INSPARK

Feedback

Purview what does this mean to me as an organization?

Innovate to Accelerate

24-01-2025 Erwin de Kreuk Data Community Austria

INSPARK

We Are InSpark

"We help organizations accelerating their digital transformation with impactful Microsoft solutions & expertise"

THANK YOU

Data Community Austria Day 2025

Awesome Partner

Microsoft

Platinum

redgate

ACP IT for innovators.

Gold

b.telligent
smart data. smart decisions.

Lucient

Bronze

Tabular Editor

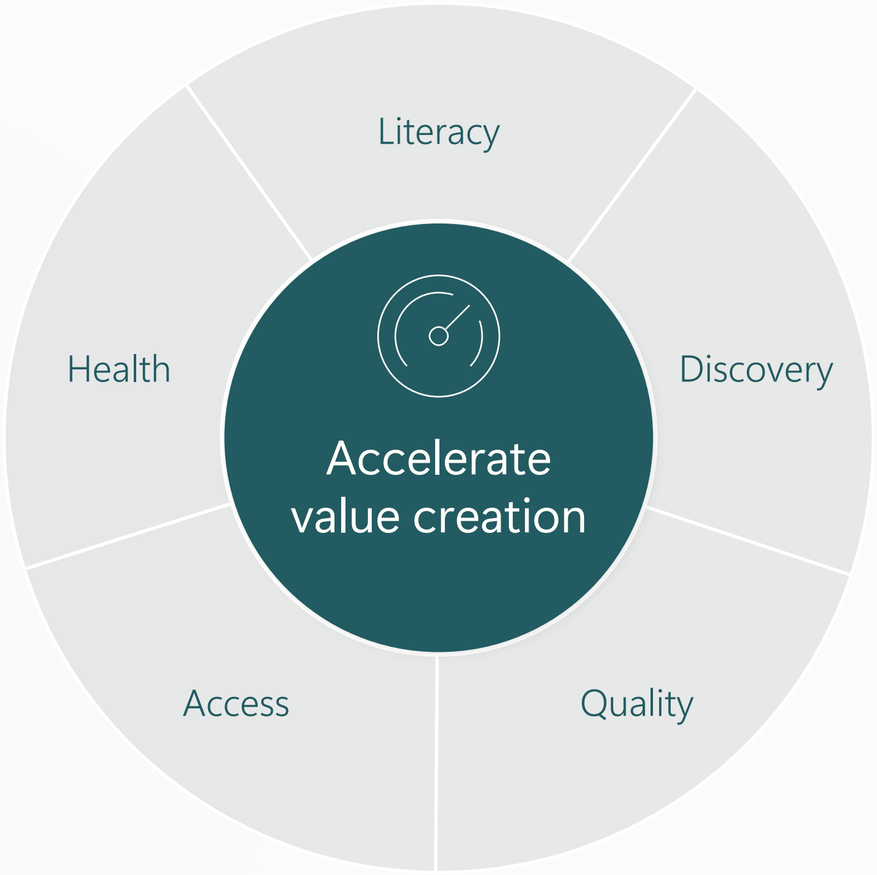
paiqo
The Platform & AI Company

Power BI Camp
www.linearis.at

Data governance and security is evolving

INSPARK

Responsible offense



Essential defense



INSPARK

Erwin de Kreuk

- Principal Consultant
- Lead Data & AI InSpark

Let's connect



@erwindekreuk

linkedin.com/in/erwindekreuk

erwindekreuk.com

github.com/edkreuk

https://sessionize.com/erwin-de-kreuk/

MVP Microsoft® Most Valuable Professional

Objectives



- **Microsoft Purview**
- **Data Security**
- **Risk/Compliance**
- **Data Governance**
  - Governance Domains
  - Data Product
  - Data Quality
  - Data Health
  - Pricing
- **Q & A**

Data Governance for the Age of AI

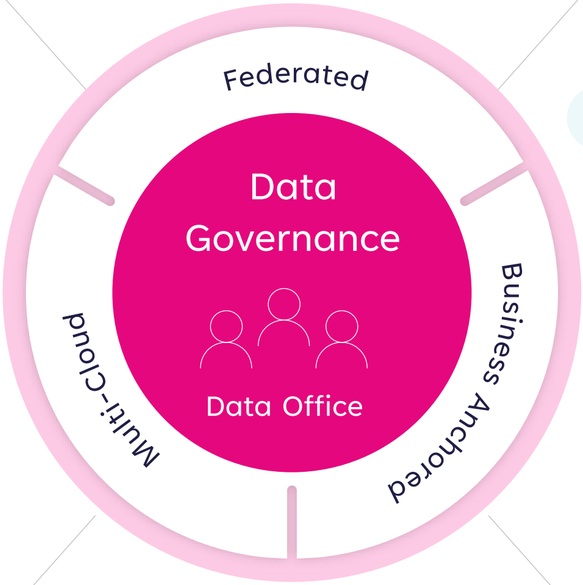
The whole is greater than the sum of the parts

INSPARK

Essential Defense

- Data Security
- Data Compliance
- Risk Management

Chief Information Security/ Compliance/Privacy/Risk Officers



Responsible Offense

- Data Activation
- Data Products
- AI and Business Transformation

Organization-wide Data and Business Teams

Accelerate responsible value from data

Data Transformation

CISO | CCO | CPO

CDOs and organization-wide data teams

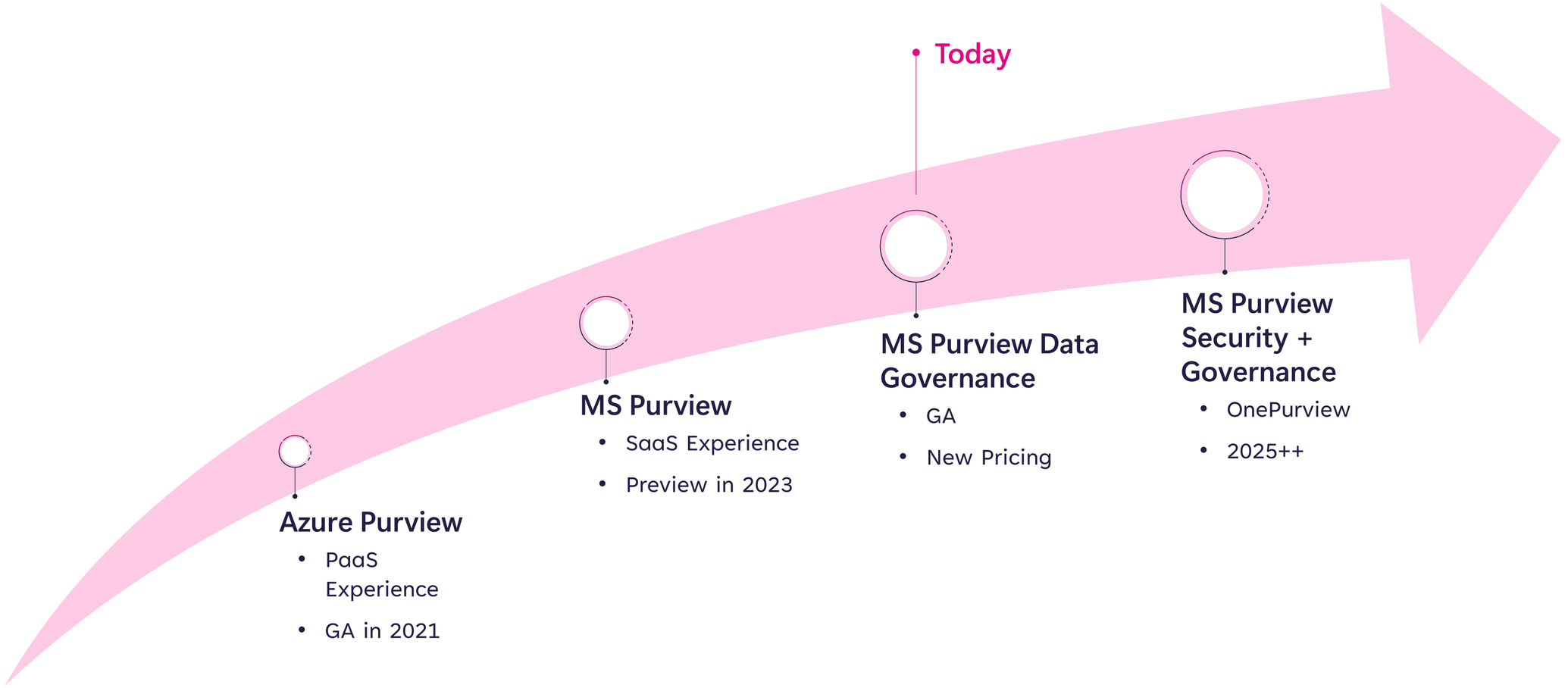


Evergreen Security and Compliance

Responsibly Democratizing Data Curation, Health Management, Discovery, Understanding, and Self-serve Access

Responsibly Democratizing Analytics and AI

Growing Responsible and Transformative Business Value Outcomes with Data


Today
Azure Purview
PaaS Experience
GA in 2021
MS Purview
SaaS Experience
Preview in 2023
MS Purview Data Governance
GA
New Pricing
MS Purview Security + Governance
OnePurview
2025++

Bringing data security, governance, and compliance together

INSPARK

Microsoft Purview

Data Security



For information and cybersecurity teams

- Data Loss Prevention
- Insider Risk Management
- Information Protection

Data Governance



For data consumers, data owners, data officers

- Data Discovery
- Data Curation & Management
- Data Estate Health

Risk & Compliance



For risk, compliance, and legal teams

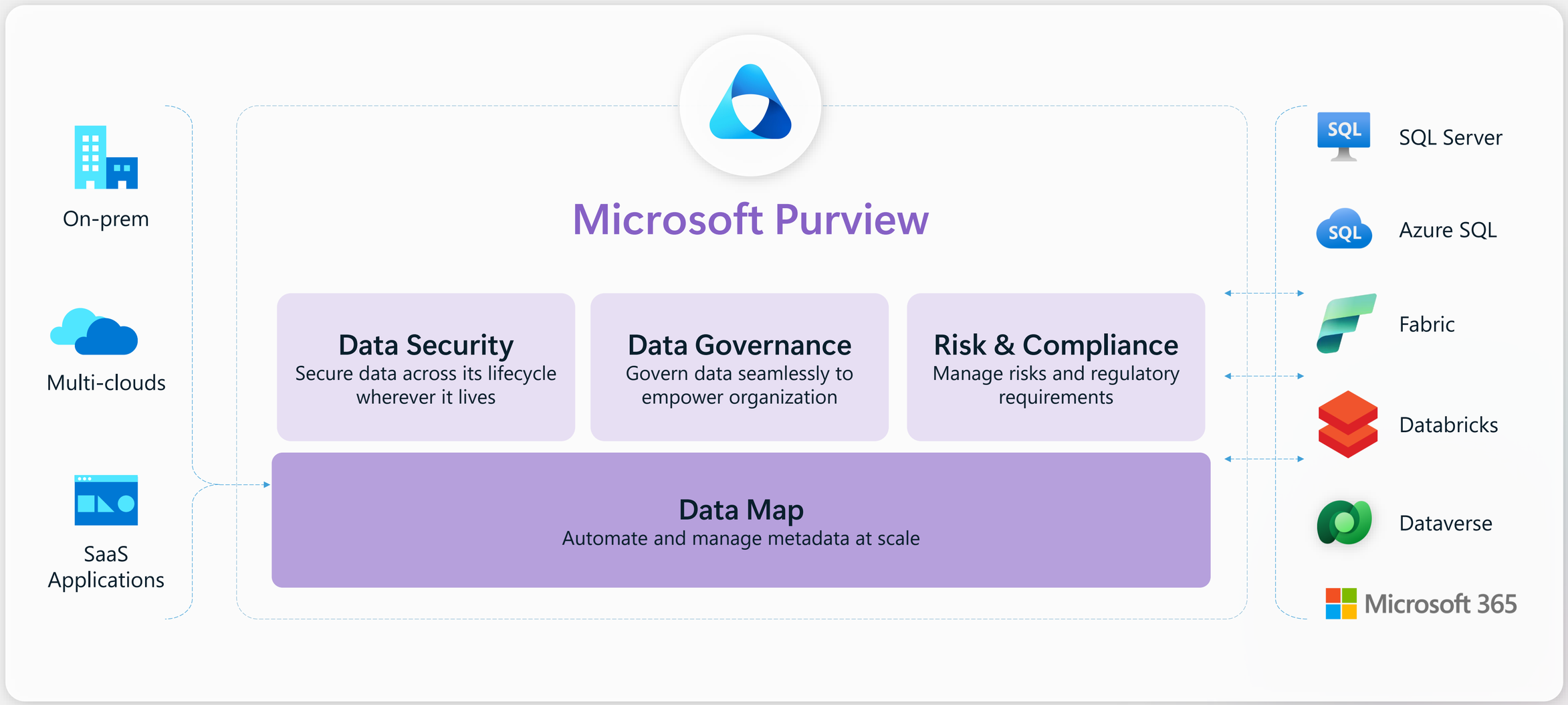
- Data Lifecycle Management
- eDiscovery & Audit
- Communication Compliance

