliance forms
- Employee information forms for Human Resources departments
- Custom forms created specifically for your organization

Ideally, your organization already has an established business practice of using certain forms to transmit sensitive information. After you upload an empty form to be converted to a document fingerprint and set up a corresponding policy, the DLP detects any documents in outbound mail that match that fingerprint.

Classify data using trainable classifiers

Training includes several steps:

1. Items of a type added to a SharePoint library.
2. Train until the classifiers do not request additional training documents.
3. Review items to improve the classifier accuracy.



Classify data using trainable classifiers (continued)

Classifiers separate into default classifiers, custom trainable classifiers, and retrained classifiers.

Built-in or default classifiers

Classification for basic use or testing: Offensive Language, Resumes, Source Code, Targeted Harassment, Profanity and Threat.

Custom classifiers

- Must be activated, which requires 7-14 days for basic analytics.
- Training requires 50-500 positive samples of seed data.
- Up to 24 hours of seed data top be processed.
- Up to 10,000 positive and negative samples for testing.

When to use trainable classifiers versus sensitive information types

Sensitive information types match specific types of data according to their schemes
Trainable classifiers look for exact matches they are trained via machine learning algorithms
Both features require E5 licensing plans.

Comparison	Sensitive information types	Trainable classifier
Detecting	Scheme	Algorithm / machine learning
Detecting content	Information in items matching the scheme	Files, emails, information, Teams Chats with similar information to the training documents
Custom	Yes, via programming	Yes, via training
Examples	Credit Card, IP Address, Azure Keys	Invoices, Offensive language, Tread harassment, Employee Agreements, NDA Forms

Train a custom trainable classifier

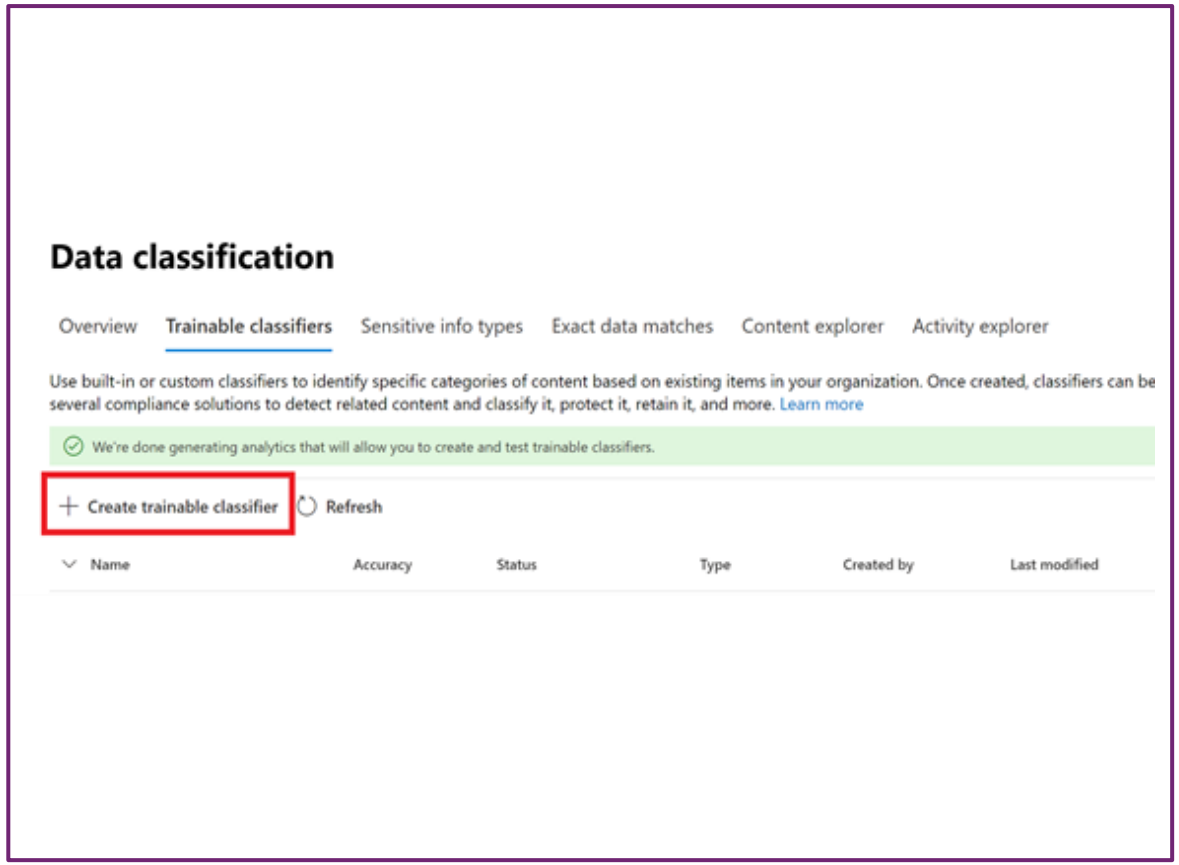
Training the algorithm to detect files.

Libraries for training the classifier should not include unwanted documents not relevant for the trained classifier.

Multiple types of information require multiple trainable classifiers.

Training requires approximately 1 hour for every 1000 items.

After publishing, the trainable classifier can be used in sensitivity labels or retention labels.



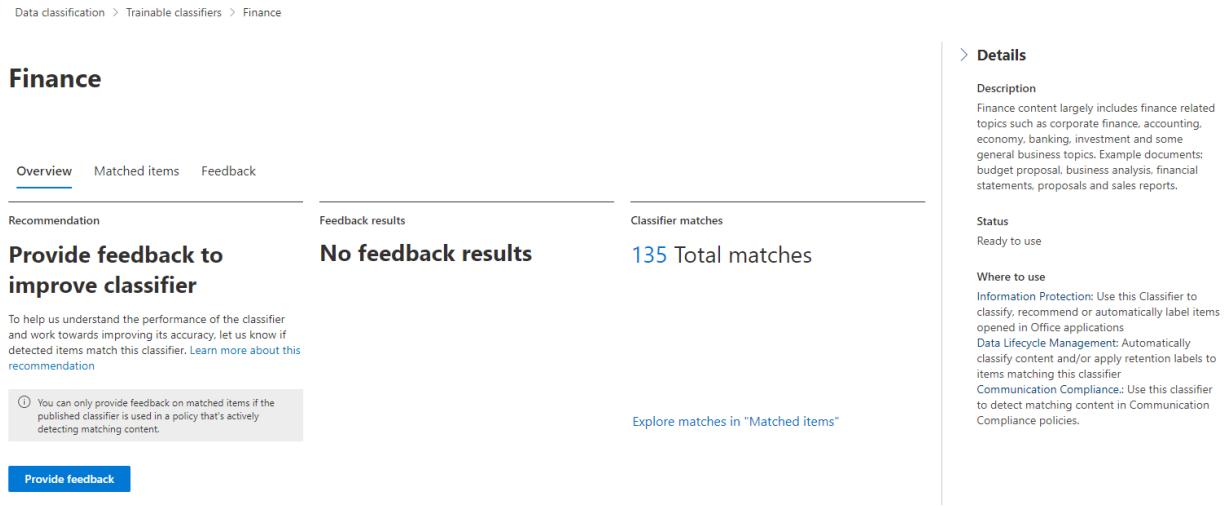
Verify a trainable classifier is performing properly

Trainable classifiers need to be monitored to determine if they work as expected.

Some additional indicators for its performance are:

- True and false positives
- Low number of matches
- No matches
- A very long time for a match

Any of these can indicate, that a refinement of the item base and a retraining of the classifier is required to make it work as expected.



Apply sensitivity labels to Microsoft Teams, Microsoft 365 groups, and SharePoint sites

Labels may be published for use in Microsoft 365 Groups, Microsoft Teams, Yammer communities and SharePoint sites.

You must activate labels in the tenant level via the Azure AD PowerShell module before they can be used.

Sensitivity labels can be manually assigned to SharePoint Sites and Teams sites through the creation settings.

Labels can be changed through the properties of existing SharePoint sites or Microsoft Teams.

A sensitivity label is applied to a Group in the Azure portal through Azure services > Groups > properties.

Create or extend existing sensitivity labels to Azure Purview

Data type	Sources
Automatic labeling for files	Azure Blob Storage Azure Files Azure Data Lake Storage Gen 1 and Gen 2 Amazon S3
Automatic labeling for schematized data assets	SQL server Azure SQL database Azure SQL Database Managed Instance Azure Synapse Analytics workspaces Azure Cosmos Database (SQL API) Azure database for MySQL Azure database for PostgreSQL Azure Data Explorer

Sensitivity labels created before Purview release, don’t include Purview Data Map locations by default.

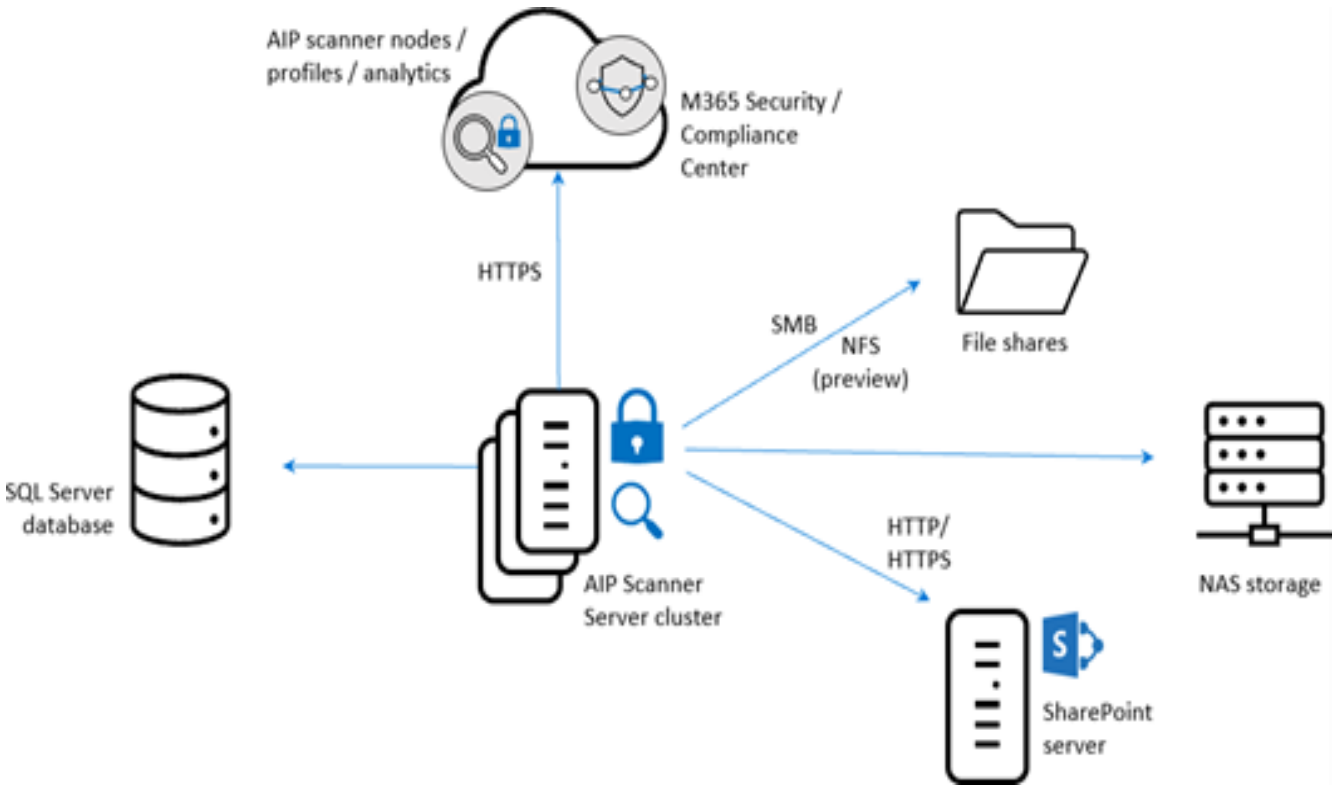
Extending of labels is required to integrate new Azure Purview source locations.

Afterwards use the Get-PolicyConfig cmdlet to check if “PurviewLabelConsent” is “True”

Labeling for SQL database columns in SQL Server Management Studio (SSMS) has limitations, because Microsoft Purview and labeling in SSMS are separate processes that don't interact with each other.

Use labeling through schematized data assets in Microsoft Purview and not local labeling in SSMS.

Plan on-premises labeling



The sensitivity labels functionality is natively only available in the user context or in the Microsoft 365 services for auto-labeling

This can be expanded to on-premise through hybrid configuration of the Unified Labeling Scanner.

The Unified Labeling Scanner has the following
requirements for setup:

- Windows server 2016 / 2019 with UI
- A SQL server installation
- An ~~Azure AD~~ / Microsoft Entra ID token
- AD Service accounts

Configure on-premises labeling for the unified labeling scanner

The Unified labeling scanner can be used for different operational scenarios, some of them include:

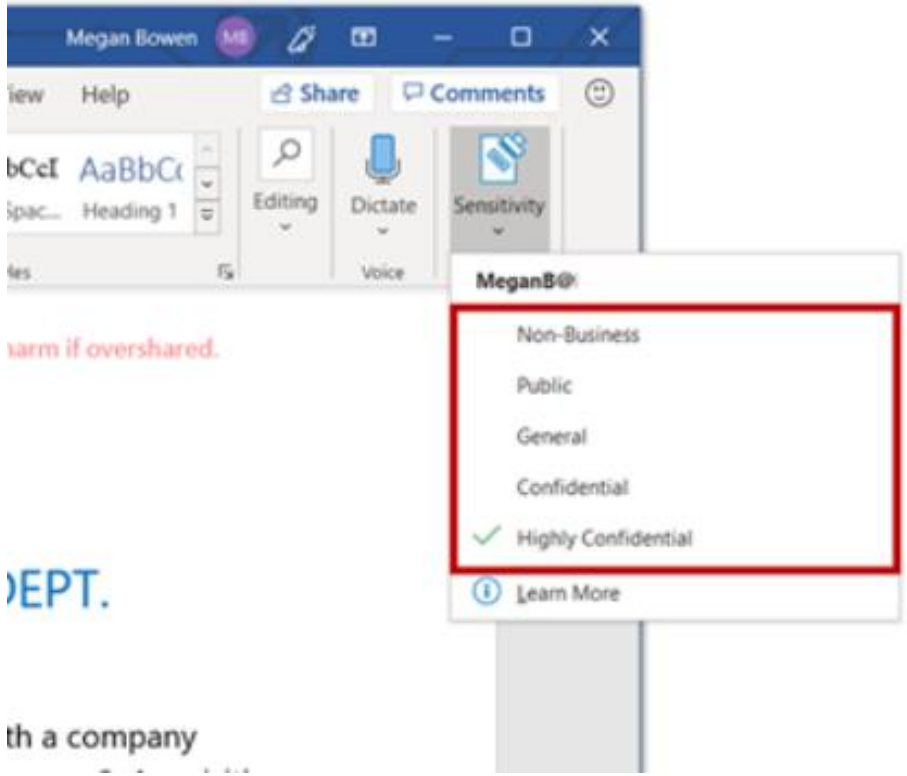
Scan for a report only to know your data: Run the scanner in discovery mode only to create reports that check to see what happens when your files are labeled.

