



Microsoft Purview Cursus

Manage data governance in Azure (Module C)



Samuel Gaston-Raoul
Cloud Solution Architect

[in](https://aka.ms/Samuel)



Philippe Dudognon
Cloud Solution Architect

[in](#)



Frederic Gisbert
Cloud Solution Architect

[in](#)



Ali Bouhaddou
Cloud Solution Architect

[in](#)



Hélène Saulet
Partner Technology Strategist

[in](#)



Jessica Grenat
Partner Marketing Advisor

[in](#)



Charles Tostain
Cybersecurity Global Black Belt

[in](#)



Thierry Matusiak
Information Security Architect

[in](https://aka.ms/Thierry)



Saida Laga
Information Security Architect

[in](#)



Guillaume Lagache
Information Protection & Data Risk Specialist

[in](#)



Nabil Senoussaoui
Information Protection Specialist

[in](#)



Laurent Cardon
Principal Consultant

[in](#)

Microsoft Purview Cursus



<https://aka.ms/MicrosoftPurviewCursus>



Gauthier Lawny
Cyber Security Consultant

[in](#)



Cyril Beurel
FastTrack Architect

[in](#)

Agenda

- Introduction
- What's new in Microsoft Purview
- Microsoft Purview overview
- Microsoft Purview features
- Microsoft Purview REST API
- Microsoft Purview CLI | Lineage Connector Solution Accelerator
- Microsoft Purview and enterprise services integration
- Microsoft Purview deployment and security
- Roadmap
- Demo (Purview Studio, Synapse integration, REST API)



Frederic Gisbert
Cloud Solution Architect



Ali Bouhaddou
Cloud Solution Architect



Samuel Gaston-Raoul
Cloud Solution Architect



aka.ms/Samuel



Introduction



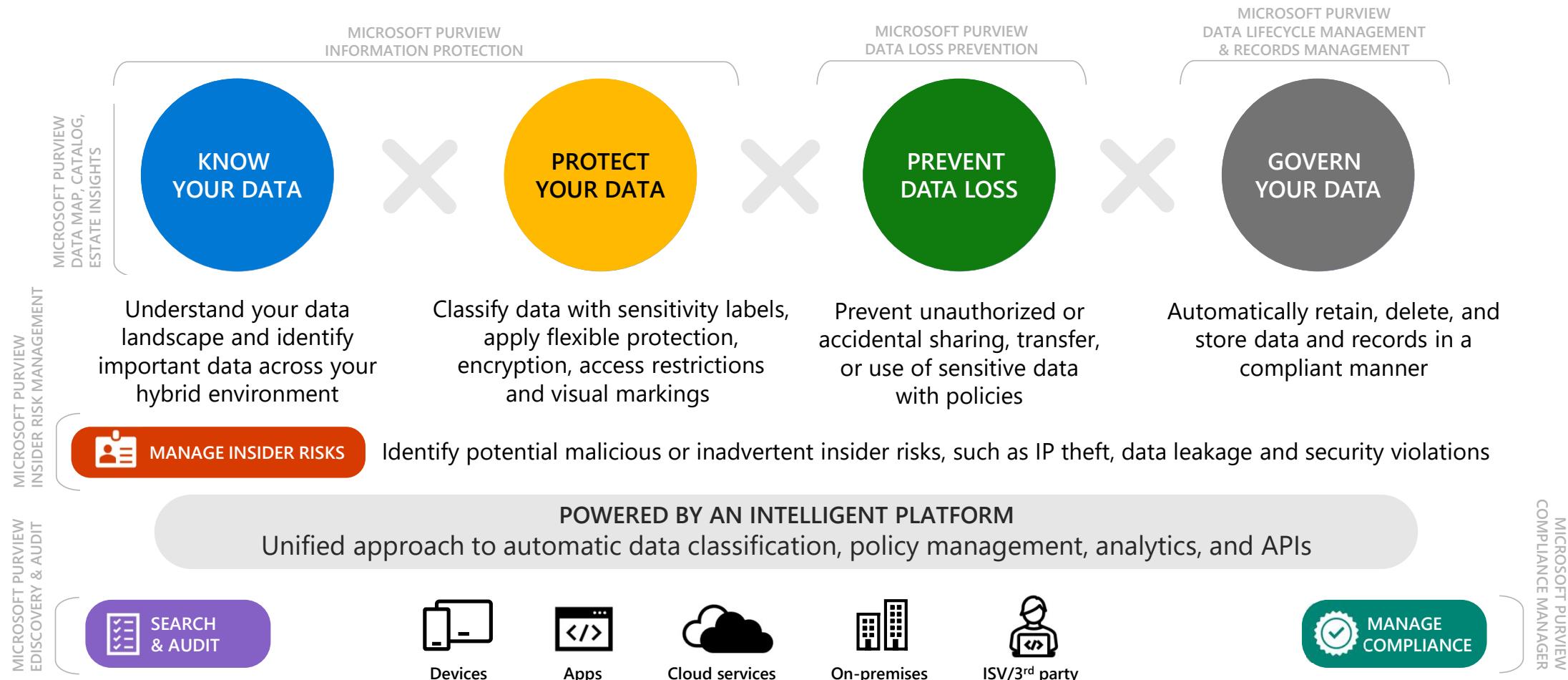
Microsoft Purview



**Comprehensive
solutions to help
govern, protect, and
manage your risks
and data estate.**

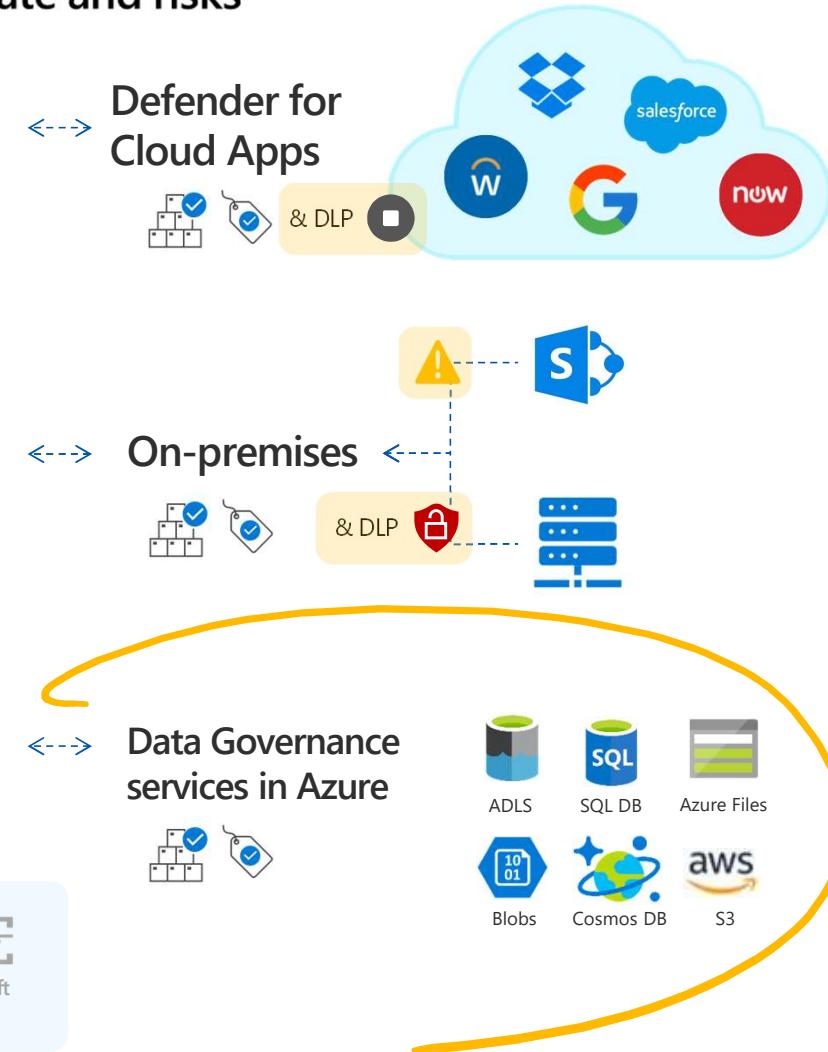
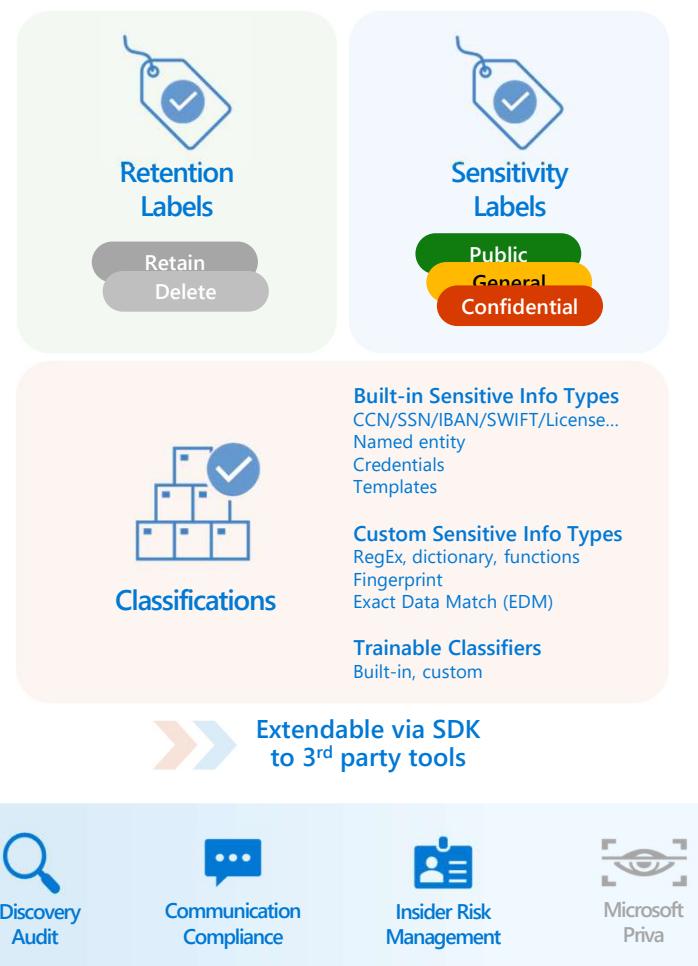
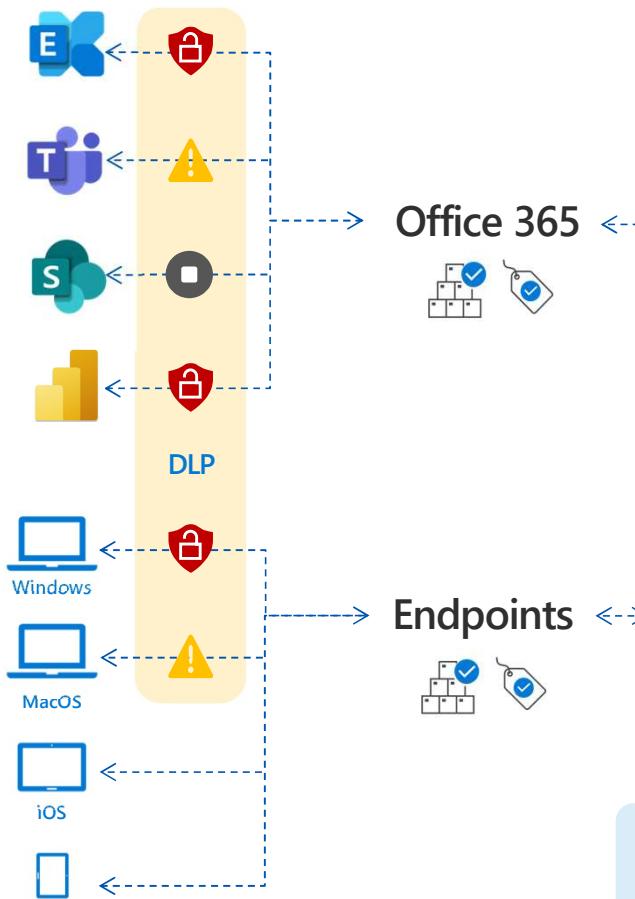
Microsoft Purview | Information protection, governance & risk management

Understand, protect and govern data wherever it lives, manage insider risks and compliance



Microsoft Purview

Comprehensive solutions to help **govern**, **protect**, and **manage** your data estate and risks





Microsoft Purview Cursus

Cursus de développement des connaissances et compétences

Module 1

Understand and cartography data

Capacités de **découverte**, de **classification** et de **cartographie** des données :

- ✓ **Sensitive Information Types**,
- ✓ **Exact Data Matching**,
- ✓ **Trainable classifiers**,
- ✓ **Content Explorer**.

Module 2

Label and Protect information

Capacités **d'étiquetage** et de **protection** de l'**information** (Information Protection) :

- ✓ **Sensitivity labels (AIP/MIP)**,
- ✓ **Microsoft RMS (IRM)**,
- ✓ **Data Loss Prevention (DLP)**,
- ✓ **Activity Explorer**.

Module 3

Govern data, search and audit

Capacités de **gouvernance** (**archivage**, **rétention**, **suppression**), **d'audit** et **recherche** des données :

- ✓ **Data Lifecycle Management** (Information Governance, retention policies & retention labels),
- ✓ **Records Management**,
- ✓ **eDiscovery**,
- ✓ **Audit**.