On-prem and multi-cloud

Unstructured & structured data

Across IaaS, and SaaS

INSPARK

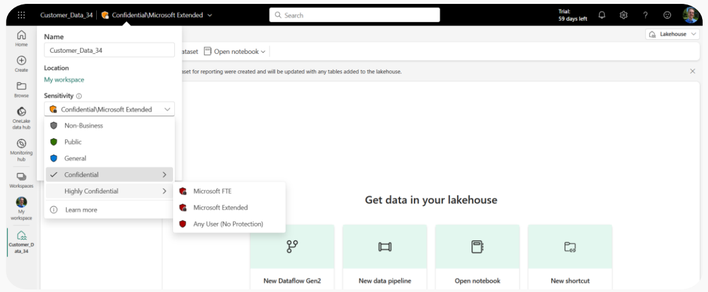

On-prem
Multi-clouds
SaaS Applications
Microsoft Purview
Data Security
Secure data across its lifecycle wherever it lives
Data Governance
Govern data seamlessly to empower organization
Risk & Compliance
Manage risks and regulatory requirements
Data Map
Automate and manage metadata at scale
SQL Server
Azure SQL
Fabric
Databricks
Dataverse
Microsoft 365

Microsoft Purview

Purview security and compliance features to protect your data in Fabric

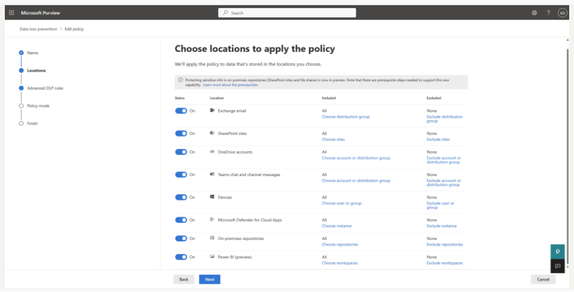
*Additional Microsoft Purview purchase required

Secure and protect data across your organization



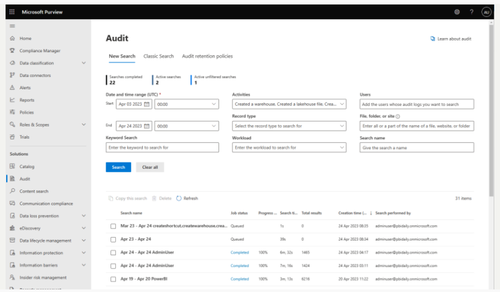
Information Protection labels

Classify sensitive Fabric data using the same sensitivity labels that are used in Microsoft 365—enforced even when the data is exported



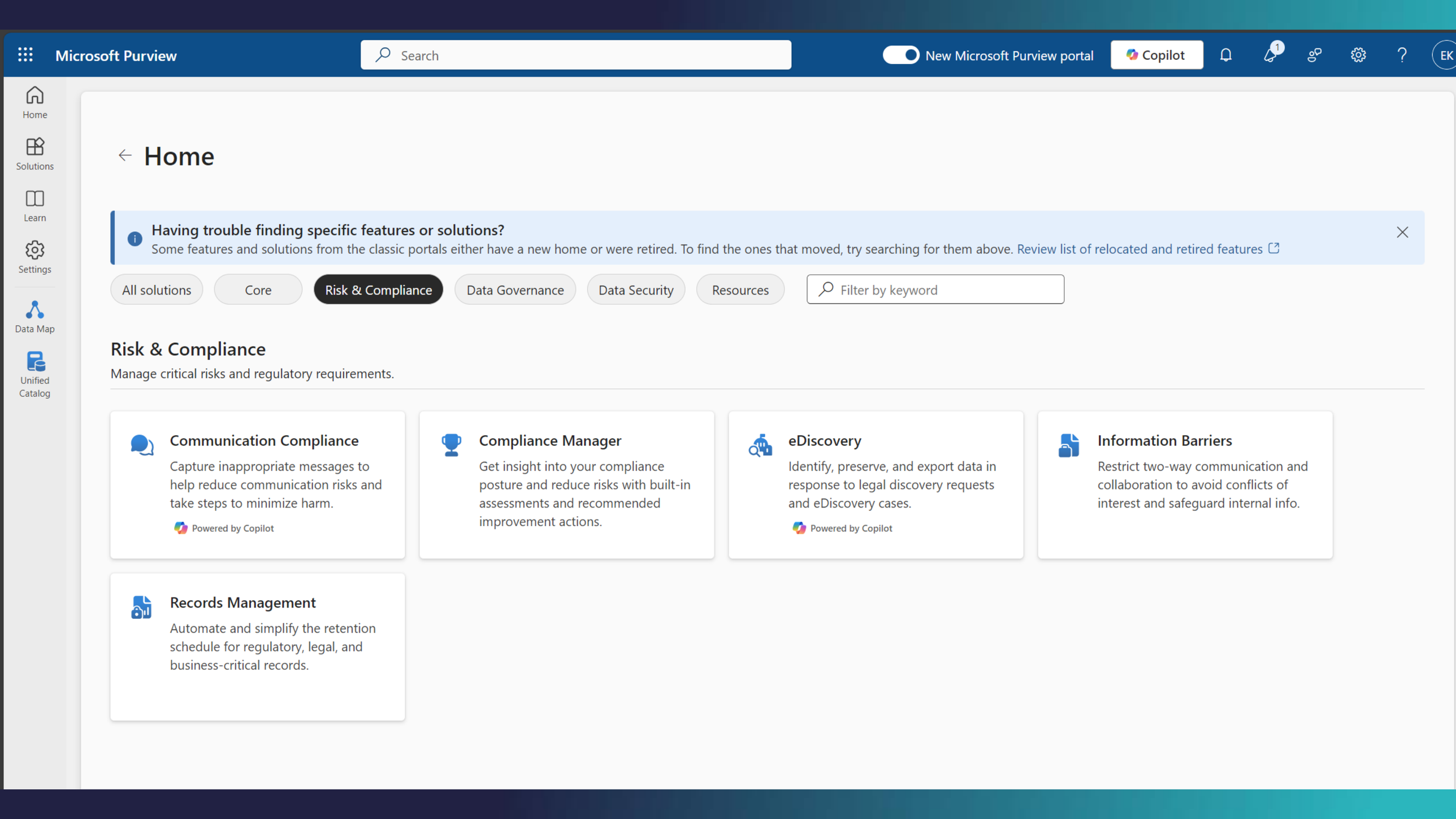
Data Loss Prevention policies

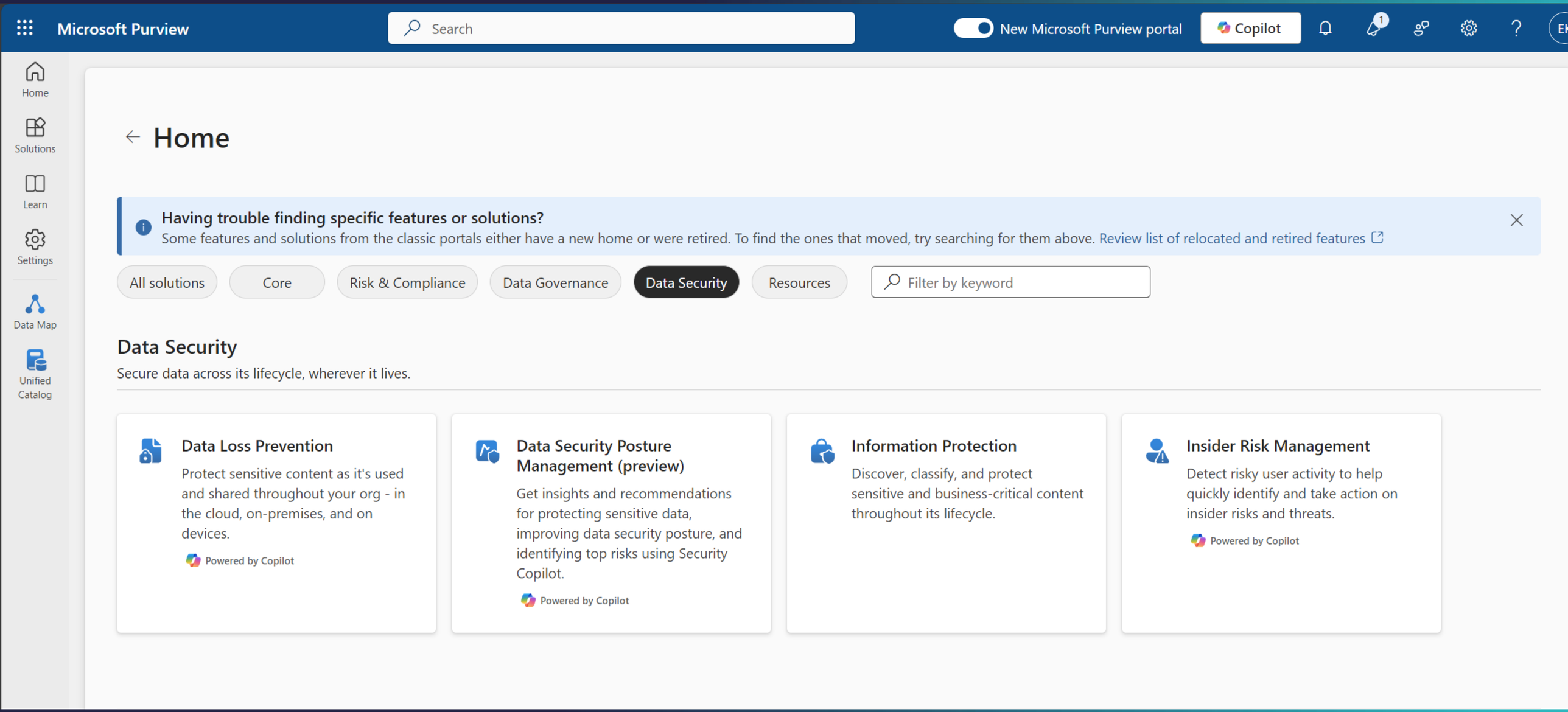
Automatically detect the upload of sensitive data such as PII and trigger automatic risk remediation actions such as alerts



Audit

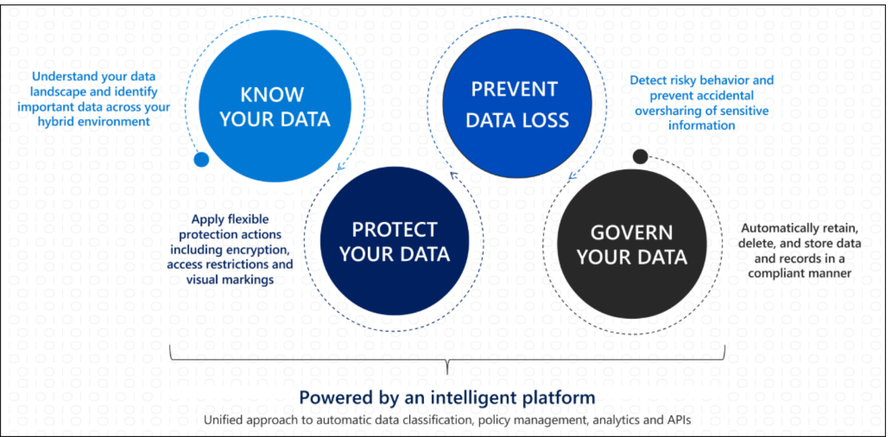
Log user activities from Microsoft Fabric to support security, forensic, and internal investigations


Microsoft Purview
Search
New Microsoft Purview portal
Copilot
EK
Home
Solutions
Learn
Settings
Data Map
Unified Catalog
Home
Having trouble finding specific features or solutions?
Some features and solutions from the classic portals either have a new home or were retired. To find the ones that moved, try searching for them above. Review list of relocated and retired features
All solutions
Core
Risk & Compliance
Data Governance
Data Security
Resources
Filter by keyword
Risk & Compliance
Manage critical risks and regulatory requirements.
Communication Compliance
Capture inappropriate messages to help reduce communication risks and take steps to minimize harm.
Powered by Copilot
Compliance Manager
Get insight into your compliance posture and reduce risks with built-in assessments and recommended improvement actions.
eDiscovery
Identify, preserve, and export data in response to legal discovery requests and eDiscovery cases.
Powered by Copilot
Information Barriers
Restrict two-way communication and collaboration to avoid conflicts of interest and safeguard internal info.
Records Management
Automate and simplify the retention schedule for regulatory, legal, and business-critical records.


Microsoft Purview
Search
New Microsoft Purview portal
Copilot
Home
Solutions
Learn
Settings
Data Map
Unified Catalog
Home
Having trouble finding specific features or solutions?
Some features and solutions from the classic portals either have a new home or were retired. To find the ones that moved, try searching for them above. Review list of relocated and retired features
All solutions
Core
Risk & Compliance
Data Governance
Data Security
Resources
Filter by keyword
Data Security
Secure data across its lifecycle, wherever it lives.
Data Loss Prevention
Protect sensitive content as it's used and shared throughout your org - in the cloud, on-premises, and on devices.
Powered by Copilot
Data Security Posture Management (preview)
Get insights and recommendations for protecting sensitive data, improving data security posture, and identifying top risks using Security Copilot.
Powered by Copilot
Information Protection
Discover, classify, and protect sensitive and business-critical content throughout its lifecycle.
Insider Risk Management
Detect risky user activity to help quickly identify and take action on insider risks and threats.
Powered by Copilot

Data loss prevention

Data Security



Know your data

Classify and identify sensitive data across your hybrid environment, and use insights from this information to inform and monitor your policies.

1. Describe the categories of sensitive information you want to protect
2. Discover and classify sensitive data
3. View your sensitive items

Protect your data

Use sensitivity labels to make the sensitivity level of items visible to admins and users, and apply protection actions that include encryption, access restrictions, and visual markings.

