

Azure Purview

Microsoft's answer to Data Governance and Data Lineage



Erwin de Kreuk

Lead Data and AI



@erwindekrek



<https://erwindekrek.com>

DATA TOBOGGAN



Azure Purview

Microsoft's answer to Data Governance and Data Lineage



Erwin de Kreuk

Lead Data and AI



@erwindekrek



<https://erwindekrek.com>

We Are InSpark

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

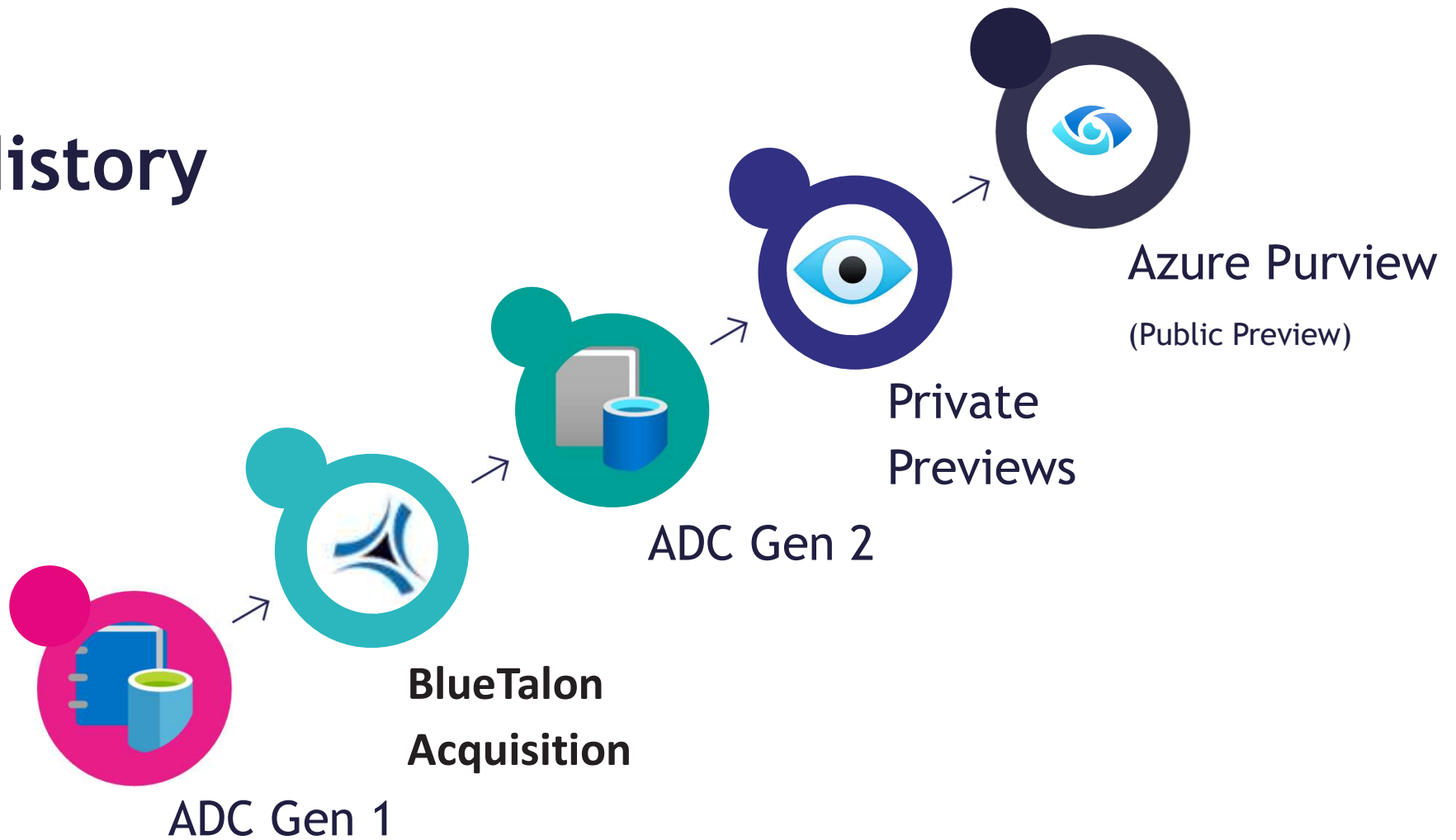
Azure Purview



Unified **data governance** to
maximize the business value of data



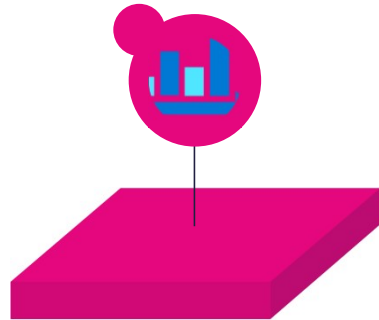
History



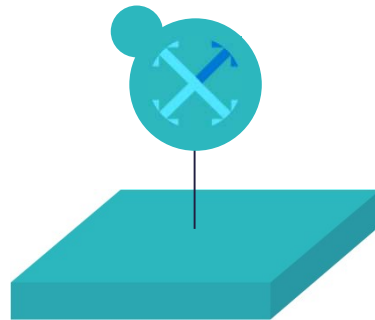
Data governance is becoming increasingly interdisciplinary



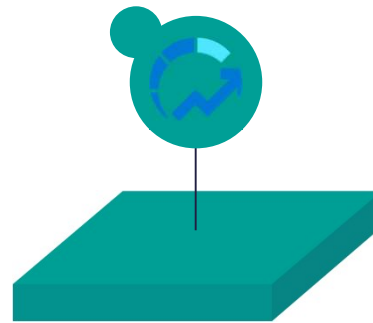
Elements of successful data governance



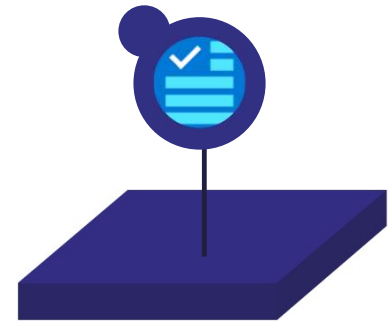
Manage growing
data landscape



Overcome
operational silos



Increase data
agility



Comply with
industry
regulations

Data Insights

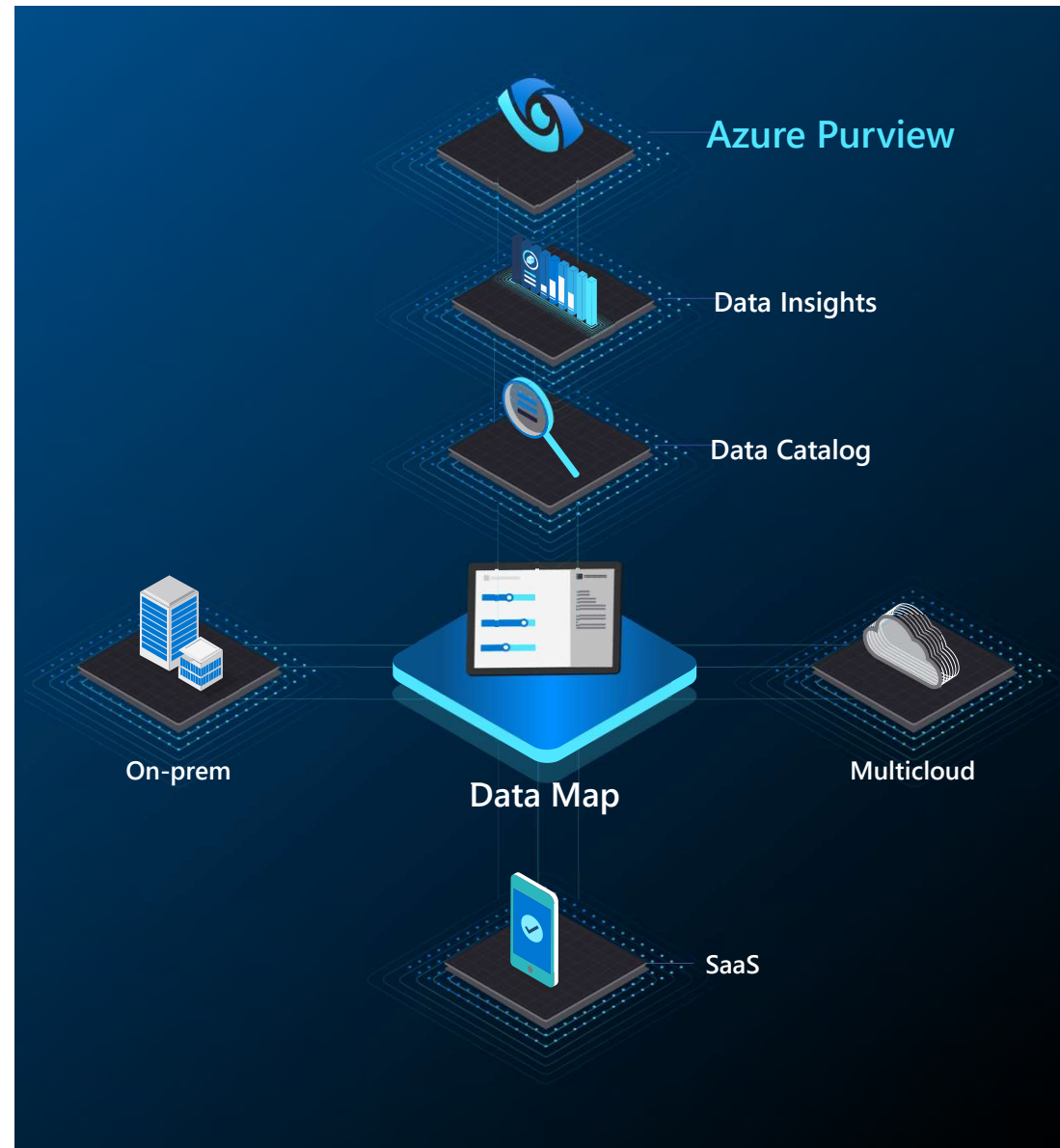
- Assess data usage across your organization

