

Azure Purview



Microsoft's answer to Data Governance and Data Lineage



datasaturdays.com #0001
pordenone, 27 febbraio 2021



Erwin de Kreuk

Lead Data and AI



@erwindekreuk



<https://erwindekreuk.com>



#datasatpn



Room 6
13:30 CET

Azure Purview



Microsoft's answer to Data Governance and Data Lineage



datasaturdays.com #0001
pordenone, 27 febbraio 2021



Erwin de Kreuk

Lead Data and AI



@erwindekreuk



<https://erwindekreuk.com>



#datasatpn





Azure Purview

Microsoft's answer to Data Governance and Data Lineage



Erwin de Kreuk



@erwindekreuk

Lead Data and AI



<https://erwindekreuk.com>

We Are InSpark

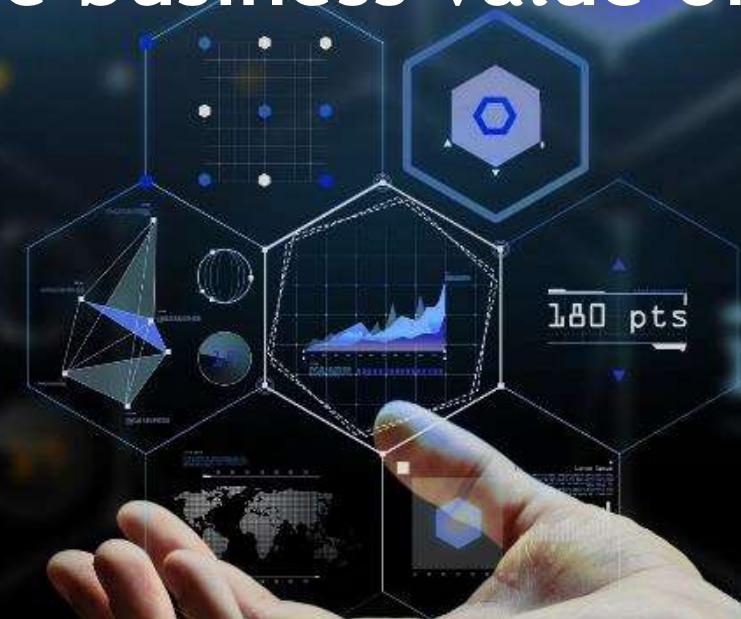
We help organizations
accelerating their digital
transformation with impactful
Microsoft solutions & expertise



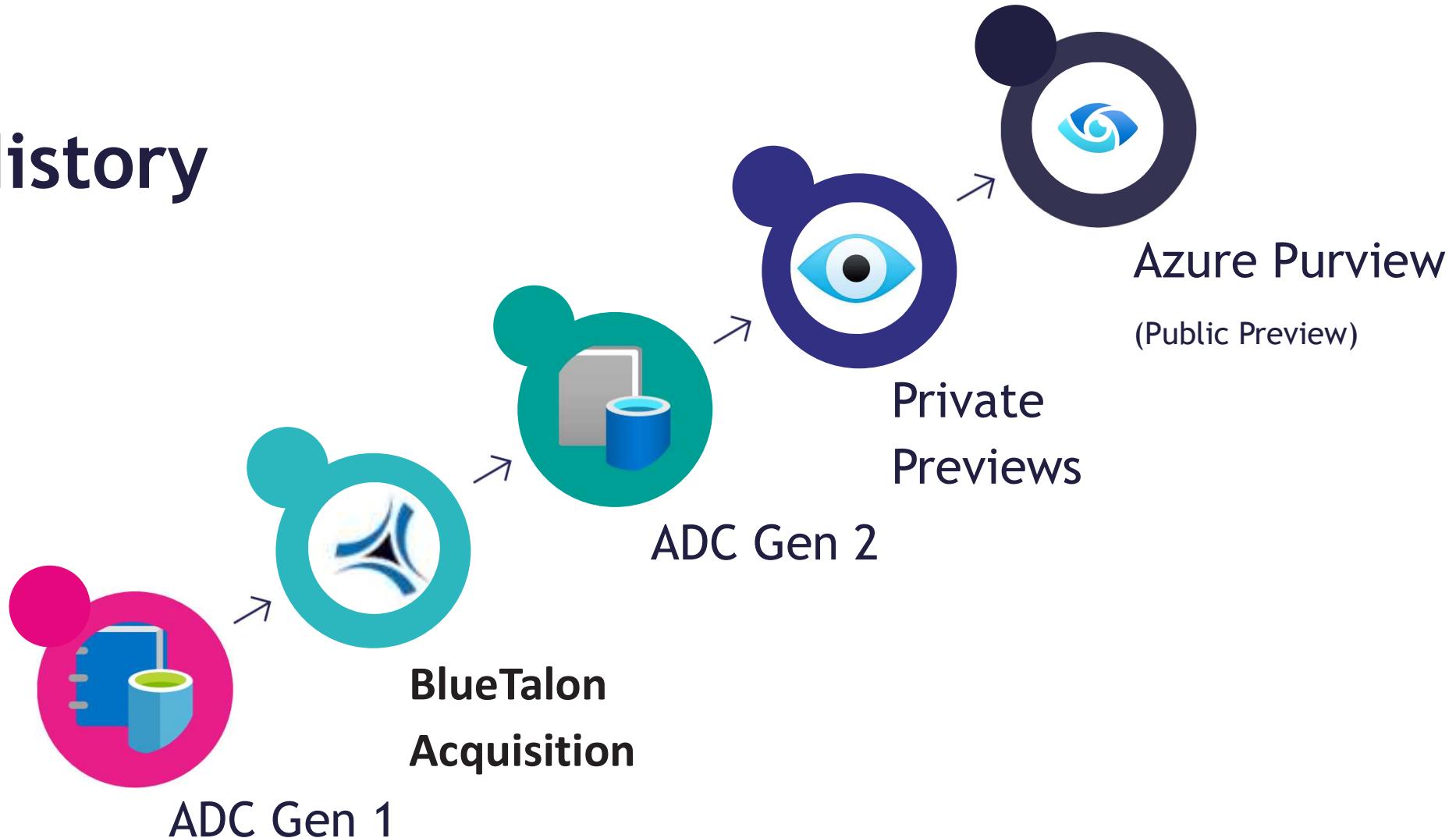
Azure Purview



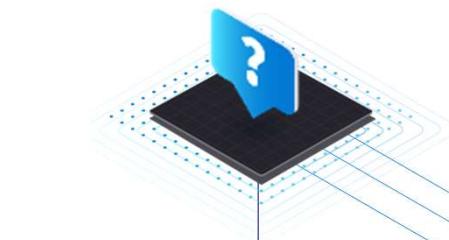
Unified **data governance** to
maximize the business value of data



History

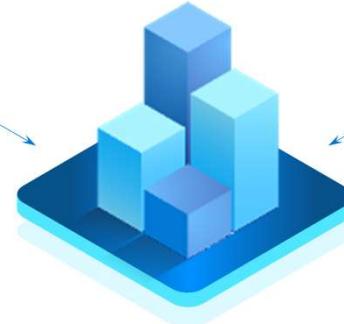


Data governance is becoming increasingly interdisciplinary



DISCOVERY

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

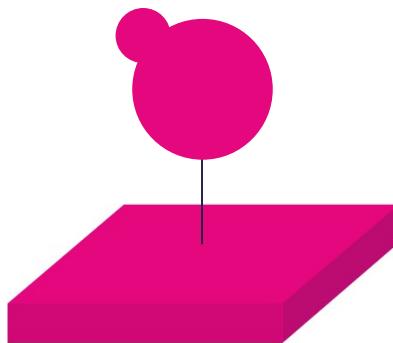


COMPLIANCE

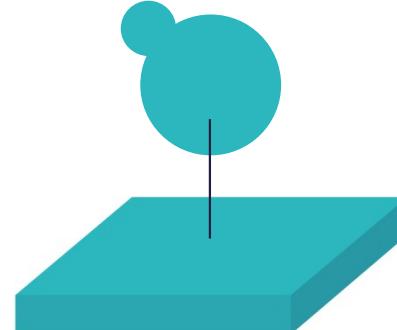
What's my exposure to risk?
Is my usage compliant?
How do I control access & use?
What is required by regulation X?