1. Define your sensitivity labels
2. Apply labels using Microsoft 365 apps and services
3. Use Microsoft Defender for Cloud Apps to apply labels in the cloud
4. Use the scanner to apply labels on premises
5. Use Microsoft Purview Data Map to apply labels to schematized data assets in Azure

Prevent data loss

INSPARK

Apply consistent policies to detect risky user behavior and prevent accidental oversharing of sensitive information.

1. Learn about DLP
2. Plan your DLP implementation
3. Design and create a DLP policy
4. Tune your DLP policies

Sensitivity labels

Data Security

Collapse Navigation

Sensitivity labels

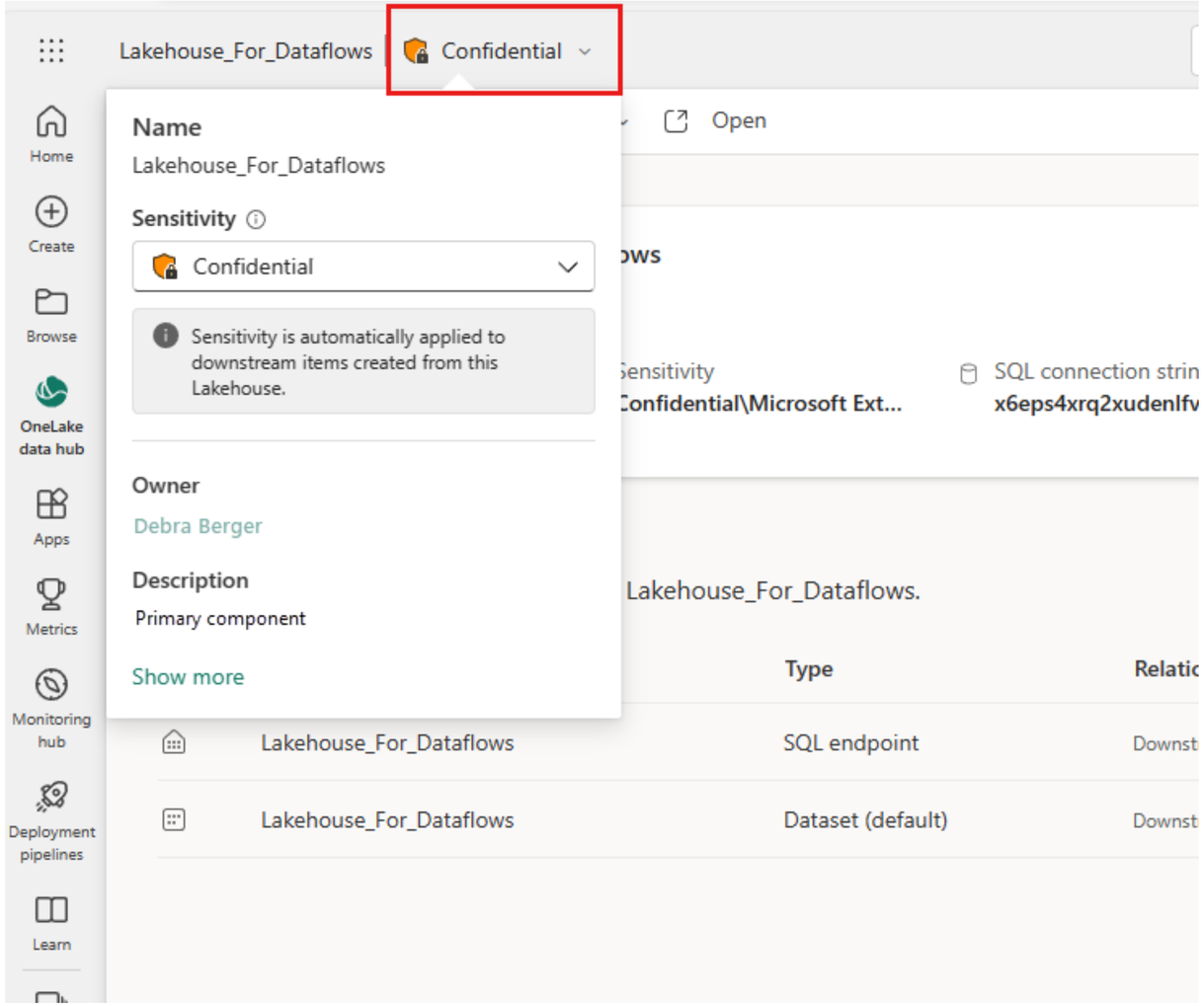
Sensitivity labels are used to classify email messages, documents, sites, and more. When a label is applied (automatically or by the user), the content or site is protected based on the settings you choose. For example, you can create labels that encrypt files, add content marking, and control user access to specific sites. Learn more about sensitivity labels

+ Create sublabel | Create auto-labeling policy | Publish label | Edit label | Reprioritize | Delete label | Refresh | 1 of 8 selected

Name	Priority	Scope	Created by	Last modified
Sebastiaan - Purview - Creditcard	0 - lowest	Files & other data assets, Email, Site, ...	1aa6c213-c0de-47b2-a6ff-446f4a6c...	Jan 10, 2025 1:52:55 PM
Purview Demo	1	Files & other data assets, Email, Site, ...	Erwin De Kreuk(Labs)	Jan 10, 2025 2:59:49 PM
Purview Public	2	Files & other data assets	Erwin De Kreuk(Labs)	Jan 10, 2025 2:59:49 PM
Purview Confidential	3	Files & other data assets	Erwin De Kreuk(Labs)	Jan 16, 2025 4:43:39 PM
Outdated	4	Files & other data assets	Peter Hofstee (ADMIN)	Jan 13, 2025 10:33:55 AM
High Secure - Enforce Compliant device and MFA	5	Files & other data assets, Email, Site, ...	20403043-19b5-43fe-9f9b-6cbca32e...	Jan 10, 2025 3:10:11 PM
InSpark Labs intern	6	Site, UnifiedGroup	Leon Bruggeman (ADMIN)	Jan 10, 2025 3:10:11 PM
InSpark Labs Extern	7	Site, UnifiedGroup	Leon Bruggeman (ADMIN)	Jan 10, 2025 3:10:11 PM
Outdated 2	8	Files & other data assets	Bert van Draanen (ADMIN)	Jan 15, 2025 10:16:02 AM
BlockContent	9 - highest	Files & other data assets	Bert van Draanen (ADMIN)	Jan 16, 2025 12:33:02 PM

Sensitivity labels

Data Security



Information protection

- Allow users to apply **sens**itivity labels for content
  Enabled for the entire organization
- Apply **sens**itivity labels from data sources to their data in Power BI
  Disabled for the entire organization
- Automatically apply **sens**itivity labels to downstream content
  Enabled for the entire organization
- Allow workspace admins to override automatically applied **sens**itivity labels
  Enabled for the entire organization
- Restrict content with protected labels from being shared via link with everyone in your organization
  Enabled for the entire organization
- Increase the number of users who can edit and republish encrypted PBIX files (preview)
  Disabled for the entire organization
- Domain admins can set default **sens**itivity labels for their domains (preview) New
  Disabled for the entire organization

Challenges for Data Consumers

- There’s no central location to register data sources.
- Data-consumption experiences require users to know the connection string or path.
- Data sources and documentation might live in several places.
- There's no explicit connection between the data and the experts that understand the data's context.



Challenges for Data Producers

- Annotating data sources with descriptive metadata is often a lost effort. Client applications typically ignore descriptions that are stored in the data source.
- Creating documentation for data sources can be difficult and it's an ongoing responsibility to keep documentation in sync with data sources. Users might not trust documentation that's perceived as being out of date.
- Creating and maintaining documentation for data sources is complex and time-consuming.
- Restricting access to data sources and ensuring that data consumers know how to request access is an ongoing challenge.



Challenges for Security Administrators

- Everything above and:
- An organization's data is constantly growing and being stored and shared in new directions. The task of discovering, protecting, and governing sensitive data is one that never ends.
- How to ensure that the organization's content is being shared with the correct people, applications, and with the correct permissions.
- Understanding the risk levels in an organization's data requires diving deep into the content, looking for keywords, RegEx patterns, and sensitive data types.
- Constantly monitor all data sources for sensitive content, as even the smallest amount of data loss can be critical to your organization.
- Ensuring that an organization continues to comply with corporate security policies is a challenging task as the content grows and changes.



DATA GOVERNANCE FOR THE AGE OF **AI**

Your AI is only as good as your data

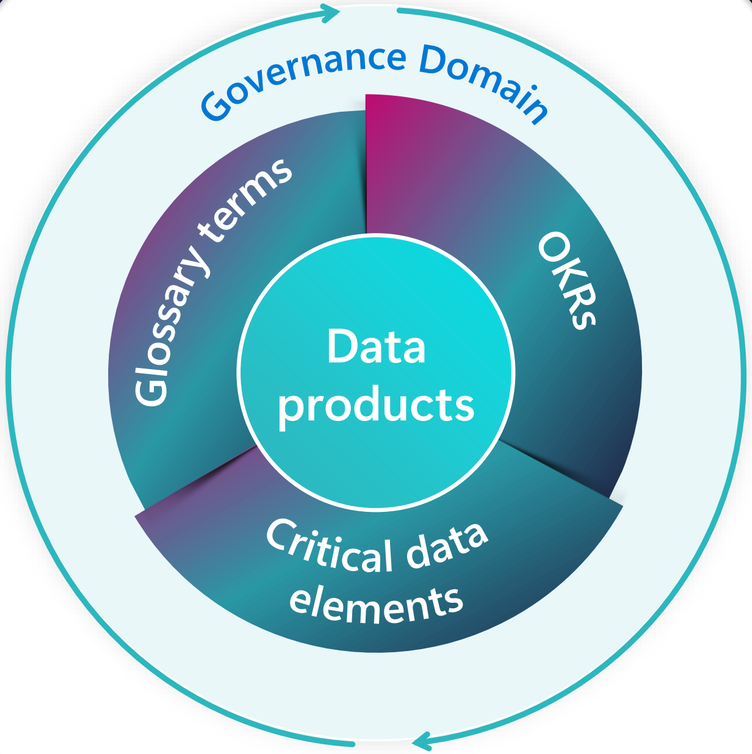
Your AI potential anchors on your people being able to responsibly use your data

Data Governance makes your data ready for your people and your AI

Know your data with business concepts

INSPARK

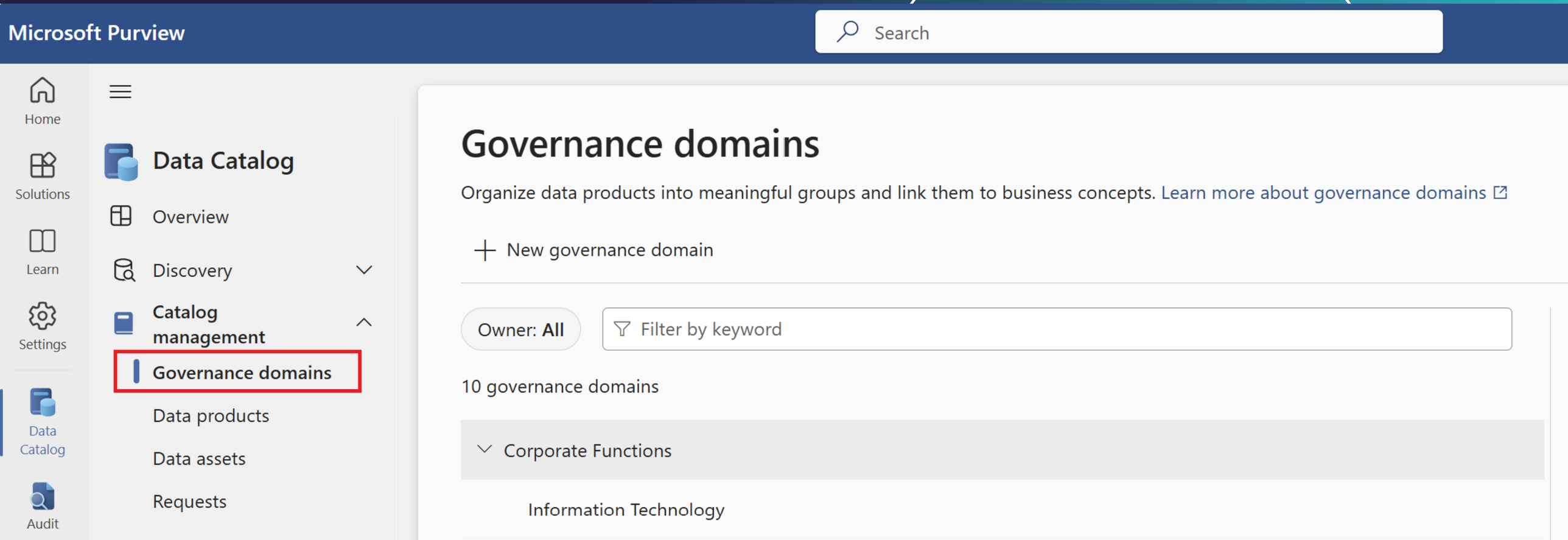
Data stewards
(Curation)



Data consumers
(Consumption)

Microsoft Purview

Governance Domains



INSPARK

Concept and Importance of Governance Domains

Definition of Governance Domains

- Governance domains are specific areas of oversight that help manage data effectively within an organization.

Structuring Governance Efforts

- These domains structure governance efforts, ensuring systematic management and clarity in data governance.