Data Catalog

- Enable effortless discovery for data consumers


Data Map

- Automate and manage metadata at scale



Azure Purview


Unified data governance to
maximize the business
value of data





Reimagine **data**
governance in the cloud



Set the foundation for
effective **data governance**



Maximize business value
of **data** for data
consumers



Gain insight into **data** use
across the estate



Reimagine **data** **governance** in the cloud

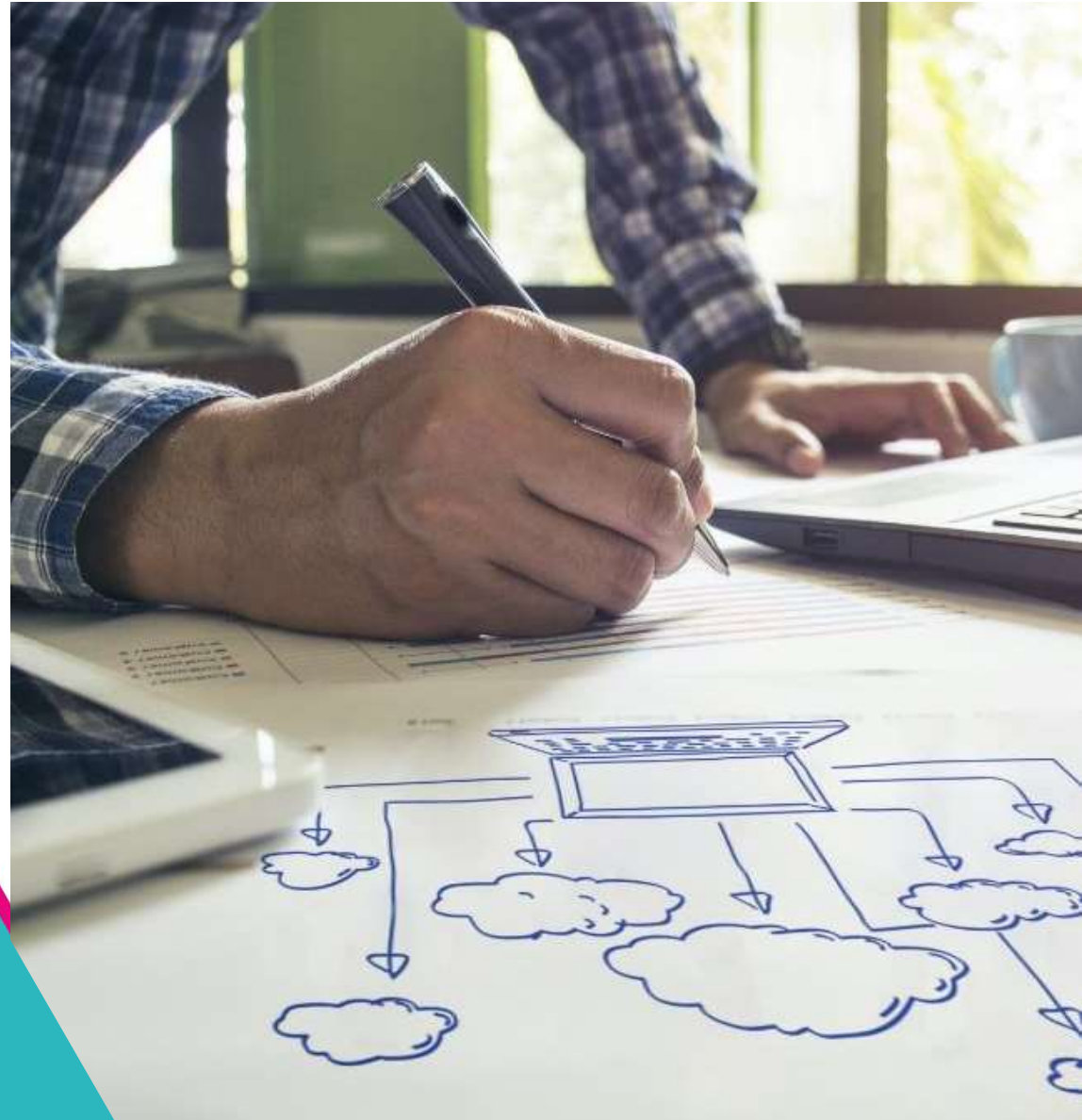
- ❑ Manage and govern operational, transactional and analytical data
- ❑ Cloud-native, purpose-built service to address discovery and compliance needs
- ❑ Fully managed, serverless, PaaS service
- ❑ Eliminate manual, ad-hoc and homegrown solutions





Set the foundation for effective **data governance**

- ❑ Automate discovery of data in on-premises, multicloud 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





Maximize business value of **data** for data consumers

- ❑ 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





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



Azure Purview Features

Azure Purview Studio

Azure Purview Catalog included with Platform (C0)

- Search, browse, asset details
- Automated meta-data and lineage extraction
- Automated classification based on content inspection
- Private Endpoint
- Management center

Azure Purview Platform

Data Map

- Dedicated per customer on shared infra
- Provisioned default capacity with option to add-on capacity

Automated Scanning & Classification

- Serverless, pay per use
- Includes connectors, scanning of sources, processing into data assets, lineage capture, classification

Power BI

SQL Server on-prem

Azure Synapse

Azure Data Services

M365 Compliance Center

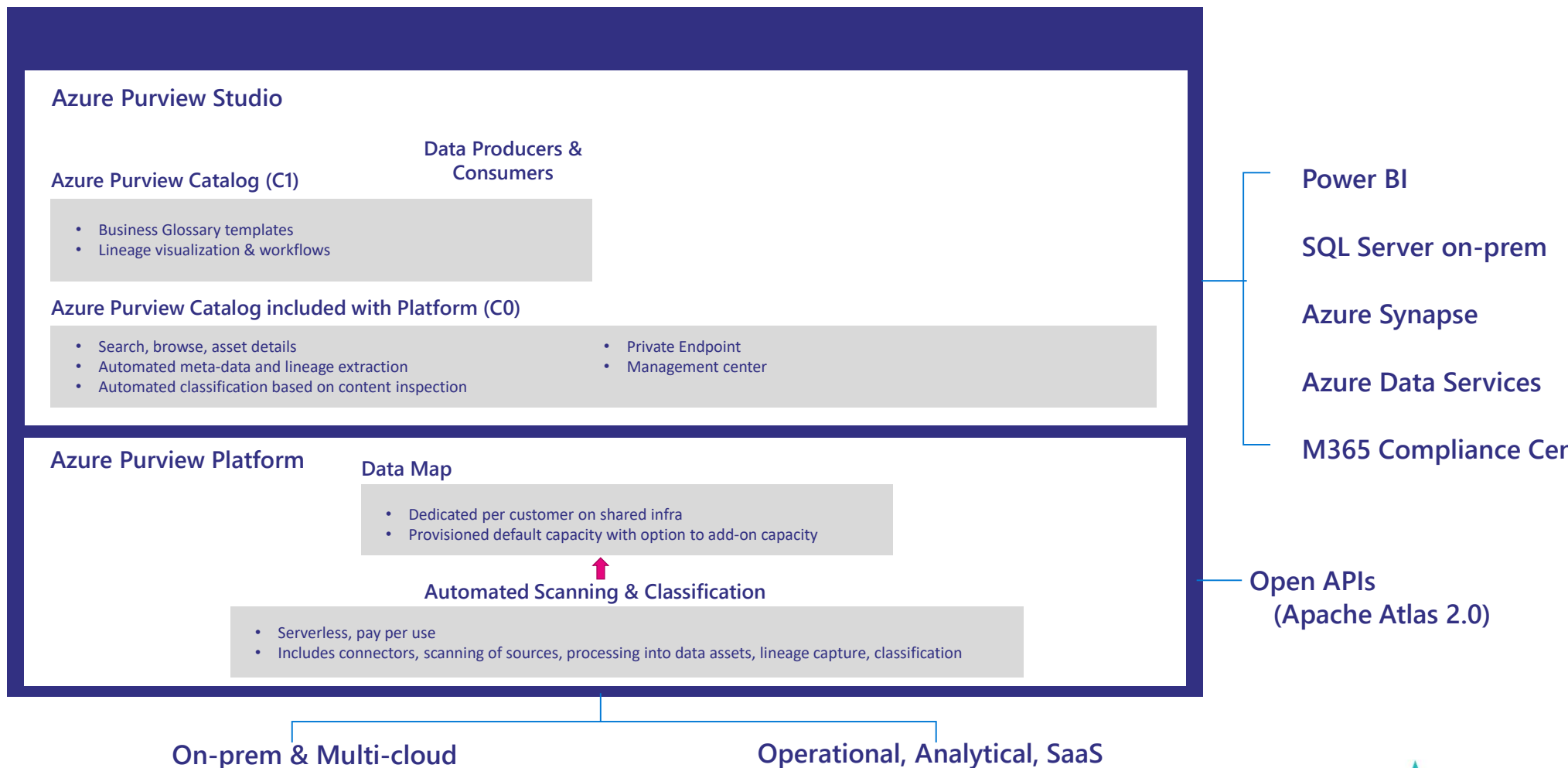
Open APIs
(Apache Atlas 2.0)