Chief Data Officer

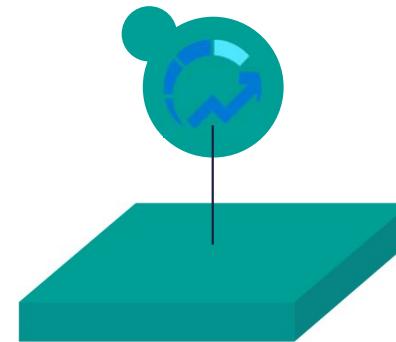
Elements of successful data governance



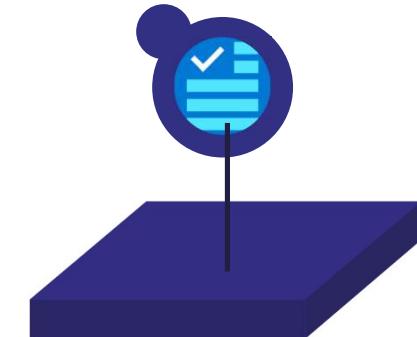
Manage growing
data landscape



Overcome
operational silos



Increase data
agility



Comply with
industry
regulations

Data Insights

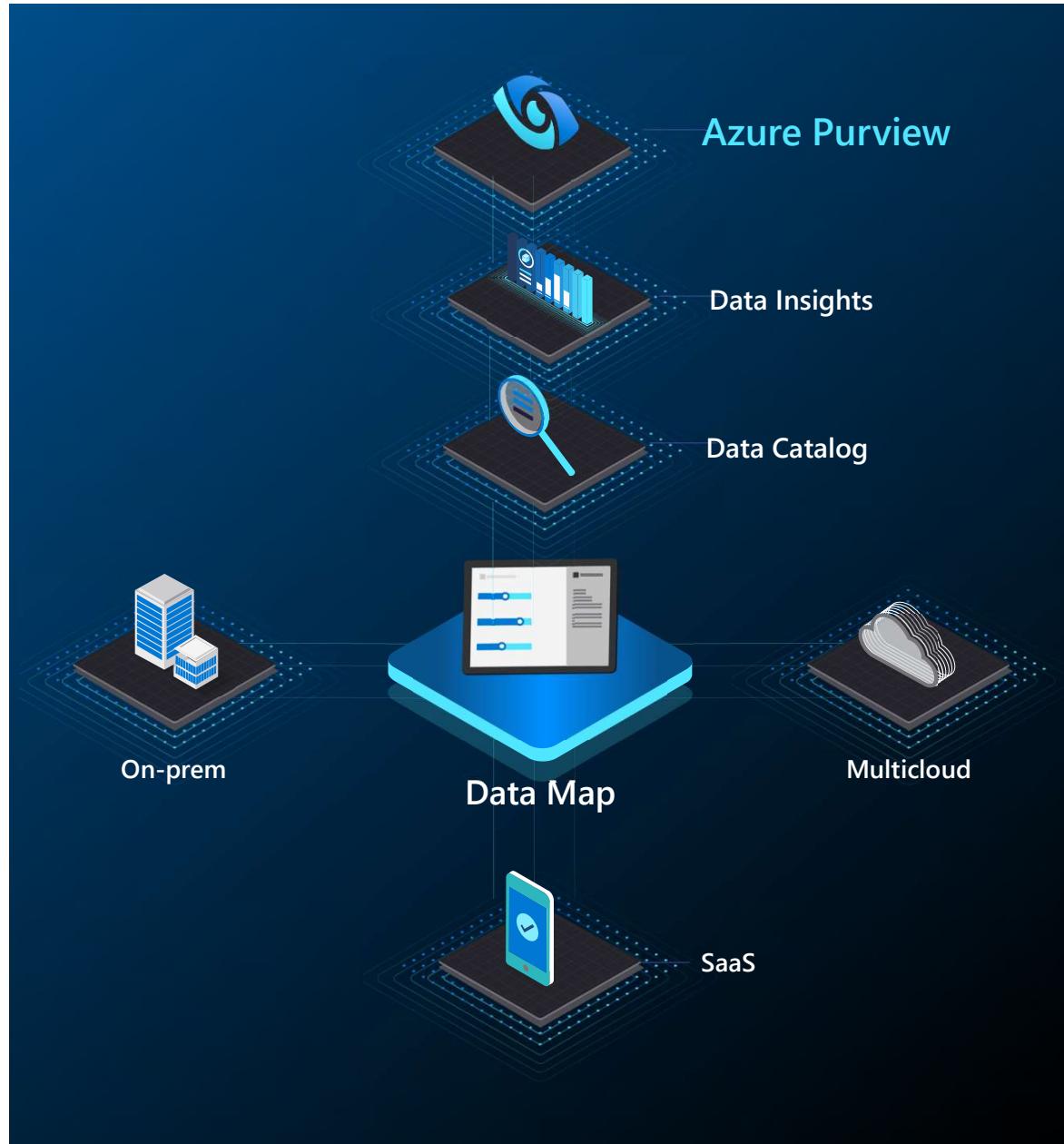
- Assess data usage across your organization

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

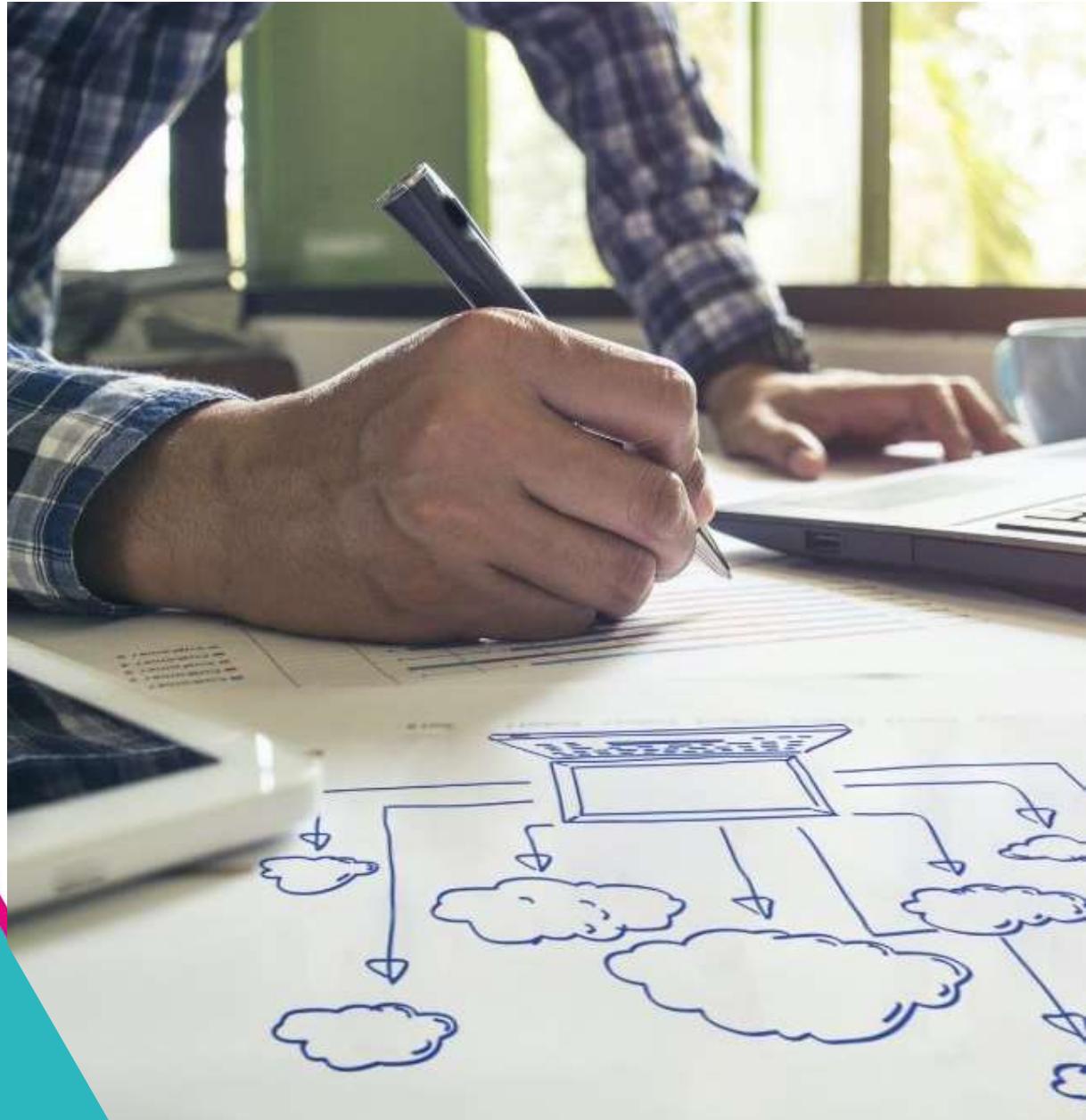
- ❑ Manage and govern operational, transactional and analytical data
- ❑ Cloud-native, purpose-built service to address discovery and compliance needs
- ❑ Fully managed, serverless, PaaS service
- ❑ Eliminate manual, ad-hoc and homegrown solutions