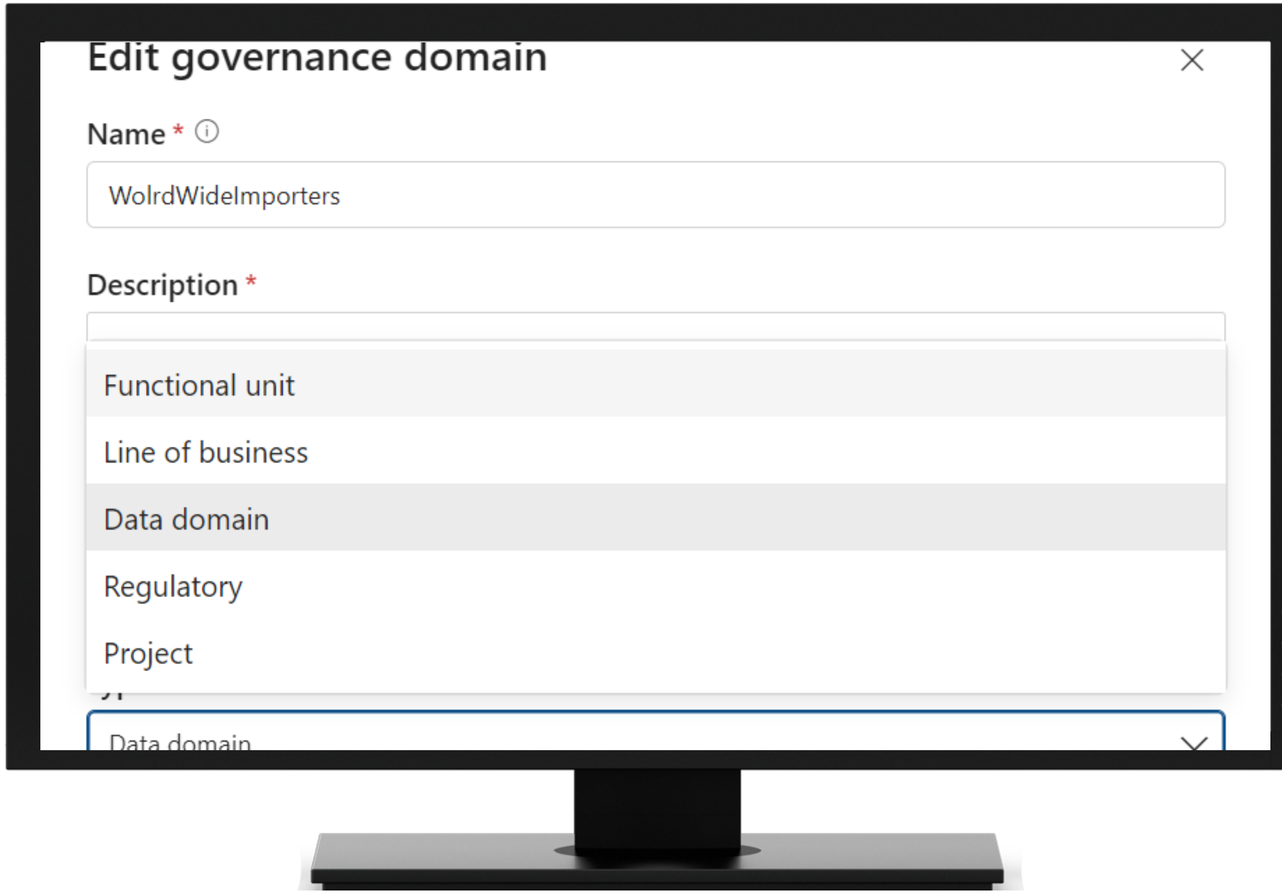
Ensuring Accountability

- Governance domains ensure accountability in data management, creating clear roles and responsibilities across teams.



Types of governance domains

- Functional unit – organizations or business units such as Sales, Marketing, or Finance
- Line of business – products or services being sold such as Xbox, Office, or Azure and different markets or subsidiaries
- Data domain – key organization-wide entities such as customers or employees
- Regulatory – compliance related such as GDPR, SOX, or HIPPA
- Project – collaborative programs across the organization



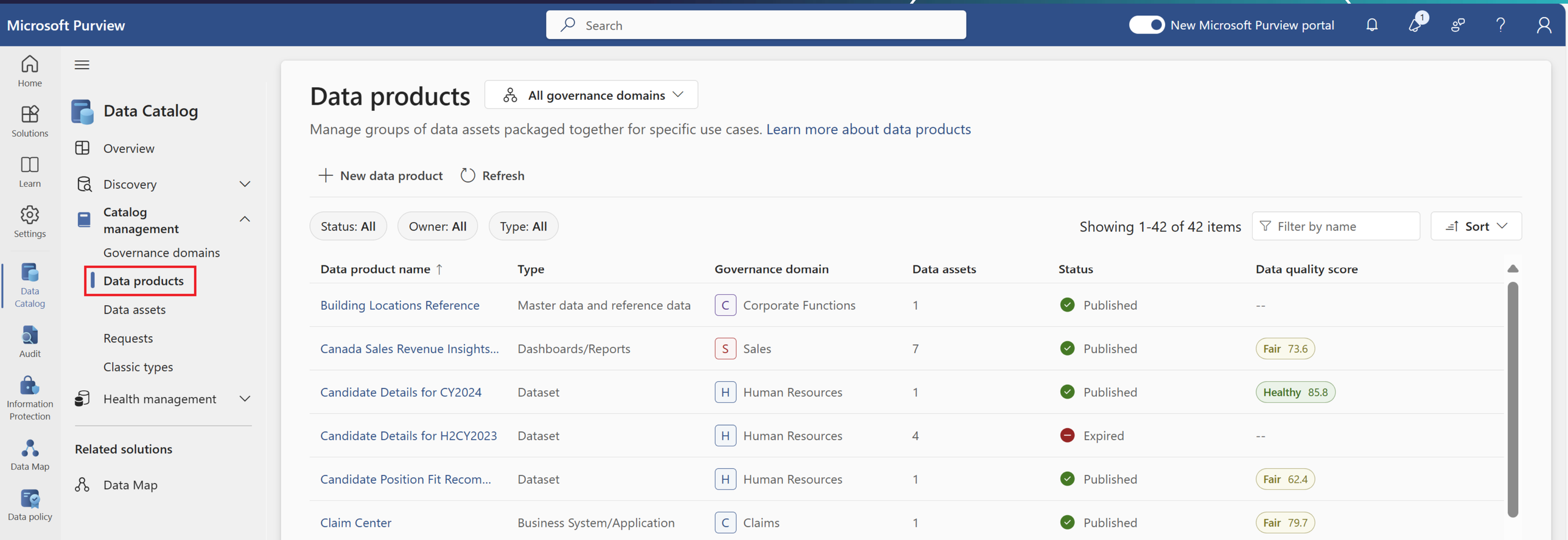
Parts of a governance domain

- Data products
- Glossary terms
- Objectives and key results(OKRs)
- Critical data



Microsoft Purview

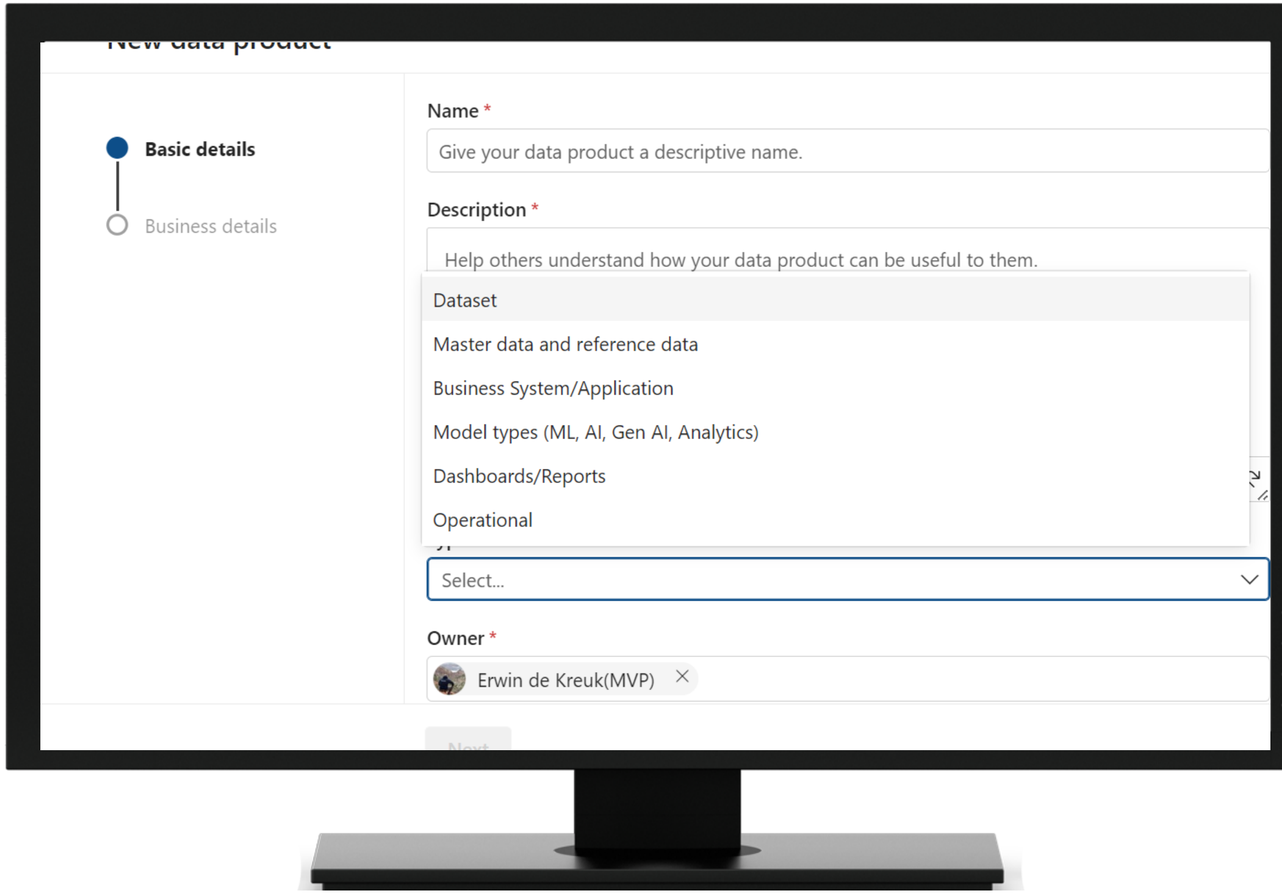
Data Products



Data products

- A data product is a group of data assets (like tables, files, and reports) packaged together for enterprise use, providing context and use cases for data consumers. In Microsoft Purview, data governance not only ensures data security and compliance but also accelerates business value by cataloging data, making it easier to manage and discover.

- A data product includes a name, description, owners, and a list of associated data assets, grouped under a use case. It helps data consumers recognize valuable data in their everyday language and streamlines ownership responsibilities.



INSPARK

Ensure Data Products Meet Organizational Standards

Adhering to Standards

Data products must comply with the established organizational standards to ensure consistency and quality.

Regular Assessments

Conducting regular assessments helps to maintain the relevance and effectiveness of data products over time.

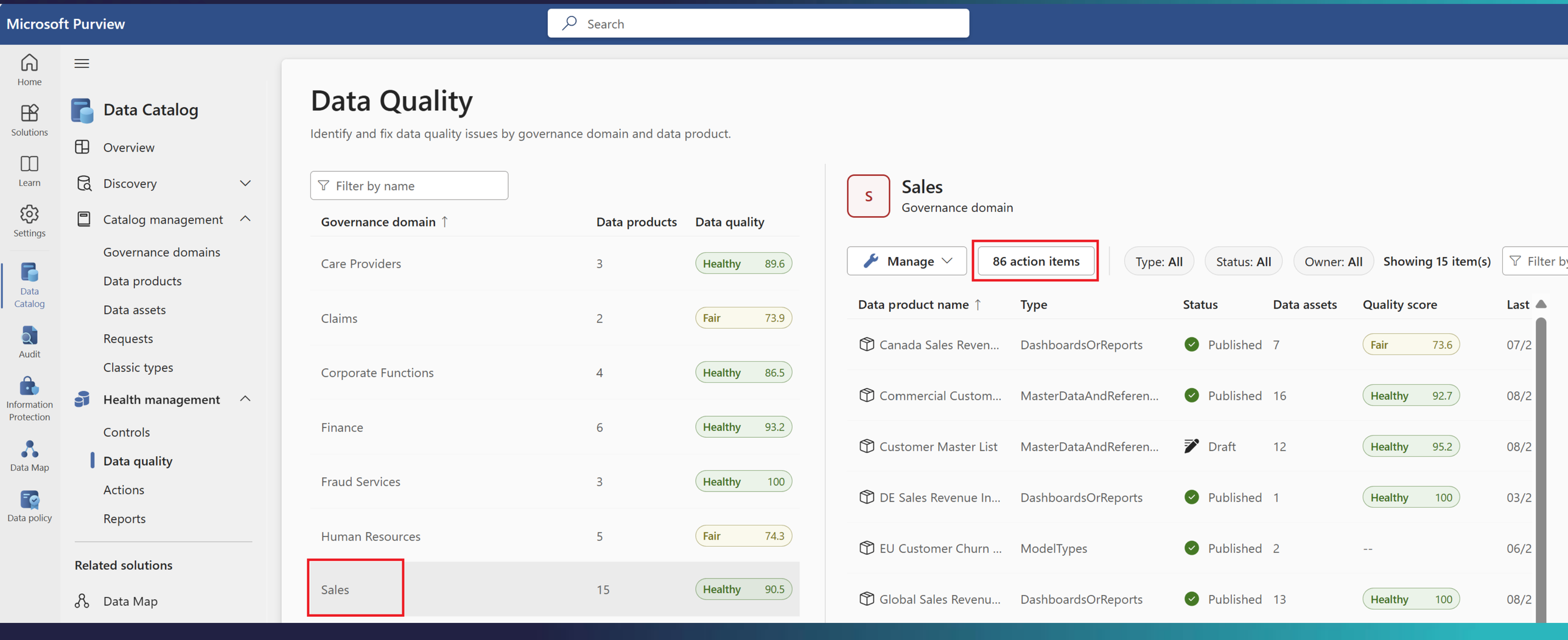
Providing Value

Ensuring compliance and regular updates on data products guarantees they deliver value to the organization.