Module 4

Ask the Experts

Q&R et **discussions** : échanges sur les bonnes pratiques ou démonstrations techniques plus spécifiques, en référence aux fonctionnalités des trois premiers modules.

Cursus « Essentiel »

Module A Encryption offers and keys management in Microsoft 365

Capacités de **chiffrement** de l'information ainsi que de **gestion des clés** :

- ✓ **Microsoft RMS (IRM / Azure RMS)**,
- ✓ **Office 365 Message Encryption (OME)**,
- ✓ **Double Key Encryption (DKE)**,
- ✓ **Service encryption with Customer Key**.

Cursus « Optionnel »

Module B Manage Risk & Compliance

Capacités de **gestion des risques** et de la **conformité** :

- ✓ **Communication Compliance**,
- ✓ **Insider Risk Management**,
- ✓ **Compliance Manager**.

Module C

Manage data governance in Azure

Capacités de gestion la **gouvernance des données** dans **Azure** :

- ✓ **Data Map**,
- ✓ **Data Catalog**,
- ✓ **Data Estate Insights**.

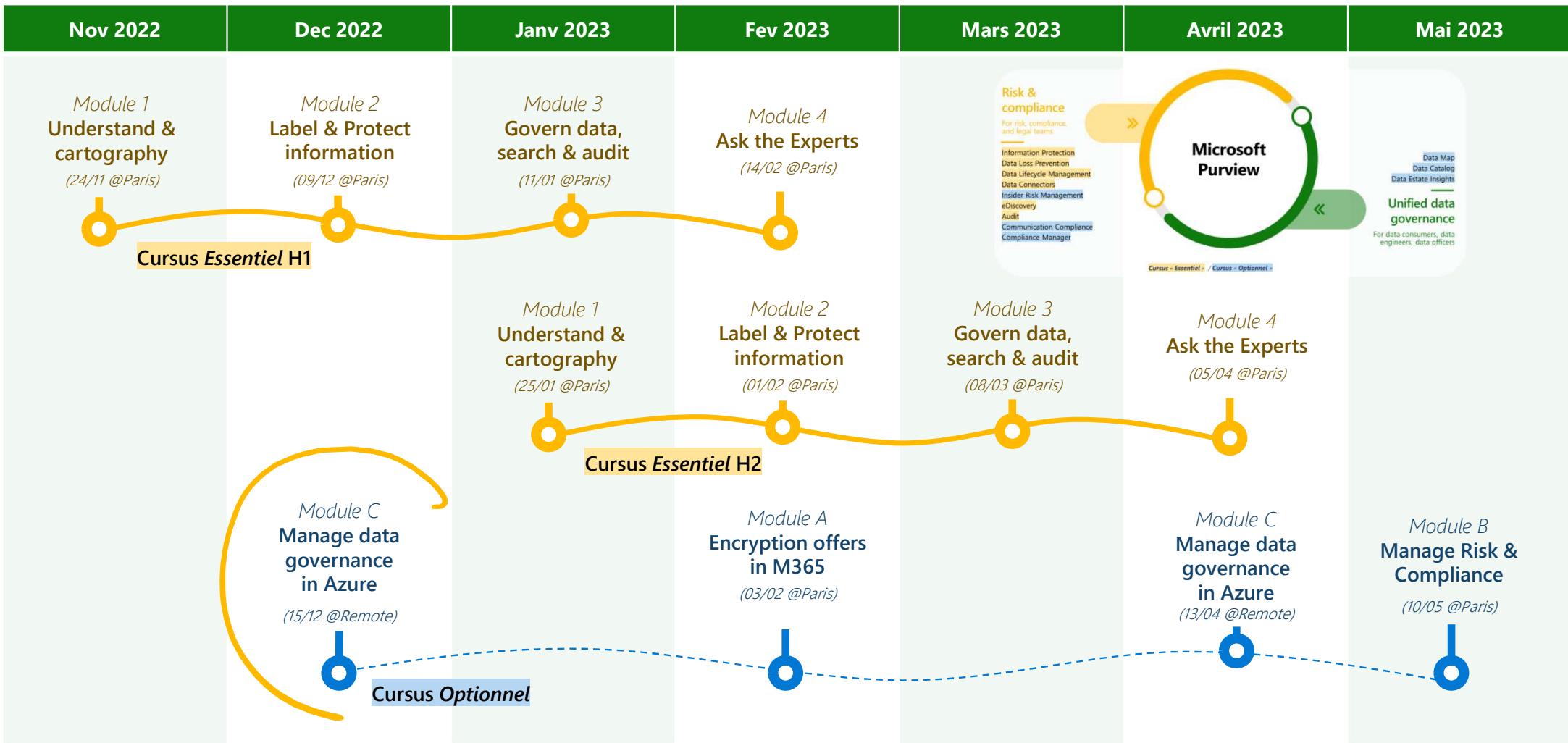
Manage data governance in Azure



➡ <https://aka.ms/MicrosoftPurviewCursus>



Microsoft Purview Cursus



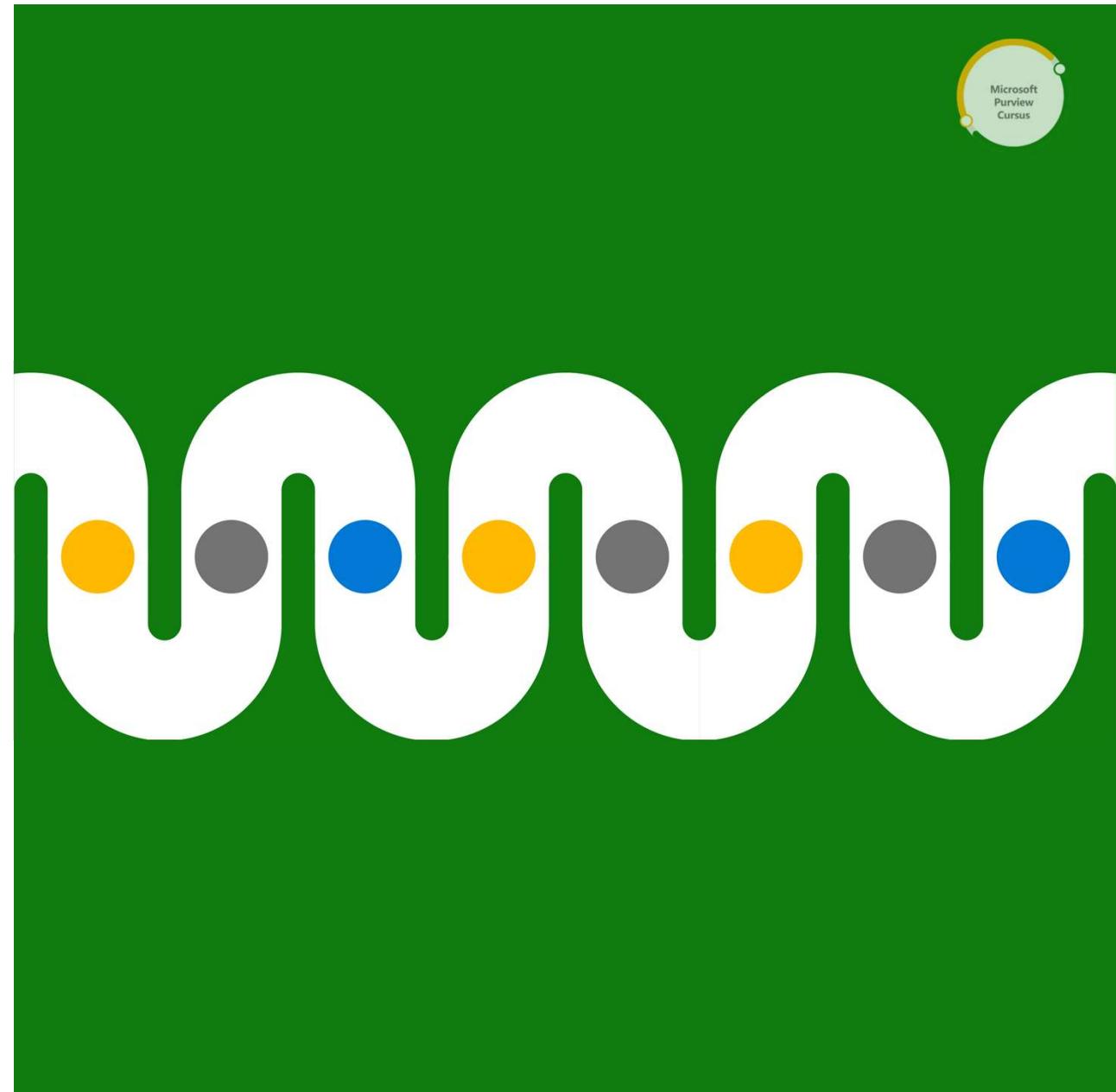


Workshop Microsoft Purview Data Governance

Agenda (2h00)

- What's new in Microsoft Purview Data Governance in Azure
- Data Governance in Azure overview
- Data Governance in Azure features
- Data Governance in Azure REST API
- Data Governance in Azure CLI | Lineage Connector Solution Accelerator
- Data Governance in Azure and enterprise services integration
- Data Governance in Azure deployment and security
- Roadmap
- Demo (Purview Studio, Synapse integration, REST API)

What's new in Microsoft Purview Data Governance in Azure



Microsoft Purview metamodel

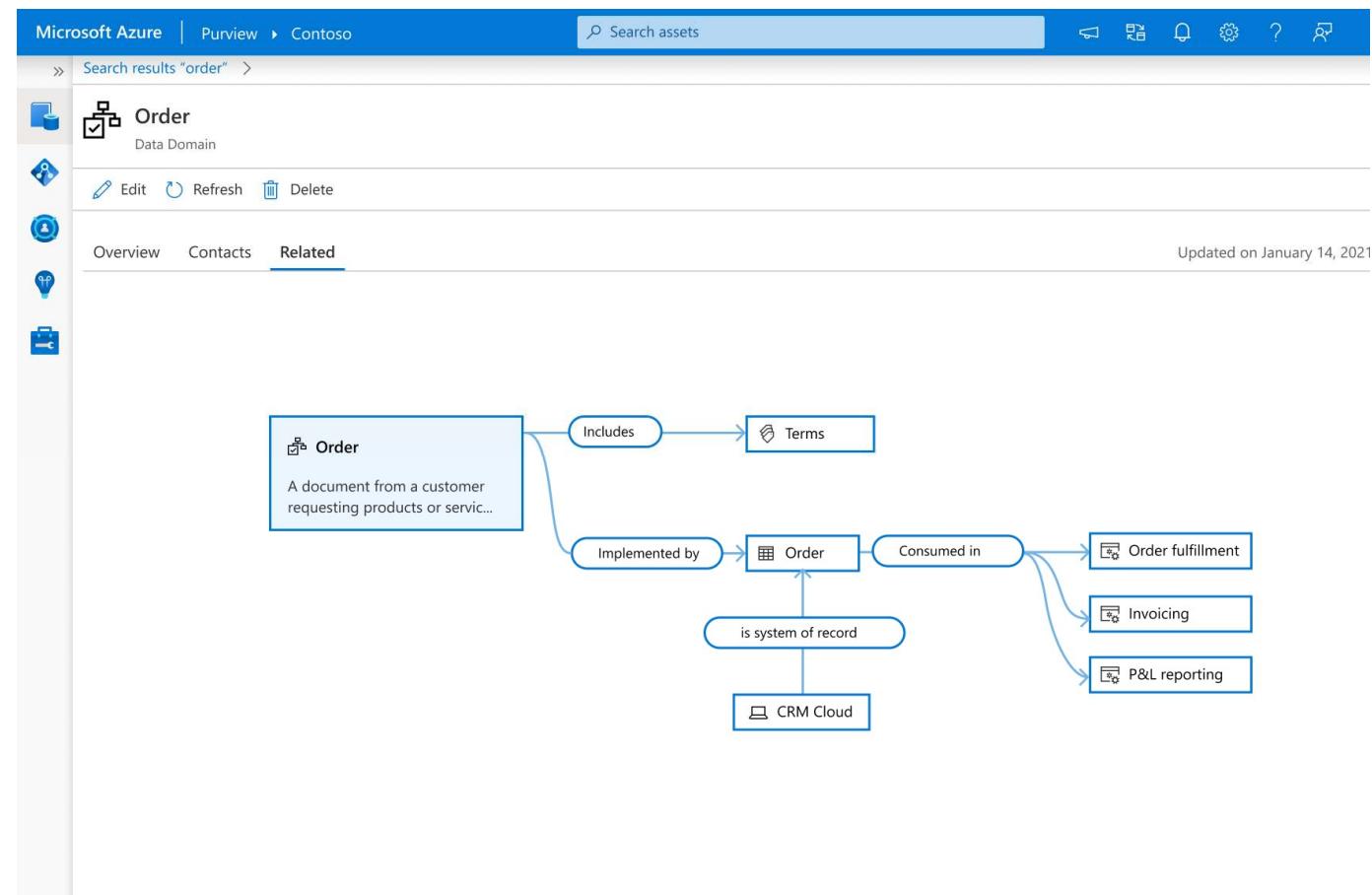
Purview Metamodel

Helps you describe and organize your domains using business assets and their relationships in Purview.