On-prem & Multi-cloud

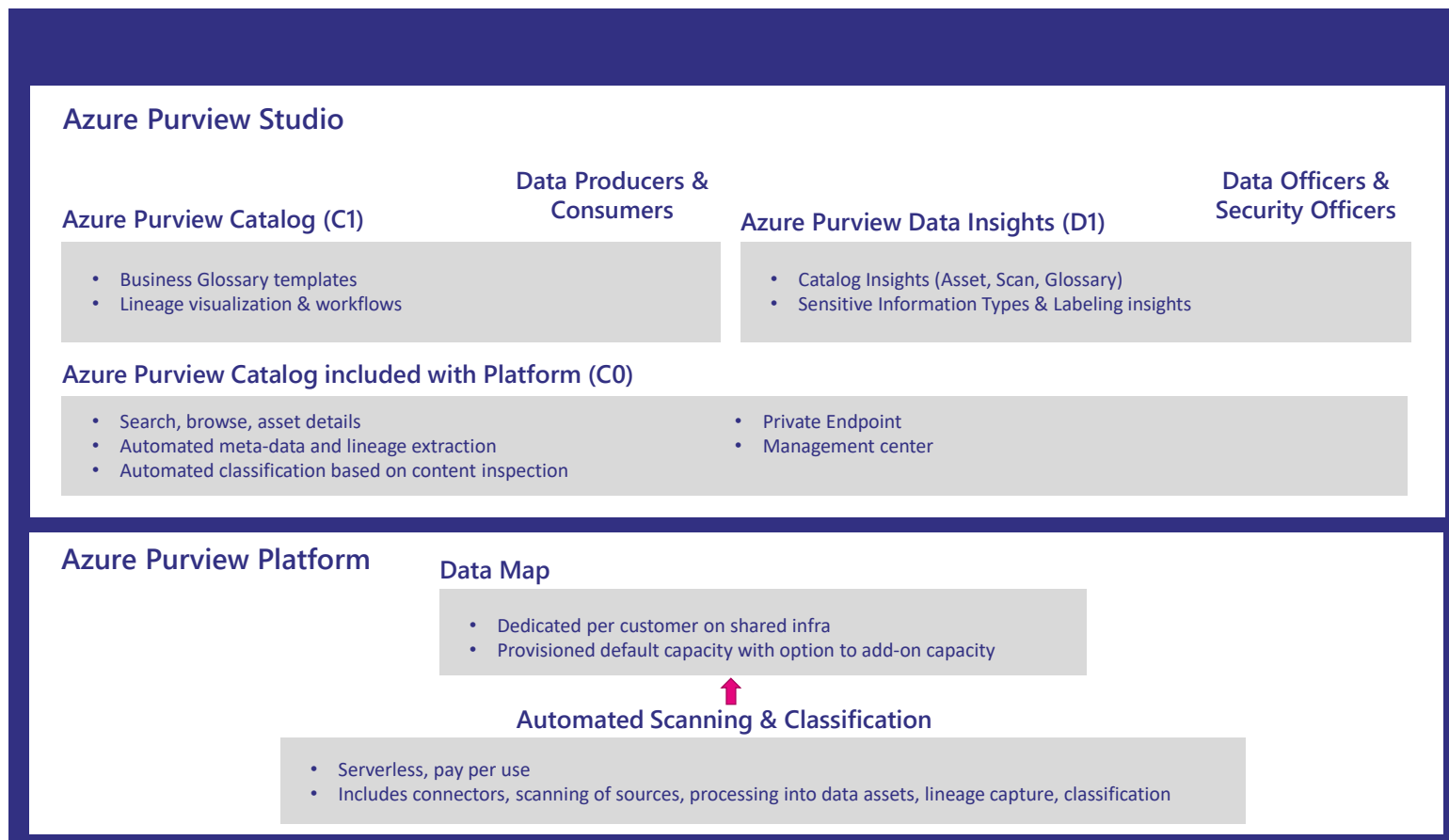
Operational, Analytical, SaaS



Azure Purview Features



Azure Purview Features



Power BI

SQL Server on-prem

Azure Synapse

Azure Data Services

M365 Compliance Center

Open APIs
(Apache Atlas 2.0)

On-prem & Multi-cloud

Operational, Analytical, SaaS



Azure Purview - Roles

Data Source Administrator

- No access to Purview Portal
- Can Manage all aspects of Scanning
- Ideal role for programmatic processes, such as service principals
- Can register Data Sources



Select a Purview account

Azure Purview is a unified data governance service that enables a foundational understanding of your on-premises, multi-cloud, operational and SaaS data. [Learn more](#)

Azure Active Directory

Account name

✖ You don't have access to this Purview account. Contact an Owner or User Access Administrator (UAA) for access in [Access control \(IAM\)](#)

Continue

Azure Purview - Roles

Data Source Administrator

Data Reader

- Has access to Purview Portal
- Can read all content in Azure Purview

The screenshot shows the 'Account information' page in the Azure Purview portal. The left sidebar contains a navigation menu with the following items: General, Account information (selected), Metrics, Metadata management, Classifications, Classification rules, External connections, Data Factory, Data Share, Security and access, and Access control. The main content area is titled 'Account information' and features a blue information banner at the top stating: 'By participating in Purview Public Preview, you acknowledge that you have read and you agree to the Supplemental Terms of Use for Microsoft Azure Previews, including additional terms for Purview Public Preview. You can find those terms [here](#).' Below this banner is an 'Overview' section with two fields: 'Account name' with the value 'labseuwdvimpurviewoxgn01' and 'Friendly name * ⓘ' with the value 'Azure Purview Demo'.

The screenshot displays the 'Azure Purview Demo' portal interface. At the top, it shows '1,138 assets' and '1 glossary term'. Below this is a search bar labeled 'Search catalog'. The main area features four interactive tiles: 'Knowledge center' (Discover learning materials and tutorials), 'Browse assets' (Explore assets by source type), 'View glossary' (View business glossary managed by your team), and 'View insights' (Get insights on your data). The background of the interface is decorated with a network diagram of data flows and server icons.

Azure Purview - Roles

Data Source Administrator

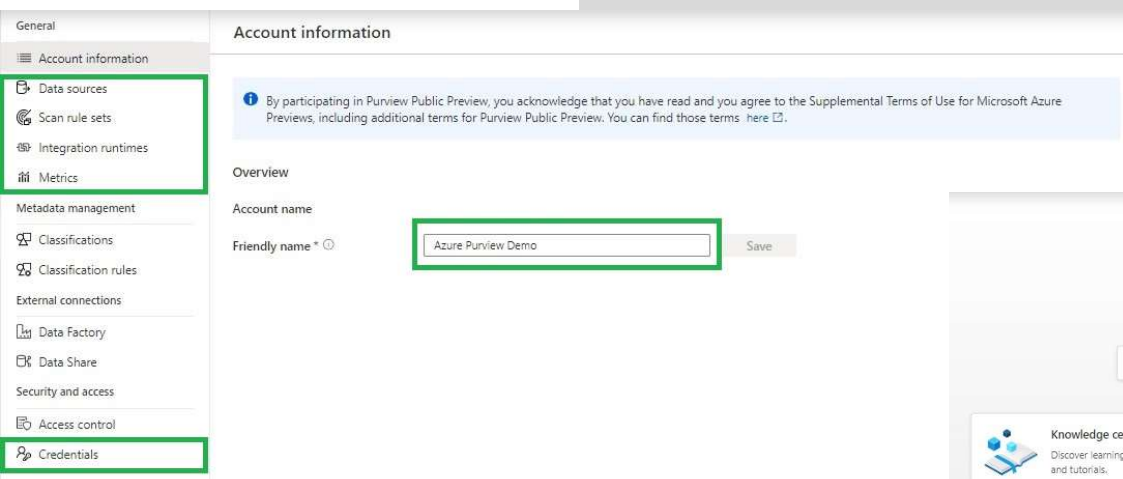
Data Reader

Data Curator

- Has access to Purview Portal
- Can read all content in Azure Purview
- Can edit assets, classification and glossary terms
- Can apply classifications and glossary terms to assets.
- Can not Register Data Sources, only read

 Error

Request failed with status code 403
Not authorized to access account



General

Account information

Data sources

Scan rule sets

Integration runtimes

Metrics

Metadata management

Classifications

Classification rules

External connections

Data Factory

Data Share

Security and access

Access control


Credentials

Account information

By participating in Purview Public Preview, you acknowledge that you have read and you agree to the Supplemental Terms of Use for Microsoft Azure Previews, including additional terms for Purview Public Preview. You can find those terms [here](#).