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



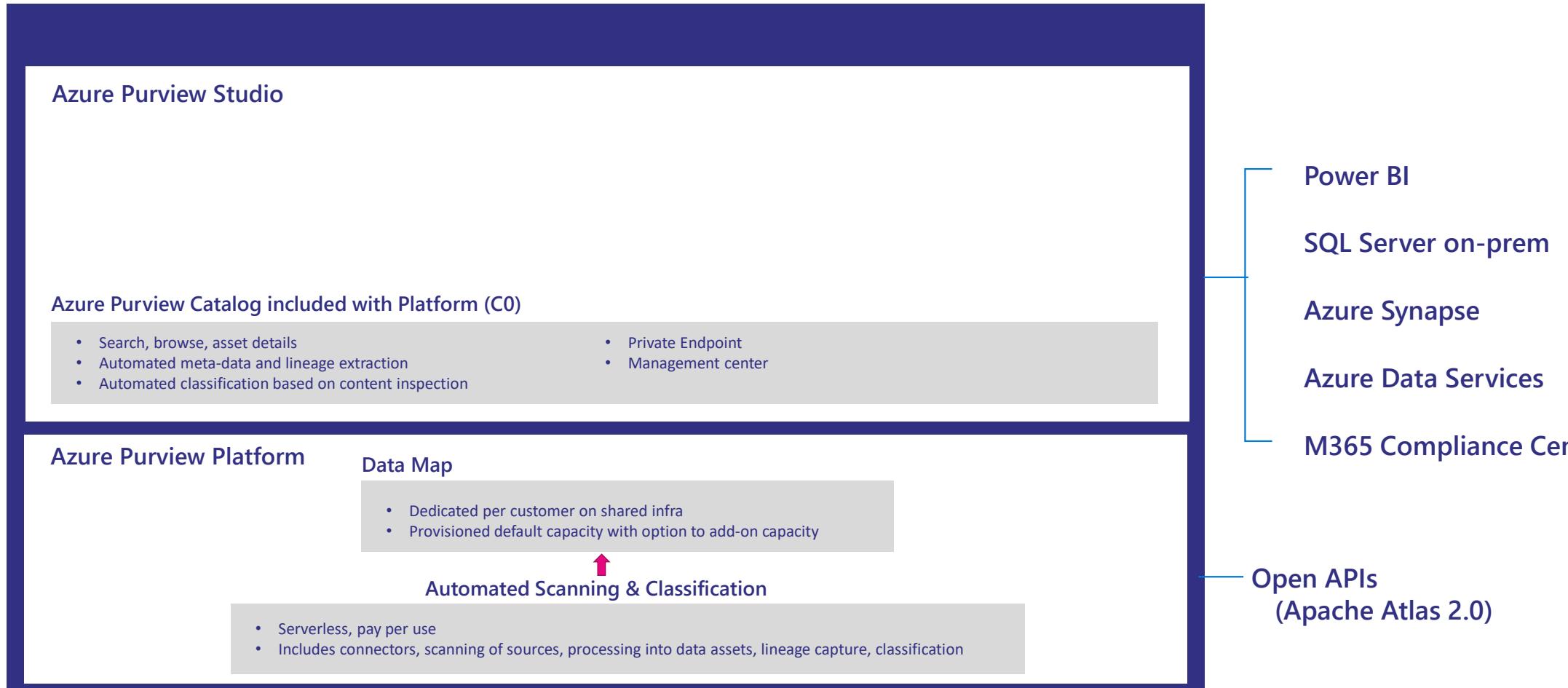


Gain insight into **data use**
across the estate

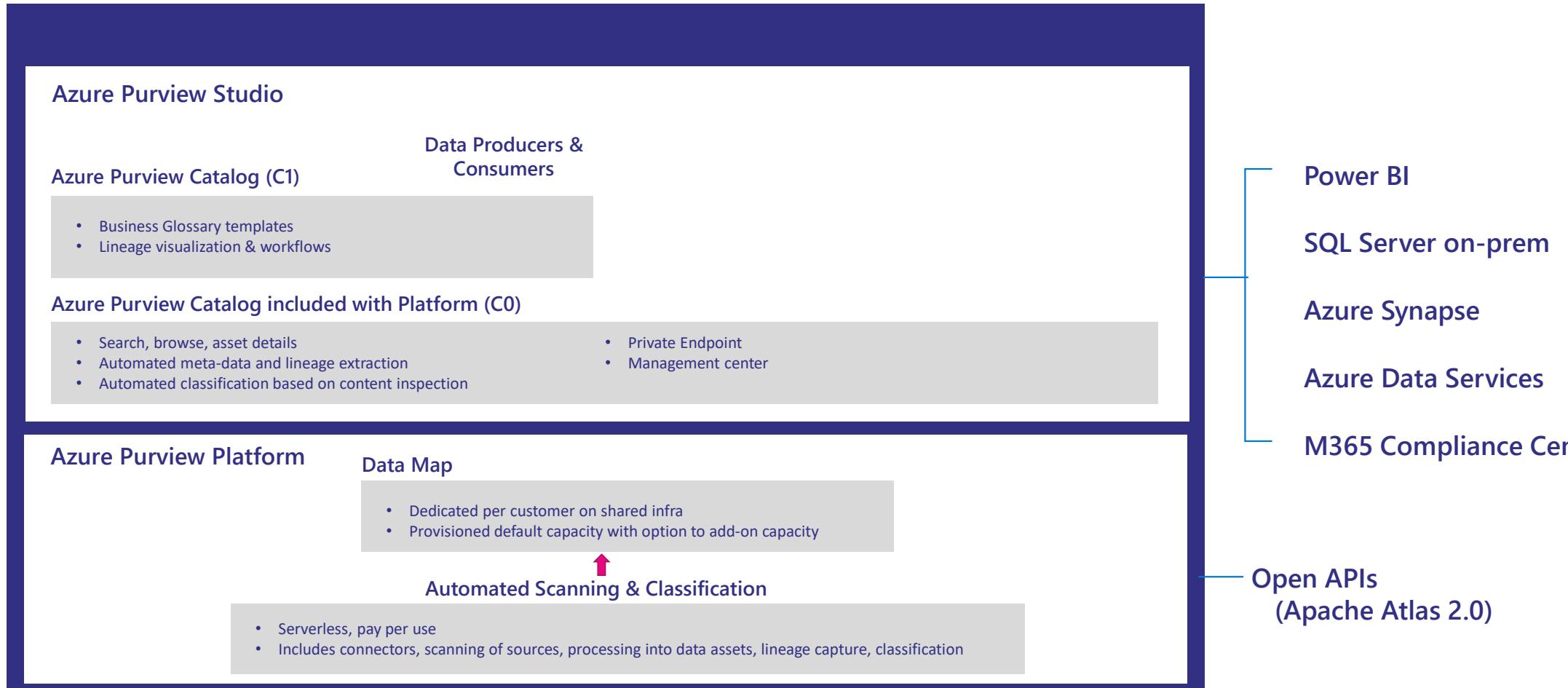
- Understand at a glance how data is being created and used across your data estate
- Visually assess the state of data assets, scans, business glossary and sensitive data



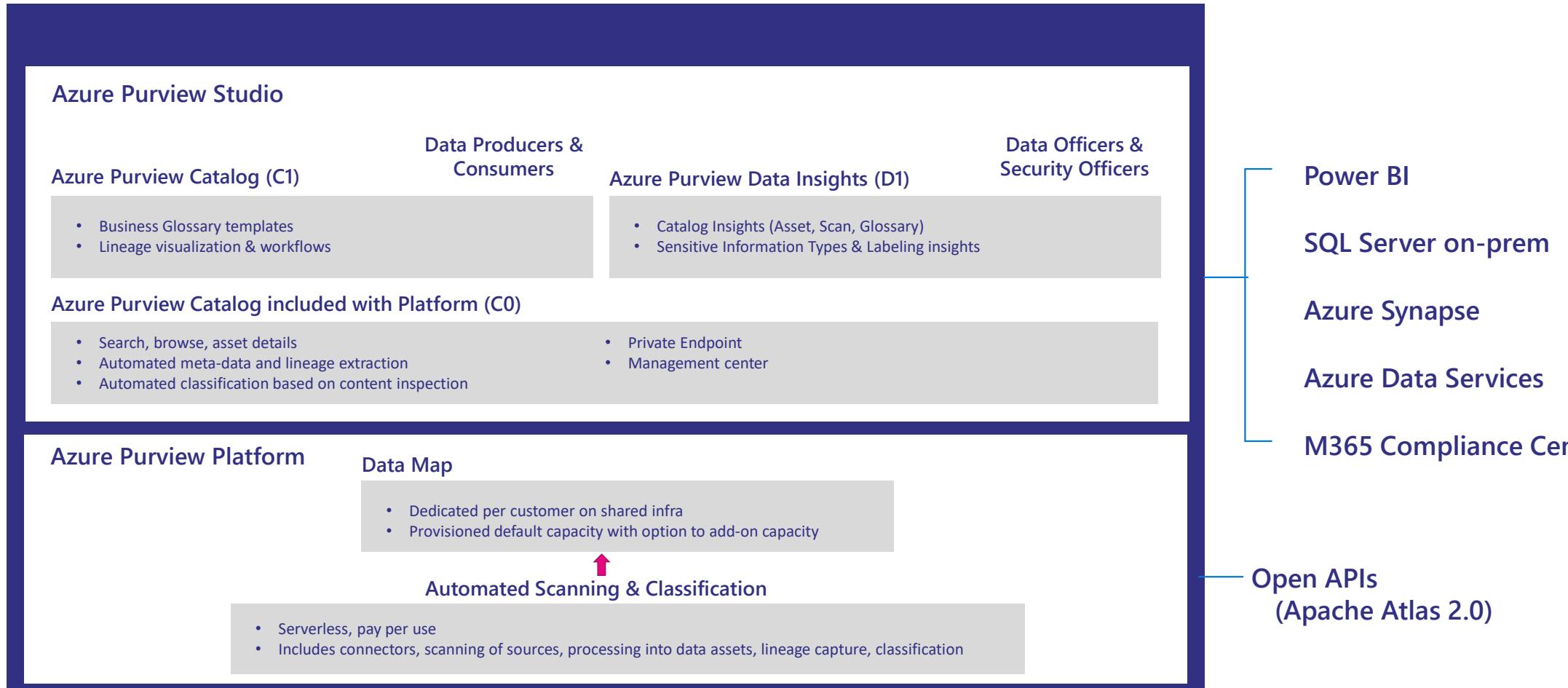
Azure Purview Features



Azure Purview Features



Azure Purview Features



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

General

Account information

Metrics

Metadata management

Classifications

Classification rules

External connections

Data Factory

Data Share

Security and access

Access control

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 labseuwvdvImpurviewoxgn01

Friendly name * Azure Purview Demo

Azure Purview Demo

1,138 assets 1 glossary term

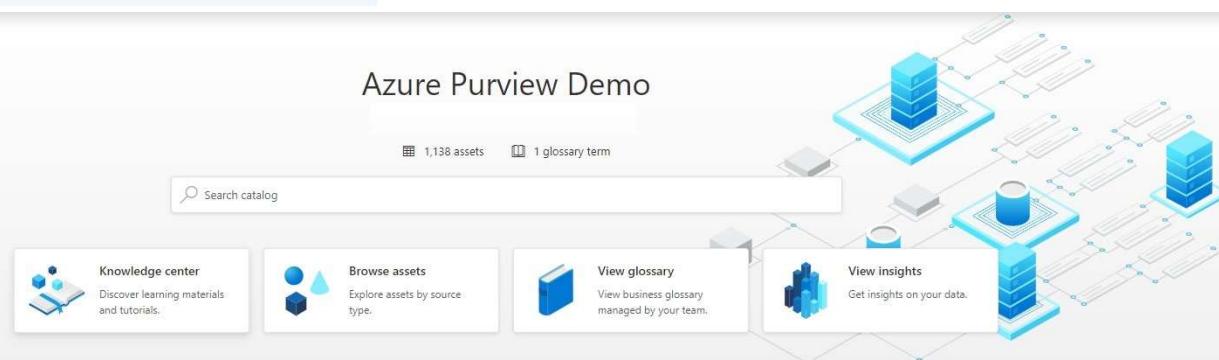
Search catalog

Knowledge center Discover learning materials and tutorials.