Run the scanner to find and discover files with sensitive information: Run the scanner to discover files with sensitive information, without configuring labels that apply automatic classification.

Run and apply labels: Run the scanner automatically to apply labels as configured.

Specific scan only a few files types: Define a file types list to specify specific files to scan or to exclude.

Apply protections and restrictions to email and files



It is important to apply a sensitivity label to an email as well as the contained files.

The email is only the container of the files with it´s own separate protection configuration for the attachments.

You can apply a sensitivity label to an email manually in the Outlook Desktop app. Create an email and before sending the email an option is available in the ribbon menu to assign a sensitivity label.

You can create a default sensitivity label for both the documents and emails using auto-apply publishing through the Compliance center.

Apply protections and restrictions to email and files (continued)

It´s possible to apply a sensitivity label to a file through the following options:
Unified Labeling Client (Windows, Mac), Native Office Desktop Apps (limited), Mobile Office Apps (iOS, Android)
You can:

- choose a label for external and internal sharing
- Assign a label without a Microsoft Office client
- Apply a default sensitivity label to a SharePoint document library
- Set a default label via sensitive information
- Apply sensitivity labels with Microsoft Defender for Cloud Apps

Monitor label performance using label analytics

The sensitivity labels reports are available in the Microsoft Purview portal
The Label analytics tool is provided through the Azure Portal and requires an Azure subscription.
Monitoring and analysis is also available through Microsoft Sentinel.

PASS ETT | SESSION ONE [B]

Configuration, Eecution, Review and Analysis of Retention Labels

Data Lifecycle management overview (continued)

Category	Examples
Personal data	Name, e-mail address, telephone number
Sensitive personal data	HR data, Health data, ethical origin, Union membership
Product data	Brand, pictures, tenders, product descriptions, internal IP
Authentication data	System generated data like username, MAC address, IP-address
Log files with system accesses	License data, log files, Telemetry, diagnostic data

Retention Strategy

The goal of a retention strategy is to fulfill the requirements of laws, internal compliance policies and business regulations.

Creating a strategy includes several steps:

- Know the data for an organization.
- Know the requirements an organization must fulfill.
- Create a retention plan.

Retention Tools in Microsoft 365

Different tools to retain data:

Retention Labels

Label Policies / Retention Label Policies

Auto-apply Retention labels

Retention Polices

Record Management

Retention Tags and Retention Policies

In-place eDiscovery & hold

eDiscovery

To understand the right tools, know:

Where is the data stored today and in the future?

Which tools are used in the Microsoft 365 environment?

Which Retention requirements exist?

Which requirements does the Adoption Team have?

Are there other Systems like a DSM on SharePoint to retain and fulfill all requirements?

Which requirements do you want to fulfill in Microsoft 365

Legal Hold to fulfill the requirements

Legal hold for retention requirements

Combination of location-based Retention with Retention Policies and Retention Labels required

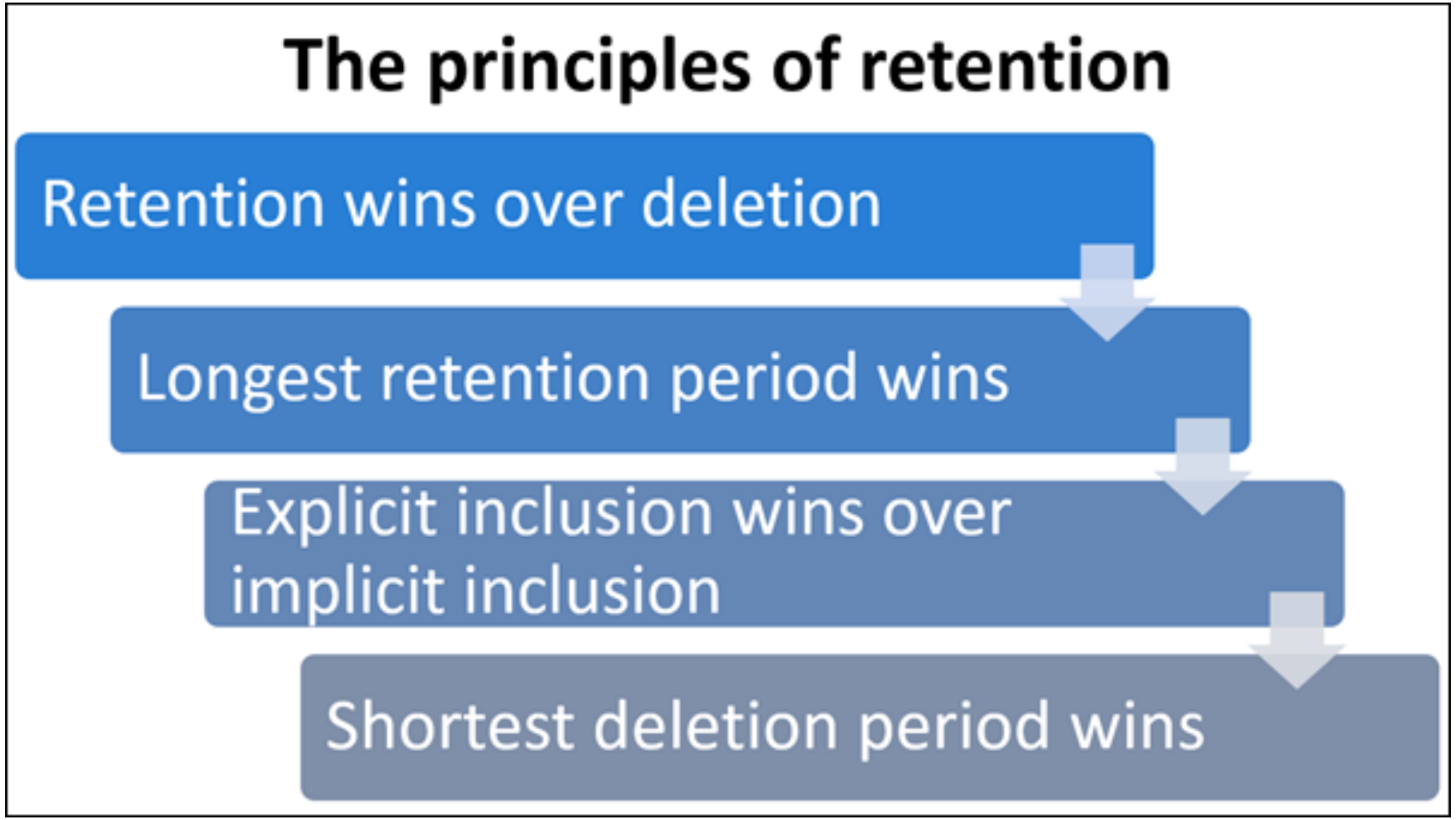
Legal Hold used to preserve stored data for legal cases across the environment.

Requirements like **SEC 17a4 (US)** or **GobD (Germany)** need the use of regulatory records or a regular record to prevent a deletion and store files for a certain period.

Important to consider data of leaver and movers.

Recommendation for adoption: Keep it simple.

Data Lifecycle management overview



Configure retention policies

Type	Based on	Travel with the document	Data lifecycle	examples
Retention Policy	Location, product	no	yes	Teams Chat, Junk Folder, SharePoint Site
Retention Label	Single file or email, library, or list	yes	yes, better	Email, document in a library

Differences between retention policies and retention labels

Retention Policies are focused on service locations.

Retention Labels are focused on individual items.

- Retention labels are created for certain kinds of business data.
- Label policies are used to publish retention labels.

Configure retention labels

Document	Label	Method	Deletion	False
Invoice	6 years + delete	Last modified on 1. May 2020	Delete after 6 years in May 2026	The file will be deleted 6 months too early.
Invoice	7 years + delete	Last modified on 1. May 2020	Delete after 7 years in May 2027	The file will be deleted 5 months later. But it´s stored for the full time to fulfill the legal requirement.

Retention period calculation

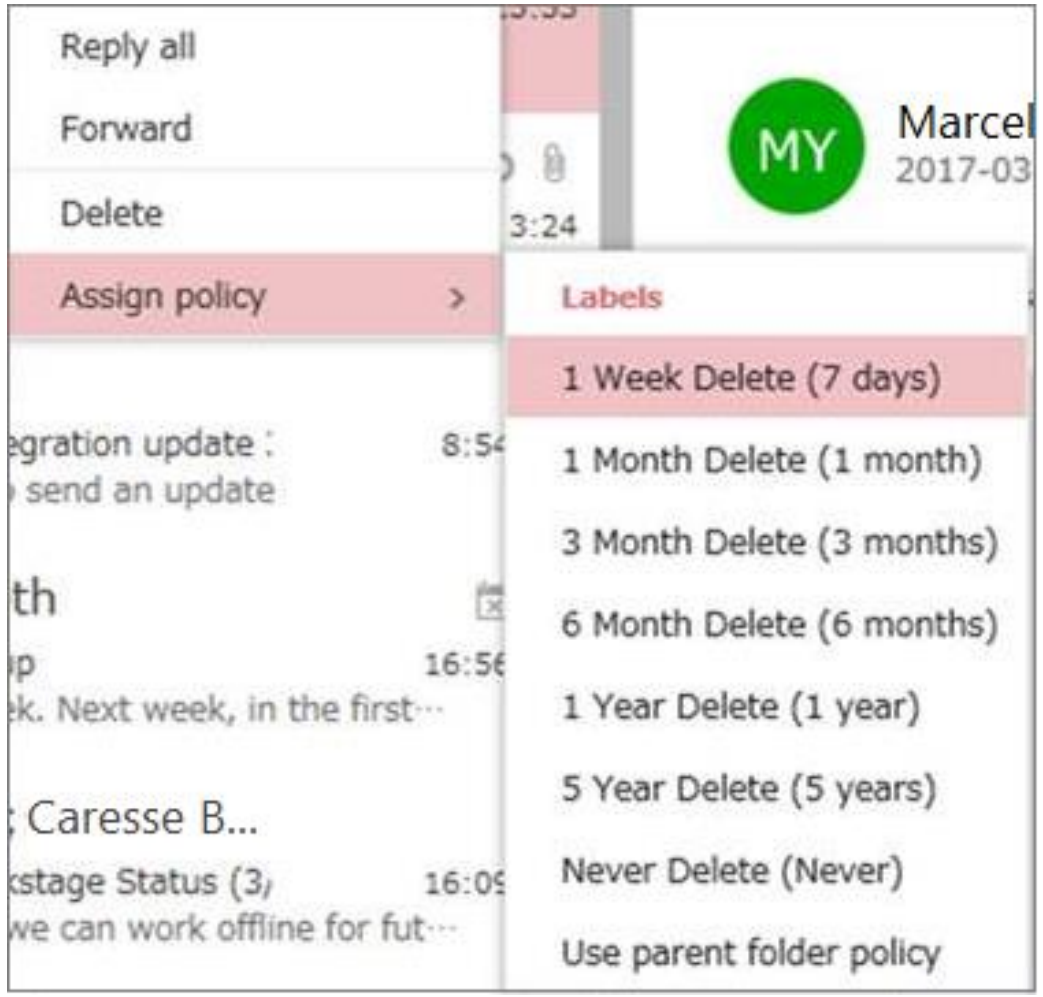
Retention label count starts when a label is applied.

Several legal and regulatory requirements start at the end of a year.

Exact requirements must be known to fulfil them with required accuracy.

A combination of retention labels and retention policies is required.