Business assets will be stored like any other asset in the catalog so they can be searched and retrieved just like glossary terms and data assets.

MVP asset types:
-Data domains
-Business processes
-Systems

More asset types will be added in future phases.



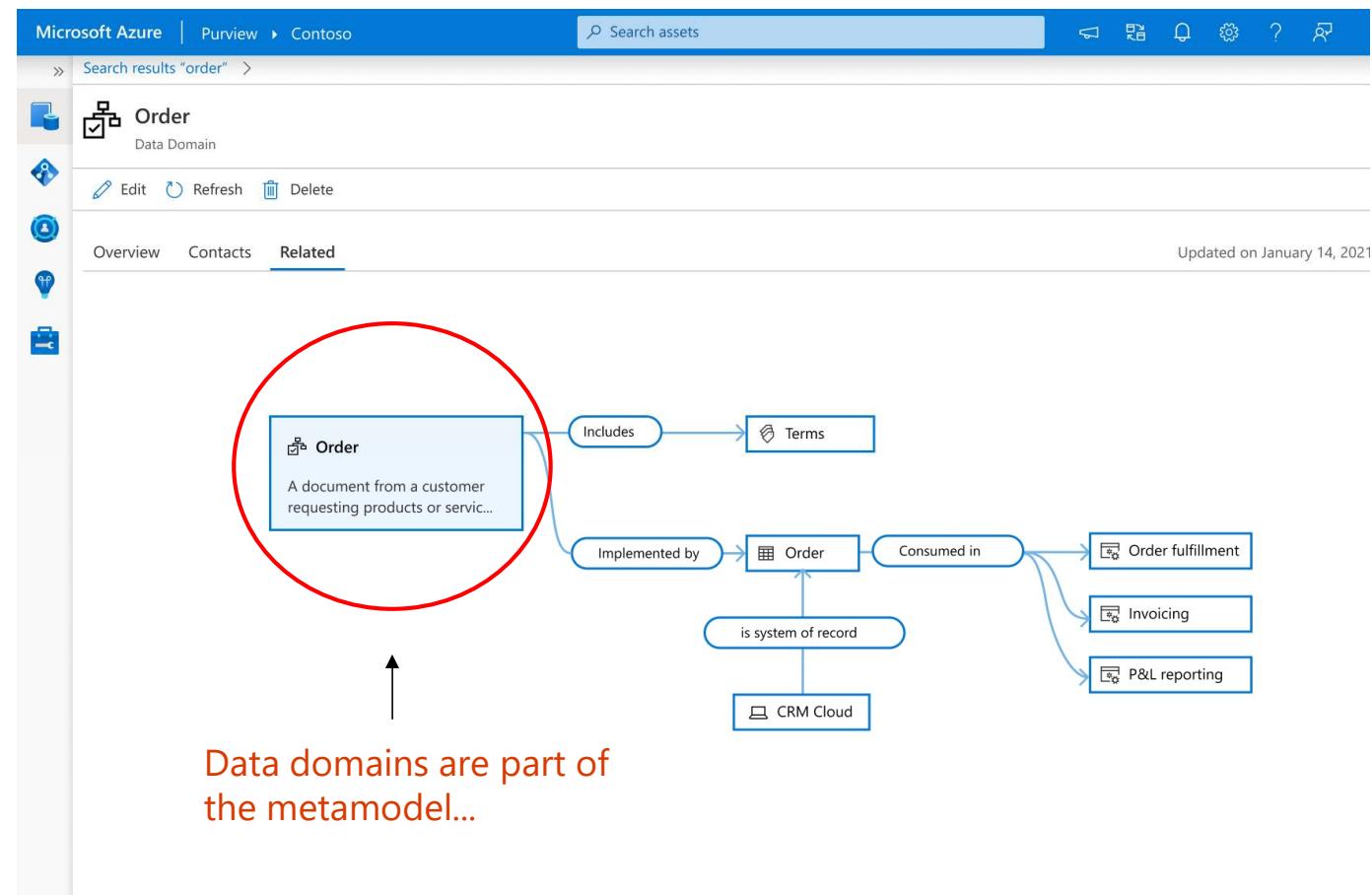
Purview Metamodel

Helps you describe and organize your domains using business assets and their relationships in Purview.

Business assets will be stored like any other asset in the catalog so they can be searched and retrieved just like glossary terms and data assets.

MVP asset types:
-Data domains
-Business processes
-Systems

More asset types will be added in future phases.



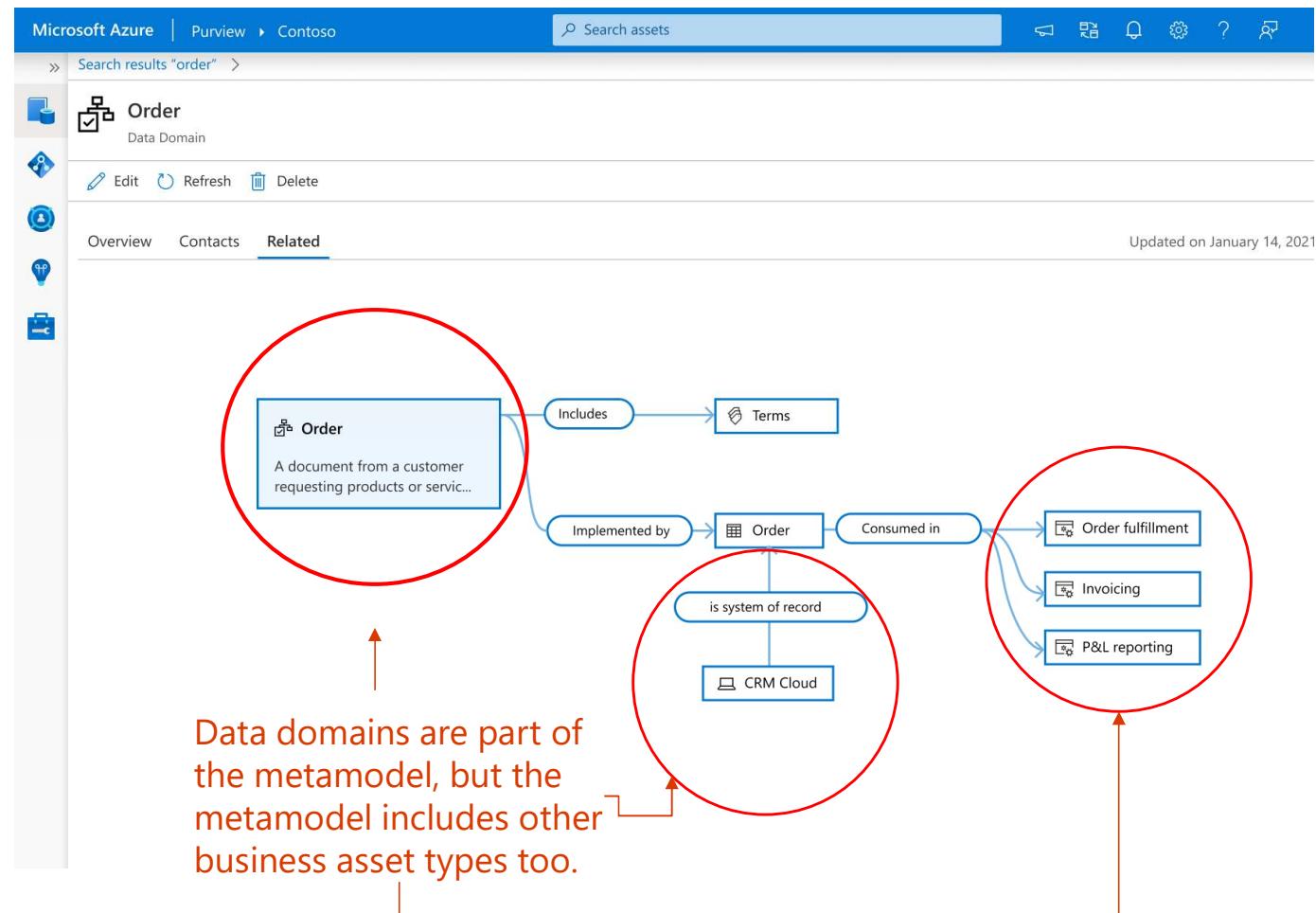
Purview Metamodel

Helps you describe and organize your domains using business assets and their relationships in Purview.

Business assets will be stored like any other asset in the catalog so they can be searched and retrieved just like glossary terms and data assets.

MVP asset types:
-Data domains
-Business processes
-Systems

More asset types will be added in future phases.



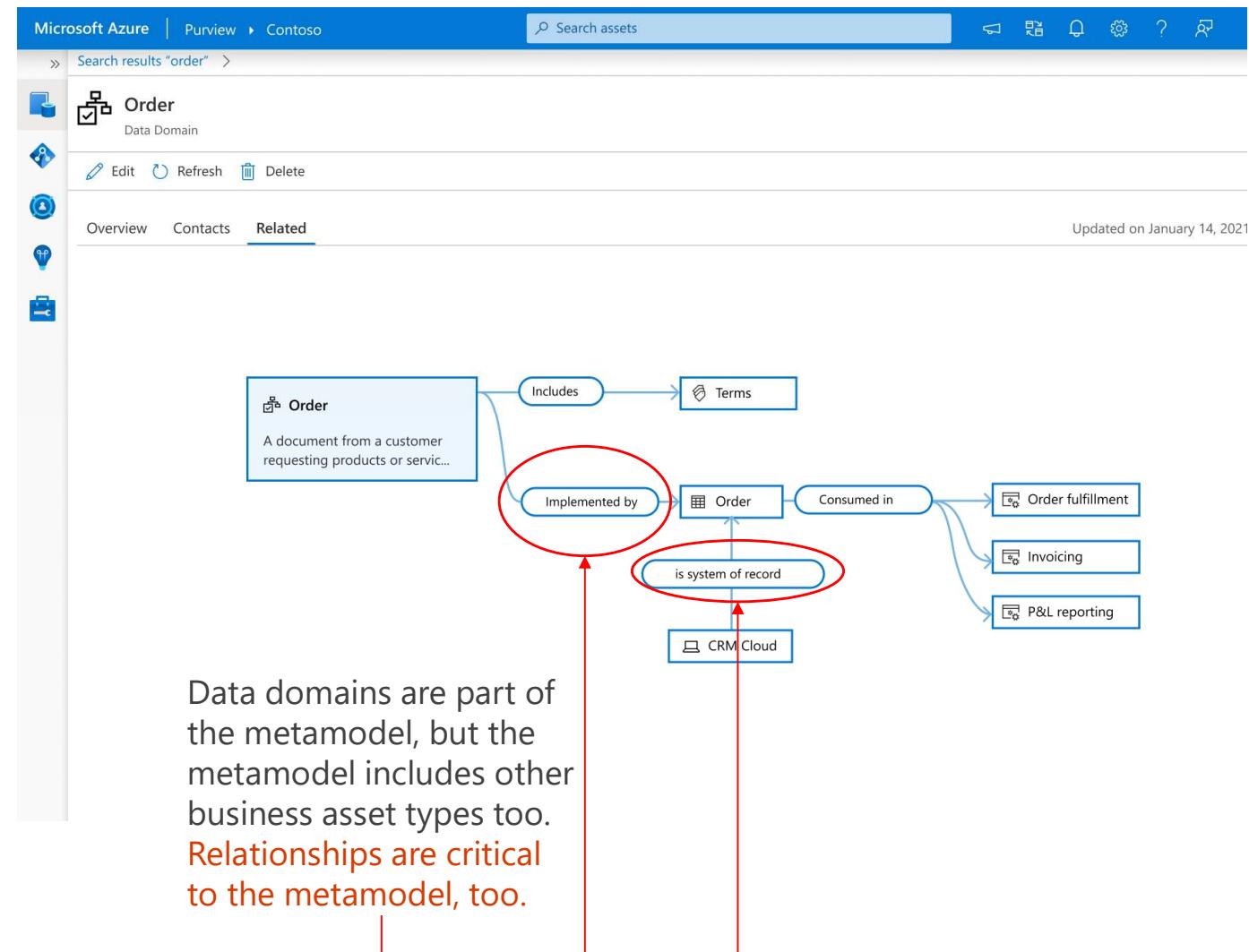
Purview Metamodel

Introduces a semantic layer in Purview by allowing customers to describe business assets and their relationships in Purview.

Business assets will be stored like any other asset in the catalog so they can be searched and retrieved just like glossary terms and data assets.