Browse assets Explore assets by source type.

View glossary View business glossary managed by your team.

View insights Get insights on your data.



Azure Purview - Roles

Data Source Administrator

- 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**

Data Reader

Data Curator

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

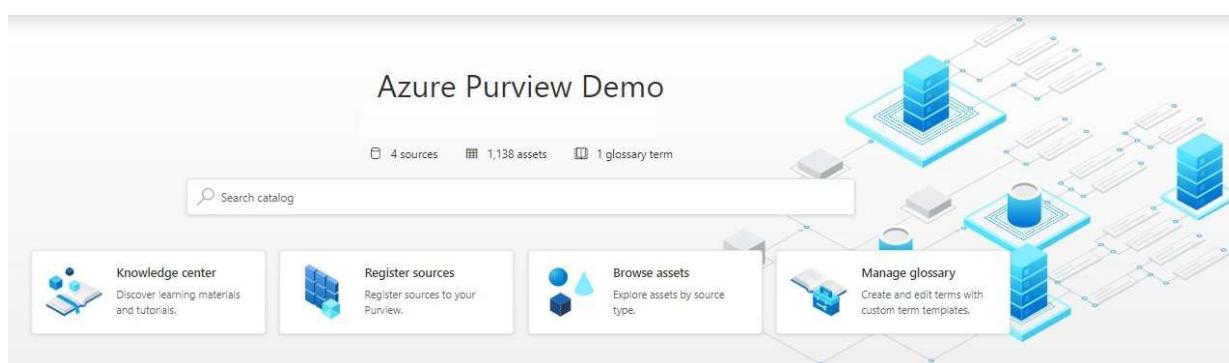
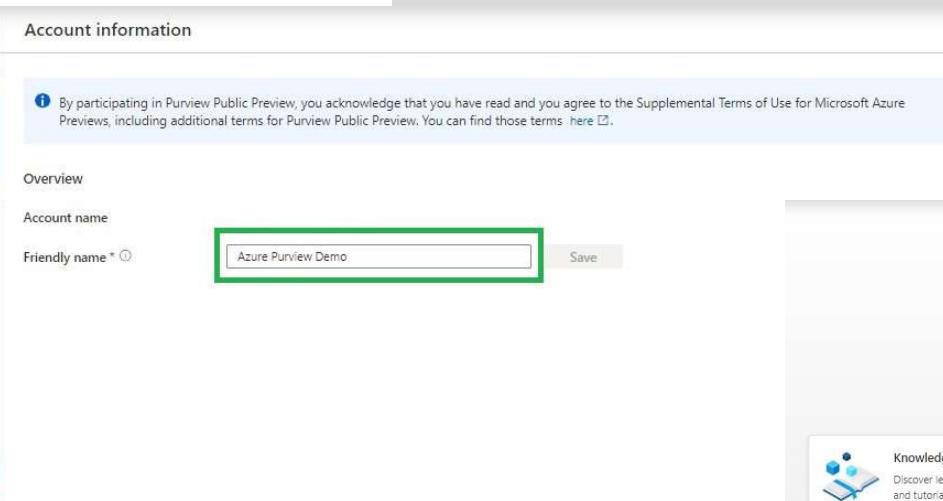
Account information

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 * Save



Error
Request failed with status code 403
Not authorized to access account

Azure Purview - Roles

Data Source Administrator

Data Reader

Data Curator

Quick action buttons on
Homepage

All Features (C0 + C1 + D1)

C0

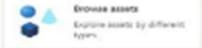
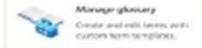
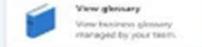
(Sources registration, automated scanning and classification, data discovery.)

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	 Discover learning materials and tutorials.		 Explore assets by different types.	 Create and edit terms with custom term templates.	 Get insights on your data.
 Data reader Read	 Discover learning materials and tutorials.		 Explore assets by different types.	 View business glossary managed by your team. <small>READ-ONLY</small>	 Get insights on your data.
 Data source admin + Data curator	 Discover learning materials and tutorials.	 Register sources to your Purview.	 Explore assets by different types.	 Create and edit terms with custom term templates.	
 Data source admin + Data reader	 Discover learning materials and tutorials.	 Register sources to your Purview.	 Explore assets by different types.	 View business glossary managed by your team. <small>READ-ONLY</small>	

Azure Purview - Pricing

Azure Purview Data Map

- Capacity Unit
 - €0.289 per 1 Capacity Unit Hour
 - Provisioned API throughput. 1 capacity unit = 1 API/sec
 - Includes 4 capacity units for free until February 28, 2021.
- Metadata Storage
 - Free in preview

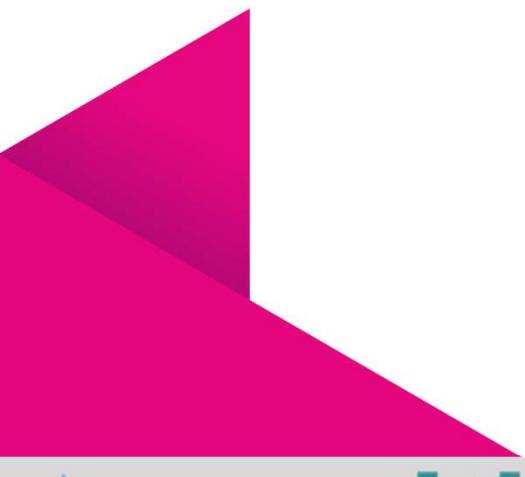


Azure Purview - Pricing

Azure Purview Data Map

- Power BI Online
 - Free in Preview
 - SQL Server On Prem
 - Free in Preview
 - Other Data Sources
 - Free in Preview
 - €0.532 per 1 vCore Hour
- Includes 16 vCore-hours for Free every month until February 28, 2021

Scanning and Classification



Azure Purview - Pricing

Azure Purview Data Map

Scanning and Classification

Azure Purview Data Catalog

- C0
 - Included with the Data Map
Search and browse of data assets
- C1
 - Free in preview
 - Business glossary, lineage visualization and catalog insights
- D0
 - Free in preview
Sensitive data identification insights

<https://azure.microsoft.com/en-us/pricing/details/azure-purview>

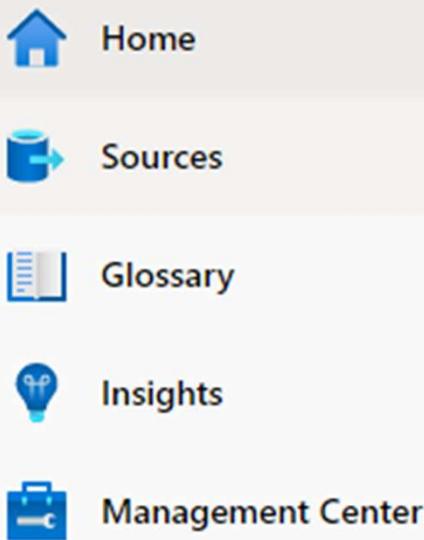
Azure Purview Studio

The screenshot shows the Azure Purview Demo interface. Several UI elements are highlighted with colored arrows:

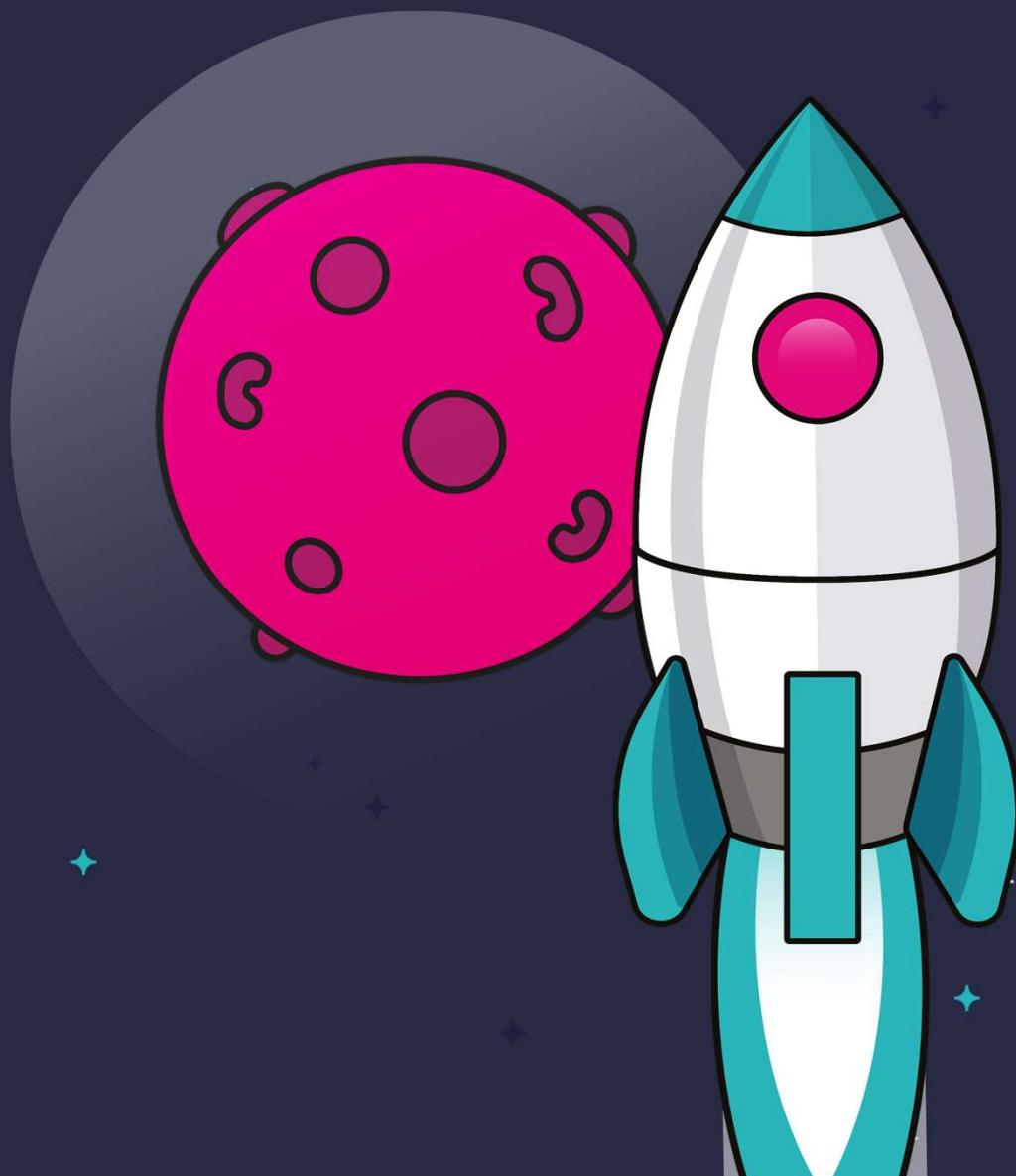
- Metrics**: A teal box pointing to the Metrics section in the top navigation bar.
- Search Bar**: A teal box pointing to the Search catalog input field.
- Recently Accessed Entities**: A teal box pointing to the Recently accessed entities section.
- Updates**: A pink box pointing to the Updates button in the top navigation bar.
- Accounts**: A pink box pointing to the Accounts button in the top navigation bar.
- Notifications**: A pink box pointing to the Notifications button in the top navigation bar.
- Feedback**: A pink box pointing to the Feedback icon in the top right corner.
- Key Activities**: A teal box pointing to the Manage glossary activity.
- Usefull Links**: A teal box pointing to the Purview overview link in the Links section.