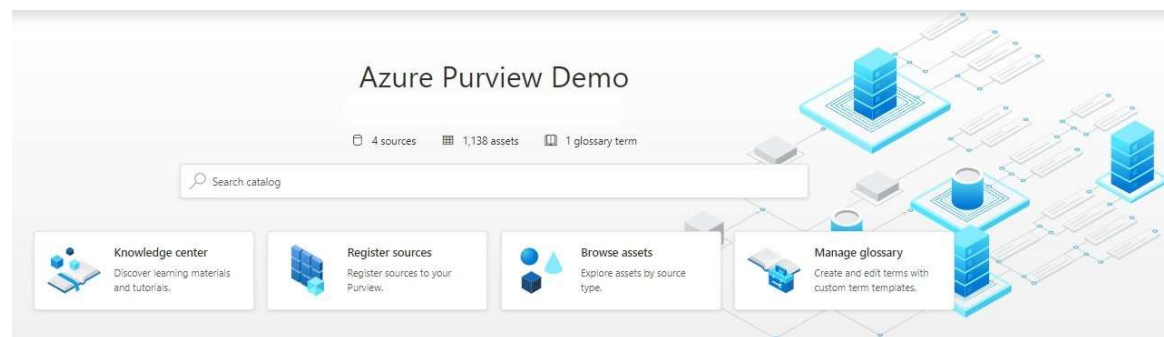
Overview

Account name

Friendly name * 

Azure Purview Demo

Save



Azure Purview - Roles































Data Source Administrator

Data Reader

Data Curator

Quick action buttons on
Homepage

All Features (C0 + C1 + D1)		
C0 (Sources registration, automated scanning and classification, data discovery.)	C1 (Business glossary and lineage visualization)	D1 (Catalog insights and sensitive data insights)

	Knowledge center	Register sources	Browse assets	Manage/View glossary	View insights
 Data source admin CRUDScanBinding, ScanClassification, CRUDSHIR, ManageDataCenterScopes					
 Data curator Read, Mutate, ApplyClassification, ApplyGlossaryTerm	 Knowledge center Discover learning materials and tutorials.		 Browse assets Explore assets by different types.	 Manage glossary Create and edit terms with custom term templates.	 View insights Get insights on your data.
 Data reader Read	 Knowledge center Discover learning materials and tutorials.		 Browse assets Explore assets by different types.	 View glossary View business glossary managed by your team. READ-ONLY	 View insights Get insights on your data.
 Data source admin + Data curator	 Knowledge center Discover learning materials and tutorials.	 Register sources Register sources to your Purview.	 Browse assets Explore assets by different types.	 Manage glossary Create and edit terms with custom term templates.	
 Data source admin + Data reader	 Knowledge center Discover learning materials and tutorials.	 Register sources Register sources to your Purview.	 Browse assets Explore assets by different types.	 View glossary View business glossary managed by your team. READ-ONLY	

Azure Purview Studio

Updates

Accounts

Notifications

Metrics

Search Bar

Recently Accessed Entities

Azure Purview Demo

Search catalog

Knowledge center

Register sources

Browse assets

Manage glossary

Recently accessed

My items

Name	Last update
COPY_ADLS_DATASTORE	---

Links

Purview overview

Discover the capabilities offered by Purview and learn how to make the most of them.