Configure manual retention label policies (continued)



- Within Outlook a retention label can be assigned to an email through the *assign policy* ribbon
- The retention label is then visible within the email
- You can set the default retention label for an entire Outlook folder through the folder's properties
- Within SharePoint you can set the retention label to a file through the file/folder/document libraries properties
- Retention labels for Onedrive can be set through the web version through the file properties

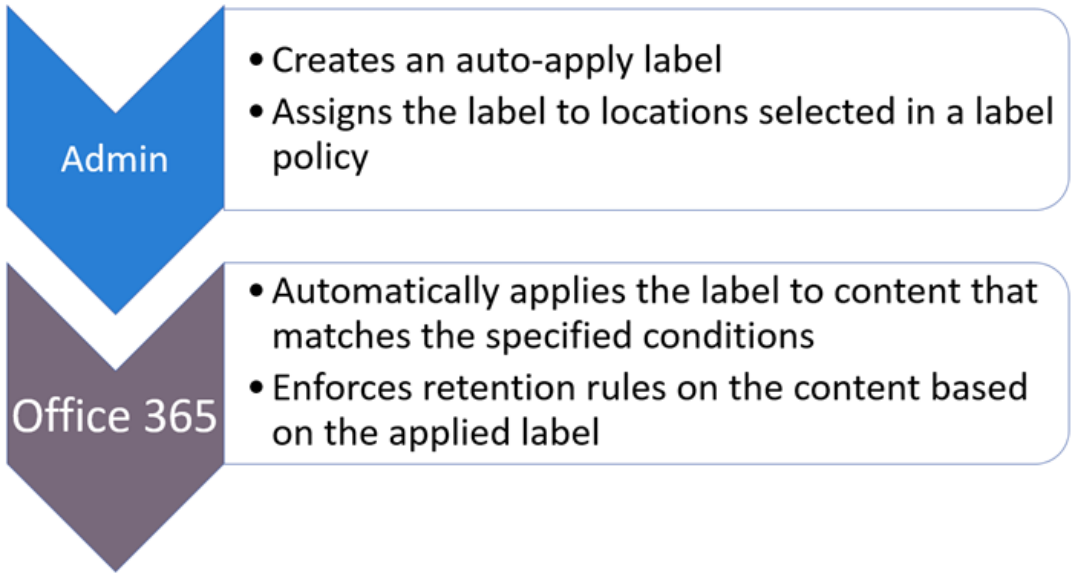
Configure auto-apply retention label policies

Auto-apply labels can apply retention labels automatically to:

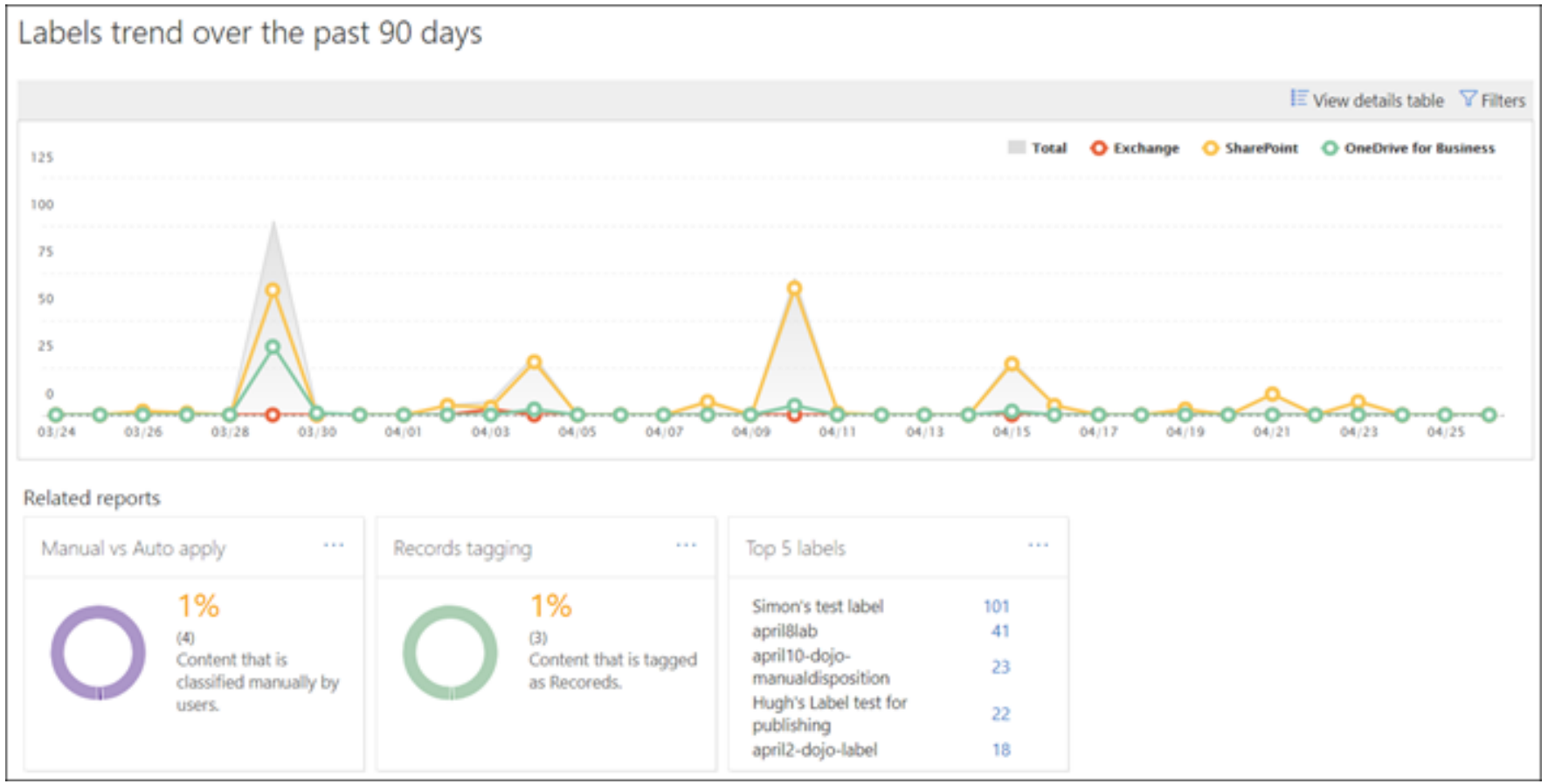
Content that contains sensitive information that match sensitivity templates.

Content that contains specific words or phrases, or properties. This is specified by the Keyword query language.

Or trainable classifiers such as source code, offensive language or resumes.



Manage and remediate Data Lifecycle Management



Data Lifecycle Management cards on the Data classification page:

Top sensitive info types

Top retention labels applied to content

Locations where retention labels are applied

Configure preservation locks

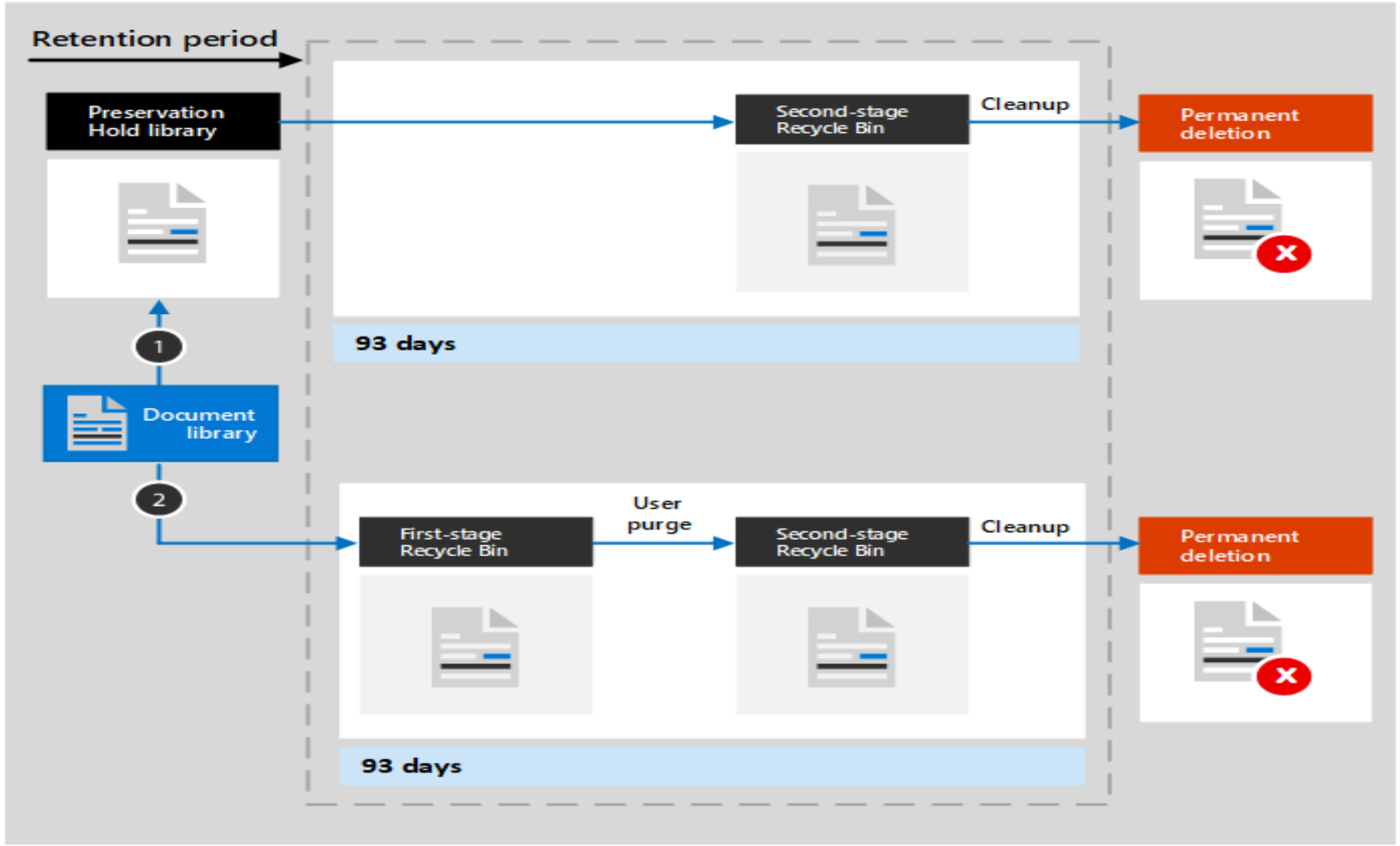
- When a retention policy is locked:
  - No one can disable the policy or delete it
  - Locations can be added but not removed
  - You can extend the retention period but not decrease it

- When a retention label policy is locked:
  - No one can disable the policy or delete it
  - Locations can be added but not removed
  - Labels can be added but not removed

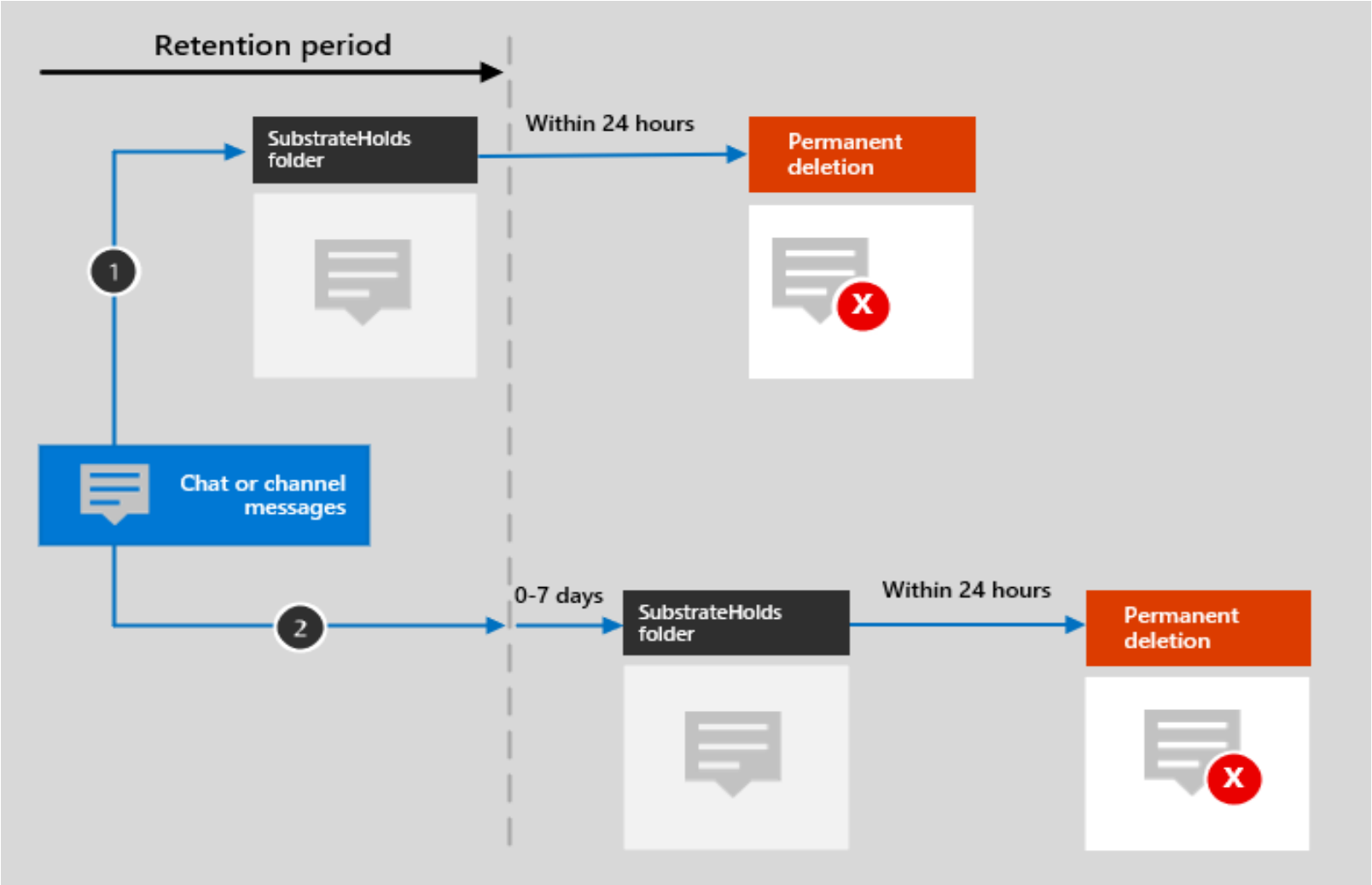
To prevent accidentally locking a policy, use PowerShell configure a preservation lock.

Important: A preservation lock cannot be reversed.
You will never be able to delete the locked policy again!