Microsoft Purview

Data Quality



Poor data quality can lead to:

- Poor decision-making
- Business Inefficiencies
- Mistrust
- Missed Opportunities
- Siloed Information
- Redundancy
- Lost Reputation
- Lost Revenue

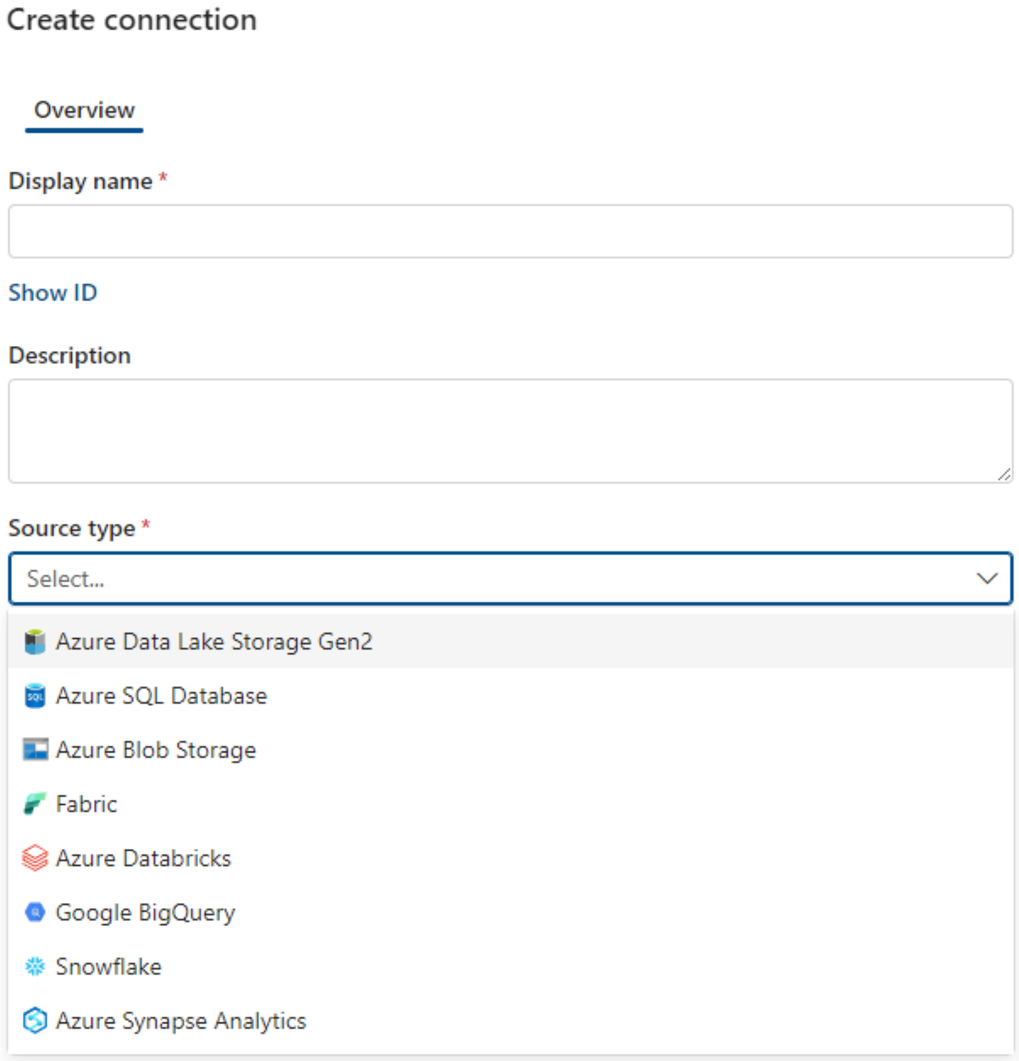
INSPARK

Data quality features

- Data source connection configuration
- Data profiling
- Data quality rules
- Data quality scanning
- Data quality job monitoring
- Data quality scoring
- Data quality for critical data elements (CDEs)
- Data quality alerts
- Data quality actions
- Data quality managed virtual network

Data source connection configuration

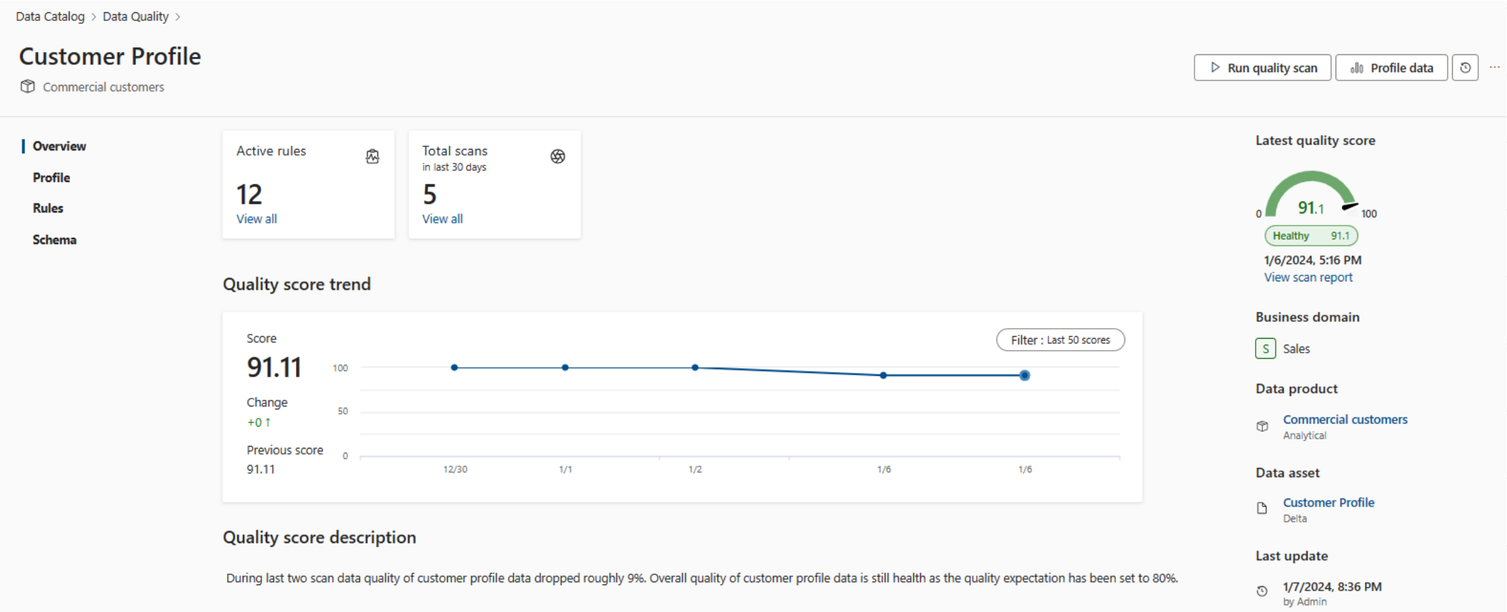
- **Azure Data Lake Storage Gen2**
  - File Types: Delta Parquet and Parquet
- **Azure SQL Database**
- **Fabric data estate in OneLake including shortcut and mirroring data estate.** **Data Quality scanning is supported only for Lakehouse delta tables and parquet files.**
  - Mirroring data estate: Cosmos DB, Snowflake, Azure SQL
  - Shortcut data estate: AWS S3, GCS, AdlsG2
- **Azure Synapse serverless and data warehouse**
- **Azure Databricks Unity Catalog**
- **Snowflake**
- **Google Big Query (Private Preview)**

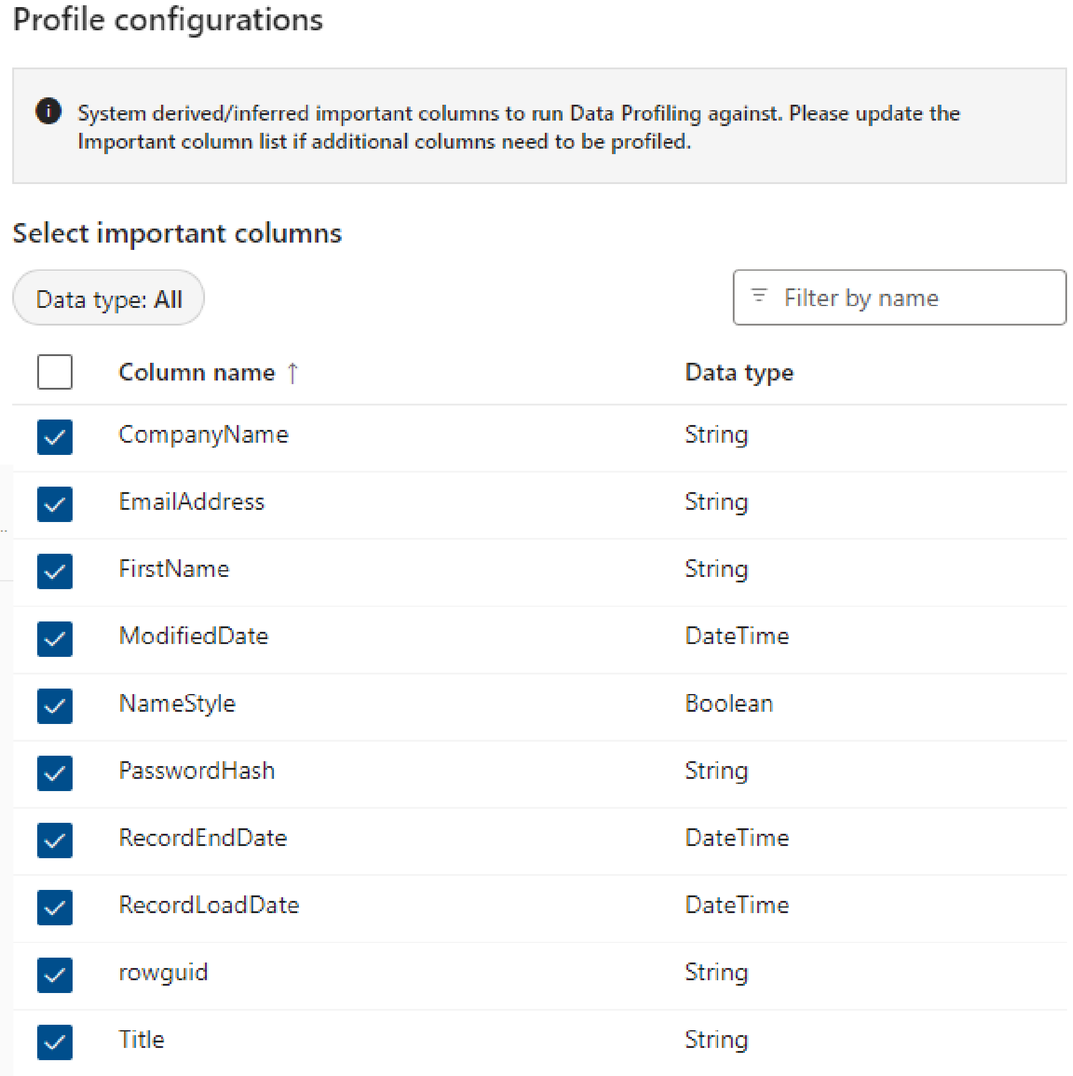


Microsoft Purview

Data Profiling

- The AI recommendation engine suggests potentially important columns to run data profiling against