Feedback

Key Activities

Usefull Links

Azure Purview Studio - Activity hubs



Home



Sources



Glossary



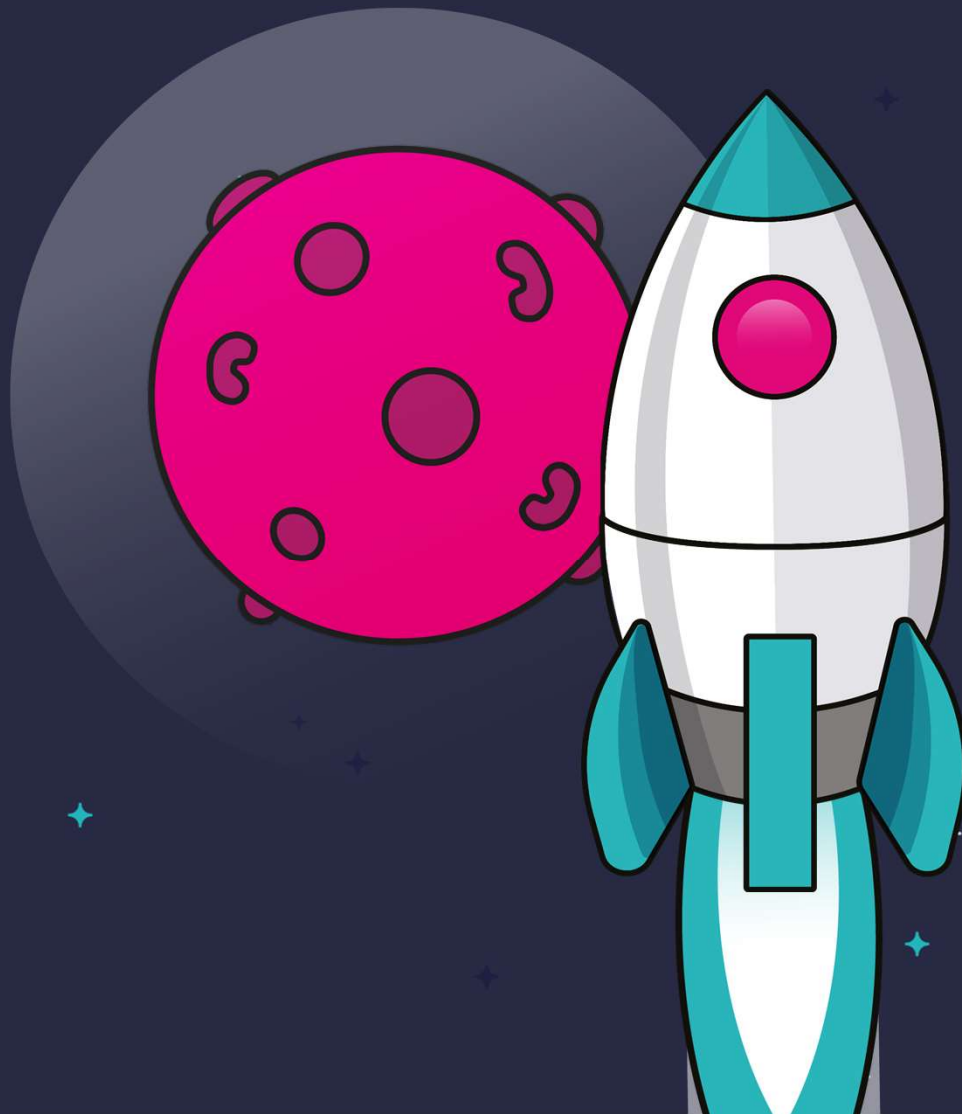
Insights

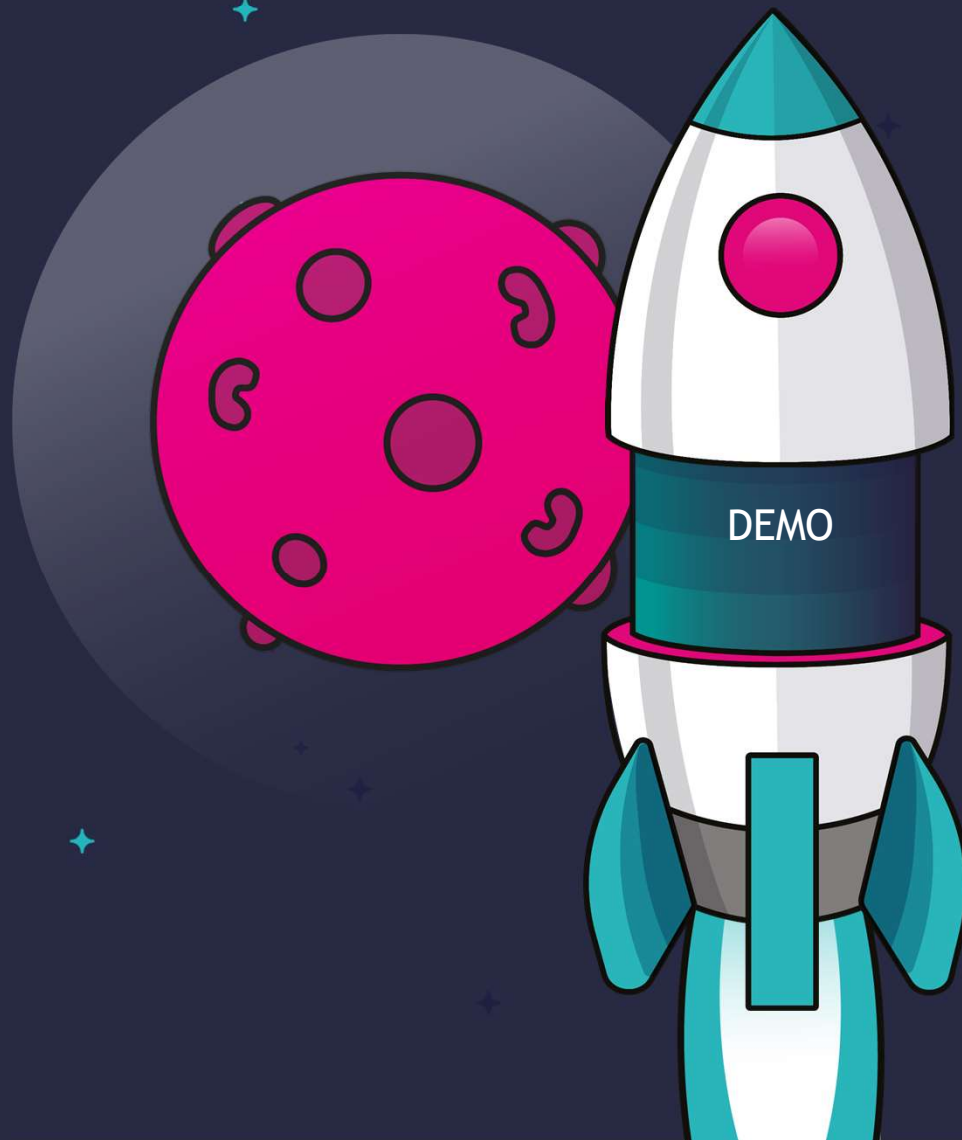


Management Center

- Quick Actions, recently accessed items, owned Items, search bar and Documentation
- Create collections, register data sources and setup Scans
- Manage Glossary Items, search, manage terms templates and custom attributes, import and export Terms using csv
- Insights on your data
- Meta Data Management-classifications-resource sets, data sources, integration runtime, Alerts, Security, ADF and data share Connections

 **INSPIRE**





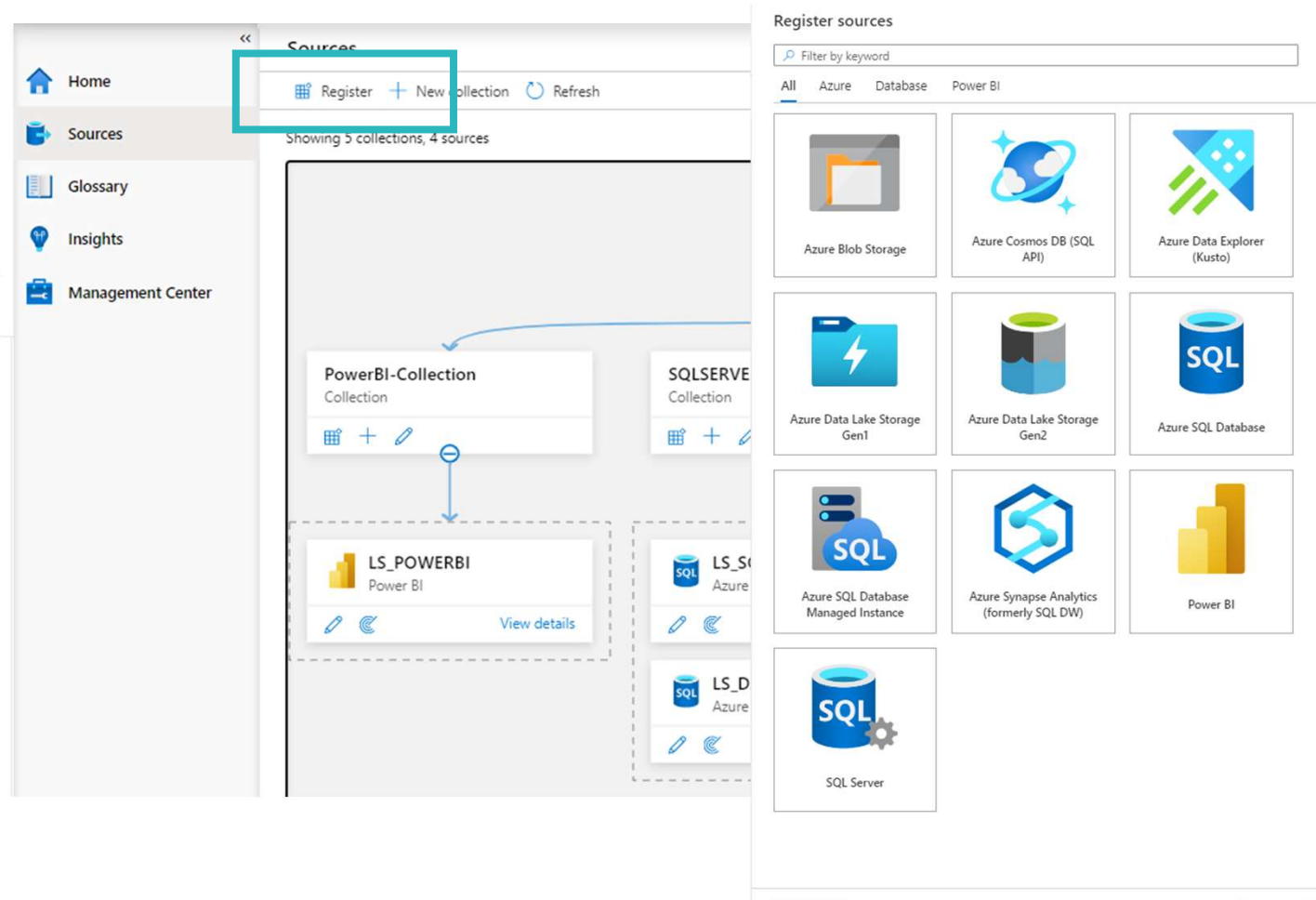
Innovate to accelerate

DEMO

Purview Data Map

Unify and make data meaningful

- ❑ Automated metadata scanning and lineage identification of hybrid data stores
- ❑ 100+ built-in and custom classifiers
- ❑ Microsoft Information Protection sensitivity labels



Purview Data Map

Unify and make data meaningful

- ❑ Automated metadata scanning and lineage identification of hybrid data stores
- ❑ 100+ built-in and custom classifiers
- ❑ Microsoft Information Protection sensitivity labels



Sources > LS_ADLS > Weekly_Scan run history

Weekly_Scan run history

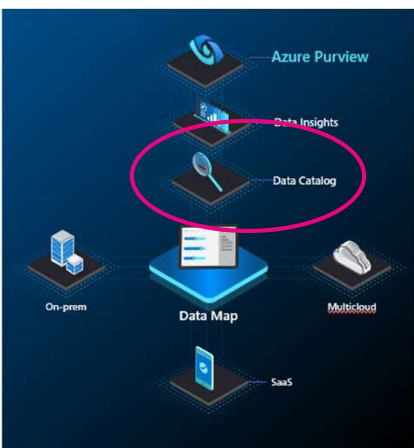
Run scan now Edit scan Delete scan Refresh

Status	Schedule	Assets classified	Assets discovered	Start time	End time
Completed	Scheduled	<div>Scope your scan</div> <div>Refresh</div> <p>All future assets under a certain parent will be automatically selected if the parent is fully or partially checked.</p> <div>Search</div> <div><div>LS_ADLS</div><div><div>azure-webjobs-hosts</div><div>azure-webjobs-secrets</div><div>intermediate</div><div>marts</div><div>raw</div><div>sandbox</div><div>temp</div><div>tempstaging</div><div>transfer</div></div></div>			
Completed	Scheduled				
Completed	Scheduled				
Completed	Manual				
Completed	Manual				
Completed					

Azure Purview Data Catalog

Enable effortless discovery

- ❑ Semantic search and browse
- ❑ Business glossary and workflows
- ❑ Data lineage with sources, owners, transformations, and lifecycle



Azure Purview Demo

4 sources 1,138 assets 1 glossary term

Knowledge c
Discover learn
and tutorials.

Manage glossary
Create and edit terms with
custom term templates.