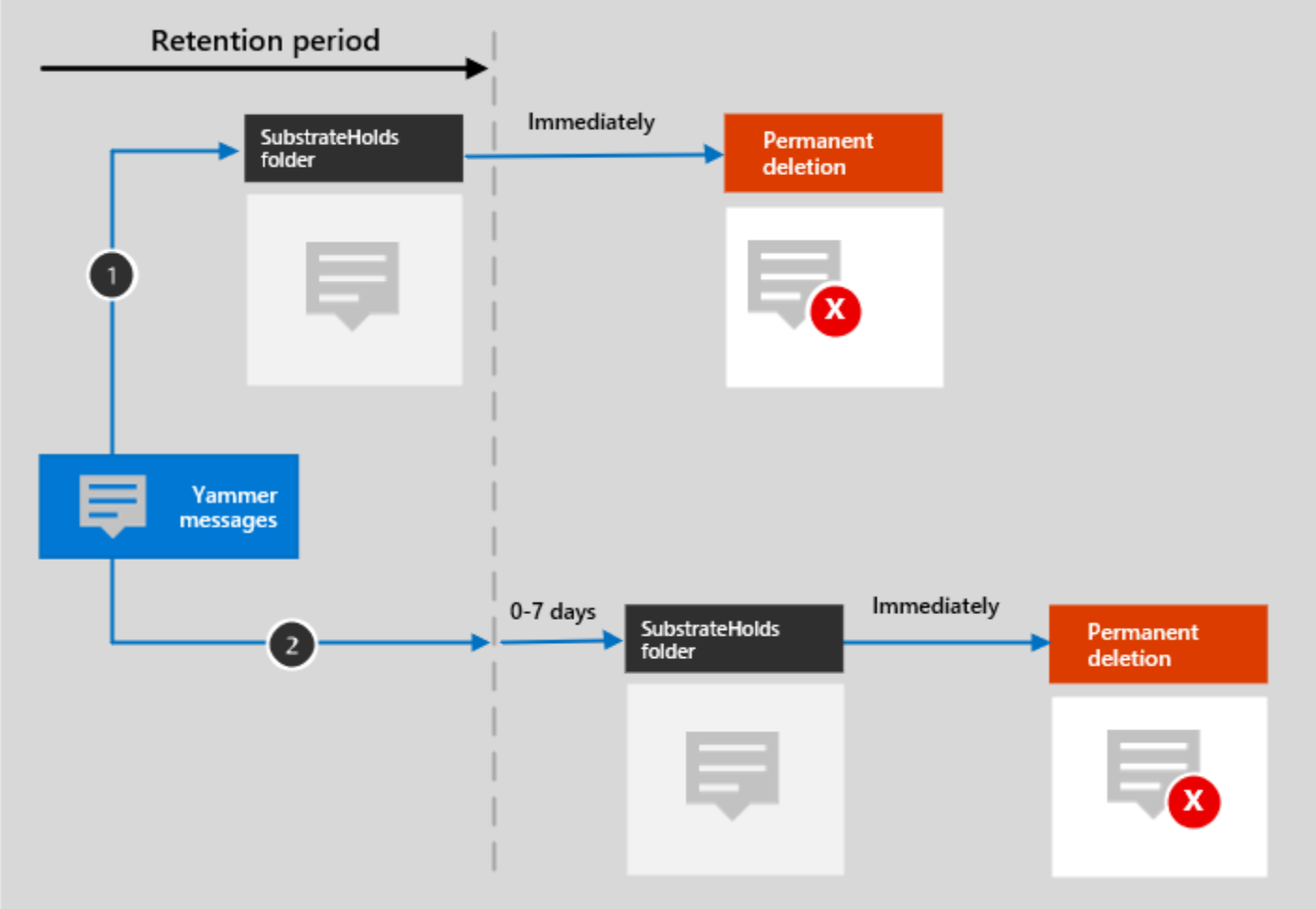
Explain retention in SharePoint Online and OneDrive



Explain retention in Microsoft Teams



Explain retention in Microsoft Yammer



Recover content in Microsoft 365 workloads

Recovery options for users and administrators

- Users can recover documents and files in OneDrive via their recycle bin.
- Users can restore items on SharePoint Online document libraries via the recycle bin.
- Users can restore versions in the SharePoint Online and OneDrive portal.
- Users can restore versions via Office Online or via Microsoft 365 apps for enterprise.
- Users can restore their entire OneDrive content.
- Administrators can access the preservation hold libraries and recover data.

Activate archive mailboxes in Microsoft Exchange

Messaging records management (MRM)

- Former feature for retention in Exchange
- Used for archiving of mailbox content
- Consists of different features:
  - **Retention policies** to publish retention tags to use.
  - **Retention policy tags (RPTs)** for default folders.
  - **Default policy tags (DPTs)** for all untagged items.
  - **Personal tags** for manual assignment.
- Mailbox Folder Assistant (MFA) processes retention tag actions.

Archive Mailboxes

- Additional mailbox for archiving.
- Retention tags move items to archive.

Apply mailbox holds in Microsoft Exchange

Mailbox Holds target Exchange Mailboxes

Protect content of Exchange mailboxes against deletion.

Two types of mailbox holds available:

- Litigation Holds set on a mailbox level to protect all content from deletion.
- eDiscovery Holds created in cases to prevent mailbox content matching search criteria from deletion.

Recover content in Microsoft Exchange

Exchange Mailbox content recovery with eDiscovery cases
Recovery of content from active and inactive mailboxes.
Results can be previewed and exported in .pst files.
Different eDiscovery permissions required for case creation, preview and export.

Records management overview

A record is a document or other electronic or physical entity in an organization that serves as evidence of an activity or transaction performed by the organization and requires retention for some time period.

Different actions that possibly needs to be prevented:

- change the assigned Retention Label of an item
- change of the content of an item or it's metadata
- deletion of a file or remove a retention label from a file
- moving a file between containers (SharePoint libraries for example)

Four steps to plan and decide:

- Decide to add a Record to a label or not
- Decide to add a Record or a Regulatory Record to a label
- Configure the Record to a label
- Decide to have an auto labeling functionality and configure it

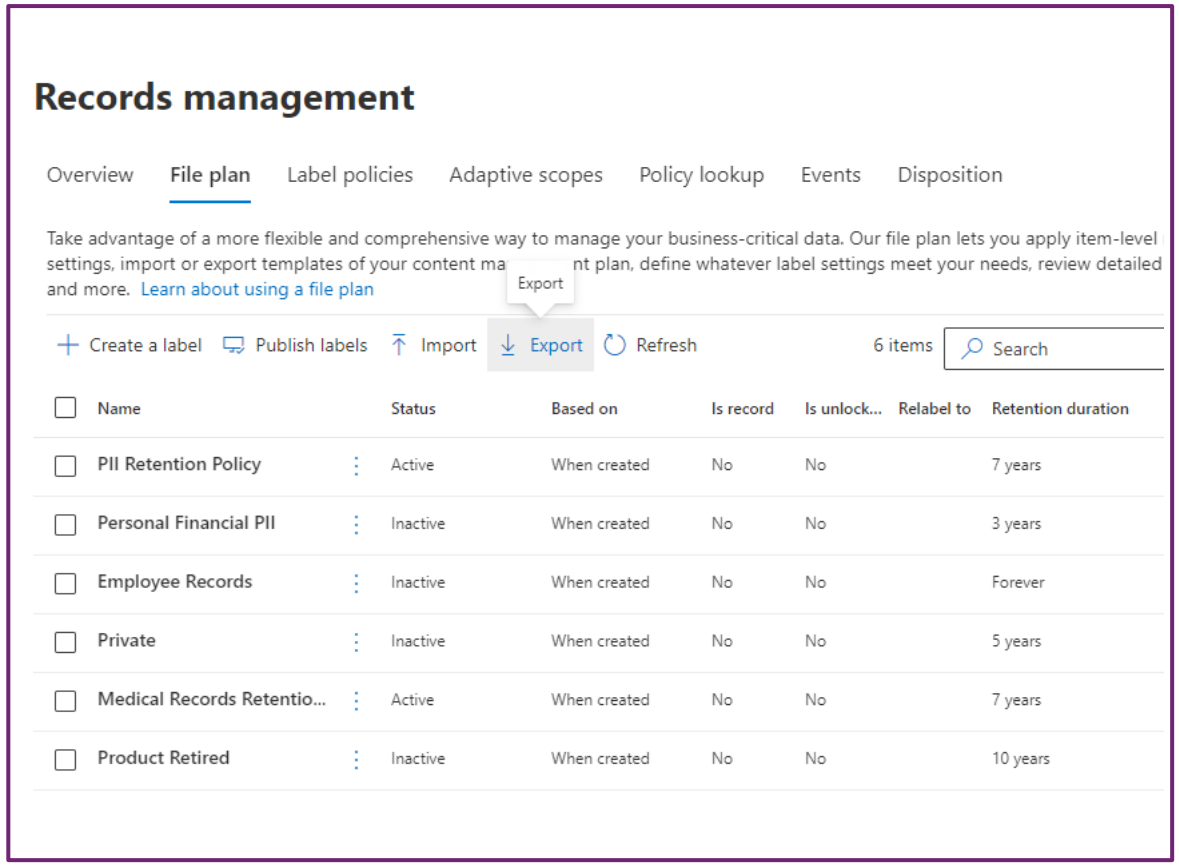
Import a file plan

File plans group the creation of labels, auto-apply label policies and additional metadata tags together.

A file plan can help you manage how to dispose of files after the retention period

Export file plans as CSV-files and import them into other tenants

Use trainable classifiers to identify content it should match



Configure retention labels

Retention Labels used for Records Management.
Activation of records management via PowerShell required.
Different options available to configure retention labels after activation.

Option	Use to..
Without a Record	Does not add a record to a retention label.
With a Record	A simple record on a retention label restricts options for users to modify labeled items.
With a Regulatory Record	A regulatory record restricts options for users to modify labeled items more strictly. A warning will be displayed when a regulatory record is created and when the retention label will be applied to a file.

Demo Time

Fika (rast) [15 min]

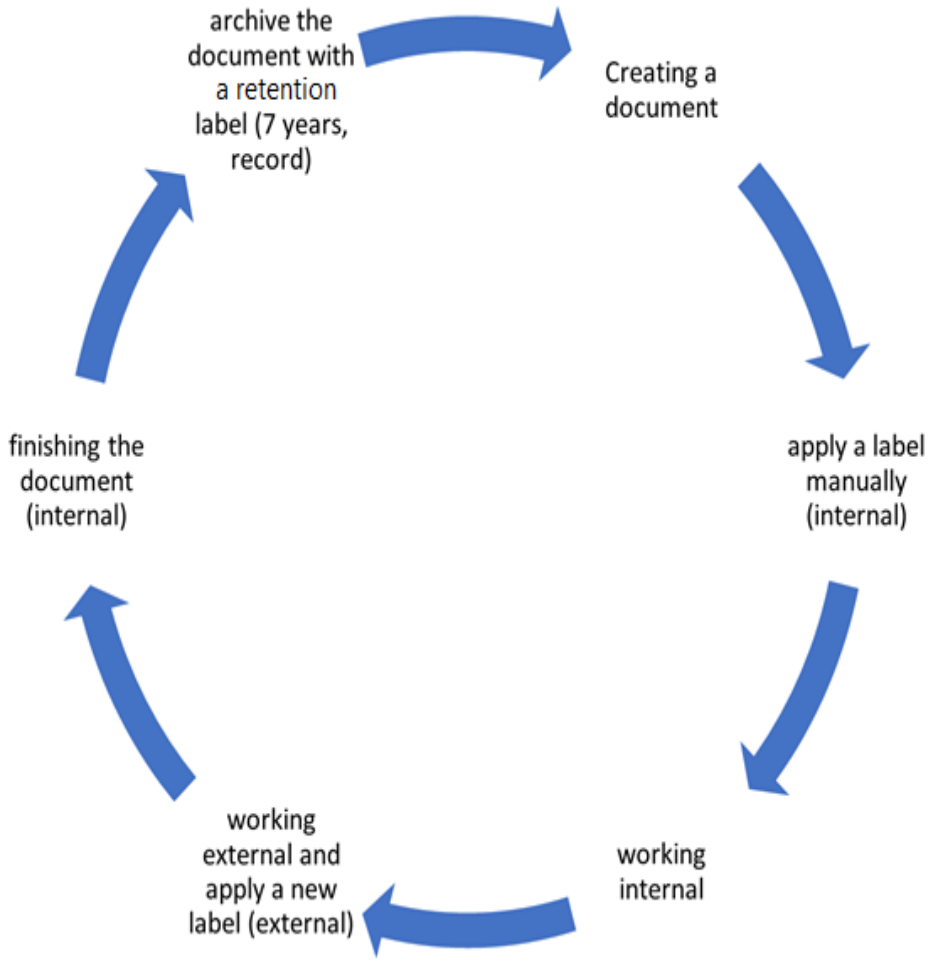
Create image of Swedish Fika with Kaffe Te and Cinnamon bun

Time for Lunch Break / Middag [45 min]

PASS TVÅ | SESSION TWO

Analysis of Sensitivity Content and the Role of Activity Explorer

Basics of sensitivity labels

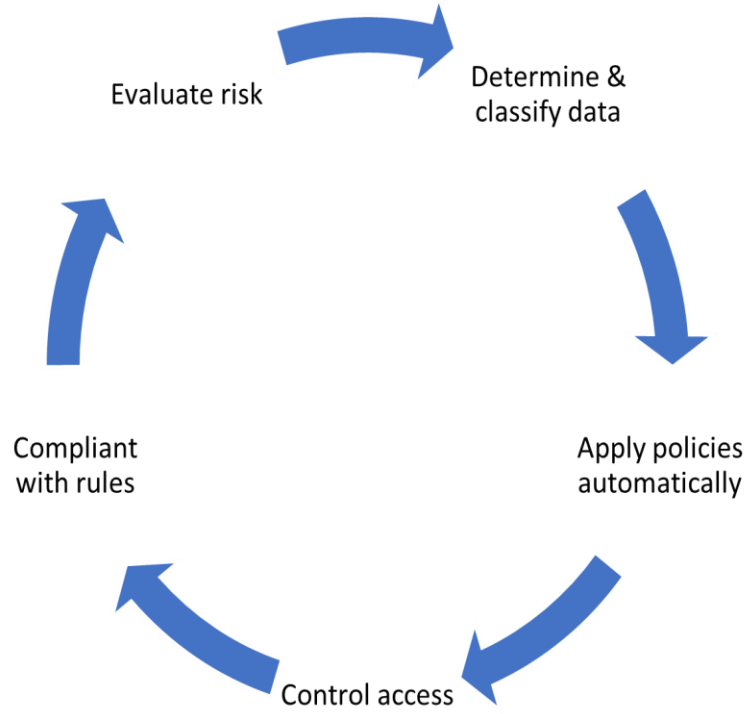


Sensitivity Labels are a solution to handle the lifecycle of company data with classify, encryption, and access management in Microsoft 365.