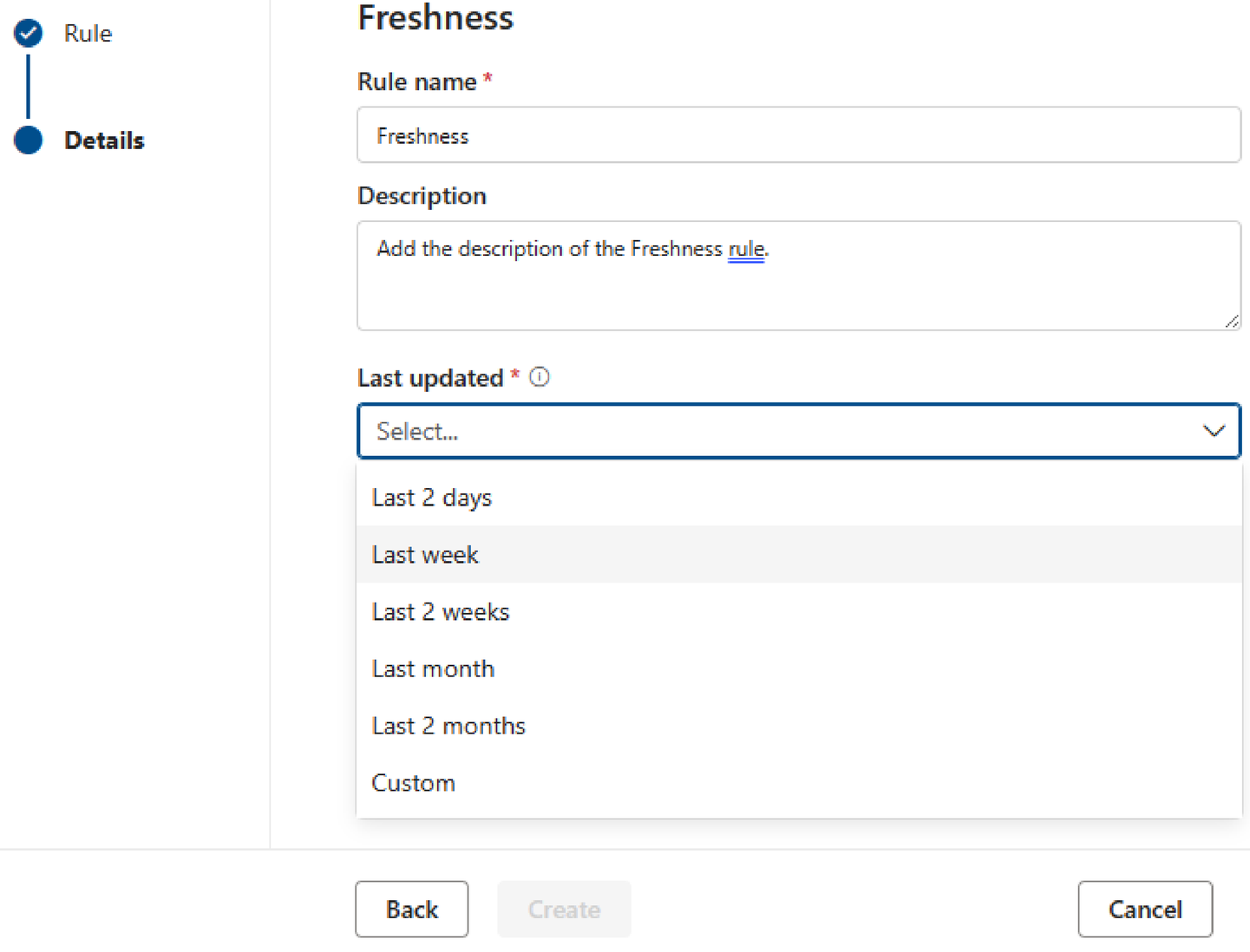




Microsoft Purview

Data quality Rules

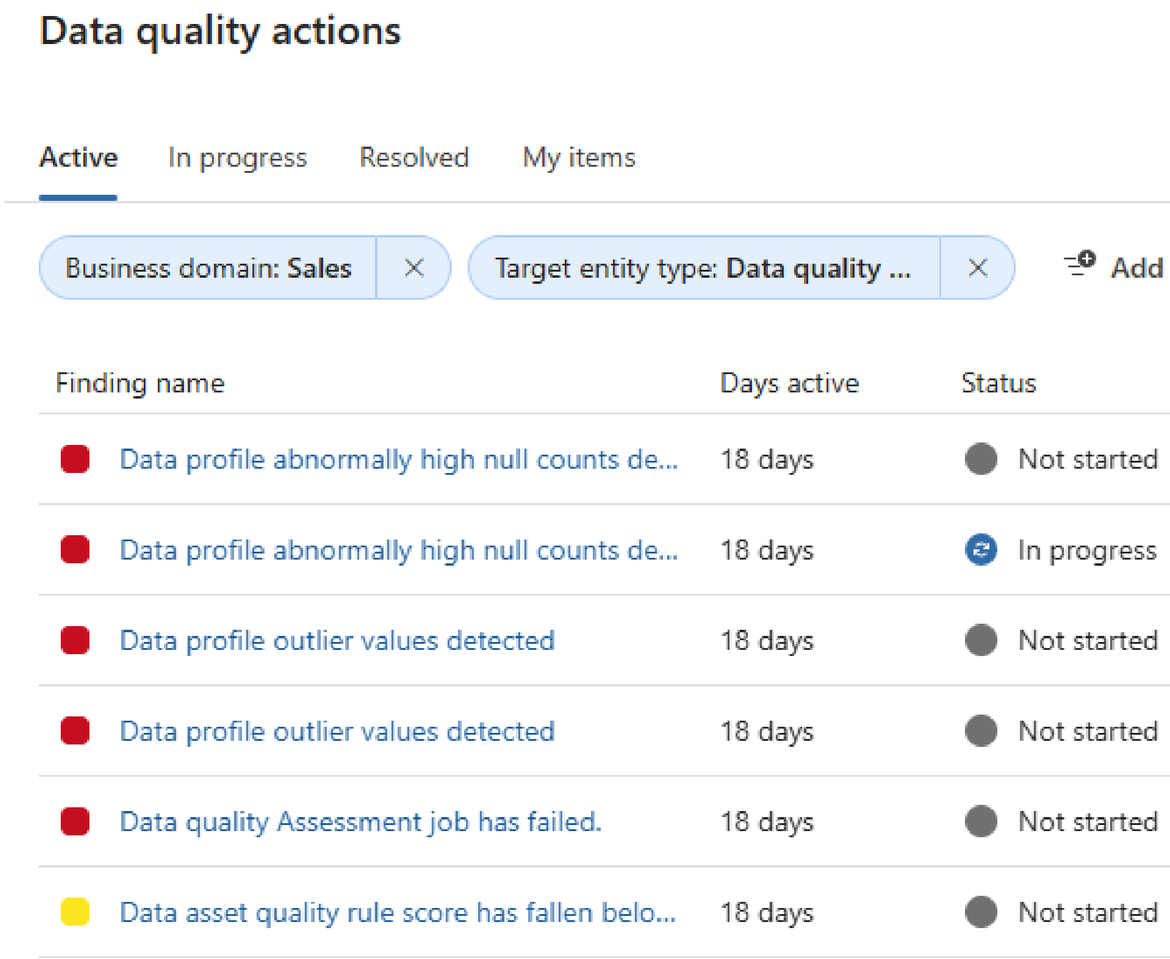
- Freshness
- Unique values
- String format match
- Data type match
- Duplicate rows
- Empty/blank fields
- Table lookup
- Custom



Microsoft Purview

Data quality Actions

- Rule has fallen below default thresholds
- Global data quality score has fallen below threshold
- data quality scanning and/or data quality profiling job has failed
- data quality scanning and/or data quality profiling job has skipped due to no changes on data since last run
- Outlier presence in profiled columns
- Too many nulls in the profiled data asset

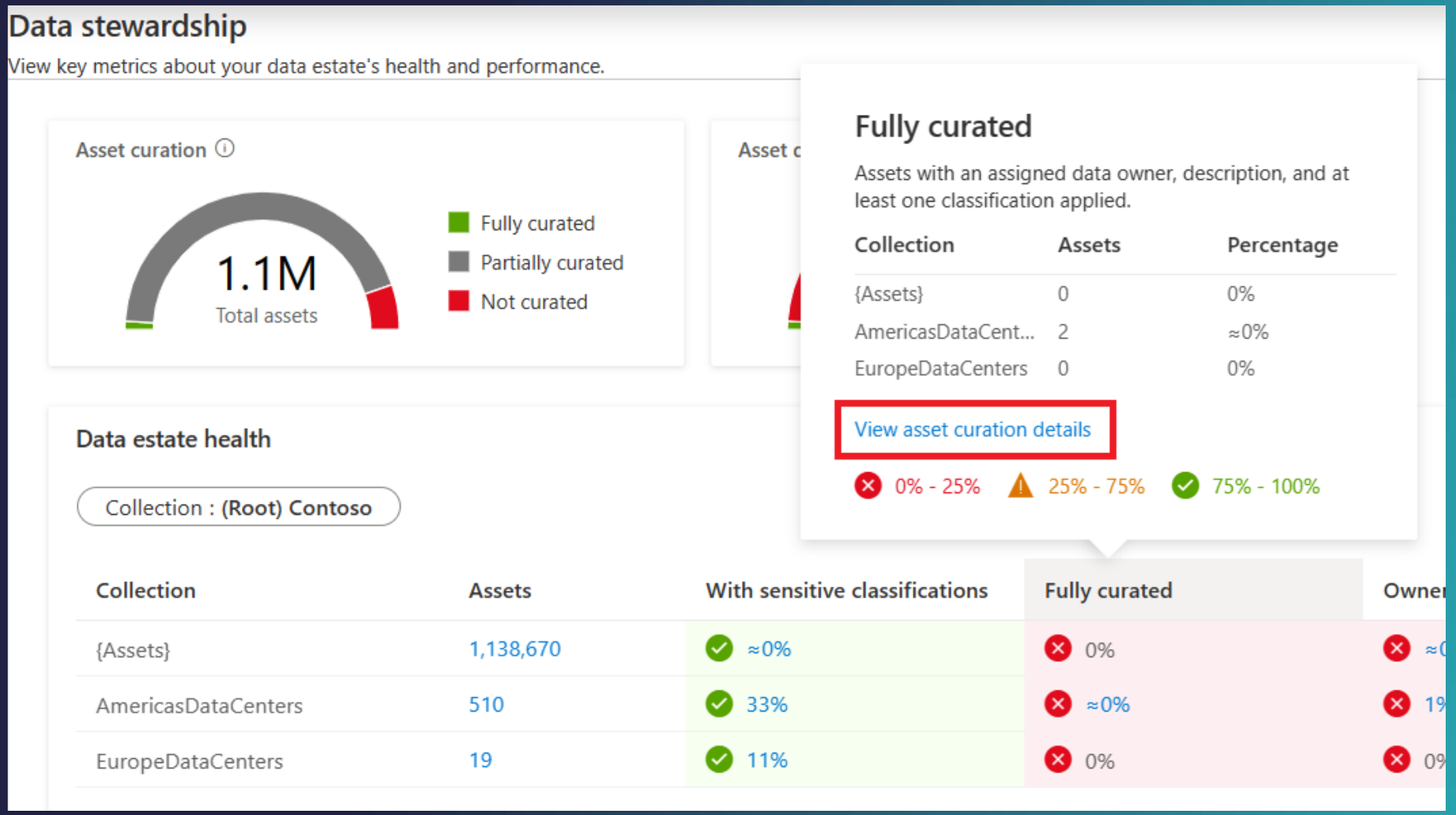




Demo

Microsoft Purview

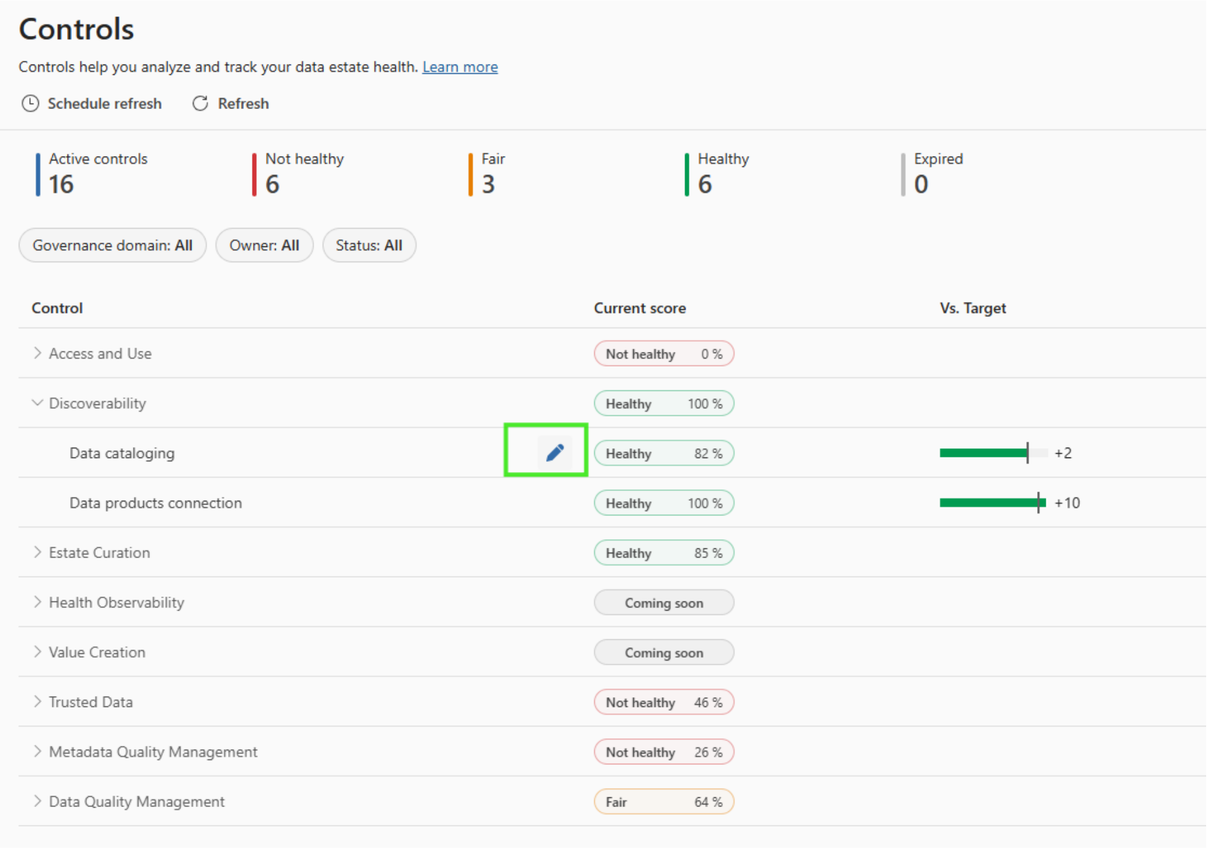
Data Estate Health



Microsoft Purview

Controls

- Health controls are rolled up into several categories that can be opened to view their parts. Category scores are an aggregate of the controls underneath them, and show their current score and their target, to pinpoint areas for growth.