MVP asset types:
-Data domains
-Business processes
-Systems

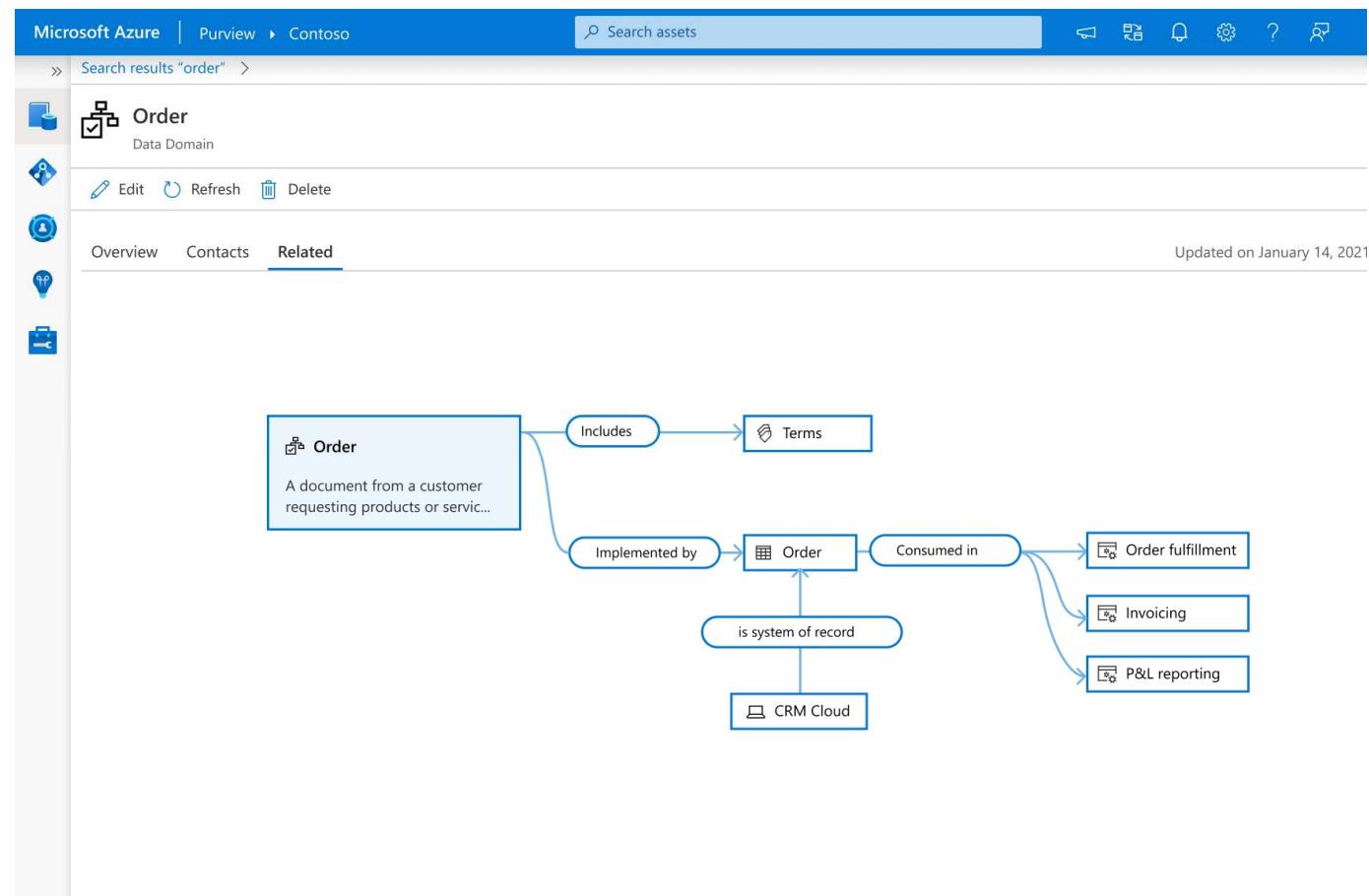
More asset types will be added in future phases.



What you can do with the metamodel

1. Organize data into domains
2. Establish ownership and data *use*: Show how data is reused to support business activities, departments, and lines of business
3. Describe your data mesh.
4. Describe your data from a business perspective so it can be more easily discovered and understood.

More...



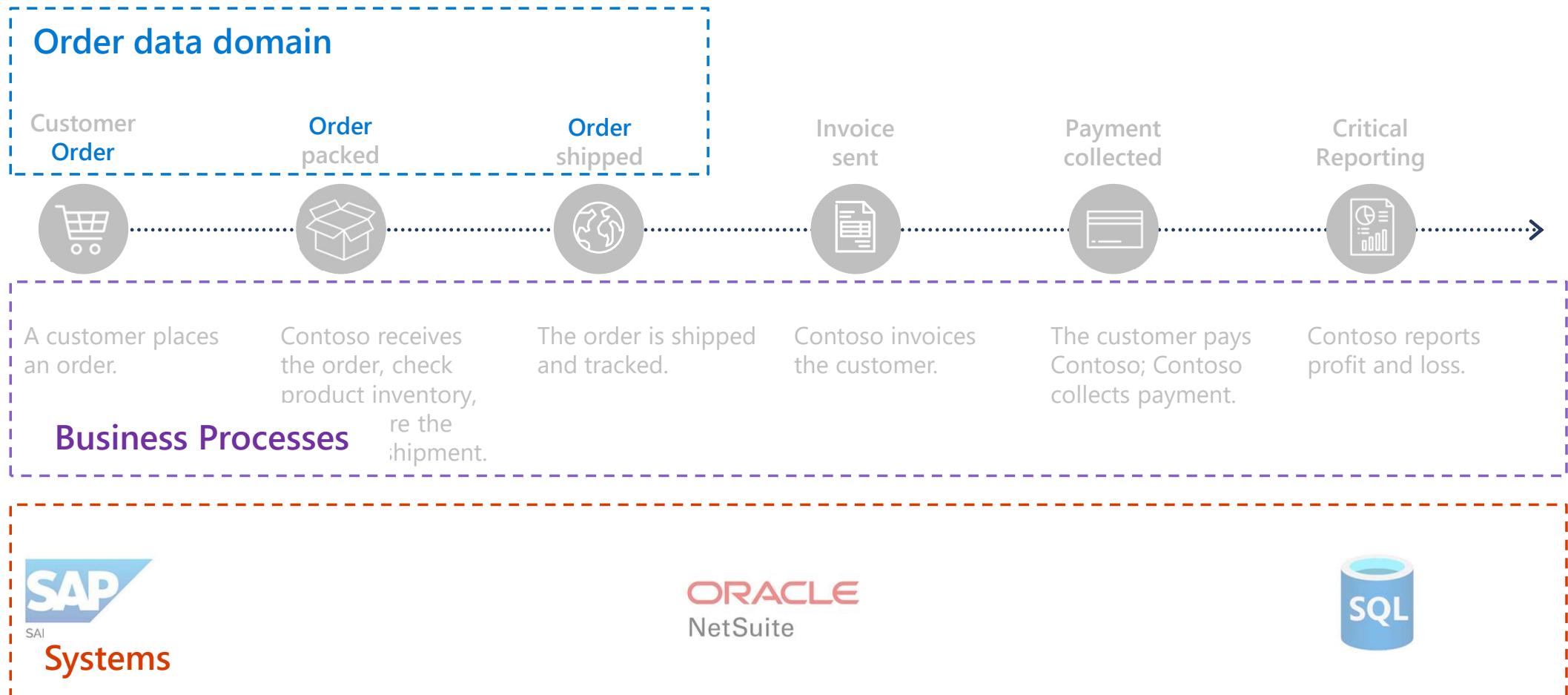
Purview Metamodel has 3 experiences

1. Authoring experience: We'll give you a set conceptual object templates you can customize to describe things and ideas important to your organization. E.g. *system, business process, data domain, project*, etc.
2. Curation experience: How you'll link physical data assets to the concepts you define. This extends the business metadata you can use to describe your physical data.
3. Discovery experiences: how these connections are surfaced in the catalog in search, asset detail pages, and relationship graphs.

Order to cash: A data story



Stars of the story: data domain, processes, systems



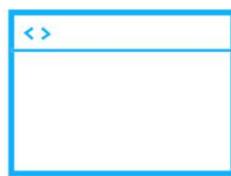
Where does this story live at Contoso?



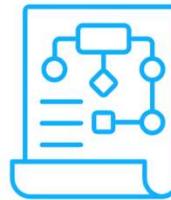
This Excel file
that Sarah made.



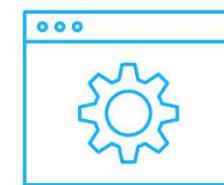
A photo that
Sagar took of a
whiteboard from
our MDM kickoff
meeting.



A custom-built
report catalog
Andy in Finance
owns.



Emeka's
architecture
diagrams. Ask
him to share his
OneDrive with
you.



Aris (Business
architecture team)

ServiceNow
(Engineering IT)



Jamal in BI
services. He
knows everyone!

Let's tell this story with Purview.

SCENARIO

1

Authoring with data domains



Claudia
Purview admin

1. To get started, Claudia navigates to the Governance Center in Purview to manage her organization's metamodel
2. Claudia sees a list of her asset templates and an example metamodel to help her get started.

Microsoft Azure | Purview > Contoso

Search assets

Metamodels

Build a metamodel by adding concepts to the canvas and defining the relationships between them. To create logical assets from a concept, go to Collections.

+ New concept + New relationship Refresh Export Import Create assets

Components

Filter by Show all available concepts

Concepts Relationships

Business

- Business process
- Data domain
- System
- Glossary term

Governance

Data

- Column
- Dataset
- Schema

Metamodel Diagram:

```
graph TD; BP[Business process] --> S[System]; BP --> DS[Dataset]; DS -- Consumes --> BD[Data domain]; DS -- Has --> SC[Schema]; SC -- Implemented by --> C[Column]; GT[Glossary term] -- Is stored in --> C
```

Business process

A series of steps or activities that produce a product.

System

is system of record for

Dataset

Consumes

Data domain

Has

Schema

Implemented by

Has

Glossary term

Is stored in

Column

Overview Attributes Relationships

Name	Business process
Description	A series of steps or activities that produce a product.
Type	Business



Claudia
Purview admin

Claudia also wants to help her organization understand *who* depends on data to complete key activities in her organization, so she'll define her organization's departments.

There's no template in Purview to do this yet, so she creates her own.

She creates a new concept.

Microsoft Azure | Purview > Contoso

Search assets

Metamodels

Build a metamodel by adding concepts to the canvas and defining the relationships between them. To create logical assets from a concept, go to the Data catalog.

+ New concept + New relationship Refresh Create assets

Components

Show only concepts in use

Filter by keywords

Concepts

Business

- Business process
- Project
- System
- Glossary term
- Data domain

Data

- Column
- Dataset
- Schema
- Database
- Report

Metamodels

Business process

A series of steps or activities that produce a product.

System

is system of record for

Business process

Consumes

Dataset

Data domain

Implemented by

Schema

Glossary term

Has

Has

Column

Overview Attributes Relationships

Name	Business process
Description	A series of steps or activities that produce a product.
Type	Business

```
graph TD; BP[Business process] -- "is system of record for" --> S[System]; BP -- Consumes --> DS[Dataset]; DS -- Has --> DD[Data domain]; DD -- Implemented by --> SC[Schema]; SC -- Has --> C[Column]; GT[Glossary term] -- Has --> C
```



Claudia
Purview admin

She names the new concept *department* and gives it a description.

This will provide the template for capturing Contoso's departments.

Microsoft Azure | Purview > Contoso

Search assets

Metamodels

Build a metamodel by adding concepts to the canvas and defining the relationships between them.

New concept

Name * Department

Description Divisions of the company.

Type * Business

Parent Null

Components

Concepts Relationships

Filter by keywords

Business

- Business process
- Project
- System
- Glossary term
- Data domain
- Geographic region
- Line of business

Governance

- Assessment
- Policy

Data

- Column
- Dataset
- Schema
- Database
- Report

System Has Project

Overview Attributes Relationships

Add relationship

Concept (Head)

Business process

Business process

Project

Create Cancel



Claudia
Purview admin

She creates a new relationship to show which departments own each business process.

Microsoft Azure | Purview > Contoso

Search assets

Metamodels

Build a metamodel by adding concepts to the canvas and defining the relationships between them. To create logical assets from a concept, go to the Data catalog.

+ New concept + New relationship Refresh Create assets

Components

Filter by keywords

Concepts Relationships

Business

- Business process
- Project
- System
- Glossary term
- Department**

Data

- Column
- Dataset
- Schema
- Database
- Report

Metamodels

Business process

A series of steps or activities that produce a product.

Department

Divisions of the company.

System

is system of record for

Dataset

Consumes

Data domain

Implemented by

Schema

Has

Glossary term

Column

Overview Attributes Relationships

Name	Department
Description	Divisions of the company.
Type	Business

The screenshot shows the Microsoft Purview Metamodels interface. On the left, there's a navigation sidebar with icons for Metamodels, Policies, Sensitivity labels, DAG rules, and Alerts & metrics. The main area has tabs for Concepts and Relationships. Under Concepts, there are sections for Business (Business process, Project, System, Glossary term, Department) and Data (Column, Dataset, Schema, Database, Report). The 'Department' concept is selected. The main canvas displays a diagram of various business and data concepts: Business process, Department, System, Dataset, Data domain, Schema, Column, and Glossary term. Relationships shown include 'is system of record for' (Business process to System), 'Consumes' (Dataset to System), 'Implemented by' (Data domain to Schema), 'Has' (Schema to Column), and another 'Has' relationship (Schema to Column). Below the diagram, there are tabs for Overview, Attributes, and Relationships, with 'Relationships' currently active. A table provides details for the selected 'Department' concept: Name (Department), Description (Divisions of the company.), and Type (Business). The top right corner shows the user's email (robert@contoso.com) and name (CONTOSO).