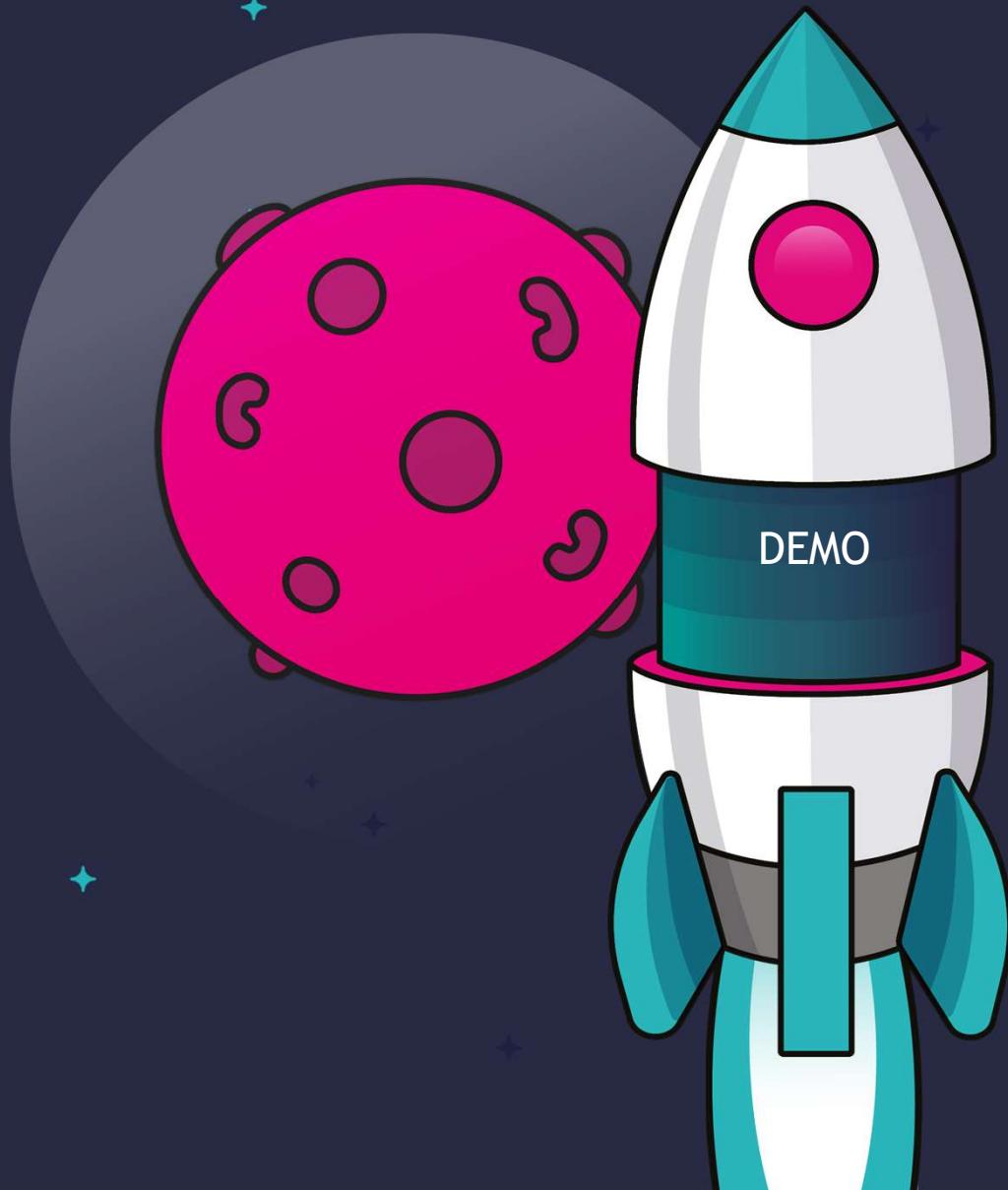
The interface includes a search bar, a dashboard with four main activities (Knowledge center, Register sources, Browse assets, Manage glossary), a recently accessed entities list, and a links section with a Purview overview link.

Azure Purview Studio - Activity hubs



- 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





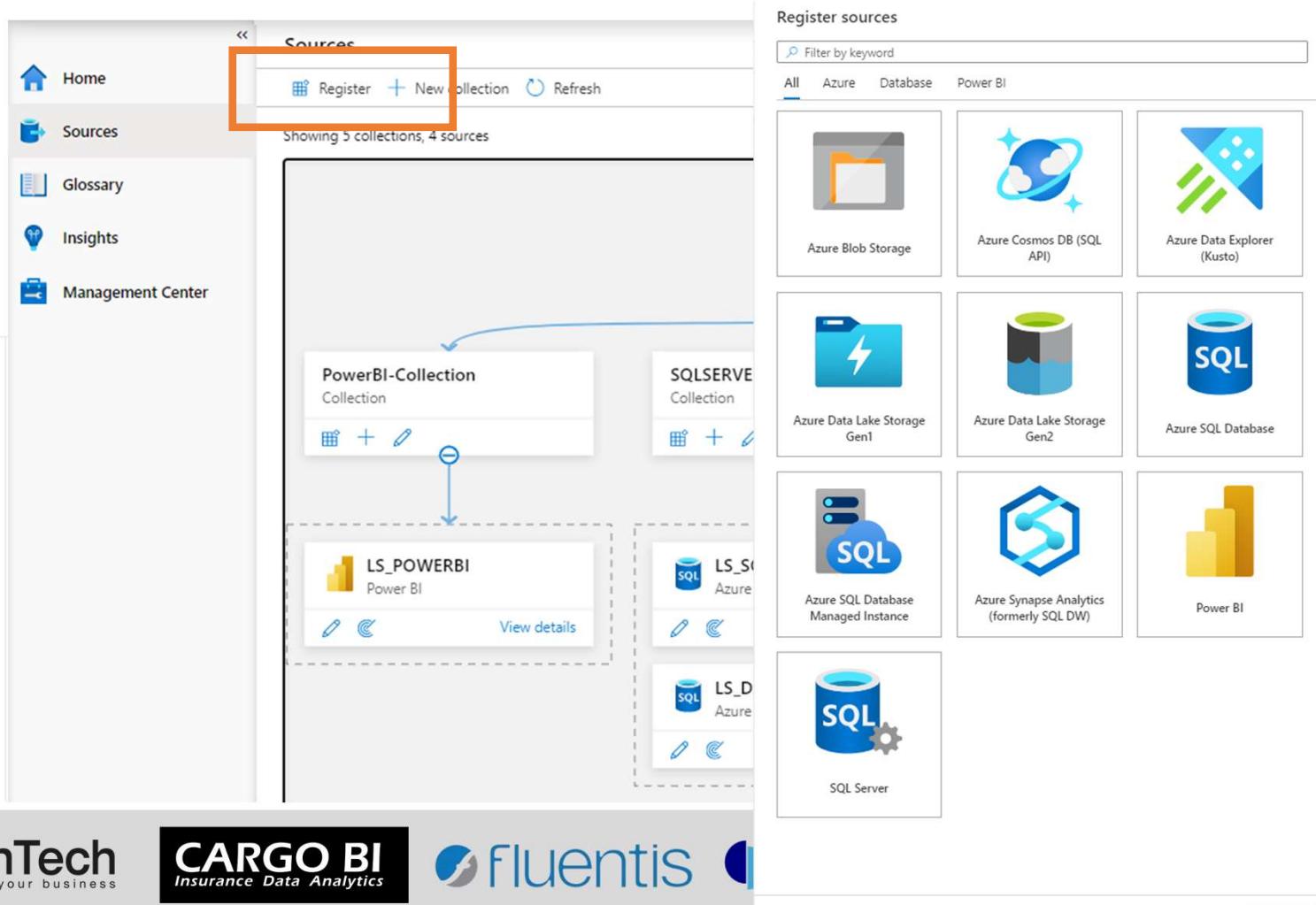
Innovate to accelerate

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

The screenshot shows the Purview Data Map interface with the following details:

- Top Navigation Bar:** Home, Sources (highlighted), Glossary, Insights, Management Center.
- Header:** Sources, Register, New collection, Refresh.
- Main Area:** Showing 5 collections, 4 sources. A "PowerBI-Collection" is highlighted with a dashed box. It contains a "LS_POWERBI" Power BI source, which has a "View details" button.
- Right Panel:** Register sources section with a search bar and filter for All, Azure, Database, Power BI.
- Source Options:** Grid of icons for registering various sources:
 - Azure Blob Storage
 - Azure Cosmos DB (SQL API)
 - Azure Data Explorer (Kusto)
 - Azure Data Lake Storage Gen1
 - Azure Data Lake Storage Gen2
 - Azure SQL Database
 - Azure SQL Database Managed Instance
 - Azure Synapse Analytics (formerly SQL DW)
 - Power BI
 - SQL Server

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				
Completed	Scheduled				
Completed	Scheduled				
Completed	Manual				
Completed	Manual				

Scope your scan

Refresh

All future assets under a certain parent will be automatically selected if the parent is fully or partially checked.