Microsoft Purview

Actions

Data quality actions

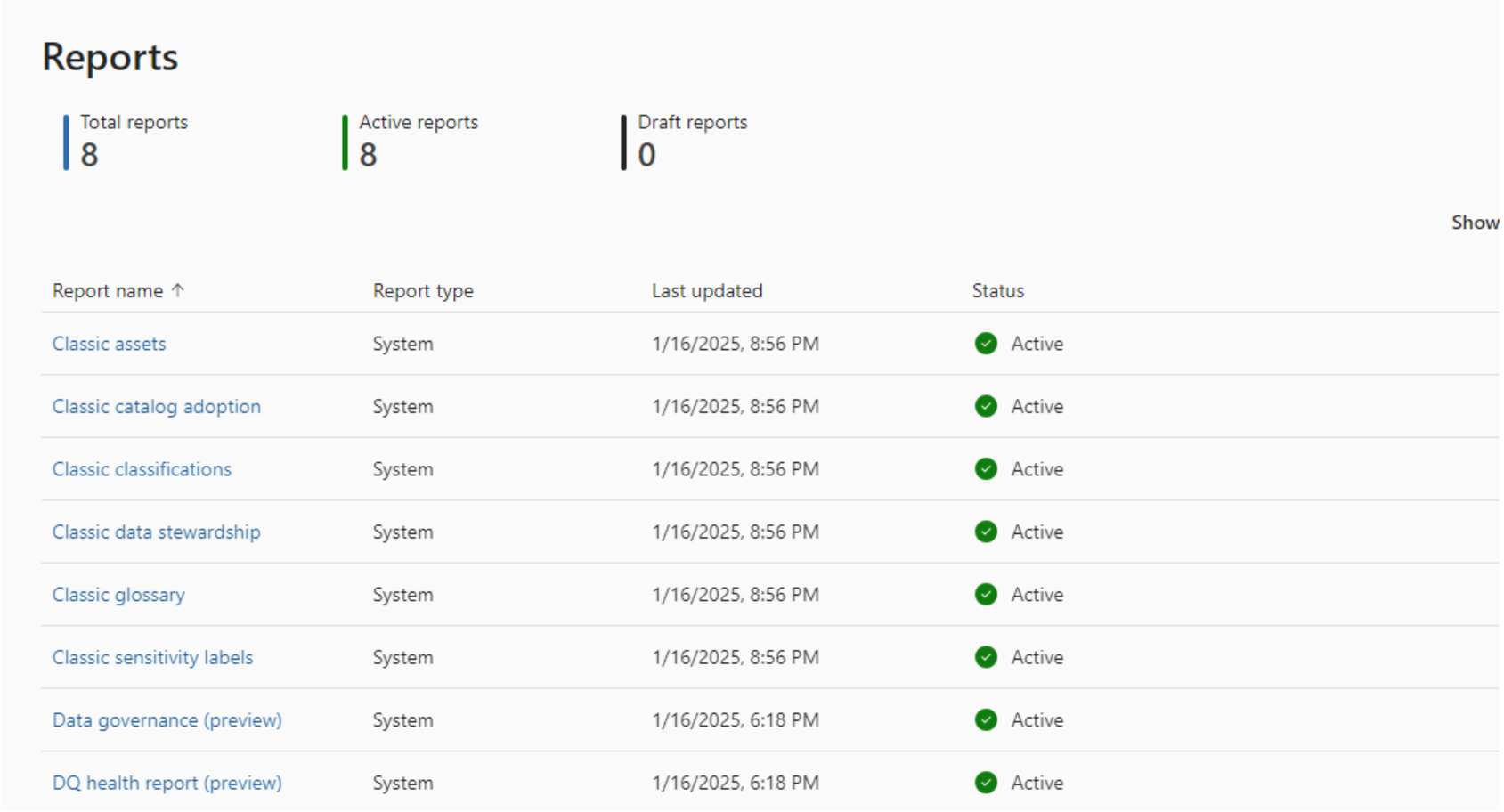
Active | In progress | Resolved | My items

Business domain: **Sales** | Target entity type: **Data quality ...** | Add filter | Showing 6 of 6 | Filter by keyword

Finding name	Days active	Status	Finding type	Finding subtype	Target entity
Data profile abnormally high null counts de...	18 days	Not started	Data Quality Management	Data profile	Address
Data profile abnormally high null counts de...	18 days	In progress	Data Quality Management	Data profile	Customer
Data profile outlier values detected	18 days	Not started	Data Quality Management	Data profile	SalesOrderDetail
Data profile outlier values detected	18 days	Not started	Data Quality Management	Data profile	Address
Data quality Assessment job has failed.	18 days	Not started	Data Quality Management	Data quality job unsuccessful	Product
Data asset quality rule score has fallen belo...	18 days	Not started	Data Quality Management	Data asset quality rule score	Customer

Microsoft Purview

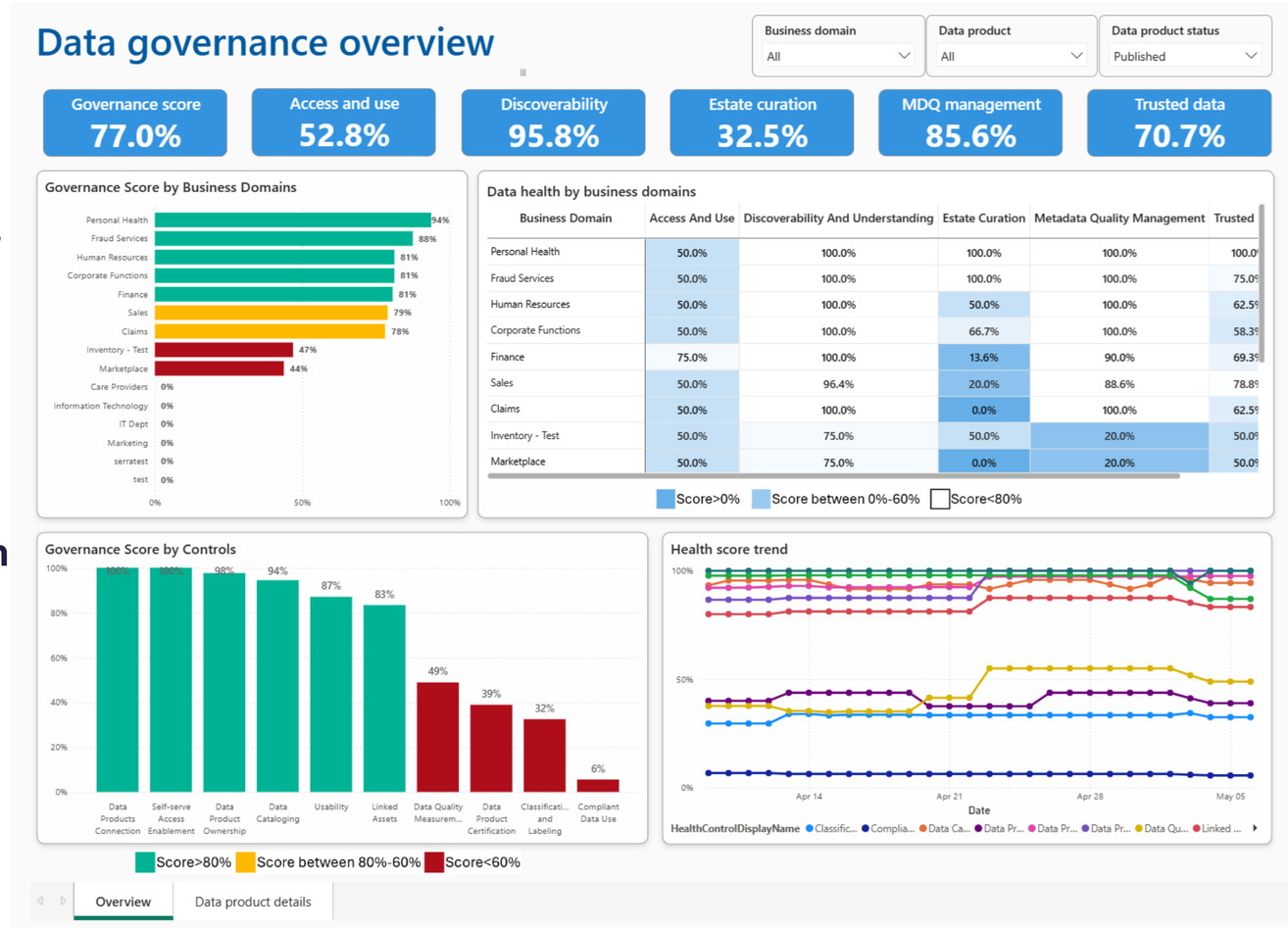
Reports



Microsoft Purview

Reports

- The Data Governance Health Report has dependencies on Data Health Controls. If customers do not use the Data Health Controls, the Data Governance Health Report will not be refreshed. Please ensure that you have setup the schedule to run Data health controls an refreshing the health controls scores and Data governance health report.

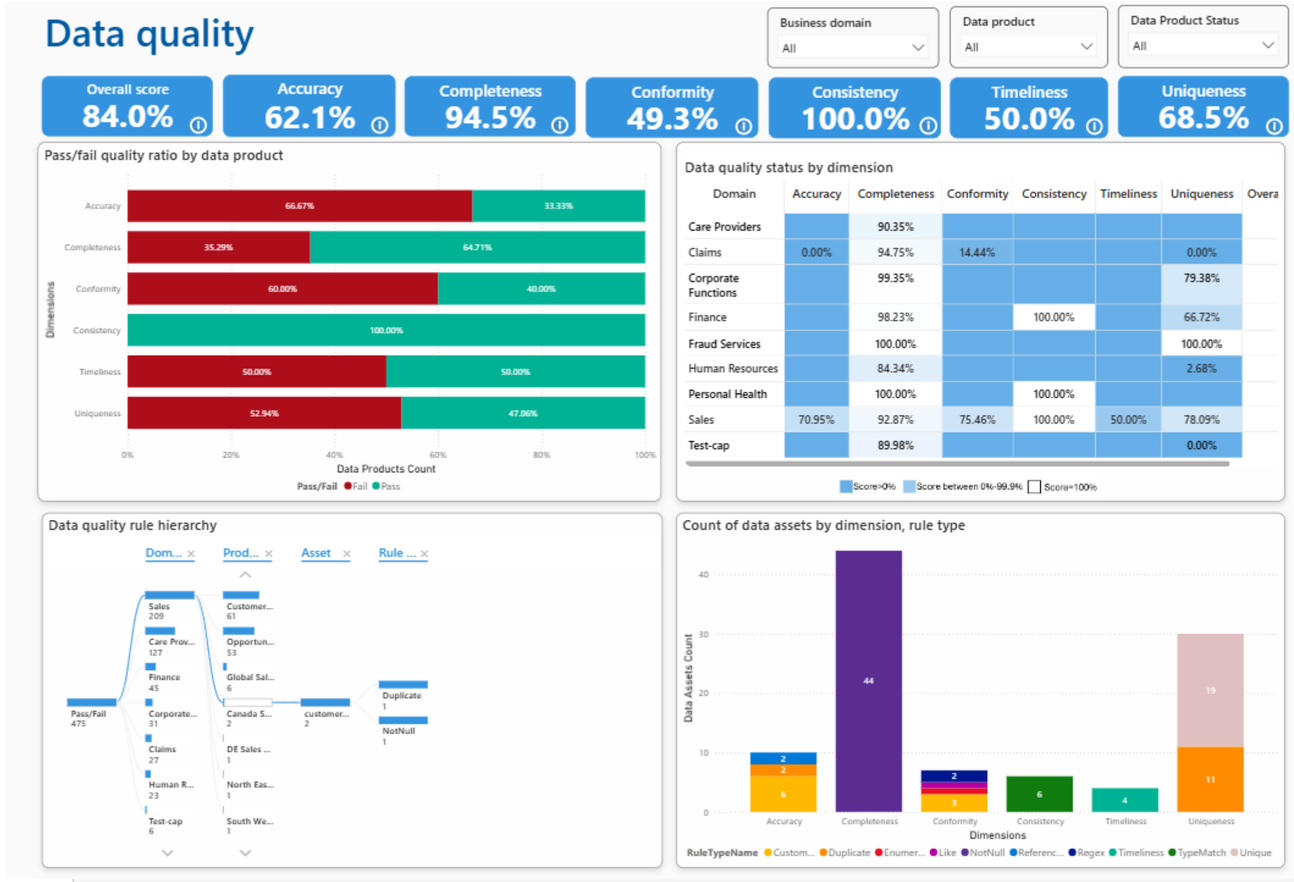


Microsoft Purview

Reports

- The Data Quality Health Report has dependencies on Data Health Controls and Purview Metadata Self-serve analytics model. If customers do not use the Data Health Controls and do not subscribe Purview Unified Catalog metadata then the Data Quality Health Report will not be refreshed. Customers either need to use Data Health Controls or need to subscribe purview metadata for self-serve analytics to refresh Data Quality health report.