Claudia
Purview admin

She creates a new relationship.

Microsoft Azure | Purview > Contoso

Metamodels

Search assets

New concept New relationship Refresh Create assets

Components

Concepts Relationships

Filter by keywords

Business

- Business process
- Project
- System
- Glossary term
- Department**

Data

- Column
- Dataset
- Schema
- Database
- Report
- Geographic region
- Line of business

Metamodels

Build a metamodel by adding concepts to the canvas and defining the relationships between them. To create logical assets from a concept, go to the Data catalog.

Department

Divisions of the company.

Business process

A series of steps or activities that produce a product.

System

Dataset

Schema

Column

Glossary term

Business process is system of record for **System**

Department Owns **Select..** **Business process**

Department Consumes **Dataset**

Dataset Has **Schema**

Schema Has **Column**

Column Has **Glossary term**

Department Owns **Select..** **Business process**

Department Consumes **Dataset**

Dataset Has **Schema**

Schema Has **Column**

Column Has **Glossary term**

Overview Attributes Relationships

Add relationship

Overview Attributes Relationships

Add relationship

Name Relationship Name

Name	Relationship	Name
Department	Owns	Select..

Cancel Save Delete

The screenshot shows the Microsoft Purview Metamodels interface. On the left, there's a navigation sidebar with icons for Metamodels, Policies, Sensitivity labels, DAG rules, and Alerts & metrics. The main area has tabs for Concepts and Relationships. Under Concepts, there are sections for Business (Business process, Project, System, Glossary term, Department), Data (Column, Dataset, Schema, Database, Report, Geographic region, Line of business), and another Data section. Under Relationships, there are tabs for Overview, Attributes, and Relationships, with an 'Add relationship' button. A search bar at the top says 'Search assets'. The main canvas shows various entities like Department, Business process, System, Dataset, Schema, Column, and Glossary term, connected by relationships like 'Owns', 'Consumes', 'Has', and 'Implemented by'. A specific relationship is being edited between Department and Business process, with 'Owns' selected as the relationship type and 'Select..' as the target entity. The status bar at the bottom right shows ' robert@contoso.com CONTOSO AC'.

Diapositive 26

A0 Change to department owns business process.

Author; 2022-01-04T16:28:54.910



Claudia
Purview admin

She creates a new relationship and connects it her business process asset so she can show departments own each business process.

Microsoft Azure | Purview > Contoso

Search assets

Metamodels

Build a metamodel by adding concepts to the canvas and defining the relationships between them. To create logical assets from a concept, go to the Data Catalog.

New concept New relationship Refresh Create assets

Components

Filter by keywords

Concepts Relationships

Business

- Business process
- Project
- System
- Glossary term
- Department
- Data domain

Data

- Column
- Dataset
- Schema
- Database
- Report

Metamodels

Business process: A series of steps or activities that produce a product.

Department: Divisions of the company.

System: is system of record for

Dataset: Consumes

Data domain: Implemented by

Schema: Has

Glossary term: Has

Column: Owns

Relationships

Add relationship

Name	Relationship	Name
Business process	Owns	Business process

```
graph TD; BP[Business process] -- "A series of steps or activities that produce a product." --- BP_desc; D[Department] -- "Divisions of the company." --- D_desc; S[System] -- "is system of record for" --- SR; DS[Dataset] -- Consumes --- D; DD[Data domain] -- Implemented by --- S; SC[Schema] -- Has --- DD; GT[Glossary term] -- Has --- C[Column]; D -- Owns --- BP; C -- Owns --- GT;
```



Claudia
Purview admin

Claudia creates new business assets in the Data catalog.

Microsoft Azure | Purview > Contoso

Search assets

Glossary terms Business assets

Data catalog > Business assets >

Assets

New asset Move Edit Delete Refresh Export Import Manage view Manage concepts

Filter by name Asset type : Business process

Showing 1 - 3 of 3 assets

Asset name ↑↓	Asset type ↑↓	Status	Consumes	System of record	Owner ↑↓
Invoice customer	Business process	Approved	cust_billing	CRM Cloud	Robert Fox
Collect payment	Business process	Approved	cust_payment	Contoso Billing	Robert Fox
Calculate profit and loss	Business process	Approved	consolidated_rev	General Ledger	Robert Fox



Claudia
Purview admin

She completes the properties.

Microsoft Azure | Purview > Contoso

Search assets

Data catalog > Business assets >

Assets

New asset Move Edit Delete Refresh Export

Filter by name Asset type : Business process

Showing 1 - 3 of 3 assets

Asset name ↑	Asset type ↑↓
Invoice customer	Business process
Collect payment	Business process
Calculate and report revenue	Business process

New asset

Data catalog > Business assets

Details

Name * Order fulfillment

Description Receive order, check inventory, ship and track.

Asset type * Business process

Status * Draft

Attributes

Naming standard * Order_fulfillment

Method * XXX

Process note

Relationships

Consumes Search for Dataset

Impacted by Search for Project

Create Cancel



Claudia
Purview admin

Completes asset creation

Microsoft Azure | Purview > Contoso

Search assets

Glossary terms Business assets

Data catalog > Business assets >

Assets

New asset Move Edit Delete Refresh Export Import Manage view Manage concepts

Filter by name Asset type : Business process

Showing 1 - 3 of 3 assets

Asset name ↑↓	Asset type ↑↓	Status	Consumes	System of record	Owner ↑↓
Invoice customer	Business process	Approved	cust_billing	CRM Cloud	Robert Fox
Collect payment	Business process	Approved	cust_payment	Contoso Billing	Robert Fox
Calculate profit and loss	Business process	Approved	consolidated_rev	General ledger	Robert Fox
Order fulfillment	Business process	Approved	Add Dataset	Add System of record	Robert Fox

SCENARIO

2

Linking technical assets to data
domains



Suki

Domain/Business
Owner/ Data curator

Search for Dataset

Microsoft Azure | Purview > Contoso

Search assets

Search for Dataset

Source type : Power BI, Azure Machine Learning Instance : All Clear all filters

Showing 1-25 out of 3,263 results Sort by: Relevance

Dataset

Collection

- Contoso 2,082
- Finance Contoso 739
- Sales Contoso 76

Classification

- Contoso Product Number 512
- Catalog_NA 6

Sensitivity label

- Confidential 726
- General 32

Contact

- Darlene Murphy 326
- Jorge Jones 203
- Morris McCoy 174
- Dustin Cooper 93

Glossary term

- Finance2020 38
- Gross Profit 2

customer

PowerBI Dataset

https://march2020demog2.dfs.core.windows.net/customer

Sarah Jones updated 10 days ago

consolidated_orders

PowerBI Dataset

adl://tradisdev.azuredatalakestore.net/VidhyaTest/Customer/Customer

Sarah Jones updated 10 days ago

CpcCustomerIncidentsReport

PowerBI Dataset

https://skypebusinessvoice.kusto.windows.net/BVLogs/CpcCustomerIncident

Sarah Jones updated 10 days ago

customer_cleaned

Azure Machine Learning Dataset

This is the primary customer dataset for Project Keto.

mssql://keto-server.database.windows.net/KETO/datasets/customer_cleaned

Sarah Jones updated 10 days ago

Customer Invoice

PowerBI Dataset

Fully qualified name: Customer.Invoice

Sarah Jones updated 10 days ago

CustomerConnectCalls

General

< Previous Page 1 of 8 Next >

Add  Cancel



Suki

Domain/Business
Owner/ Data curator

Microsoft Azure | Purview > Contoso

Search assets

Glossary terms

Business assets

Data catalog > Business assets >

Assets

New asset Move Edit Delete Refresh Export Import Manage view Manage concepts

Filter by name Asset type : Business process

Showing 1 - 3 of 3 assets

Asset name ↑↓	Asset type ↑↓	Status	Consumes	System of record	Owner ↑↓
Invoice customer	Business process	Approved	cust_billing	CRM Cloud	Robert Fox
Collect payment	Business process	Approved	cust_payment	Contoso Billing	Robert Fox
Calculate and report revenue	Business process	Approved	consolidated_rev	General Ledger	Robert Fox
Order fulfillment	Business process	Approved	consolidated_orders	Add system of record	Robert Fox



Suki

Domain/Business
Owner/ Data curator

Complete

Microsoft Azure | Purview > Contoso

Search assets

Glossary terms

Business assets

Data catalog > Business assets >

Assets

New asset Move Edit Delete Refresh Export Import Manage view Manage concepts

Filter by name Asset type : Business process

Showing 1 - 3 of 3 assets

<input type="checkbox"/> Asset name ↑↓	Asset type ↑↓	Status	Consumes	System of record	Owner ↑↓
<input type="checkbox"/> Invoice customer	Business process	✓ Approved	cust_billing	CRM Cloud	Robert Fox
<input type="checkbox"/> Collect payment	Business process	✓ Approved	cust_payment	Contoso Billing	Robert Fox
<input type="checkbox"/> Calculate and report revenue	Business process	✓ Approved	consolidated_rev	General Ledger	Robert Fox
<input type="checkbox"/> Order fulfillment	Business process	✓ Approved	consolidated_orders	CRM Cloud	Robert Fox

SCENARIO

3

Search and discover data using
data domains



Daisy

Data scientist/Purview reader

Daisy is a data scientist who's been asked to build an algorithm that recommends new products to customers based on their previous orders.

She searches for order.

Microsoft Azure | Purview > Friendly name

connie@contoso.com
CONTOSO

Friendly name

(account name)

3 source types 8 sources 425 assets 24 terms

Search

Browse assets Explore assets by source type and collection.

View glossary View your team's business glossary.

Knowledge center Discover learning materials and tutorials.

Recently accessed My items

Name	Last updated
contoso_financialdata	10 mins ago
contoso_testreport	30 mins ago
device	a week ago
customer	a month ago
customer	a month ago
customer	a month ago

→ View all recently accessed

Links

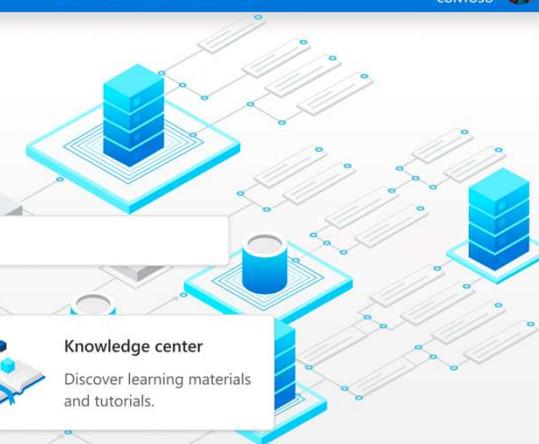
Purview overview Discover the capabilities offered by Purview and learn how to make the most of them.

Getting started Learn about how to use Purview Studio.

Documentation Visit the documentation center for quickstarts, how-to guides, and references, etc.

Purview status Learn about the Purview status.

Region availability





Daisy
Data
scientist/Purview
reader

Her search results include
logical assets, so she sees a
data domain for order

Microsoft Azure | Purview > Contoso

Order

Search results for Order

Source type : All Instance : All Clear all filters

Showing 1-25 out of 3,263 results

Sort by: Relevance

order Azure Data Lake Storage Gen2 File System https://march2020demog2.dfs.core.windows.net/order (0) ★★★★★ 26 views Sarah Jones updated 10 days ago

order Azure Data Lake Storage Gen1 Path adl://tralidsdev.azuredatalakestore.net/VidhyaTest/Order/Order 4.1 (23) ★★★★★ 32 views Sarah Jones updated 10 days ago

CpcOrderIncidentsReport Azure Data Explorer Table https://skypebusinessvoice.kusto.windows.net/BVLogs/CpcOrderIncidentsReport (0) ★★★★★ 12 views Sarah Jones updated 10 days ago

order_cleaned Confidential Azure Machine Learning Dataset This is the primary customer dataset for project keto. Columns: OrderCategoryID, OrderID, OrderModelID, OrderNumber mssql://keto-server.database.windows.net/KETO/datasets/order_cleaned 5.0 (2) ★★★★★ 3 views Sarah Jones updated 10 days ago

Order Data domain A document from a customer requesting products or services to be supplied. Order include an order ID, product ID, location data... (0) ★★★★★ 23 views Sarah Jones updated 10 days ago

OrderShipments General Azure Blob Storage Columns: Orders, OrderID Fully qualified name: adl://tralidsdev.azuredatalakestore.net/Watermarks/Devices (0) ★★★★★ 61 views Sarah Jones updated 10 days ago

Contoso / Sales / Order Glossary term From a finance perspective, a customer is an individual account holder who has purchased 1 or more products from Contoso. Approved Sarah Jones updated 10 days ago

order Certified

< Previous Page 1 of 199 Next >

The screenshot shows the Microsoft Purview Data Catalog interface. On the left, there's a sidebar with a user profile picture of a woman named Daisy and her title 'Data scientist/Purview reader'. Below that, there are sections for 'Data' (Dashboard, Report, Data pipeline, Stored procedure, Table, Folder, File), 'Business' (Glossary term, Data domain, Business process, System), 'Collection' (Contoso, Americas, Finance, Revenue, Sales, FY21), 'Classification' (Email Address, Country/Region, Canada Passport Number, Contoso Product Number, Canada Social Insurance ..., Date of Birth, Catalog_Europe, Catalog_NA, DWConfiguration, DWQueue), and 'See all'.