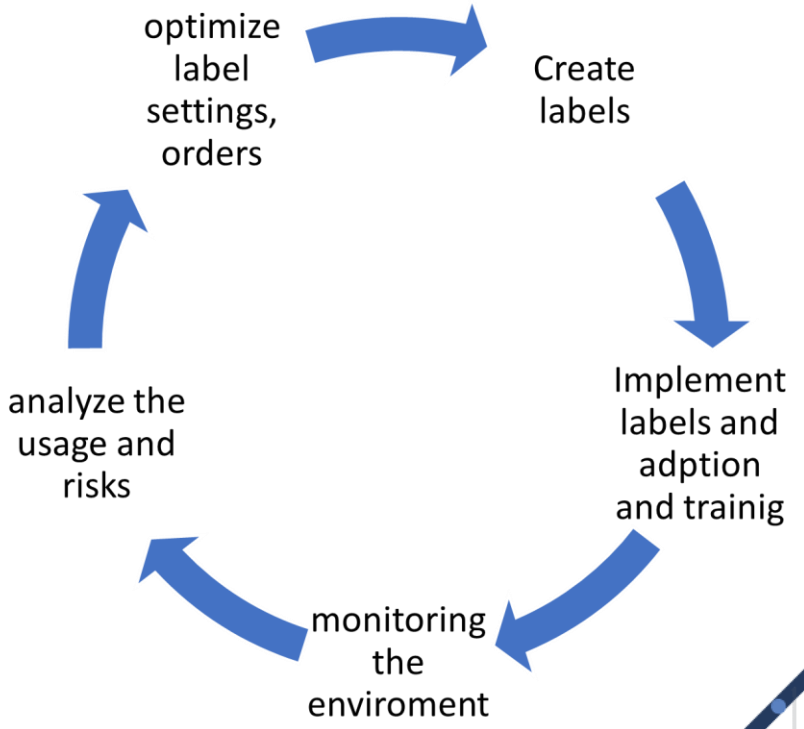
Sensitivity Labels move with a labeled document and can be applied to locations such as SharePoint sites,
Microsoft 365 groups, documents, and e-mails or may be applied to Azure Purview assets.

Administer the sensitivity labeling lifecycle

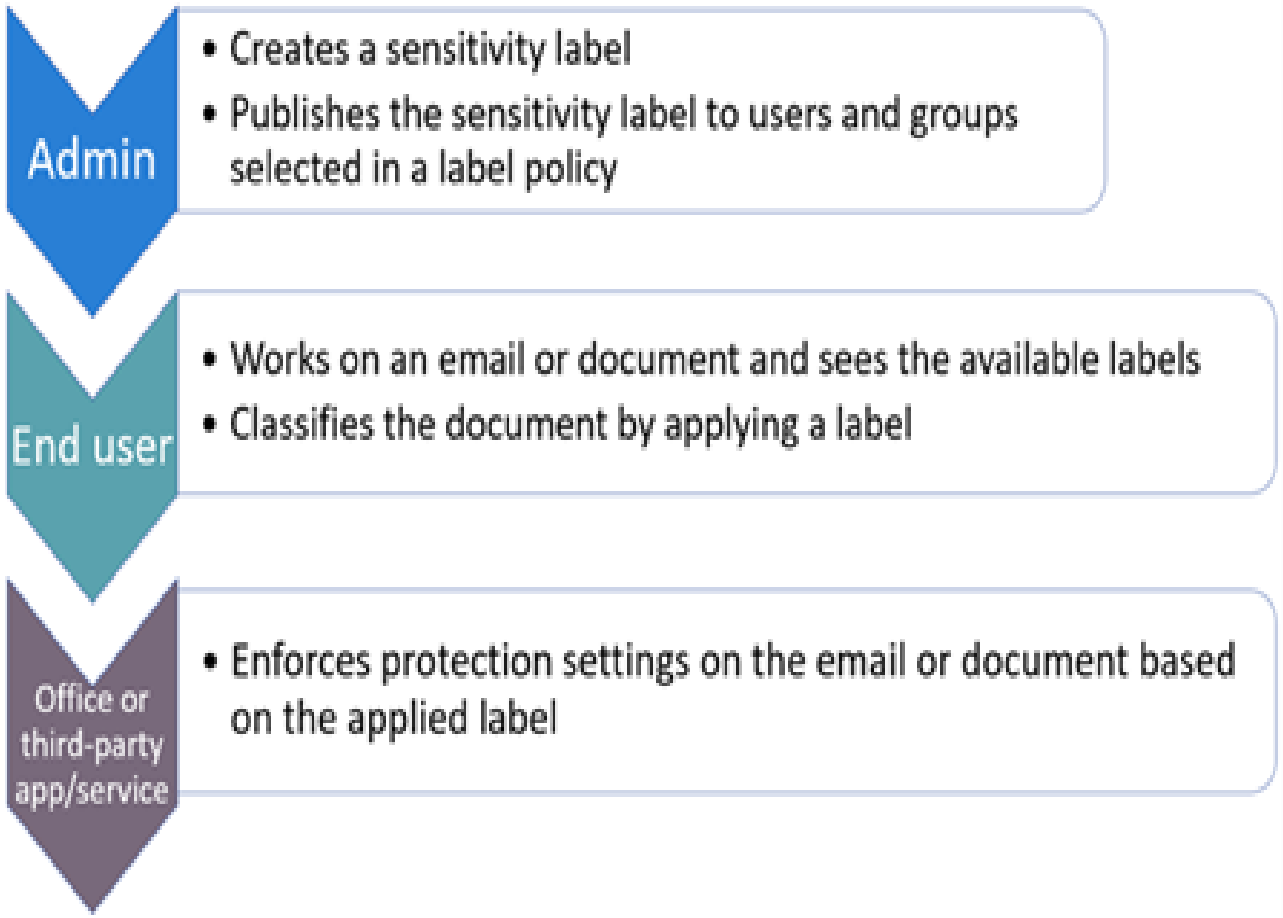
The lifecycle of the sensitivity labeling for the administrator:



The lifecycle of the sensitivity label:



Configure sensitivity labels



Sensitivity labels can be created by an admin through either the Purview portal UI or PowerShell.

Select a scope of either Files & emails, Group & sites or Azure Purview assets.

Items can be auto-labeled on either Sensitive info types or trainable classifiers.

PowerShell provides additional options for multilanguage support.

Configure sensitivity label polices

Sensitivity labels must be published before they can be applied to items.
Groups must be email enabled like security groups or distribution lists.
Microsoft Teams is a use case.
It is a best practice to name the sensitivity label using a naming guideline (like "2021-Jan-HR-HRLabels-RK").

Introduction to Microsoft 365 encryption

What is encryption in Microsoft 365?

- Encoding plain text into cipher text.
- Decryption requires encryption keys.
- Control access to authorized users or machines only.
- Differentiation between 'data at rest' and 'data in transit'

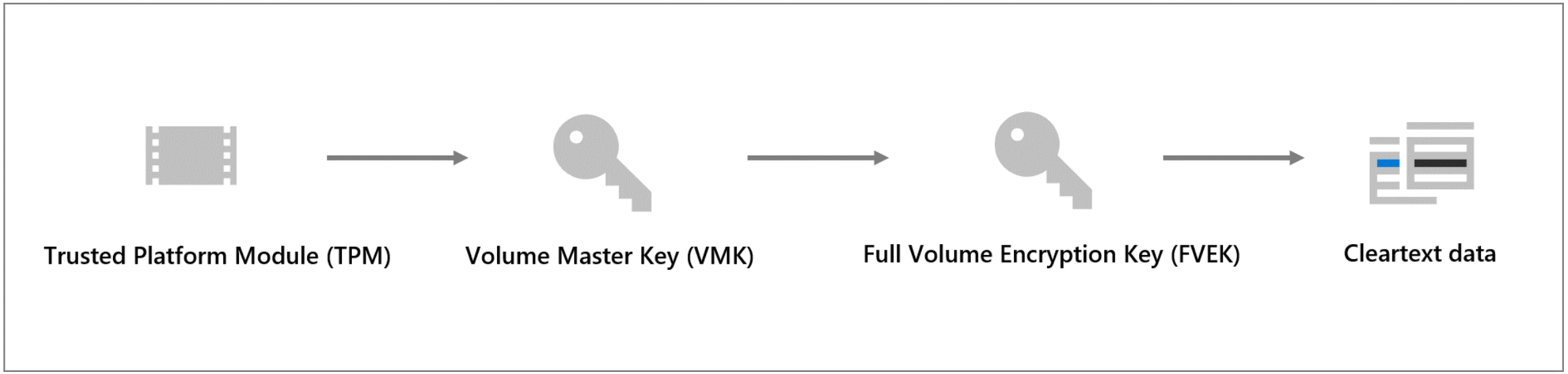
Data at rest

- Files saved on computers and mobile devices
- Documents and files saved in SharePoint Online and OneDrive
- Mails saved in Mailboxes in Exchange Online

Data in transit

- Documents and files accessed in SharePoint Online and OneDrive
- Mails transported between servers
- Shared files and conversations in Teams meetings

Learn how BitLocker encrypts data-at-rest



Service encryption in Microsoft Purview

Available encryption options:

Service Encryption, Azure RMS, and S/MIME

- Service Encryption, optionally with Customer Key
  - Encryption of all data saved in a Microsoft 365 tenant.

- Information Rights Management (IRM) with Rights Management Service (Azure RMS)
  - Protection of individual documents, files and emails

- Secure/Multipurpose Internet Mail Extensions (S/MIME):
  - Encryption and digital signing of email messages and attachments only.

Encryption Key Management:

- Service Encryption - Microsoft managed Keys or Customer Keys
- Azure RMS - Microsoft managed Keys, BYOK, DKE, or HYOK

Implement keys for Service Encryption

Service Encryption

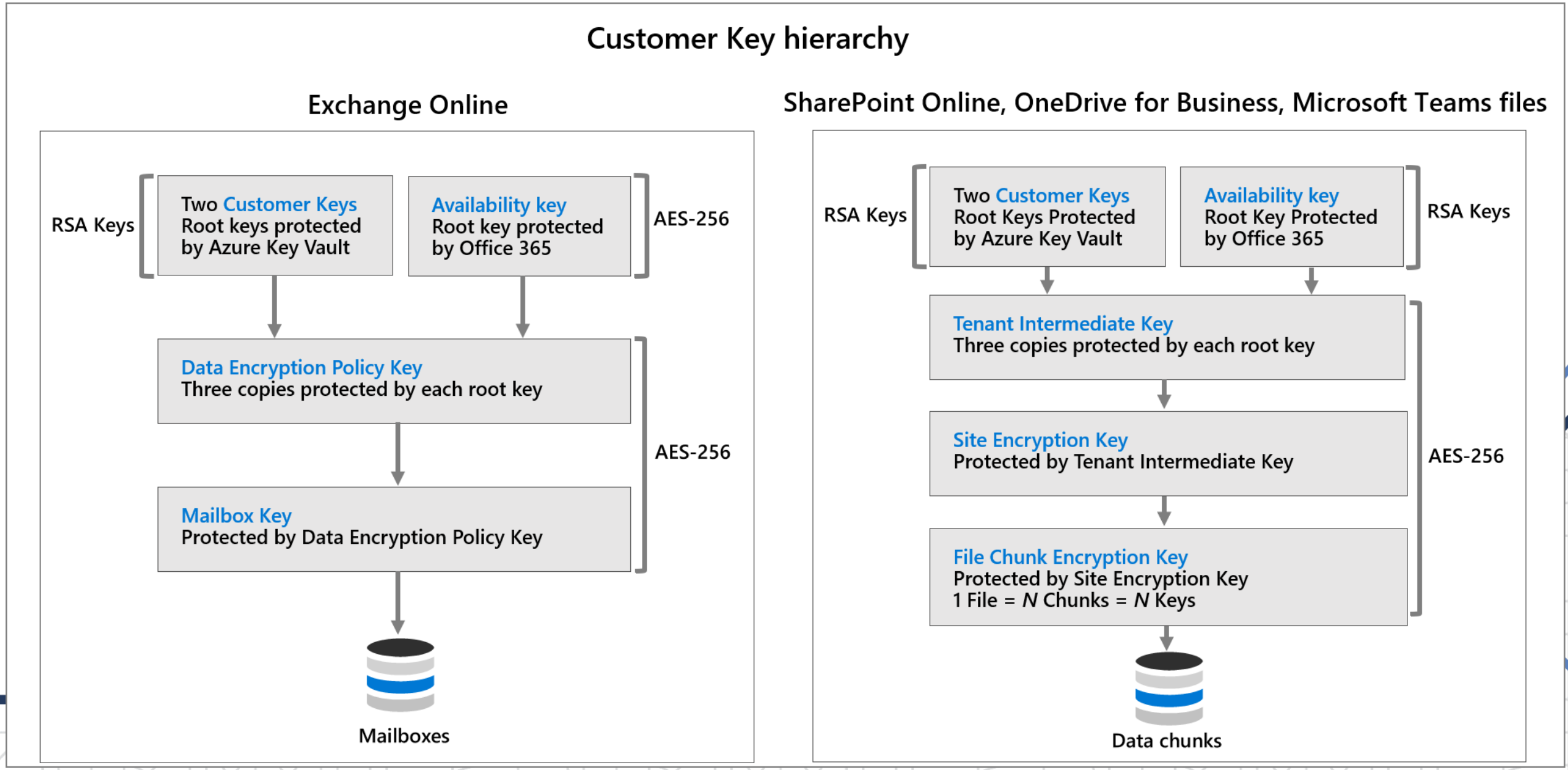
Configured via Azure Key Vault (AKV) and Data Encryption Policies (DEP)

DEPs associate customer keys from AKVs with mailboxes or services

- Exchange Online: 50 DEPs per tenant
- SharePoint Online: One DEP per geo location or tenant

Note: Activation of Service Encryption requires intervention of the Microsoft support.

Customer key management using Customer Key



Encrypted data in transit

Data-in-transit scenarios include:

When a client machine communicates with a Microsoft server.

When a Microsoft server communicates with another Microsoft server.

When a Microsoft server communicates with a non-Microsoft server (for example, Exchange Online delivering email to a third-party email server).