- If customers don't use Data Quality feature, the Data Quality Health report will show blank because Data Qualty Health report is created using data that are



Microsoft Purview

Solution Integrations (Self-serve analytics (Preview))

Unified Catalog

- Roles and permissions
- **Solution integrations (Preview)**
- Usage settings (Preview)

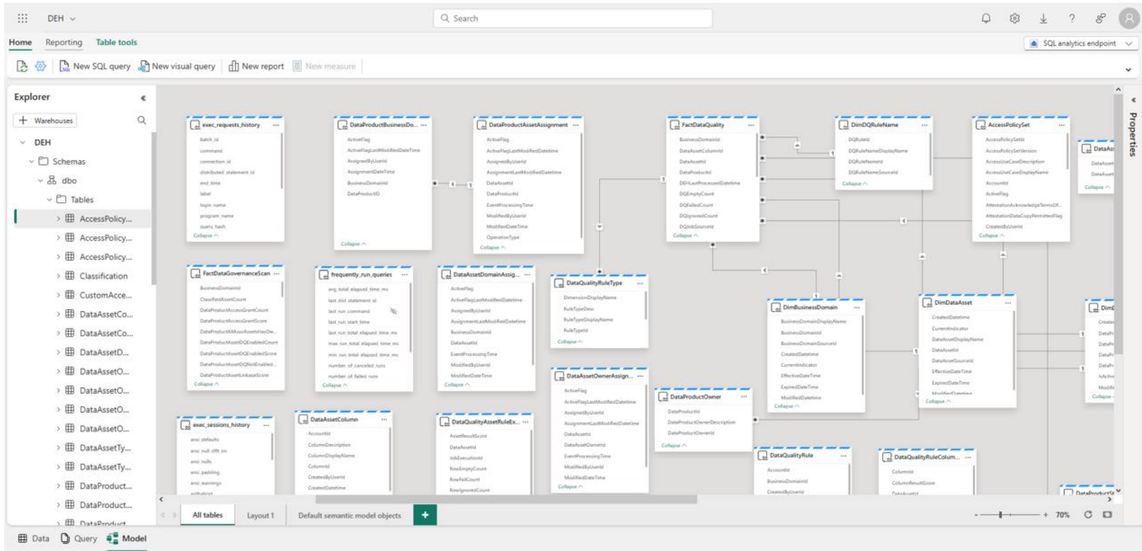
Self-serve analytics (Preview)

Enhance your data estate health with self-serve analytics and insights.

Configure storage

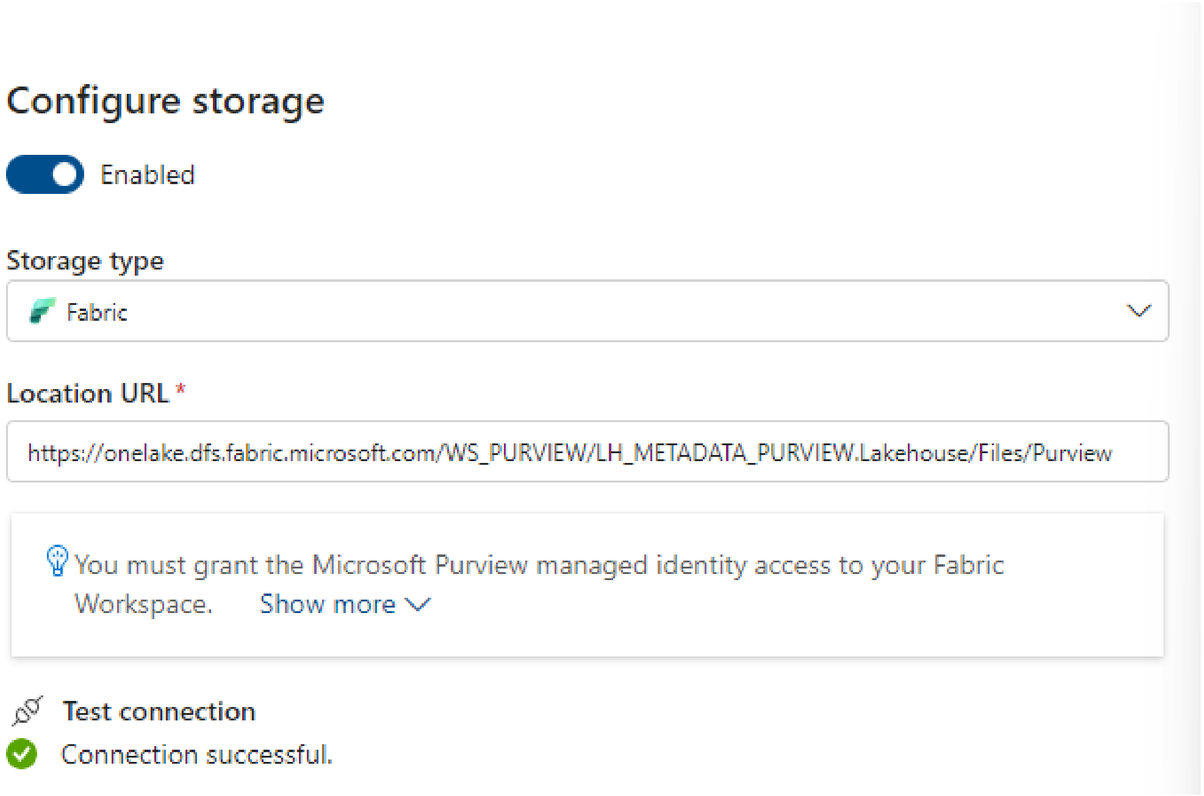
Edit

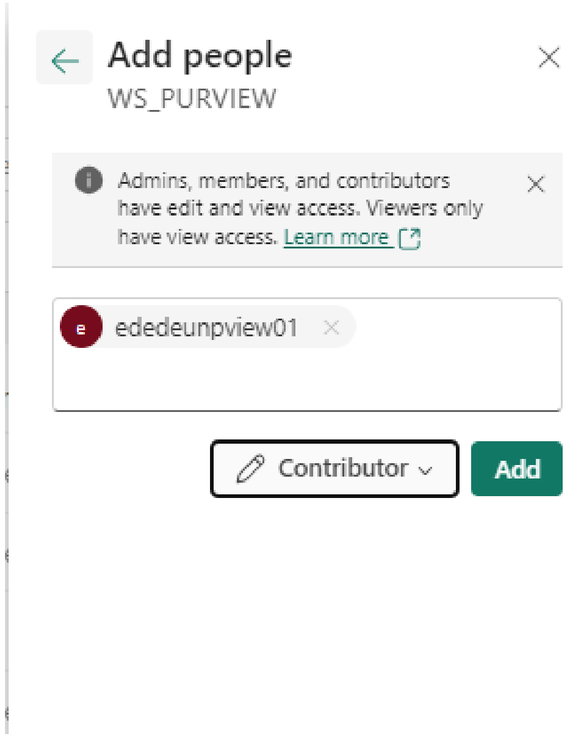
Status:	Disabled
Storage type:	-
Last published:	-



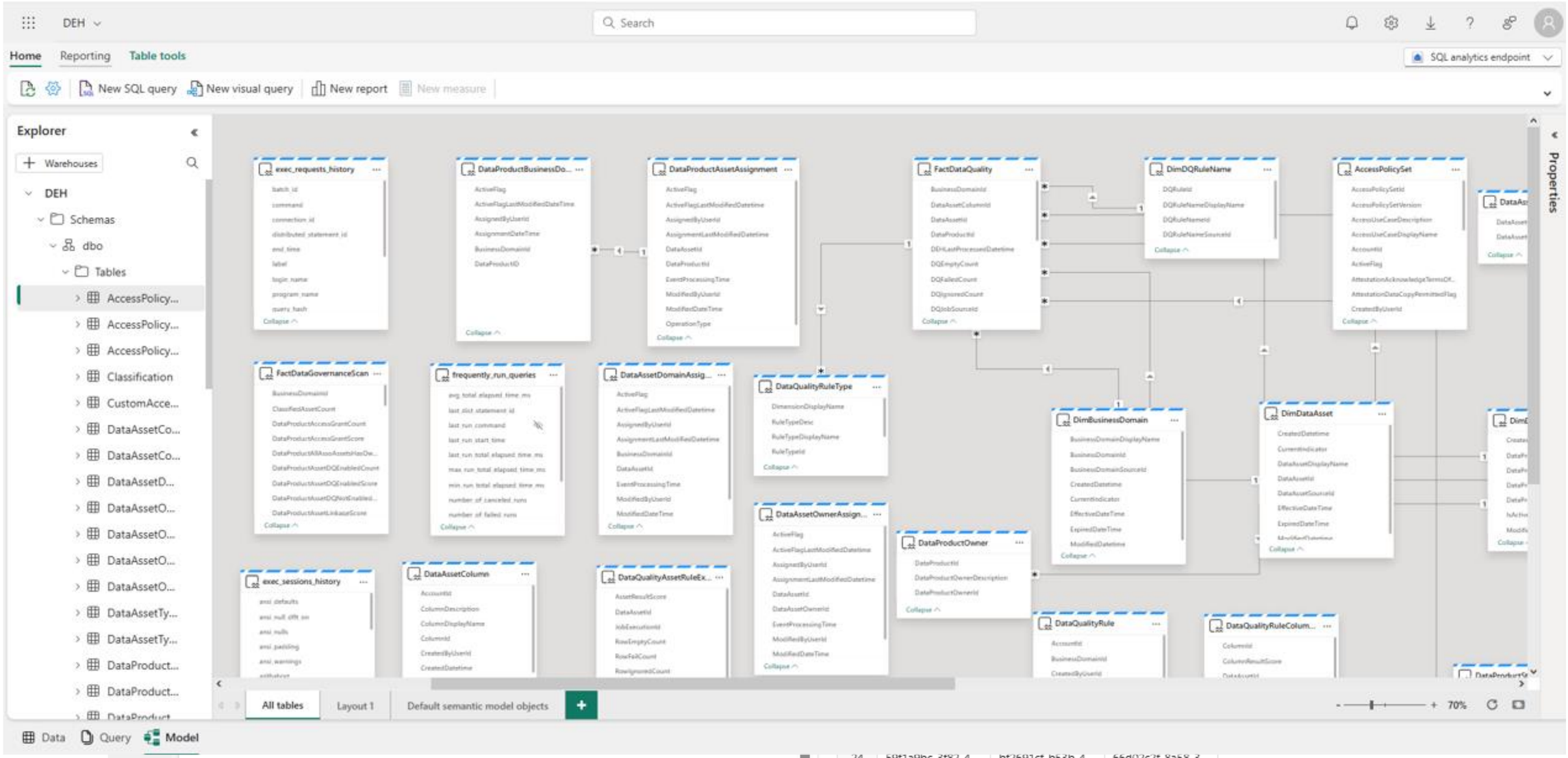
INSPARK

Solution Integrations (Self-serve analytics (Preview))





Solution Integrations (Self-serve analytics (Preview))



Microsoft Purview

Master Data Management

PARTNER INTEGRATIONS









Microsoft Purview

Pricing



Pricing

- **Number of Unique Governed Assets per Day**: This metric counts the unique technical assets, such as tables, files, datasets, and reports, that are actively managed and curated within the Microsoft Purview Unified Catalog. Only assets associated with governance concepts, like data products or critical data elements, are considered governed assets.
- **Data Governance Processing Units (DGPU) per Run**: DGPUs are fully managed compute units used for running compute-heavy capabilities, such as data quality and data health management. Each DGPU represents 60 minutes of compute time, which can be run across varying sets of nodes based on the workload needs.

INSPARK

Pricing

Microsoft Purview Data Catalog (new)

Enterprise Data Catalog

Data Curation
Data Discovery and Access
Governance Policies

$0.50 per Governed Asset per Month

Enterprise Data Management

Data Quality
Data Health Management

$ per Processing Unit

Basic $15 | Standard $60 | Advanced $240

Compute Power

Example customer profile

		Approx Monthly Costs
Average % of data assets governed (the larger the data estate, smaller % governed)	15-25%	
Average number of asset governed (orgs w/ data estates >1TB)	100,000	$50,000
Approx Data Management Rules + Controls (per month, based on 100k assets governed*)	500	$18,375

Dimensions that influence costs

Number of governed assets
Less — More
Impacts Data Catalog meter

Frequency of Data Management Actions (eg Data Quality Rules)
Less — More
Impacts Data Management units

Complexity of estate and/or Data Management Actions
Less — More
Influences Data Management power

*Assumes each Data Management Action consumes an average of .5 Data Management Processing Unit.
Mix of 50% Basic, 30% Standard, 20% Advanced. Data Management actions examples are 1) running a Data Quality rules over your data assets 2) running controls within Data Health. This example customer profile is illustrative only and does not represent a price quote – each customer environment will be different.



Demo


Feedback

Erwin de Kreuk

Feedback

Thank You



@erwindekreuk

linkedin.com/in/erwindekreuk

erwindekreuk.com

github.com/edkreuk

https://sessionize.com/erwin-de-kreuk/

Let's connect

MVP Microsoft® Most Valuable Professional



INSPARK