The main area displays a search results page for 'Order'. It shows 25 results out of 3,263. The results include:

- order**: Azure Data Lake Storage Gen2 File System. Description: https://march2020demog2.dfs.core.windows.net/order. Rating: 0 stars, 26 views, updated 10 days ago.
- order**: Azure Data Lake Storage Gen1 Path. Description: adl://tralidsdev.azuredatalakestore.net/VidhyaTest/Order/Order. Rating: 4.1 (23) stars, 32 views, updated 10 days ago.
- CpcOrderIncidentsReport**: Azure Data Explorer Table. Description: https://skypebusinessvoice.kusto.windows.net/BVLogs/CpcOrderIncidentsReport. Rating: 0 stars, 12 views, updated 10 days ago.
- order_cleaned**: Confidential Azure Machine Learning Dataset. Description: This is the primary customer dataset for project keto. Columns: OrderCategoryID, OrderID, OrderModelID, OrderNumber. Description: mssql://keto-server.database.windows.net/KETO/datasets/order_cleaned. Rating: 5.0 (2) stars, 3 views, updated 10 days ago.
- Order**: Data domain. Description: A document from a customer requesting products or services to be supplied. Order include an order ID, product ID, location data... Rating: 0 stars, 23 views, updated 10 days ago.
- OrderShipments**: General Azure Blob Storage. Description: Columns: Orders, OrderID. Fully qualified name: adl://tralidsdev.azuredatalakestore.net/Watermarks/Devices. Rating: 0 stars, 61 views, updated 10 days ago.
- Contoso / Sales / Order**: Glossary term. Description: From a finance perspective, a customer is an individual account holder who has purchased 1 or more products from Contoso. Approved. Rating: 0 stars, updated 10 days ago.
- order**: Certified. Rating: 0 stars.

At the bottom, there are navigation links for 'Previous', 'Page 1 of 199', and 'Next'.



Daisy

Data scientist/Purview reader

When she drills into it, she learns more about this concept.

- She can see the language and terms used to describe it.
- She can see the specific database tables/entities that implement this information.
- She can also see a list of systems that contain this data.
- Finally, she can see how order data is used in business processes.

Microsoft Azure | Purview > Contoso

Search assets

Order Data Domain

1 of 100 ↑ ↓

Updated on January 14, 2021, 4:36 PM by Carrier Hoffman

Collection

- Contoso Purview Collection
- Sales Purview Collection
- Order Data Domain

Description

A document from a customer requesting products or services to be supplied. Orders include an order ID, product ID, location data, activity codes.

Includes business terms

Definition
Order details
Order ID
Order type
Order activity
Product ID
Quantity product sold

Implemented by tables

- Order
- OrderActivity
- OrderLocation

Related systems

- CRM Cloud
- Contoso Billing
- General Ledger

Platform

- SAP Commerce Cloud
- Oracle NetSuite
- Oracle NetSuite

Related processes (4)

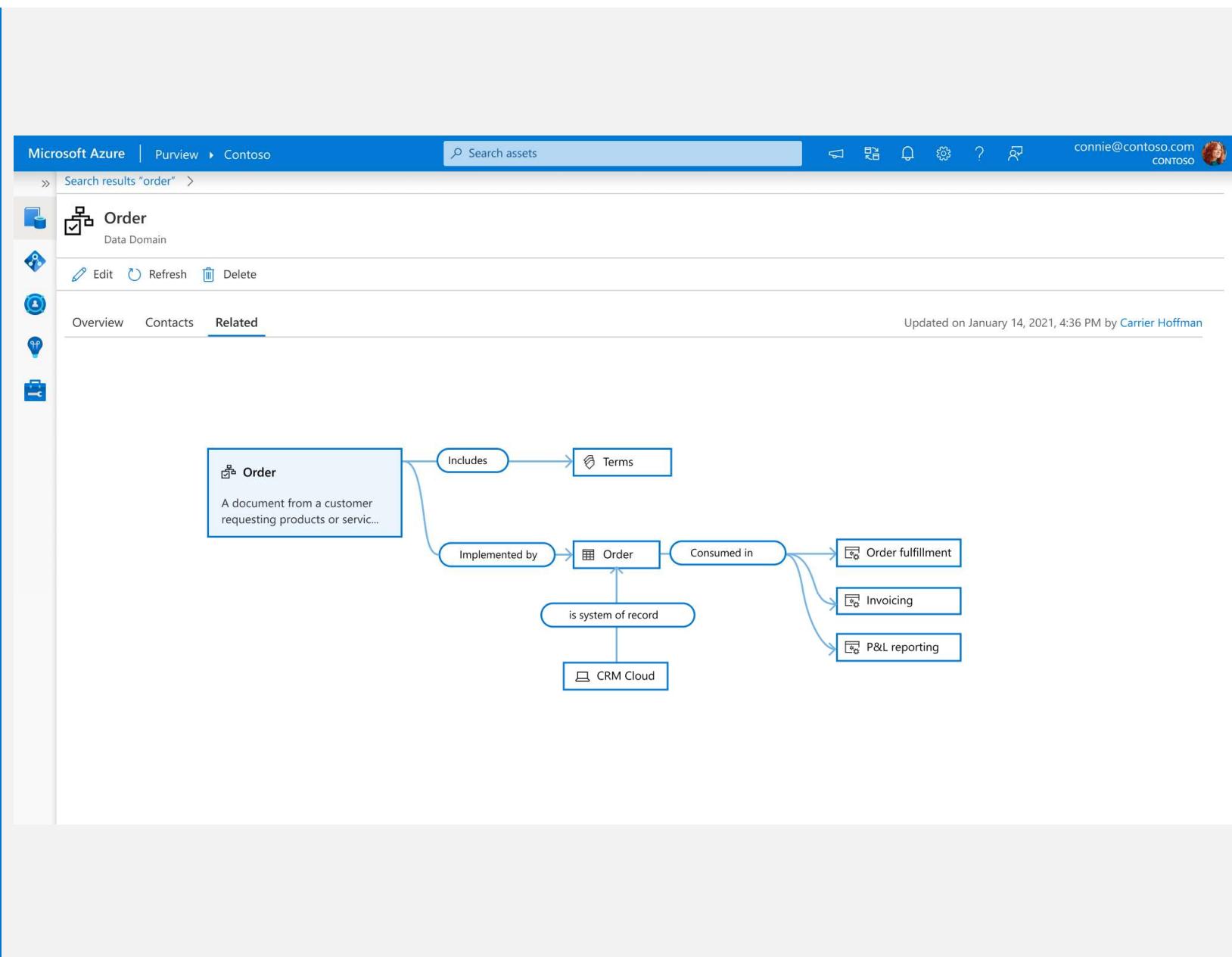
- Order fulfillment
- Invoicing
- Profit and loss reporting



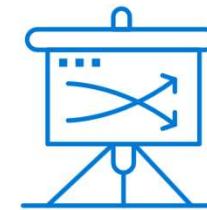
Daisy

Data scientist/Purview reader

When she selects the related tab, she can visualize the relationships between these assets in a graph.



Five approaches to scale



Bulk import & export

Users can download assets, curate in a csv spreadsheet and load back into the catalog.

Industry accelerators

Industry and compliance-specific controlled vocabularies.

Examples:
-Synapse database templates & Common Data models

Workflow/rules engines

Like Purview's Business rules.

Connectors and integrations

Examples:
-Connector for ServiceNow CMDB

Distribute the effort

Crowd-source knowledge with conversations and tags. Suggest stewards based on usage data.

Microsoft Purview Data Governance in Azure overview



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 Pettry. Mar, 2019

175 ZB

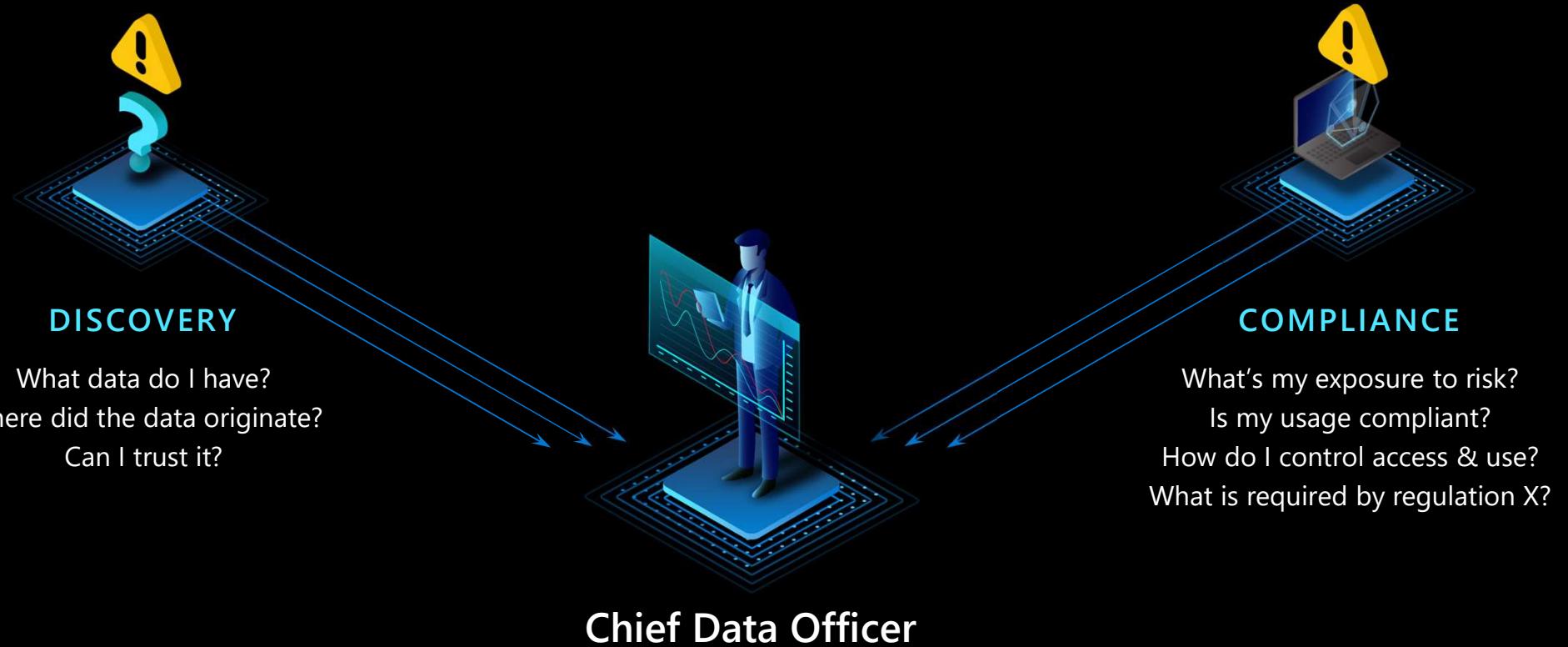
Expected global volume
of data generated
annually by 2025

IDC

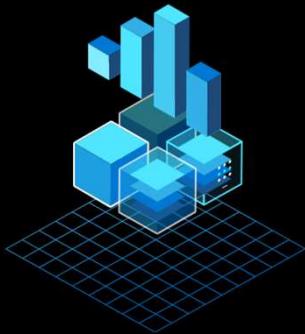
IDC Data Age 2025, Dave Reinsel,



Data governance is becoming increasingly interdisciplinary



Elements of **successful** data governance



Manage growing
data landscape



Overcome
operational silos



Increase
data agility



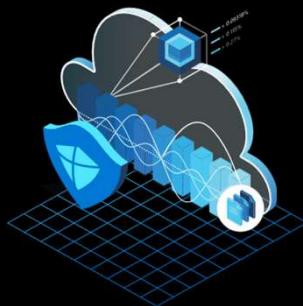
Comply with
industry regulations

Reimagine data governance with **Microsoft Purview**

Unified | Hybrid | Open



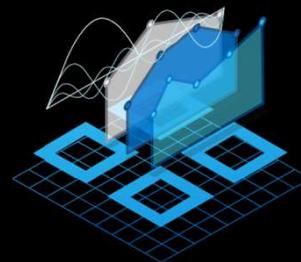
Microsoft Purview enables unified data governance



Reimagine data governance in the cloud



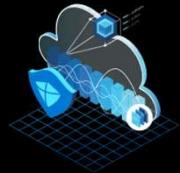
Set the foundation for effective data governance