Search

- LS_ADLS
 - > azure-webjobs-hosts
 - > azure-webjobs-secrets
 - > intermediate
 - > marts
 - > raw
 - > sandbox
 - > temp
 - > tempstaging
 - > transfer



beanTech
IT moves your business

CARGO BI
Insurance Data Analytics

Fluentis

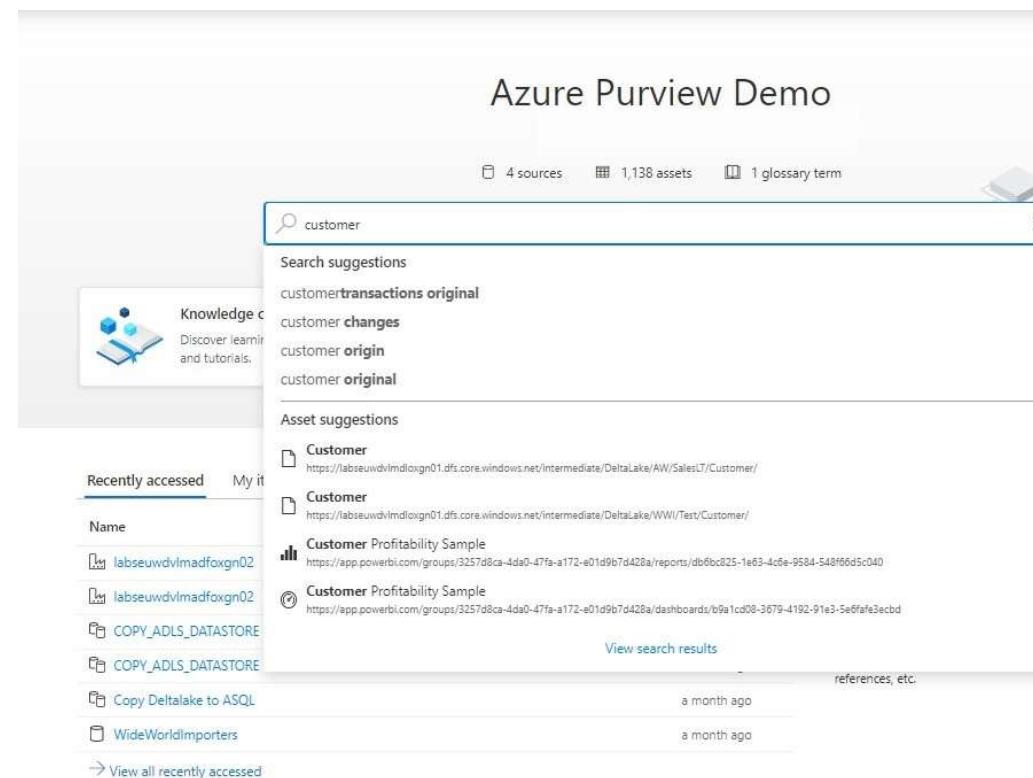
QUANTUMDATIS

Real Comm
Easy for real

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

Search suggestions:

- customer transactions original
- customer changes
- customer origin
- customer original

Asset suggestions:

- Customer <https://labseuwdvlmdfoxgn01.dfs.core.windows.net/intermediate/DeltaLake/AW/SalesLT/Customer/>
- Customer <https://labseuwdvlmdfoxgn01.dfs.core.windows.net/intermediate/DeltaLake/WWI/Test/Customer/>
- Customer Profitability Sample <https://app.powerbi.com/groups/3257d8ca-4da0-47fa-a172-e01d9b7d428a/reports/db6bc825-1e63-4c6a-9584-548f66d5c040>
- Customer Profitability Sample <https://app.powerbi.com/groups/3257d8ca-4da0-47fa-a172-e01d9b7d428a/dashboards/b9a1cd08-3679-4192-91e3-5e6faf3ecbd>

[View search results](#)



Manage glossary
Create and edit terms with custom term templates.

Learn how to make the

charts, how-to guides, and



Azure Purview Data Catalog

Unify and make data meaningful

Browse assets > Azure SQL Database

Refresh

Browse assets > Azure SQL Database > WideWorldImporters

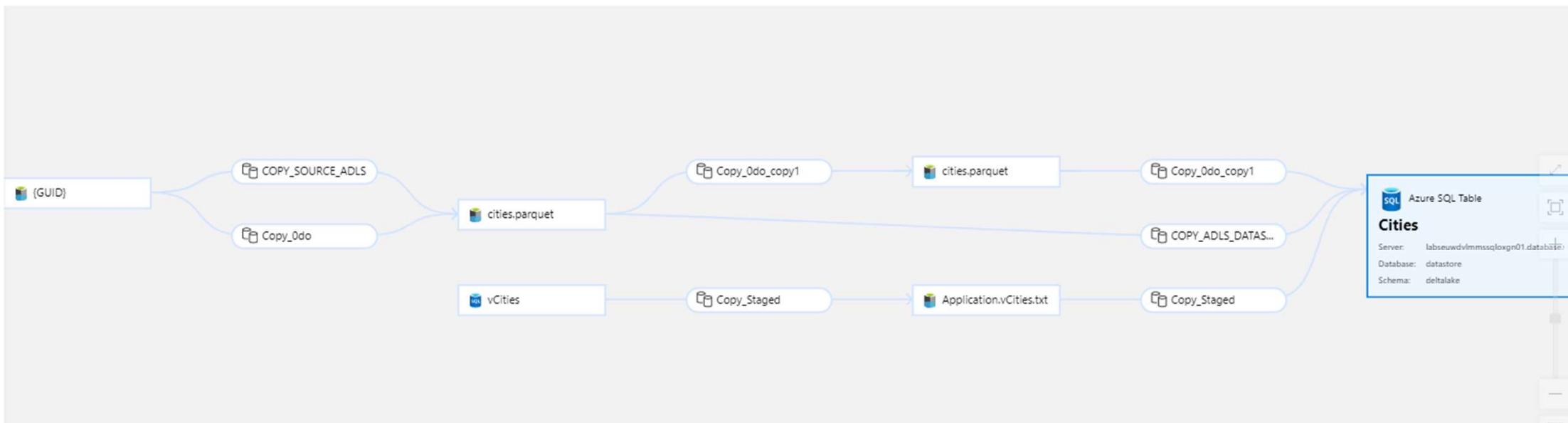
Name	Type
Adventureworks	Azure SQL Table
Datastore	Azure SQL Table
Datastore	Azure SQL Table
datastore	Azure SQL Table
oxygenconfigdb	Azure SQL Table
test_purview	Azure SQL Table
WideWorldImporters	Azure SQL Table
WideWorldImportersDW	Azure SQL Table
Azure Data Lake Storage Gen2	1 item
Azure SQL Database	8 items
Azure SQL Server	3 items
Azure Storage Account	1 item
Power BI	18 items