Implement Microsoft Purview Message Encryption

Microsoft Purview Message Encryption uses IRM and RMS templates
Default configuration

Default configuration named “OME Configuration” is available

Modifications apply to all users

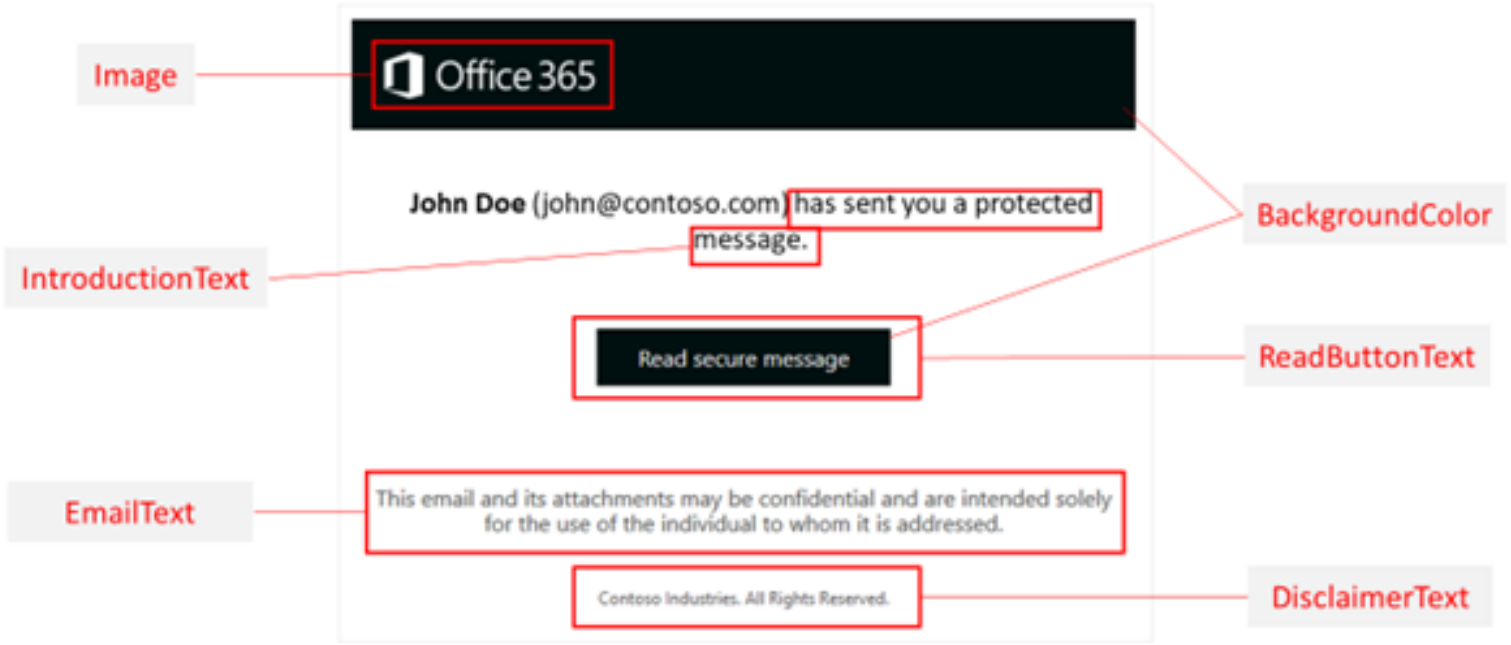
Customization of the Outlook, *OME portal,* and functionality is possible

Branding templates

Only one branding template is available for all with basic OME

Implement Microsoft Purview Advanced Message Encryption

- Additional Features
  - Multiple branding templates
  - Message Expiration added (1 – 730 days)
  - Customized templates for different groups of users



Use message encryption templates in mail flow rules

Configuration via Mail Flow Rules

Different OME configurations are assigned via mail flow rules

Possible use cases:

- Individual departments.
- Different products.
- Different geographical regions.
- Determine whether emails can be revoked.
- Determine whether emails sent to externals expire after several days.

Demo Time

PASS TRE| SESSION THREE

Microsoft Purview {Azure} P1:
Account, Data Sources, Scan, Search,
Browse, Glossary

Use Microsoft Purview for managing your data estate at large

Microsoft Purview is a unified data governance solution that helps you manage and govern your on-premises, multi-cloud, and software-as-a-service (SaaS) data.

It does data governance at scale, because it's a fully automated service that intelligently performs data discovery, data scanning and access management.

It also provides a holistic map for providing many insights of your data mesh architecture

Use Microsoft Purview for managing your data estate at large

When implementing Microsoft Purview, it's recommend not introducing too much change and complexity quickly. The technical metadata is recommended to be the foundation.

You'll need to gather and organize this before you can make sense of it. After this, start with the basics:

- business terms,
- lists of authoritative data sources and databases,
- schema information,
- data ownership and stewardship,
- and security.

Slowly scale by involving more domain owners and data stewards. Also scale by adding more classifications and sensitivity labels. This improves the search experience and allows for better data access management.

For your custom metadata attributes, such as list of domains and application metadata, you could consider creating extra type definitions in Microsoft Purview using Purview's REST APIs.

Use Microsoft Purview for managing your data estate at large

When you envision a domain-oriented or more decentralized architecture, it's recommended to align your Microsoft Purview Collections and Glossaries with your data domains.

Collections in Microsoft Purview are used to organize assets and sources. You can use a Collection as a boundary for your assets and sources and align this with a particular domain.

You can do the same for your Glossary: create hierarchy structures within your glossary and align these with your domains.

Ask your domains to take ownership for creating relationships between your glossary terms and collection attributes.

This creates transparency over data ownership and improves your data semantics.

Relationship to Metadata Management

Data Governance has a strong relationship with Metadata Management.

Metadata is data about data.

It describes other data, providing a reference that helps you to find, secure and control data. Metadata also binds data together. It can be used for validating the integrity and quality of the data, routing or replicating it to the new location, transforming the data, and knowing its meaning are all performed through metadata. Metadata is also essential in democratizing data through self-service portals.

A good metadata management strategy grows organically. It starts simple and small by first identifying the most important areas. A good metadata management strategy also is supported with services and clear processes. To get started, it's good to be aware of the different metadata categories:

Relationship to Metadata Management

Business metadata is a category of metadata that describes all aspects used for governance, finding & understanding data. Some well-known examples are business terms, definitions, data ownership information, information about data usage and origination, and so on.

Technical metadata is a category of metadata that describes the structural aspects of data at design time. Some well-known examples are schema information, information about data formats and protocols, encryption and decryption keys, and so on.

Operational metadata is a category of metadata that describes processing aspects of data at run time. Some well-known examples are process information, execution time, information about whether a process failed, ID of the job, and so on.

Social metadata is a category of metadata that describes the user perspective of the data by its consumers. Some well-known examples are use and user tracking information, data on search results, filters and clicks, viewing time, profile hits, comments, and so on.

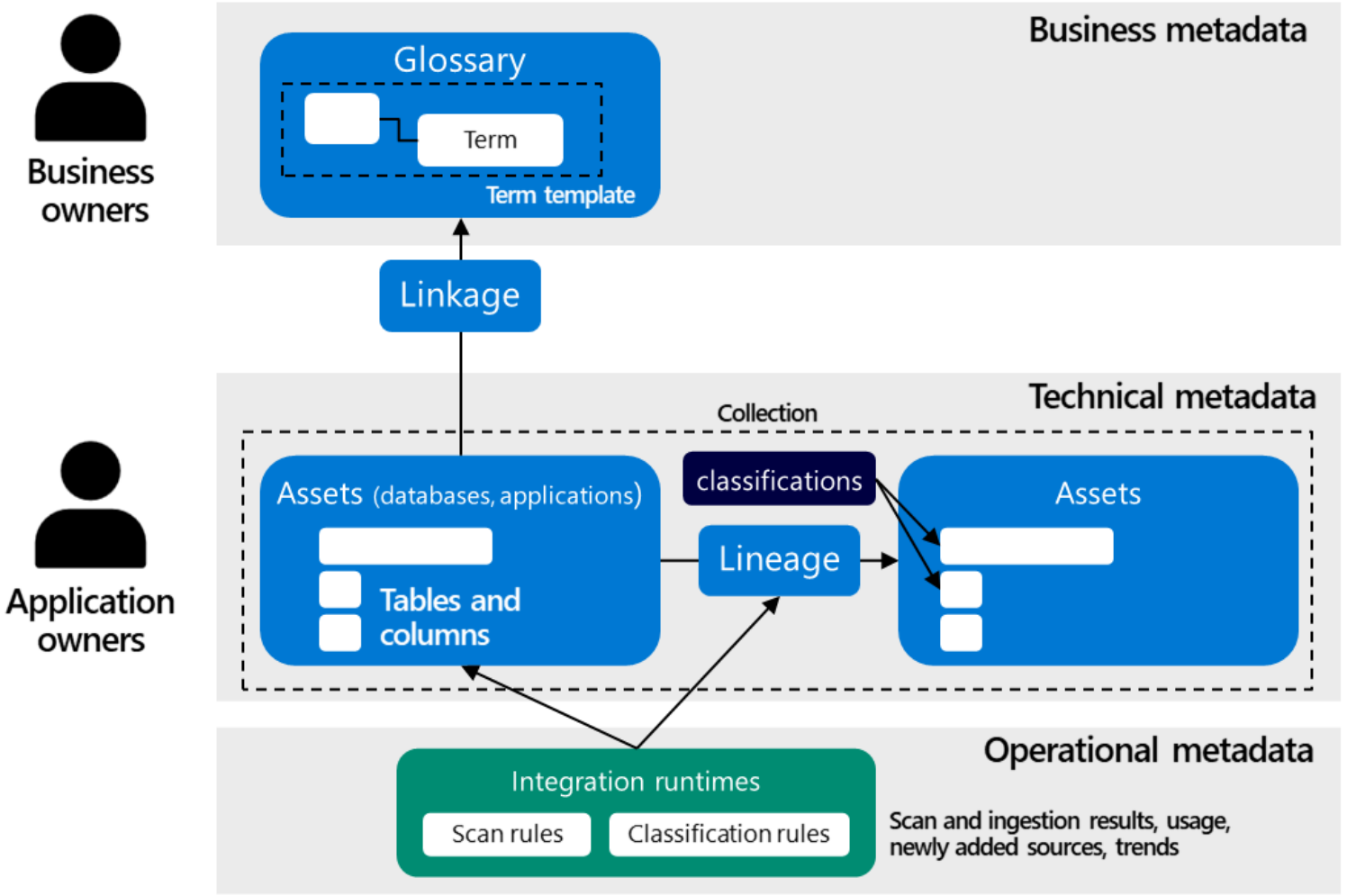
Relationship to Metadata Management

When it comes to decentralized data architecture, metadata management is mostly an organizational challenge. It's about finding the right balance between centrally managed metadata and federated managed metadata.

Metadata in this respect also has a strong relationship with **DataOps,** which is an advanced and collaborative data management practice focused on improving the efficiency of communication, integration, and automation of data flows between teams across an organization.

DataOps addresses some of the complexity associated with metadata management. It strikes a balance between central governance and domain ownership for metadata management.

Relationship to Metadata Management



After you know what metadata you need, you need to find a place for storing and processing metadata.

This brings you to Microsoft Purview.

Collections for organizing technical metadata

When planning your Microsoft Purview deployment and aligning your data governance activities, you need to define how technical metadata, such as data asset information, will be managed together. This grouping and the granularity of your technical metadata is what collections are for.

A collection is a logical container or a boundary in which your metadata, such as data sources, will be managed. When creating collections and placing them in a hierarchy, you need to make different considerations, such as your security requirements, governance structure and democratization needs

Collections for organizing technical metadata

The collection metadata, as you can see in the image from the previous section, mostly sits on a technical level. This also often implies a more technical-oriented data ownership. The contact persons assigned to your data assets are typically application- or database owners, knowing how these systems are designed. You can support these users also with additional roles in Purview:

Collection Admins can edit Microsoft Purview collections and their details and add sub-collections. They can also add users to other Microsoft Purview roles on collections where they're admins.

Data Source Admins can manage data sources and data scans.

Assigning users to roles enables them to maintain technical metadata themselves without any help from a central departments. Enabling them with these roles makes your metadata activities more self-service.

Glossary for capturing business knowledge

The glossary, contrasting to collections, sit on a business level. It is used to capture business knowledge that is commonly used, communicated, and shared in the organization. A glossary can help for improving an organizations overall business productivity and performance. It can also help to find data more easily.

Within Microsoft Purview, the Glossary supports several features:

- Supports adding business terms, including definitions, for capture relevant knowledge.
- Supports relationship between terms, such as synonyms and related.
- Supports hierarchies for better organizing your terms.
- Supports term templates for capturing additional custom attributes.
- Bulk import and exports of terms.
- Allows terms to be mapped to assets like a database, tables, columns etc.

Glossary for capturing business knowledge

Because business terms provide vocabulary for business users, it's also logical to assign business-oriented ownership. Purview uses data experts and data stewards.

Data experts are typically the knowledge holder. They know a particular domain or business unit very well.

Data stewards are often the people accountable. They oversee completeness and correctness.

You can support these users also with additional roles in Purview like:

Data curators a role that provides access to the data catalog to manage assets, configure custom classifications, set up glossary terms, and view insights. Data curators can create, read, modify, move, and delete assets. They can also apply annotations to assets.

💡 Did you know?

Capacity Units determine the size of the platform and is a **provisioned** (always on) set of resources that is needed to keep the Microsoft Purview platform up and running. 1 Capacity Unit is able to support approximately 25 data map operations per second and includes up to 10GB of metadata storage about data assets.

Capacity Units are required regardless of whether you plan to invoke the Microsoft Purview API endpoints directly (i.e. ISV scenario) or indirectly via the Microsoft Purview Governance Portal (GUI).

Note: With the introduction of the Elastic Data Map, you no longer need to specify how many Capacity Units that you need. Microsoft Purview will scale capacity elastically based on the request load.

vCore Hours on the other hand, is the unit of measure for **serverless** compute that is needed to run a scan. You only pay per vCore Hour of scanning that you consume (rounded up to the nearest minute).