Search suggestions

- customertransactions original
- customer changes
- customer origin
- customer original

Asset suggestions

- Customer
https://labseuwdvlimdioxgn01.dfs.core.windows.net/intermediate/DeltaLake/AW/SalesLT/Customer/
- Customer
https://labseuwdvlimdioxgn01.dfs.core.windows.net/intermediate/DeltaLake/WW/Test/Customer/
- Customer Profitability Sample
https://app.powerbi.com/groups/3257d8ca-4da0-47fa-a172-e01d9b7d428a/reports/db6bc25-1e63-4c6e-9584-548f66d5c040
- Customer Profitability Sample
https://app.powerbi.com/groups/3257d8ca-4da0-47fa-a172-e01d9b7d428a/dashboards/b9a1cd08-3679-4192-91e3-5e6fafa3ecbd

View search results

Recently accessed My it

Name

- labseuwdvlimdioxgn02
- labseuwdvlimdioxgn02
- COPY_ADLS_DATASTORE
- COPY_ADLS_DATASTORE
- Copy Deltalake to ASQL a month ago
- WideWorldImporters a month ago

→ View all recently accessed

references, etc.

starts, how-to guides, and

ew and learn how to make the









Azure Purview Data Catalog

Unify and make data meaningful

Browse assets > Azure SQL Database

Refresh








Showing 1 to 7 of 7 items

Name	Type
 Adventureworks	
 Datastore	
 Datastore	
 datastore	
 oxygenconfigdb	
 test_purview	
 WideWorldImporters	
 WideWorldImportersDW	

WideWorldImporters

- Application
- dbo
- Purchasing
- Sales
- Warehouse

Showing 1 to 7 of 7 items

Name	Type
 PurchaseOrderLines	Azure SQL Table
 PurchaseOrders	Azure SQL Table
 SupplierCategories	Azure SQL Table
 SupplierCategories_Archive	Azure SQL Table
 Suppliers	Azure SQL Table
 Suppliers_Archive	Azure SQL Table
 SupplierTransactions	Azure SQL Table