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

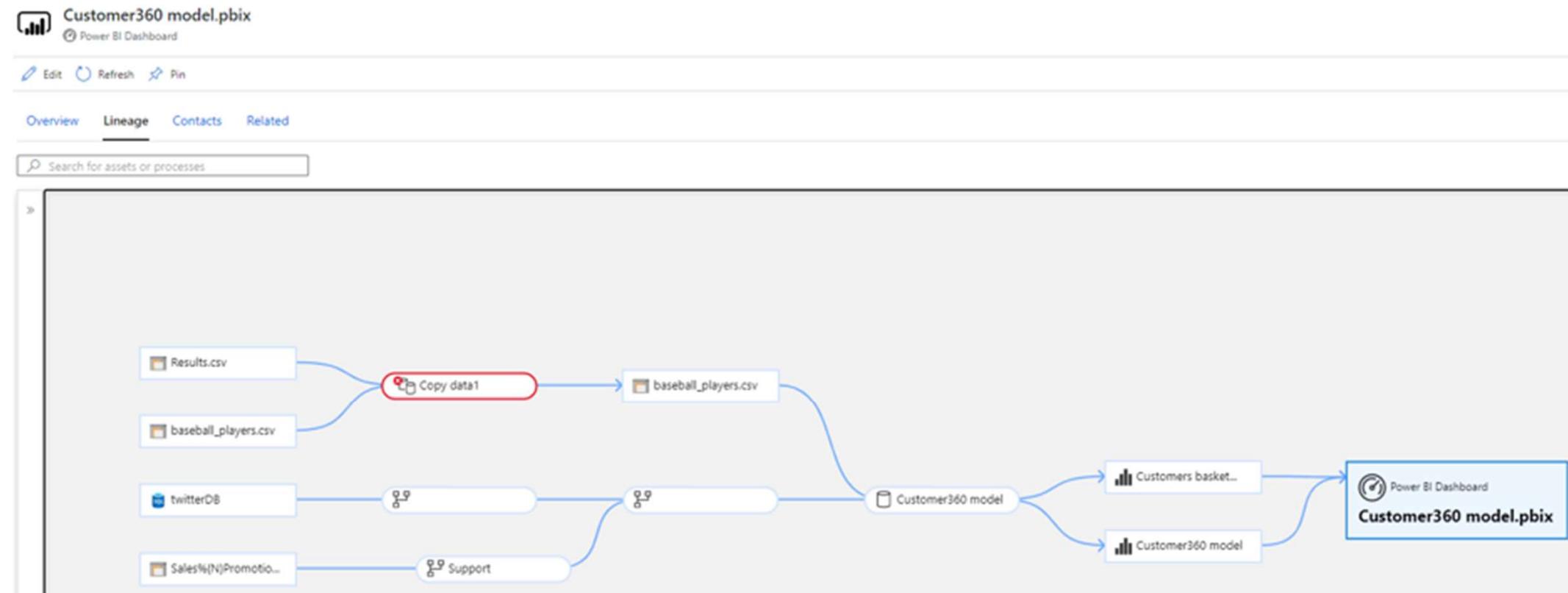
Azure Purview Data Catalog

Unify and make data meaningful



Azure Purview Data Catalog

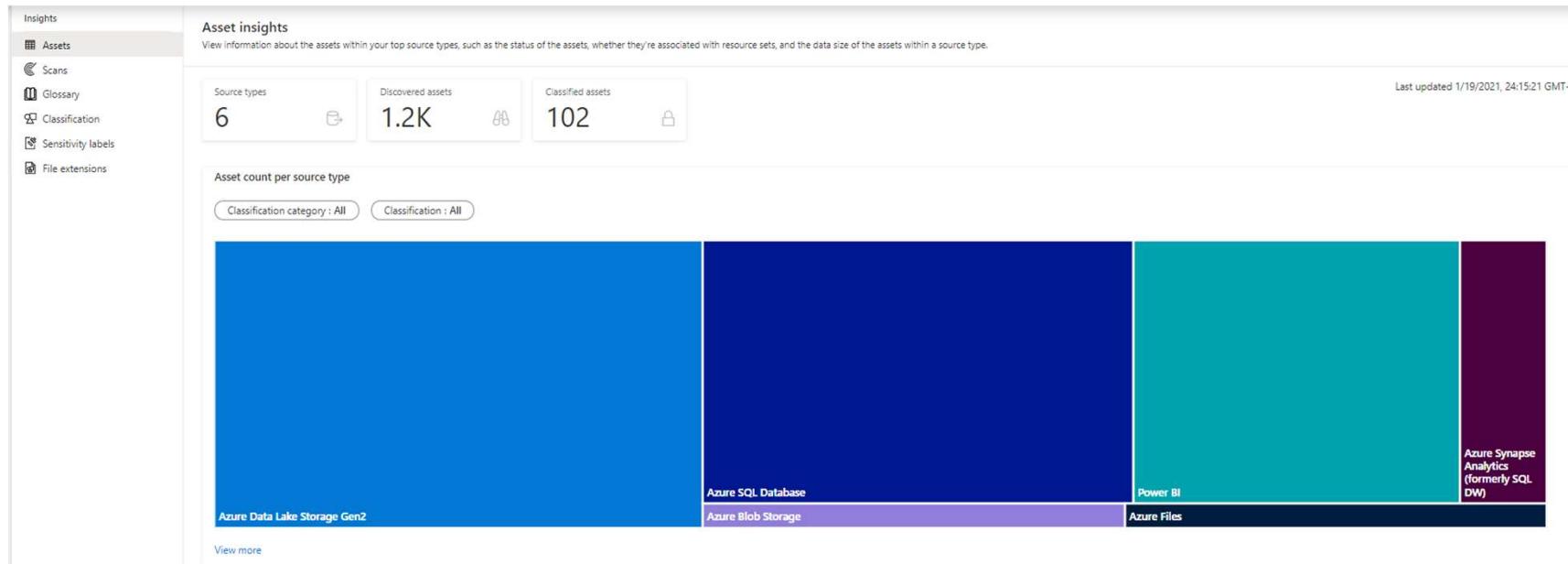
Unify and make data meaningful



Insights

Get a bird's-eye view of sensitive data

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



beanTech
IT moves your business

CARGO BI
Insurance Data Analytics

Fluentis

QUANTUMDATIS

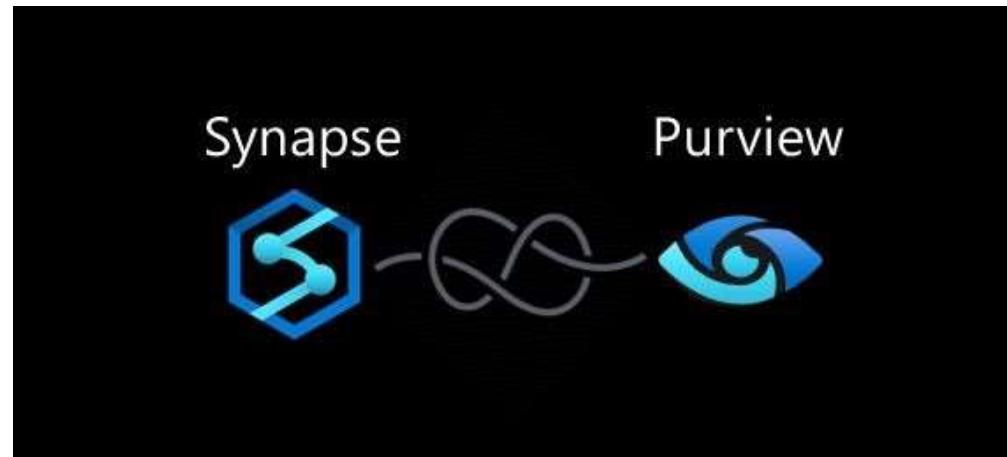
Real Comm
Easy for real

Integrate Azure Purview in Azure Synapse Analytics

hn0va
community

Discover data registered and scanned by Azure Purview

In Preview



Integrate Azure Purview in Azure Synapse Analytics

hn0va
community

Discover data registered and scanned by Azure Purview

In Preview



Integrate Azure Purview in Azure Synapse Analytics

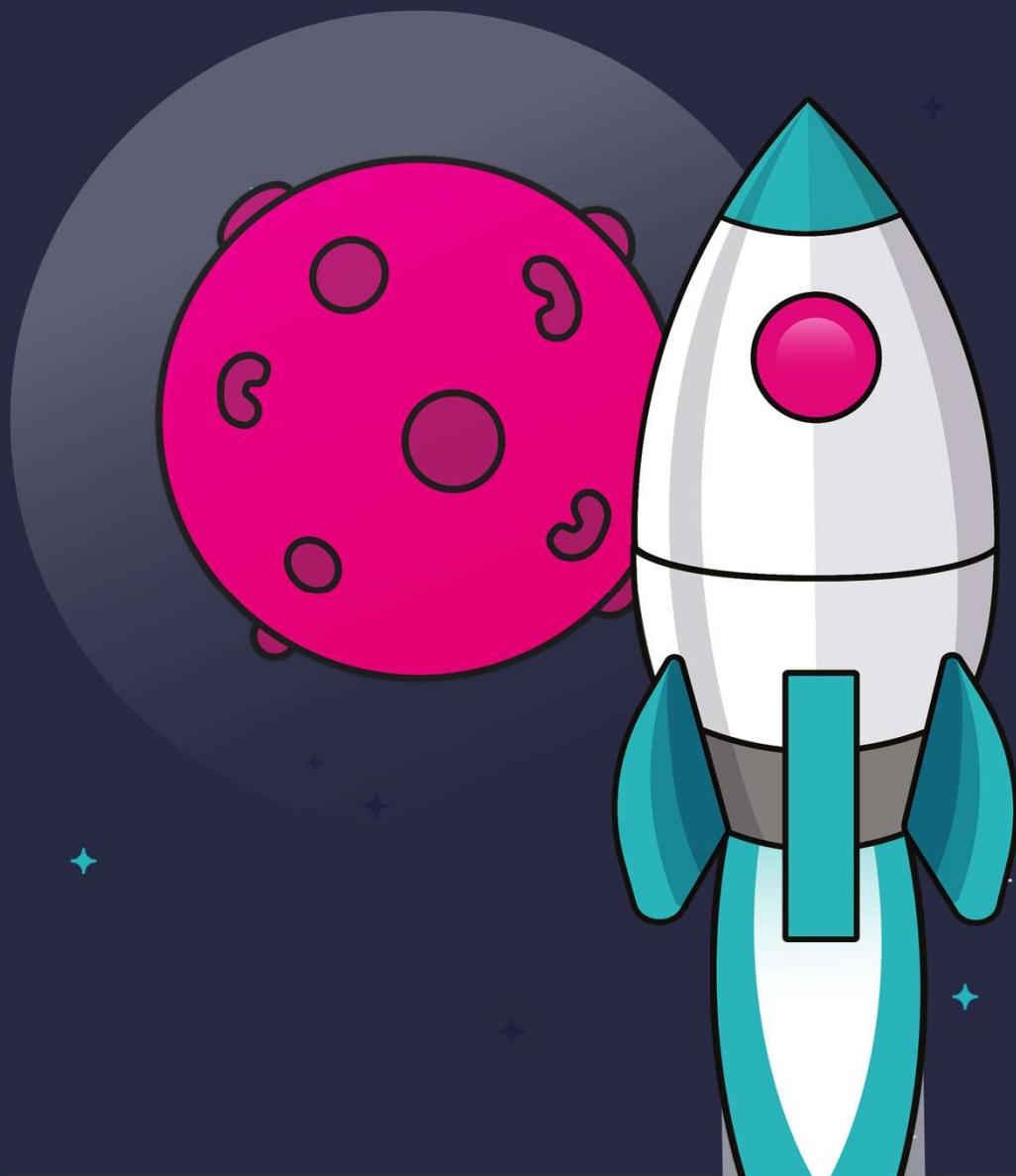
hn0va
community

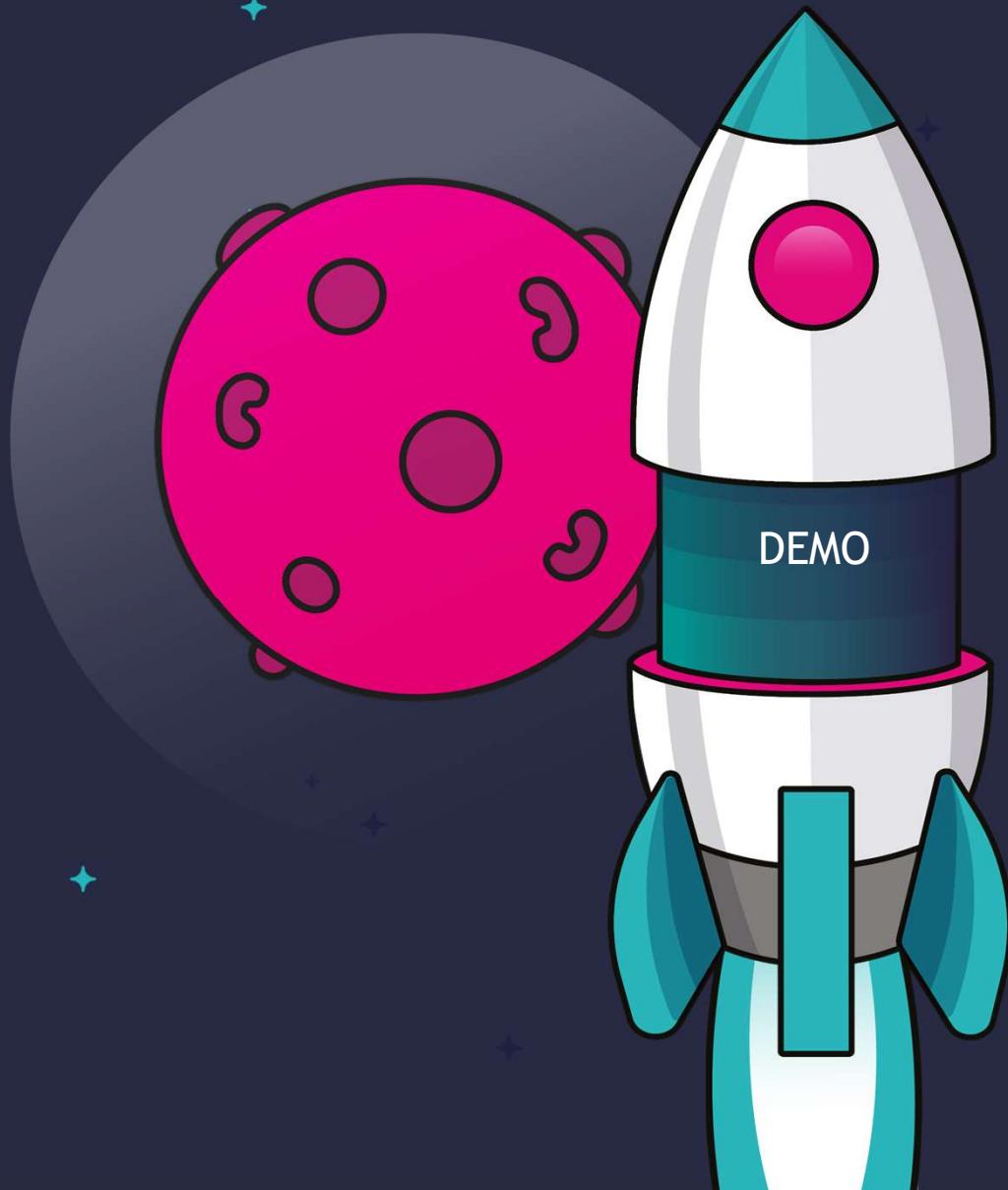
Discover data registered and scanned by Azure Purview

In Preview

The screenshot shows the Azure Purview interface with a sidebar containing various navigation items. The main area is divided into three sections:

- Register:** Shows a plug icon and the text "Register Synapse workspaces with Purview".
- Browse & Search:** Shows a search icon and the text "Search datasets in Synapse workspaces from linked Purview catalog".