Maximize business value of data for data consumers



Gain strategic insight into data use across the estate



Reimagine data governance in the cloud

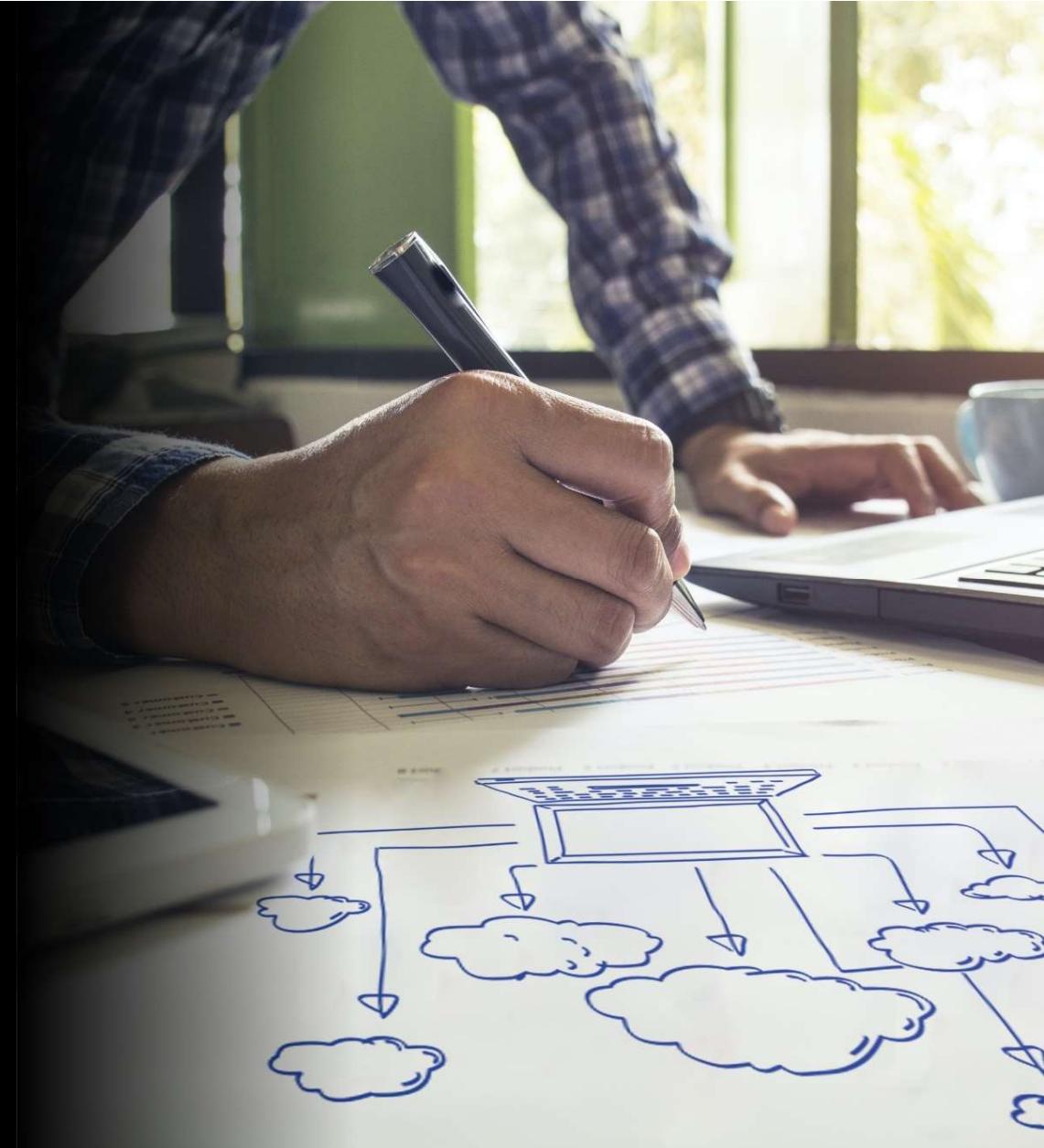
- Manage and govern operational, transactional and analytical data
- Cloud-native, purpose-built service to address discovery and compliance needs
- Fully managed, serverless, PaaS service
- Eliminate manual, ad-hoc and homegrown solutions





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





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





Gain insight into data use across the estate

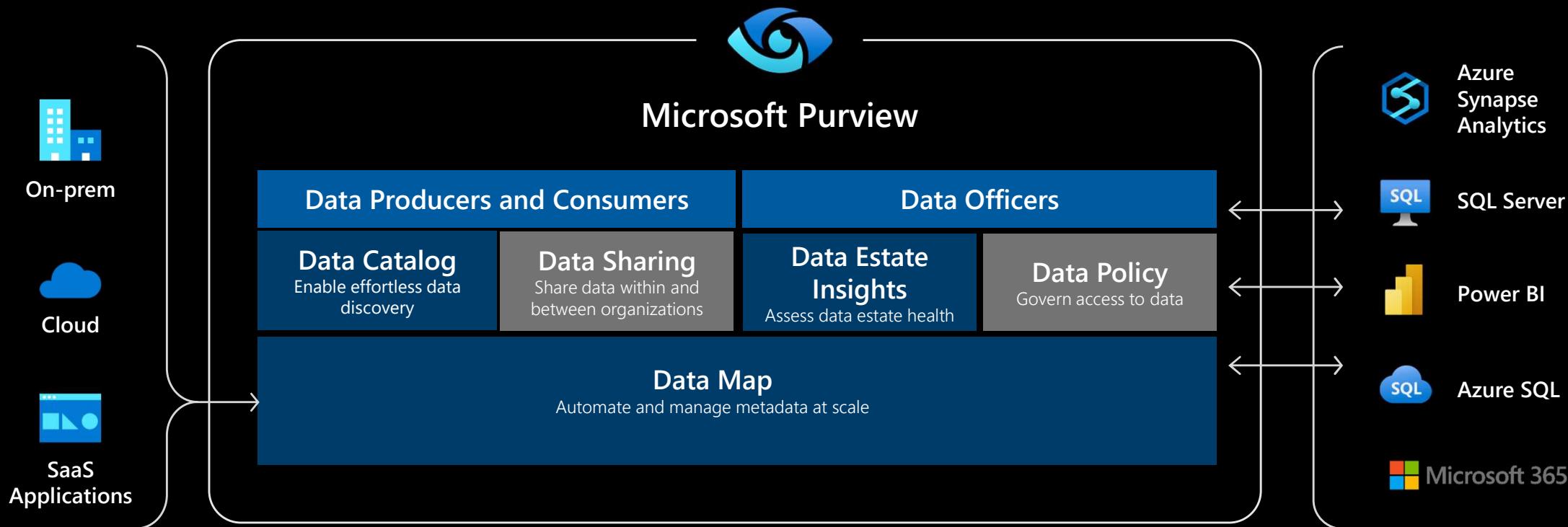
- Understand at a glance how data is being created and used across your data estate
- Visually assess the state of data assets, scans, business glossary and sensitive data



Generally Available

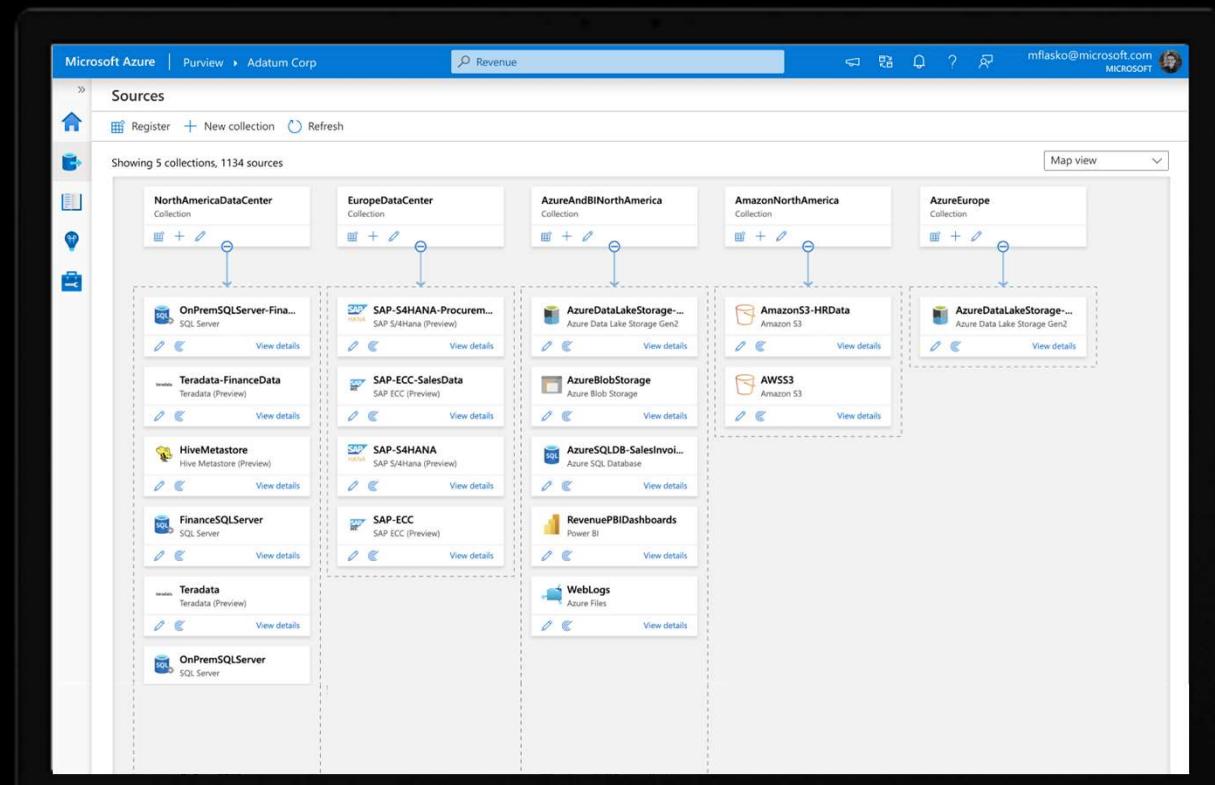
Preview

Unified Data Governance with Microsoft Purview



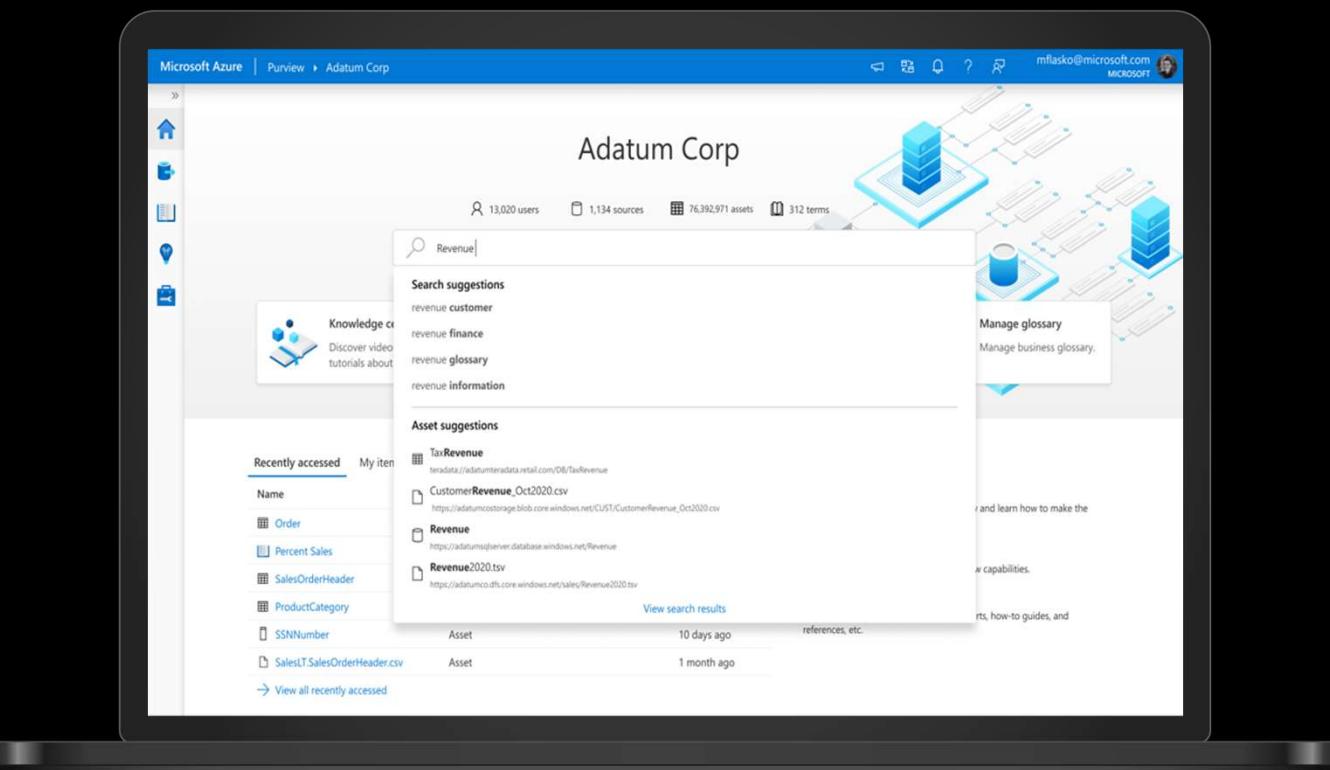
Unify and make data meaningful using Purview Data Map