Azure Data Lake Storage Gen2 1 item

Azure SQL Database 8 items

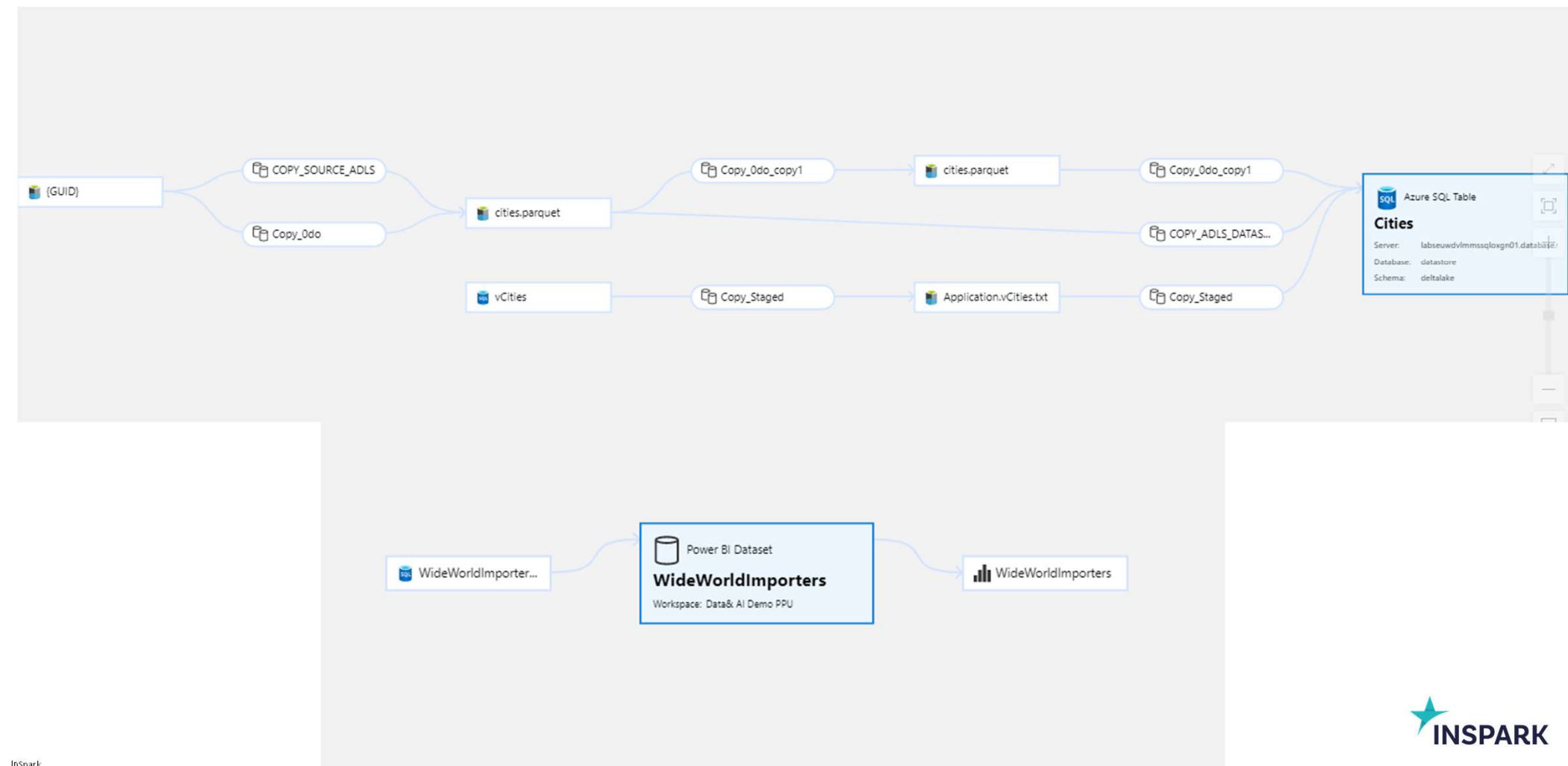
Azure SQL Server 3 items

Azure Storage Account 1 item

Power BI 18 items

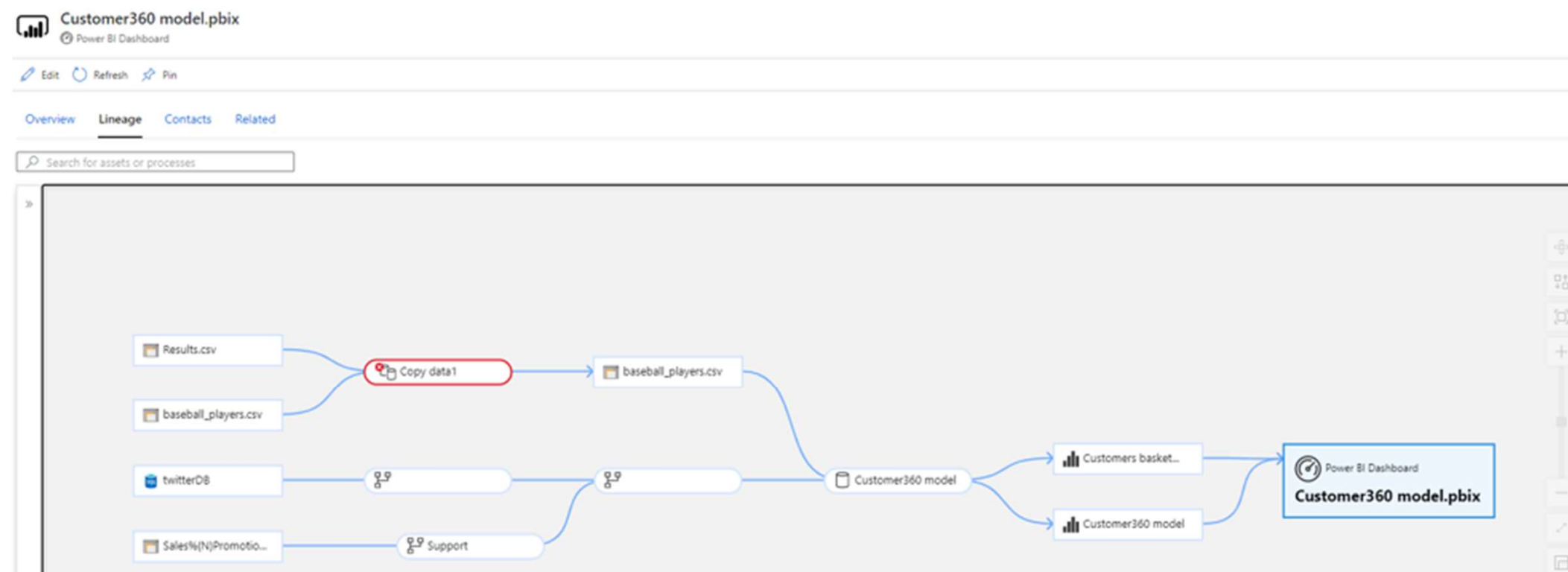
Azure Purview Data Catalog

Unify and make data meaningful



Azure Purview Data Catalog Lineage Future

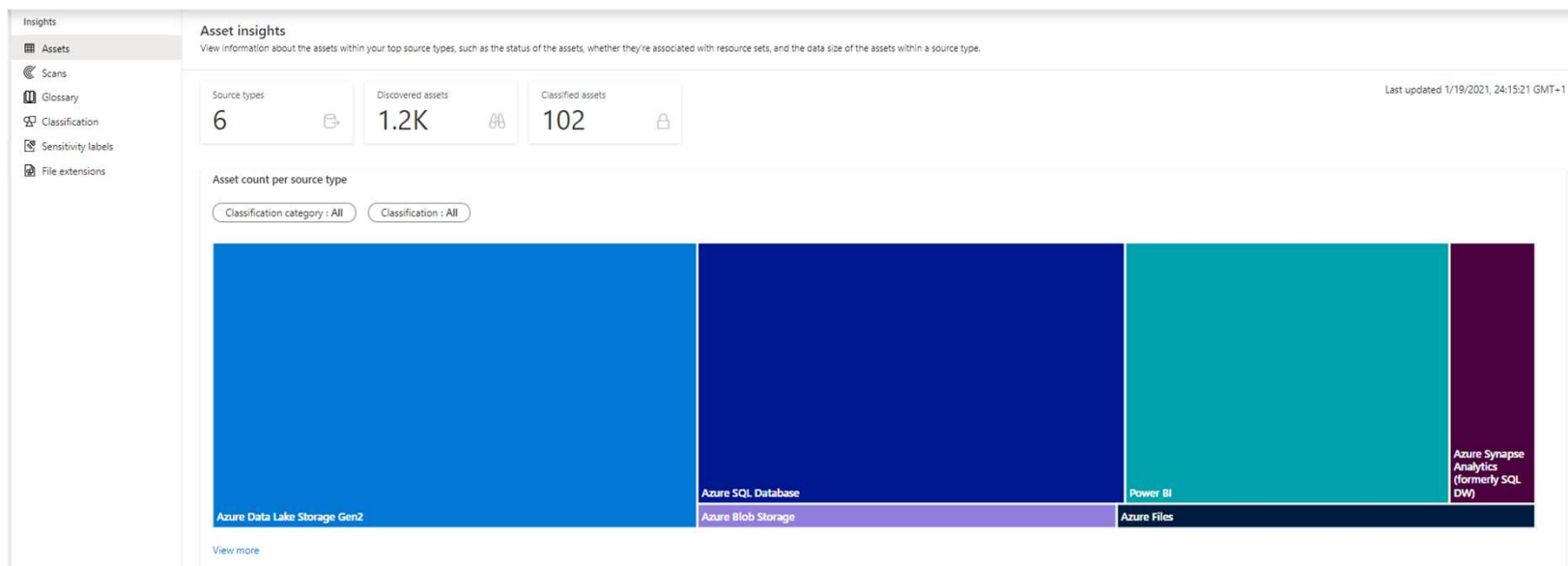
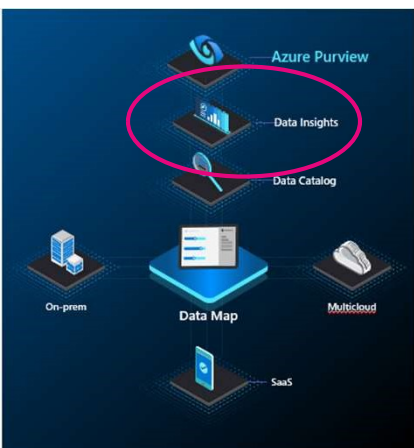
Unify and make data meaningful



Insights

Get a bird's-eye view of sensitive data

- ❑ Reports on Assets, Scans, Glossary, Classification, and Labeling



Integrate Azure Purview in Azure Synapse Analytics

Discover data registered and scanned by Azure Purview

 In Preview

Analytics pools

SQL pools

Apache Spark pools

External connections

Linked services

Azure Purview (Preview)

Integration

Triggers

Integration runtimes

Security

Access control

Credentials

Managed private endpoints

Source control

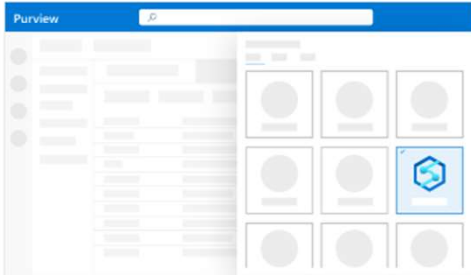
Git configuration

Azure Purview (Preview)

Connecting a Synapse workspace to Azure Purview will enable you to discover trusted and accurate data across your hybrid environment. [Learn more](#)

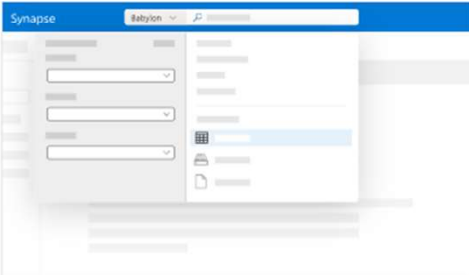
About

Purview account



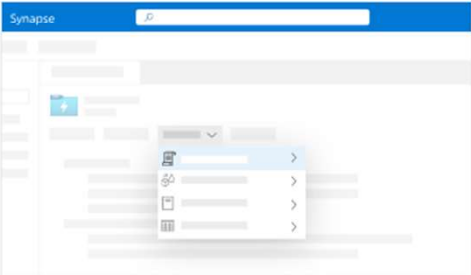
Data registration

Expand your governance and cataloging scope by registering new data. [Learn more](#)



Data discovery

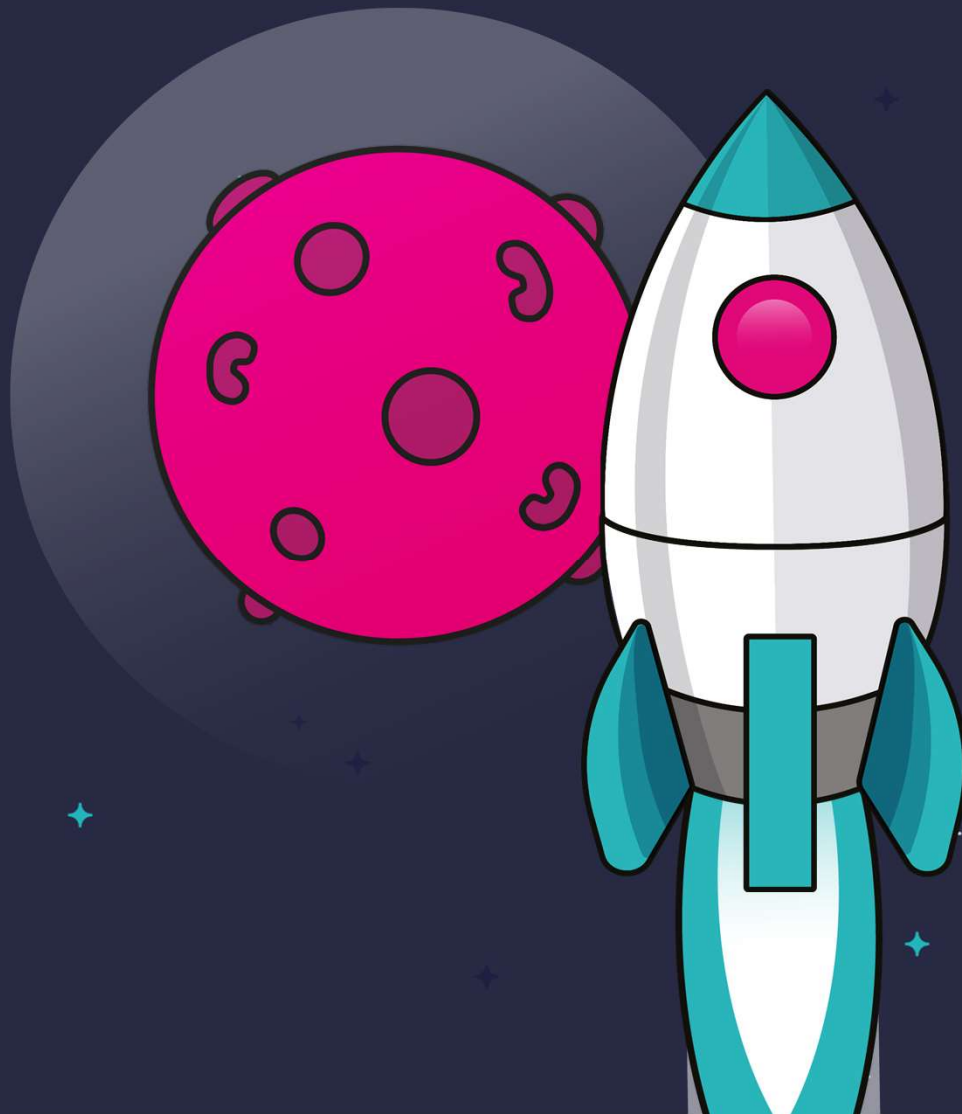
Find relevant data through the global search bar based on keywords and advanced filtering. [Learn more](#)