💡 Did you know?

- Microsoft Purview has a set of predefined data plane roles that can be used to control who can access what.

Role	Collections	Catalog	Sources/Scans	Description
Collection Admin	Read/Write			Manage collections and role assignments.
Data Reader		Read		Access to catalog (read only).
Data Curator		Read/Write		Access to catalog (read & write).
Data Source Admin			Read/Write	Manage data sources and data scans.

💡 Did you know?

To scan a source, Microsoft Purview requires a set of **credentials.** For Azure Data Lake Storage Gen2, Microsoft Purview supports the following authentication methods

- System-assigned Managed Identity (recommended)
- User-assigned Managed Identity
- Service Principal
- Account Key

💡 Did you know?

Collections in Microsoft Purview can be used to organize data sources, scans, and assets in a hierarchical model based on how your organization plans to use Microsoft Purview. The collection hierarchy also forms the security boundary for your metadata to ensure users don't have access to data they don't need (e.g. sensitive metadata).

Azure Key Vault is a cloud service that provides a secure store for secrets. Azure Key Vault can be used to securely store keys, passwords, certificates, and other secrets.

While the **Microsoft Purview Governance Portal** is provided as the default user experience, customers and partners can interface with the underlying platform using the Microsoft Purview REST API. The openness of the platform enables the possibility of integrating Microsoft Purview with custom user interfaces as well as the ability to perform read/write operations programmatically.

💡 Did you know?

Assets can be related to two different types of contacts:

Experts are often business process or subject matter experts.

Where as **Owners** are often senior executives or business area owners that define governance or business processes over certain data areas.

For assets in which you are tagged as a **Contact** these will appear on the home screen (Data catalog), under **My items**.

💡 Did you know?

A **Term Template** determines the attributes for a term. The **System default** term template has basic fields only (e.g. Name, Definition, Status, etc). **Custom** term templates on the other hand, can be used to capture additional custom attributes.

Synonyms are other terms with the same or similar definitions. Where as **Related terms** are other terms that are related but have different definitions.

Glossary terms can be related to two different types of contacts. **Experts** are typically business process or subject matter experts. Where as **Stewards** define the standards for a data object or business term. They drive quality standards, nomenclature, rules.

Demo Time

Fika (paus) [15 min]

Create image of Swedish Fika with Kaffe Te and Cinnamon bun

PASS FYRA | SESSION FOUR

Microsoft Purview {Azure} P2: Lineage, Insight, Monitor, Integration with Azure Synapse Analysis

Lineage

Lineage is a vital aspect of data governance. It is the silver bullet that helps us track where our data originated, how it was gathered, how it has been modified, and how it is consumed downstream.

It provides traceability as data flows through the enterprise. The need for such insight is driven by compliance, regulations, privacy, ethics, and reproducibility and transparency of advanced analytics models.

Lineage

To consistently capture lineage, you need to set standards like tools and methodologies to use. Users tend to follow the path of least resistance, so keep this in mind when offering services to your organization. Services like Azure Data Factory, AzureData Share, and Power BI automatically capture the lineage of data as it moves. Alternatively you create custom lineage, which is supported via Atlas hooks and REST API. Lineage in Microsoft Purview includes datasets and processes. Datasets are also referred to as nodes while processes can be also called edges:

Dataset (Node): A dataset (structured or unstructured) provided as an input to a process. For example, a SQL Table, Azure blob, and files (such as .csv and .xml), are all considered datasets. In the lineage section of Microsoft Purview, datasets are represented by rectangular boxes.

Process (Edge): An activity or transformation performed on a dataset is called a process. For example, ADF Copy activity, Data Share snapshot and so on. In the lineage section of Microsoft Purview, processes are represented by round-edged boxes.

Lineage

Lineage can also be used to provide insights in an end-to-end context.

You could for example organize your metadata in Microsoft Purview as a graph by connecting it to other metadata subjects, such as domains, data quality, data usage, business capabilities, application functions, application life cycle management information, and so on.

This approach of bringing metadata closer to data analysts and scientists is also known as "Data observability".

💡 Did you know?

Thresholds help minimise the possibility of false-positive classifications. **Minimum match threshold** is the minimum percentage of data value matches in a column that needs to be found by the scanner for the classification to be applied.

Scan Rule Sets determine which **File Types** and **Classification Rules** are in scope. If you want to include a custom file type or custom classification rule as part of a scan, a custom scan rule set will need to be created.

Ignore patterns tell Microsoft Purview which assets to exclude during scanning. During scanning, Microsoft Purview will compare the asset's URL against these regular expressions. All assets matching any of the regular expressions mentioned will be ignored while scanning.

💡 Did you know?

Dataset: A dataset (structured or unstructured) provided as an input to a process. For example, a SQL Table, Azure blob, and files (such as .csv and .xml), are all considered datasets. In the lineage section of Purview, datasets are represented by **rectangular boxes.**

Process: An activity or transformation performed on a dataset is called a process. For example, ADF Copy activity, Data Share snapshot and so on. In the lineage section of Purview, processes are represented by **round-edged boxes.**

Microsoft Purview can connect to **multiple** Azure Data Factories but each Azure Data Factory account can **only** connect to one Microsoft Purview account.

💡 Did you know?

When a user creates an Azure Data Factory connection, behind the scenes the Data Factory managed identity is added to the **Data Curator** role. This provides Azure Data Factory the necessary access to push lineage to Microsoft Purview during a pipeline execution.

Insights

<< PLEASE READ BEFORE PROCEEDING >>

Data Estate Insights can take several hours to surface post the completion of a scan.
At this point of the workshop, only a limited number of data visualisations may be populated.

To populate all reports with data, Microsoft Purview requires an environment with a variety of sources and assets to be scanned that is beyond the scope of this workshop.

The screenshots and information below, has been provided so that you can conceptualise the type ! of insights that can be gleaned from a fully populated environment.

💡 Did you know?

Using the quick filters on the **Data assets by source type** graph and drilling into the details by clicking, **View details** is a quick and easy way of identifying which sources contain certain types of data (e.g. set the **Classification category** filter to **PERSONAL**

Terms are considered **incomplete** if they are missing a definition, expert, or steward. If a term is missing more than one of these things, it is shown as **Missing multiple items**.

Sensitivity labels state how **sensitive** data is in your organization. For example, data contained within a particular asset might be **HIGHLY CONFIDENTIAL**. Classifications on the other hand indicate the **type** of data values (e.g. Driver's License Number, Email Address, SWIFT Code, etc).

Microsoft Purview's ability to apply sensitivity labels is controlled within the Microsoft Purview Compliance Portal. Note: You must have an active Microsoft 365 license that offers the benefit of automatically applying sensitivity labels.

💡 Did you know?

Monitoring Reader role can view all monitoring data but cannot modify any resource or edit any settings related to monitoring resources. This role is appropriate for users in an organization such as Microsoft Purview administrators.

Diagnostic settings can be used to send platform logs and metrics to one or more destinations (Log Analytics Workspace, Storage Account, an Event Hub).

ScanStatus tracks the scan life cycle. A scan operation follows progress through a sequence of states, from Queued, Running and finally a terminal state of Succeeded | Failed | Canceled. An event is logged for each state transition.

💡 Did you know?

One of the key benefits of integrating Azure Synapse Analytics with Microsoft Purview, is the ability to discover Microsoft Purview assets from within Synapse Studio (i.e. no need to switch between user experiences), with added abilities using Synapse specific capabilities (e.g. SELECT TOP 100).

Note*: Before we can demonstrate the ability to query external data sources from Azure Synapse Analytics, we need to ensure our account has the appropriate level of access (i.e.* ***STORAGE BLOB DATA READER****).*

When connecting a Synapse workspace to Purview, Synapse will attempt to add the necessary Purview role assignment (i.e. **DATA CURATOR**) to the Synapse managed identity automatically. This operation will be successful if you belong to the **Collection admins** role on the Purview root collection and have access to the Microsoft Purview account.

Demo Time

Fika (paus) [15 min]

Create image of Swedish Fika with Kaffe Te and Cinnamon bun

PASS FEM | SESSION FIVE

Microsoft Purview {Azure} P3: REST API, Private Endopints, Data Owner, Policies, Data Sharing

Apache Atlas

Microsoft Purview's **data catalog** is largely based on **Apache Atlas**, and therefore shares much of the same surface area that allows users to programmatically perform CRUD (CREATE/READ/UPDATE/DELETE) operations over Microsoft Purview assets.

"Apache Atlas provides open metadata management and governance capabilities for organizations to build a catalog of their data assets, classify and govern these assets and provide collaboration capabilities around these data assets for data scientists, analysts and the data governance team."

Atlas Endpoints As can be seen in the **Apache Atlas Swagger**, Atlas has a variety of REST endpoints that handle different aspects of the catalog (e.g. types, entities, glossary, etc).

Microsoft Purview

Interface

Governance Portal

Data Catalog | Data Map | Data Estate Insights | Data Policy | Management

Custom

Custom Connectors | Custom Lineage | Custom UX | ...

Platform (REST API Endpoints)

Account

Account | Collection
Resource Set Rule

Catalog

Entity | Glossary | Lineage
Relationship | Types | Search

Map & Discover

Asset | Scan

Policy Store

Metadata Policies | Data Policies

Scan

Class. Rule | Credential | Source
Scan | Filter | Ruleset | Trigger

Apache Atlas 2.2

Types

A definition of how a particular type of metadata object is stored and accessed.

Entity

An instance of an entity “type” (i.e. entityDef) is an entity.

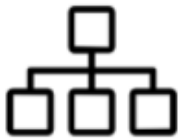
Glossary

A hierarchical set of business terms that represents your business domain.

Relationship

Relationships between entities.

Lineage



Return lineage information about an entity.

Discovery*

Data discovery using DSL or full text search.

💡 Did you know?

An Azure **service principal** is an identity created for use with applications, hosted services, and automated tools to access Azure resources.

The OAuth2 service endpoint is used to gain access to protected resources such as Microsoft Purview. The HTTP request enables us to acquire an **ACCESS TOKEN** , this will subsequently be used to query the Microsoft Purview API.

While Microsoft Purview provides a number of system built type definitions for a variety of object types, customers can use the API to create their own custom type definitions.

Securely scan sources ~ Self-Hosted Integration Runtimes

Microsoft Purview comes with a managed infrastructure component called ***AutoResolveIntegrationRuntime***.

This component is required when scanning sources and most useful when connecting to data stores and computes services with public accessible endpoints.

However some of your sources might be VM-based or can be applications that either sit in a private network (VNET) or other networks, such as on-premises. For these kind of scenarios a Self-Hosted Integration Runtime (SHIR) is recommended..

Securely scan sources ~ Self-Hosted Integration Runtimes

With the newly released features Microsoft Purview now provides three options for scanning sources:

Microsoft Purview default's integration runtime: this option is useful when connecting to data stores and computes services with public accessible endpoints.

Self-hosted integration runtimes (SHIR): this option particularly useful for VM-based data sources or applications that either sit in a private network (VNET) or other networks, such as on-premises.

Managed Virtual Network Integration Runtime: this new option supports connecting to data stores using private link service in private network environment. This ensures that data scanning process is completely isolated and secure, while also being fully managed.

Managed private endpoints

An integration runtime (IR) is compute infrastructure that Microsoft Purview uses to power data scan across different network environments. These integration runtimes come in different flavors.

One of them is using managed private endpoints, which Microsoft recently added. With this new set of features you can better manage and secure your data scanning within Purview.

As a result your metadata traffic is distributed via Azure Private Link, which eliminated any exposure to the internet. This protects you from any data exfiltration risks.

💡 Did you know?

The **Purview Integration Runtime** can also be used to scan and ingest metadata assets from Azure cloud services that are hidden behind private endpoints, such as Azure Data Lake, Azure SQL Database, Azure Cosmos DB and more.

By using managed private endpoints, you do not have to manage your own VNETs. Microsoft takes care of managing this.

Process events using Atlas Kafka topics via Event Hubs and NodeJS

Microsoft Purview is a unified data governance solution that runs on Azure. Some people say it is a data catalog, but it better to Purview as a control framework for managing and controlling your data landscape. For example, you can use Microsoft Purview as a metastore for dynamically orchestrating your data pipelines.

When Microsoft Purview is deployed, a managed event hub is created as part of your Purview account creation. This opens up many possibilities when integrating Microsoft Purview with other applications. Because of the open nature of Microsoft Purview, you can automate and integration different aspects. For example, you can trigger a workflow outside Microsoft Purview, when new data is scanned, or make an API call for fetching and storing additional metadata inside Purview.

Data owner policies (Azure Storage)

A new feature of Purview is **Policies**, which enables you to secure your data estate from within the Microsoft Purview Governance Portal. This feature is in **Preview** as of July 2022.

Data access policies can be enforced through Purview on data systems that have been registered in Purview for scanning and data use management. This feature allows Data stewards and owners to grant read, write access to various data stores from within Purview by creating a data access policy through the Policy Management app in the governance portal, enabling a single dashboard view of all access granted to all systems.

A policy is a named collection of **policy statements.** When a policy is published to one or more data systems under Purview's governance, it's then enforced by the system. A policy definition includes a policy name, description, and a list of one or more policy statements.

Data owner policies (Azure Storage)

Tmake a data resource available for policy management, the **Data Use Management (DUM)** toggle needs to be enabled. A user, who will manage the policies in Purview, will need certain **Identity Access management (IAM)** privileges on the resource and MS Purview in order to enable the DUM toggle.o

💡 Did you know?

Data Source Administator role can publish a policy.

Policy authors role can create or edit a policy

Azure SQL Database Lineage Extraction

Microsoft Purview can show us how data moves through various systems; this is referred to as **data lineage** and part of the lifecycle of any piece of data in an organization.

Lineage extraction is a complicated process because there are many ways data may be moved and transformed throughout its lifecycle. Lineage information can either be extracted by Purview during the scanning process (when supported), or lineage information can be pushed to Purview via the Apache Atlas REST API.

💡 Troubleshooting / FAQ

Do I have to assign an AD Admin to the Azure SQL Database to add the Microsoft Purview Managed Identity?