- Enhanced metadata, lineage and annotation:** Shows a table icon and the text "View enhanced metadata, lineage from Purview for Synapse tables and add annotations".





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



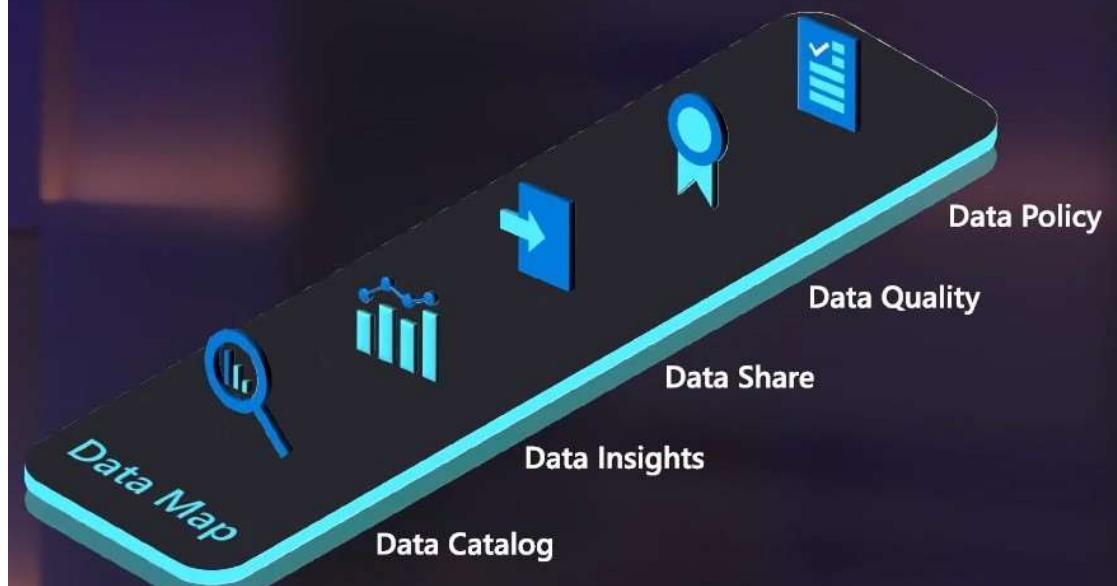
Purview Data Map	Available Now	Coming Soon
Automated scanning of hybrid sources	●	AWS S3
Classification	●	
Microsoft Information Protection Sensitivity Labels support	●	
Apache Atlas API support	●	
Purview Data Catalog		
Semantic Search and Browse	●	
Business Glossary	●	Hierarchical
Data Lineage	●	
Purview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

Roadmap

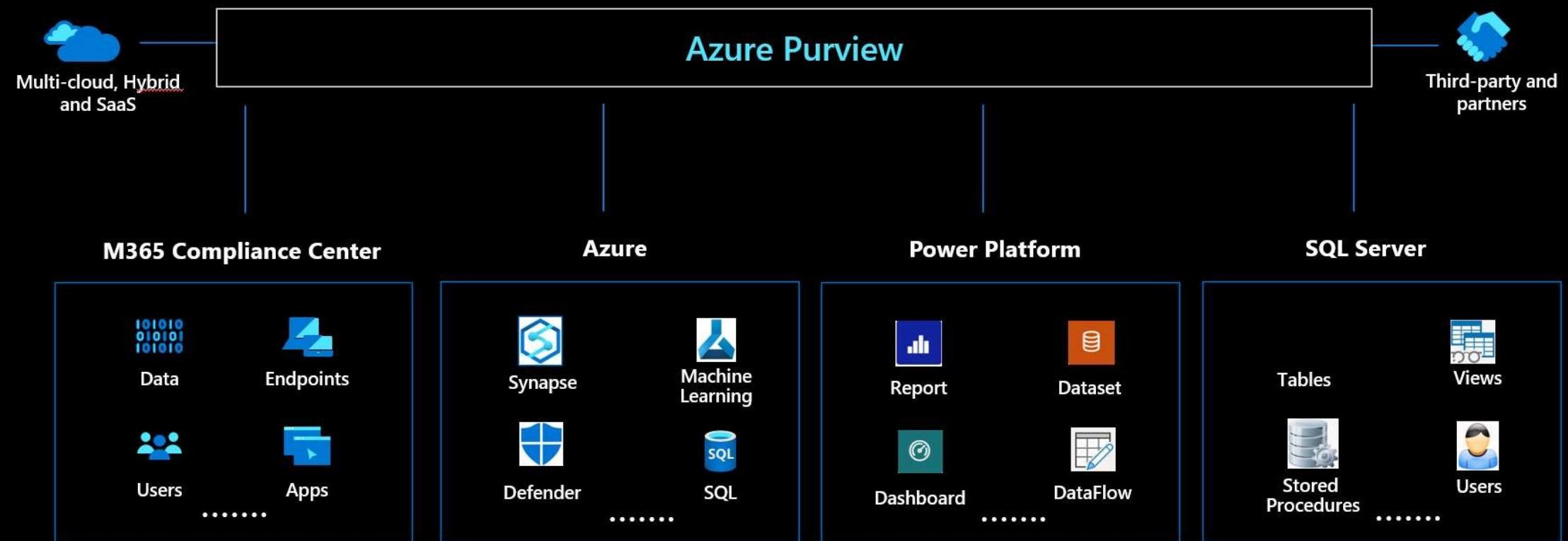
- Data Quality
- Data Privacy
- Master Data Management

Azure Purview roadmap

Unified | Open | Hybrid



Data Governance Opportunity





Microsoft Ignite

Take charge of data governance across your digital landscape

<https://myignite.microsoft.com/sessions/ee24433e-c7e9-4ef1-9b78-1d4add9231f3?source=sessions>

Enable unified data governance with Azure Purview

<https://myignite.microsoft.com/sessions/e1d2efc6-f8cc-406e-b666-9f866fe0b562?source=sessions>

Questions?



<https://erwindekreuk.com>



@erwindekreuk



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

Slides will be available on my blog