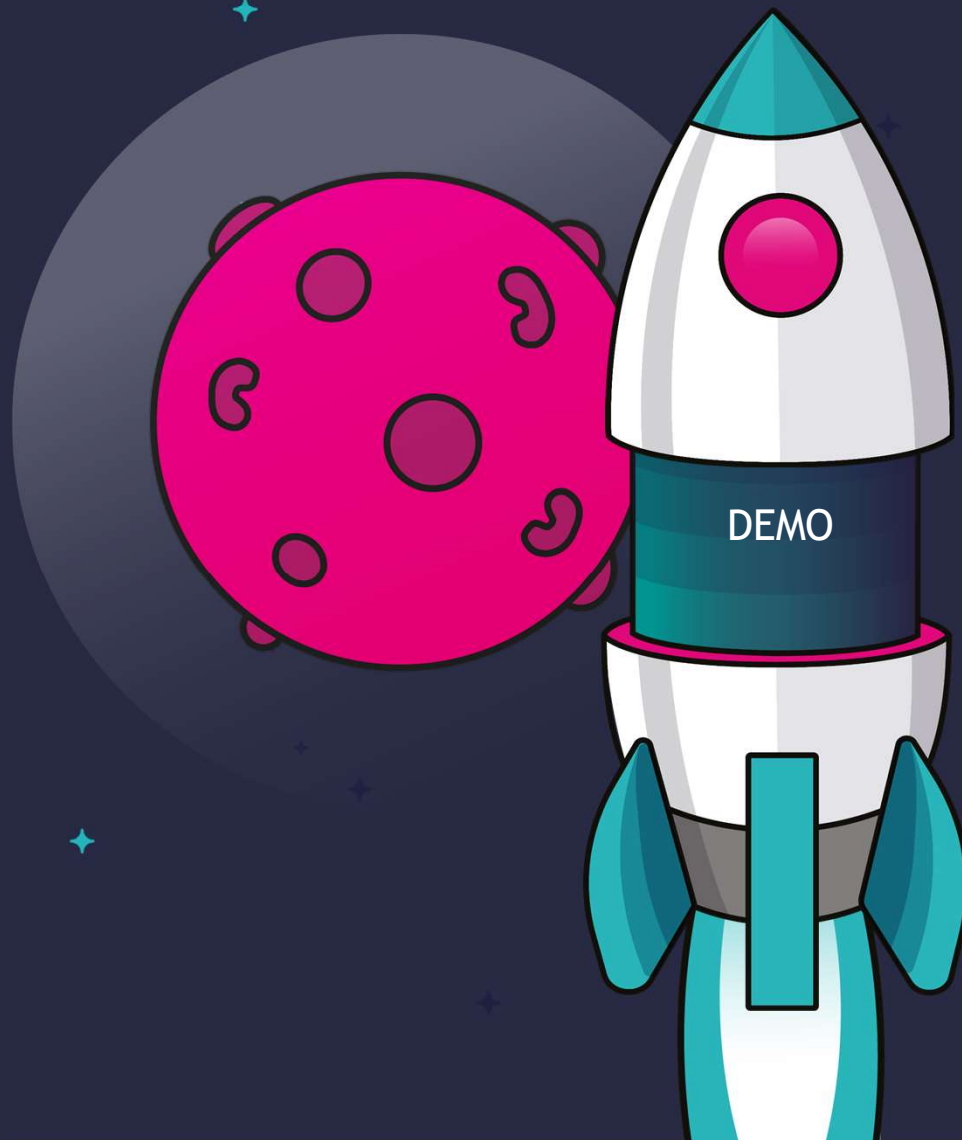


Data exploration

Connect data to your workspace and analyze those data with Synapse. [Learn more](#)

 **INSPIRE**





Innovate to accelerate

DEMO

Azure Purview

■ Features in Public Preview

Data Sources

- Azure Synapse
- Azure DataBricks
- SAP EEC / Hana
- Teradata
- Hive Metastore

Data Lineage

- Notebook support
- Delta Lake Support

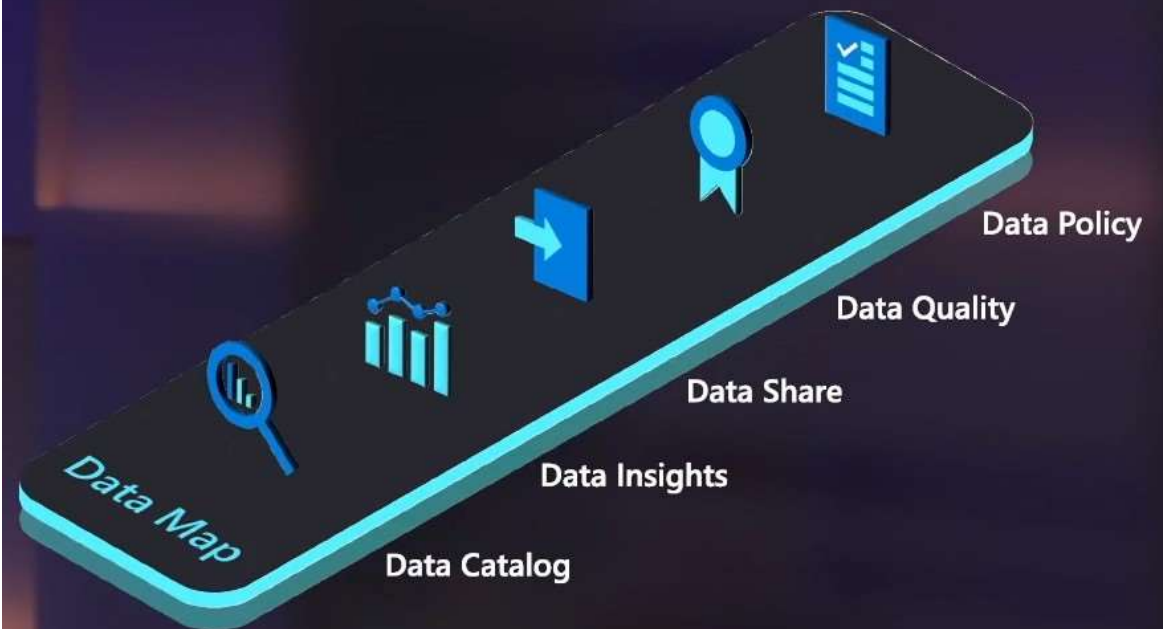
Purview Data Map	Available Now	Coming Soon
Automated scanning of hybrid sources	●	AWS S3
Classification	●	
Microsoft Information Protection Sensitivity Labels support	●	
Apache Atlas API support	●	
Purview Data Catalog		
Semantic Search and Browse	●	
Business Glossary	●	Hierarchical
Data Lineage	●	
Purview in Azure Synapse workspaces		●
Purview data insights		
Assets and Scans Reports	●	
Glossary reports	●	
Classification and Labelling Reports	●	
Asset-level drill down by sensitivity	●	

Roadmap

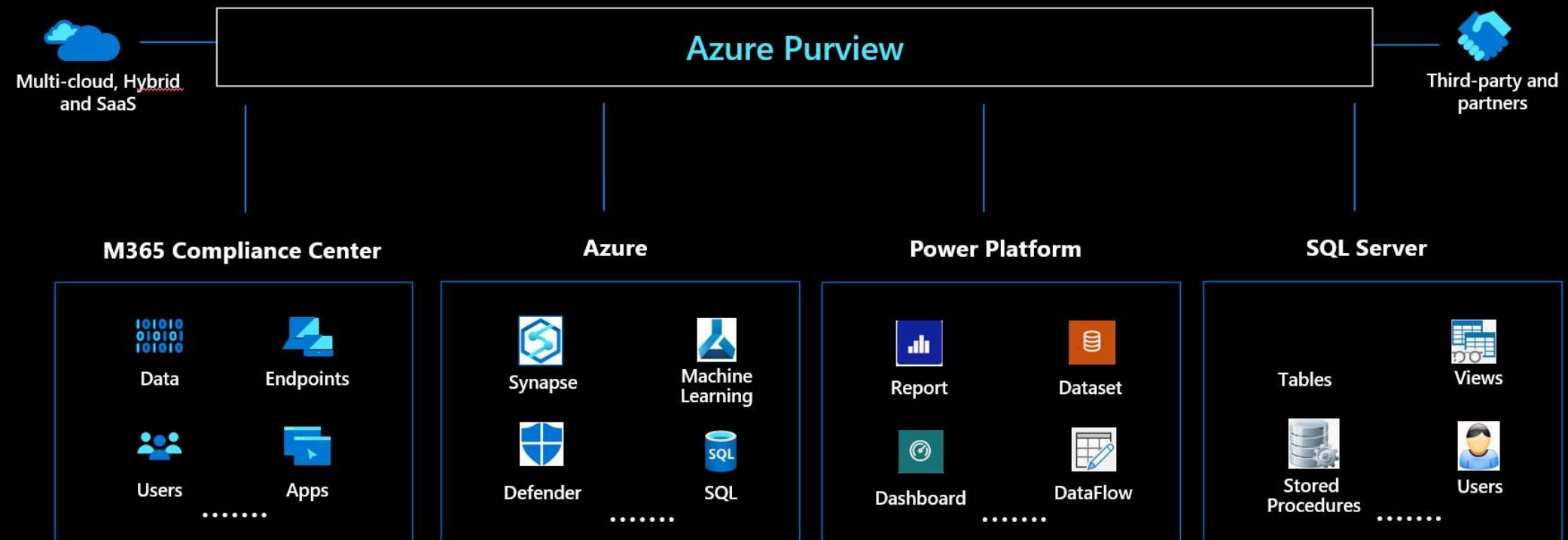
- Data Quality
- Data Privacy
- Master Data Management

Azure Purview roadmap

Unified | Open | Hybrid



Data Governance Opportunity



Questions?



<https://erwindekreuk.com>



@erwindekreuk



Use Slack



<https://www.linkedin.com/in/erwindekreuk/>

Slides will be available on my blog