Yes. Adding the Microsoft Purview Managed Identity to the Azure SQL Database requires an AD user with sufficient priviledge; you won't be able to add a Managed Identity with SQL-Auth.

Can I use a SQL-Auth account with appropriate permissions, and configure the scan to use that account instead of the Managed Identity?

You can, but the recommended approach is to use Managed Identity. If you'd like to use a SQL-Auth account or have no way to use a AD account, consider creating a different account specifically for scanning. To do so, use a tool like Azure Data Studio to connect to the master database, create a login (in the master database) and a user (in the sample database) using SQL statements, and assign the user to the appropriate db-owner role. This credential can be stored in Azure Key Vault the same way as it is done in the earlier modules.

💡 Did you know?

In order for Microsoft Purview to detect the **lineage**, it observes the actual execution of the stored procedure. Therefore, the lineage will not be detected until there is an execution of the MoveDataTest stored procedure.

Data Sharing

Organizations are increasingly looking for ways to enable seamless and secure data sharing.

Whether that be to send and receive data to external organizations or for inter-departmental use cases. However, doing so in a way where organizations are able to maintain control and visibility can be a challenge.

Even today, data continues to be shared using File Transfer Protocols (FTPs), Application Programming Interfaces (APIs), USB devices, and email attachments. These methods are traditionally not secure, challenging to govern, and generally inefficient.

Data Sharing

With Microsoft Purview Data Sharing:

Data providers can now share data in-place from Azure Data Lake Storage Gen2 and Azure Storage accounts, both within and across orgnaizations. Share data directly without data duplication and centrally manage your sharing activities from within Microsoft Purview.

Data consumers can now have near real-time access to shared data. With storage data access and transactions charged to the data consumers based on what they use, at no additional cost to the data provider.

Data Sharing

Note

Supported storage account configurations:

| Configuration | Support | | --- | -- | |

Regions | Canada Central, Canada East, UK South, UK West, Australia East, Japan East, Korea South, and South Africa North

| | Performance | Standard | | Standard | LRS, GRS, RA-GRS |.

💡 Did you know?

Note

While the shared data can be visibly seen within the data consumers target storage account, the data has been shared in-place (without data duplication).

In other words, the files are **symbolic links** pointing back to the original files that reside in the data producers source storage account. These files can be read and downloaded by the data consumer, but the data consumer is unable to delete or modify the files..

Demo Time

Home Work Labs

What is the Microsoft 365 Developer Program?

What is the Microsoft 365 Developer Program and who should join it?

The Microsoft 365 Developer Program is designed to help you build people-centric, cross-platform productivity experiences that extend Microsoft 365. Join our program to get access to the tools, documentation, training, experts, and community events that you need to build great solutions for Microsoft 365 products and technologies, including:

Excel, Outlook, Word, and PowerPoint web add-ins

SharePoint

Microsoft Teams

Microsoft Graph

Use sample data packs with your Microsoft 365 Developer Program subscription | Microsoft Learn

Try Microsoft Purview For Free [90 days]

Did you know you can try the premium versions of all nine Microsoft Purview solutions for free? Use the 90-day Purview solutions trial to explore how robust Purview capabilities can help your organization meet its compliance needs.

Microsoft 365 E3 and Office 365 E3 customers can start now at the

Microsoft Purview compliance portal trials hub.

Learn details about who can sign up and trial terms.

Register Providers

An [Azure account] with an active subscription.
* Owner permissions within a Resource Group to create resources and manage role assignments.
* The subscription must have the following resource providers registered.
  * Microsoft.Authorization
  * Microsoft.DataFactory
  * Microsoft.EventHub
  * Microsoft.KeyVault
  * Microsoft.Purview
  * Microsoft.Storage
  * Microsoft.Sql
  * Microsoft.Synapse
  * Microsoft.Insights

Products in Microsoft Purview {azure} workshop

- Azure Subscription
- Azure Resource Manager
- GitHub Repository
- GitHub Desktop Client
- Visual Studio Code (or other Markdown client)
- NodeJS

- Files in txt, csv and parquet formats

- Apache Atlas
- Apache Atlas Storage
- Atlas Kafka
- Postman

- Contoso BI DataSet
- World Wide Importers

- Azure Storages
- Azure SQL Databases
- Azure Purview Account
- Azure Data Lake Storage Gen 2
- Azure Data Factory
- Azure Pipelines
- Azure SQL Virtual Machine
- Azure Synapse Analytics
- Azure Key Vaults
- Azure Rest API
- Azure Windows Client Virtual Machine
- Azure Storage Explorer
- Azure Virtual Network Private Endpoints
- Self-Hosted Integration Runtime
- Azure Even Hub

SUMMARY

Q&A | Discussion

- Questions?

- Answers?

Q&A | Discussion

- Questions?

- Is Microsoft Purview Support TERADATA?

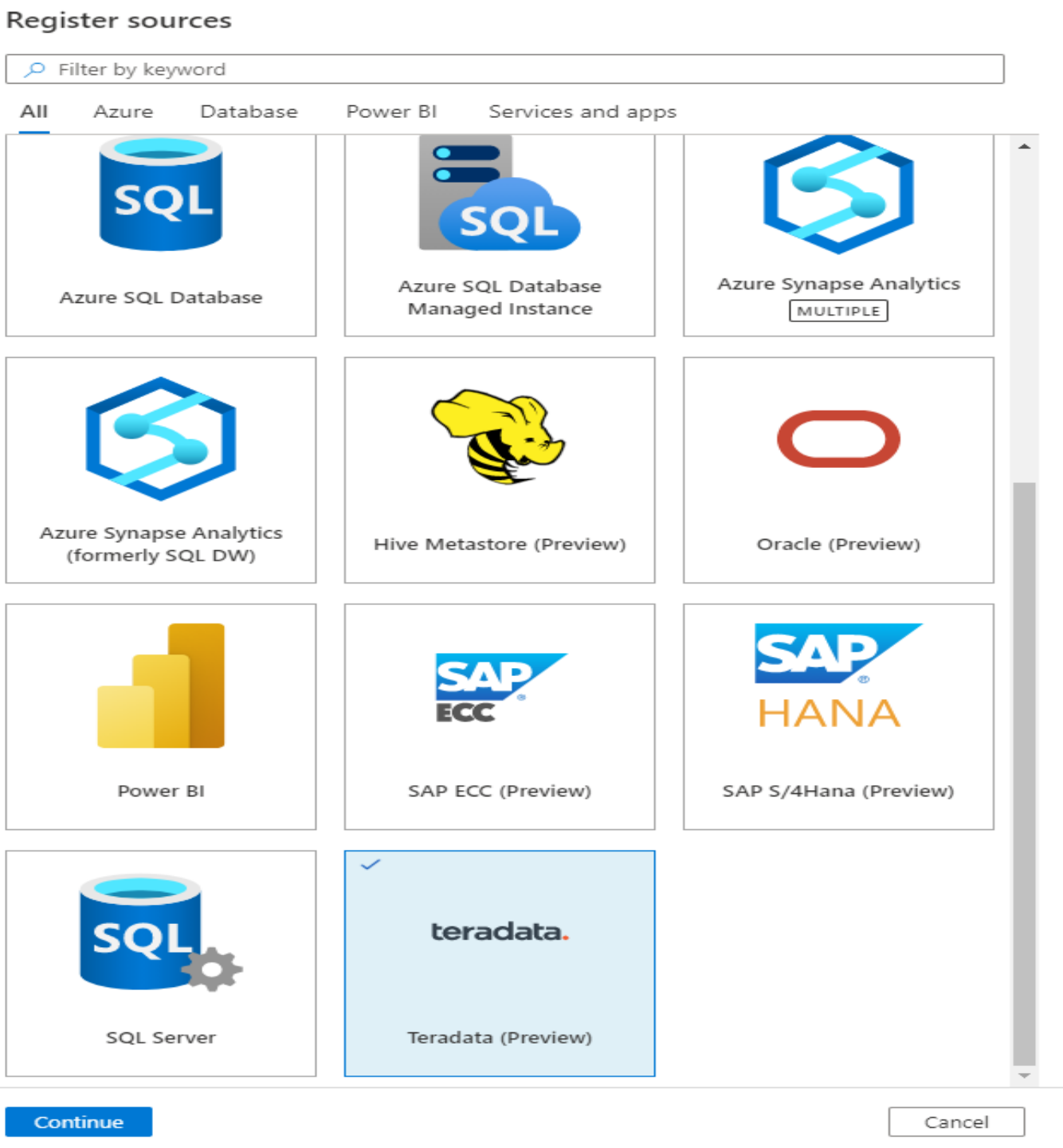
- Answers?

- YES!
- YES for: Metadata Extraction, Full Scan, Scoped Scan, Classifiction, Lineage
- NO for: Incremental Scan, Access Policy, Data Sharing

- The supported Teradata database versions are 12.x to 17.x.

Q&A | Discussion

[Connect to and manage Teradata - Microsoft Purview | Microsoft Learn]



Manage visibility and governance of data assets

- **Microsoft Purview Data Map**
- Capture metadata about analytics data, software-as-a-service, and in hybrid, on-premises, and multi-cloud environments.
- **Microsoft Purview Data Catalog**
- Find trusted data sources by browsing and searching your data assets, aligning your assets with friendly business terms and data classification.
- **Microsoft Purview Data Estate Insights**
- Gain insights into your data estate, to help you discover what kinds of data you have and where.

Protect sensitive data across clouds, apps, and devices | p1

Microsoft Purview Information Protection

Discover, classify, and protect sensitive information wherever it lives or travels.

Microsoft Purview Data Loss Prevention

Use data loss prevention to help prevent accidental sharing of sensitive information.

Microsoft Purview Message Encryption

Send and receive encrypted email messages to people inside and outside your organization.

Protect sensitive data across clouds, apps, and devices | p2

Microsoft Purview Customer Key

Help meet compliance requirements by exercising control over your organization's encryption keys.

Microsoft Purview Double Key Encryption

Uses two keys together to access protected content. Microsoft stores one key in Microsoft Azure, and you hold the other key.

Microsoft Purview Data Connectors

Import and archive non-Microsoft data so you can apply Microsoft 365 protection and governance capabilities to third-party data.

organization.

Protect sensitive data across clouds, apps, and devices | p3

Microsoft Purview Data Lifecycle Management

Retain the Microsoft 365 data that you need to keep, and delete the content that no longer has business value.

Microsoft Information Protection SDK

Extend sensitivity labels to third-party apps and services.

Microsoft Purview Customer Lockbox

Maintain control over your content with explicit access authorization for service operations.

Identify data risks and manage regulatory compliance requirements | p1

Microsoft Purview Insider Risk Management

Detect, investigate, and act on malicious and inadvertent risk activities in your organization.

Microsoft Purview Communication Compliance

Detect, investigate, and act on inappropriate and sensitive messages in your organization.

Microsoft Purview Information Barriers

Restrict communication and collaboration among specific groups of users in highly-regulated organizations.

Identify data risks and manage regulatory compliance requirements | p2

Microsoft Purview Privileged Access Management

Help protect your organization from breaches that use accounts with standing access to sensitive data and critical configuration settings.

Microsoft Purview Compliance Manager

Compliance Manager helps with taking inventory of data protection risks, managing the implementation of controls, staying current with regulations & certifications, and reporting to auditors.

Microsoft Purview Records Management

Manage high-value items for business, legal, or regulatory record-keeping requirements.

Identify data risks and manage regulatory compliance requirements | p3

Microsoft Purview Audit

Learn about premium and standard audit solutions and tools to help you log and search for audited activities in SharePoint and OneDrive.

Microsoft Purview eDiscovery

Learn about eDiscovery tools that allow you to search for content in Exchange Online, OneDrive for Business, SharePoint Online, Microsoft Teams, Microsoft 365 Groups, and Yammer teams.

Microsoft Graph APIs for compliance capabilities

Adapt, extend, integrate, accelerate, and support compliance solutions with programmatic access.

Licensing requirements

License requirements for Microsoft Purview Information Protection depend on the scenarios and features you use, rather than set licensing requirements for each capability listed on this page.

To understand your licensing requirements and options for Microsoft Purview Information Protection, see

Information Protection sections from Microsoft 365 guidance for security & compliance

related PDF download for feature-level licensing requirements.

Assessments

- *License Type:*
- Microsoft 365 or Office 365 A1/E1/F1/G1
- Microsoft 365 or Office 365 A3/E3/F3/G3
- Microsoft 365 or Office 365 A5/E5/G5
- Microsoft 365 A5/E5/F5/G5 Compliance
- Microsoft 365 A5/E5/F5/G5 eDiscovery and Audit
- Microsoft 365 A5/E5/F5/G5 Insider Risk Management
- Microsoft 365 A5/E5/F5/G5 Information Protection and Governance

- *Assessment Templates (included by default)*
- Data Protection Baseline
- Data Protection Baseline
- EU GDPR
- NIST 800-53
- ISO 27001
- CMMC Level 1-5 (only available for G5)
- Custom Assessments

Licensing for Everyone

- Enterprise & Frontline Workers | https://bit.ly/M365ModernWorkEnterprise
- Small & medium Business | https://bit.ly/M365ModernWorkBusiness
- Education | https://bit.ly/M365ModernWorkEducation

Few words about me

Tobias Koprowski

Bachelor Degree in Banking
Higher national diplomas in: European Law & Corporate Governance
Three years in personal and home insurance
Five years in consumer & corporate banking
Ten years in physical Data Center
Microsoft Certified Trainer (MCT) & Educator (MCE)
CertNexus Authorized Instructor (CAI)
ISO 27001 Lead Auditor (PCA/TÜV Nord)

Member of:
| **BCS** (The Chartered Institute of IT)
| **IAPP** (International Association of Privacy Professionals)
| **ISSA** (Information Security System Association)
| **ISACA** (Information Systems Auditing & Control Association)
| **ISC²** (International Information System Security Certification Consortium)
| **CSA** (Cloud Security Alliance) – AI Usage Policy Working Group

STEM Ambassador | Royal Voluntary Service

Social Media: KoprowskiT @ [TW|LI|BS|FB]