- 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



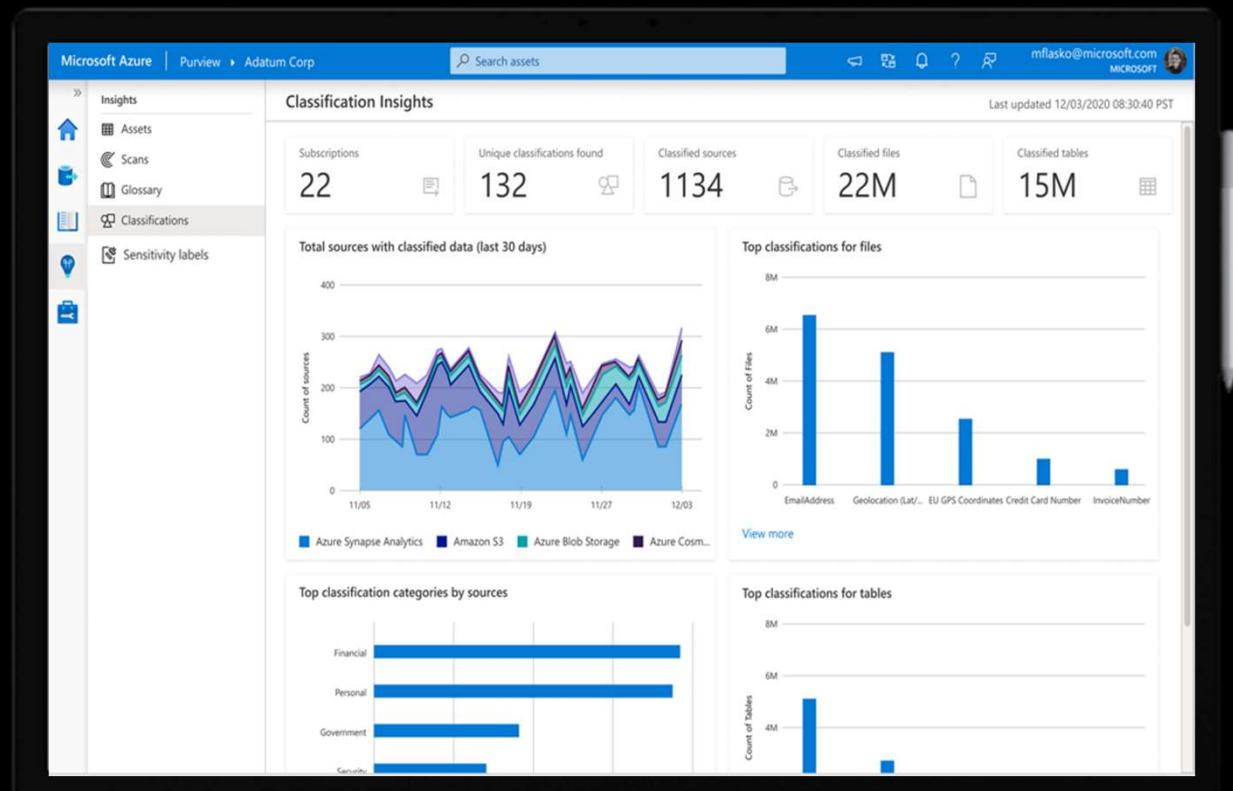
Enable effortless discovery of trusted data with Purview Data Catalog

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

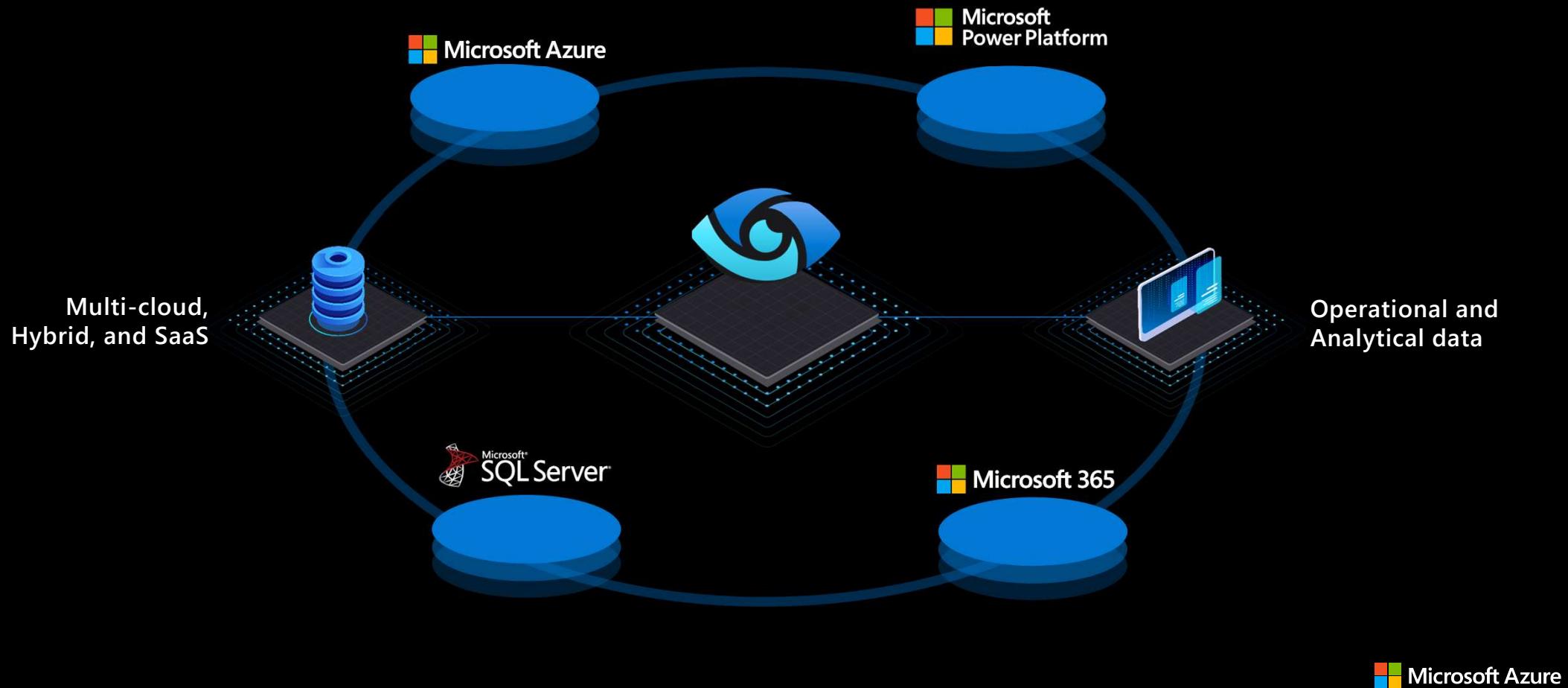


Get a bird's-eye view of your data landscape with Purview Data Insights

- Data Asset distribution
- Sensitive Data
- Data Scanning Coverage
- Business Glossary Utilization

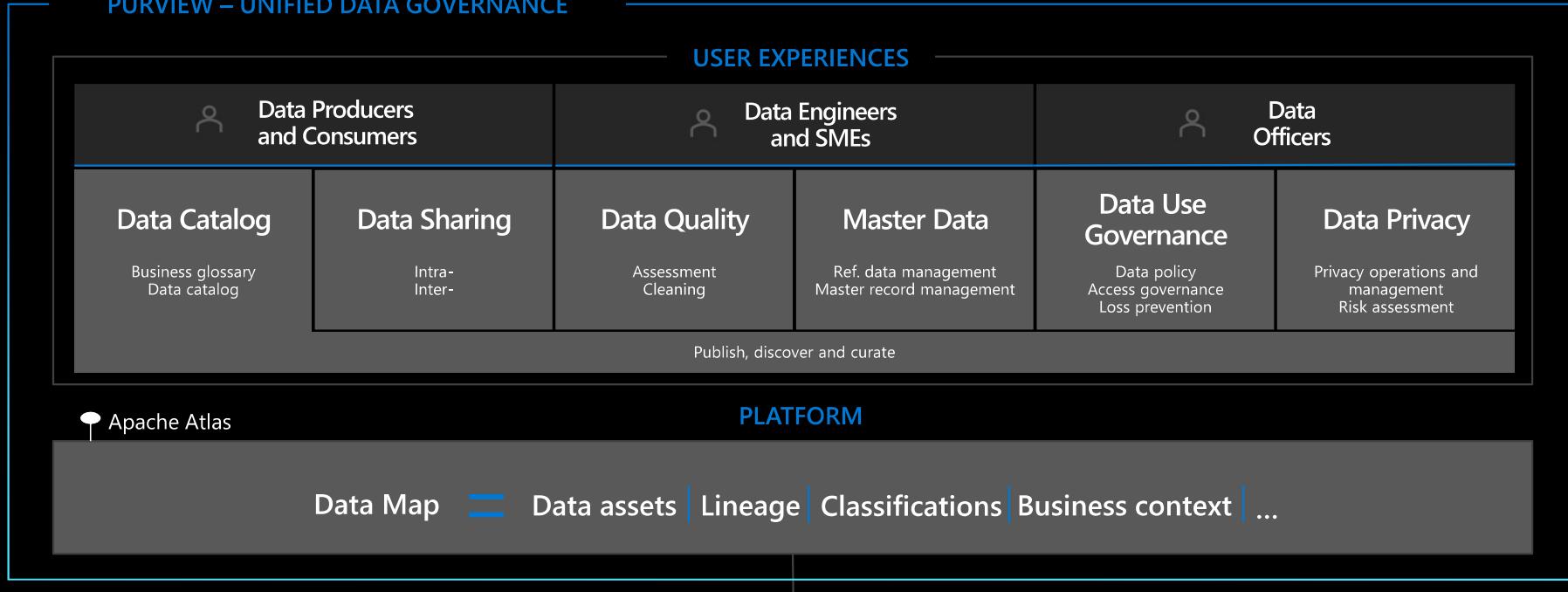


Make the most of your Microsoft investments





PURVIEW – UNIFIED DATA GOVERNANCE



On-prem, multi-cloud, SaaS

Operational & Analytical

Synapse

Power BI

AML

Sentinel

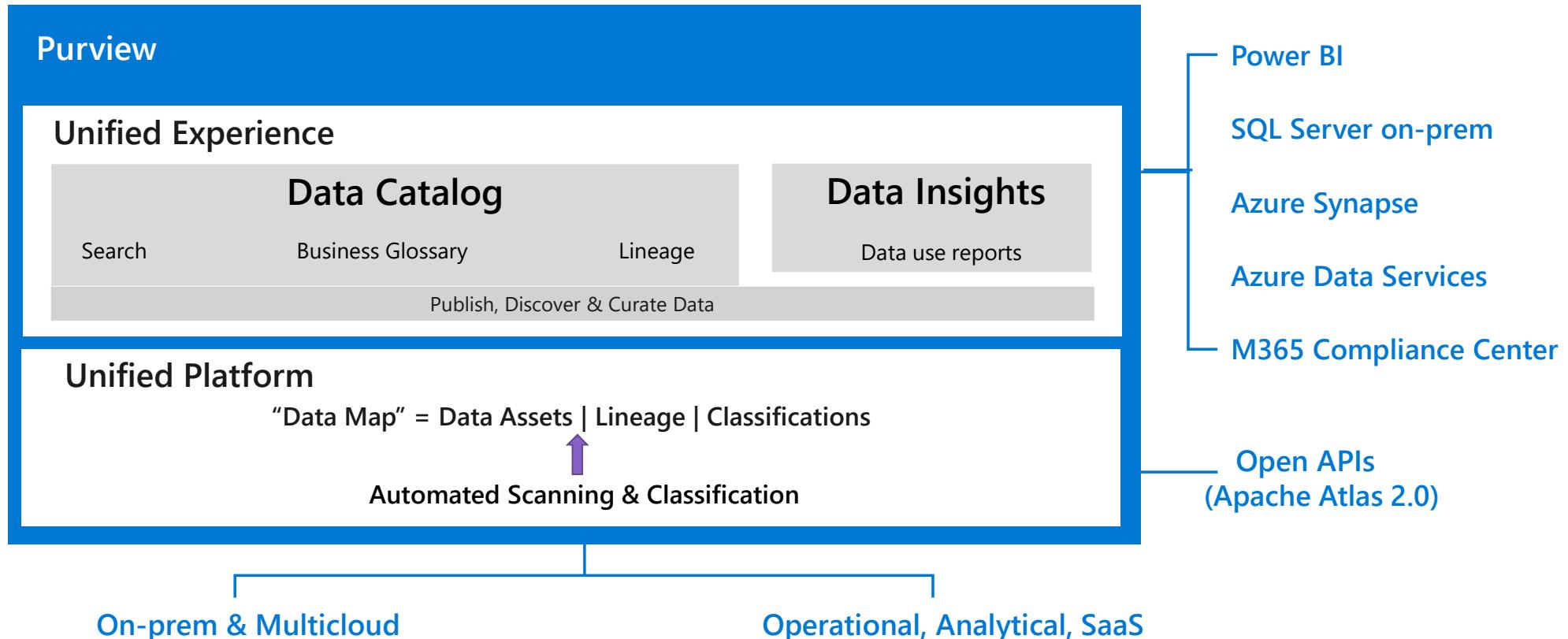
Microsoft 365

and more...

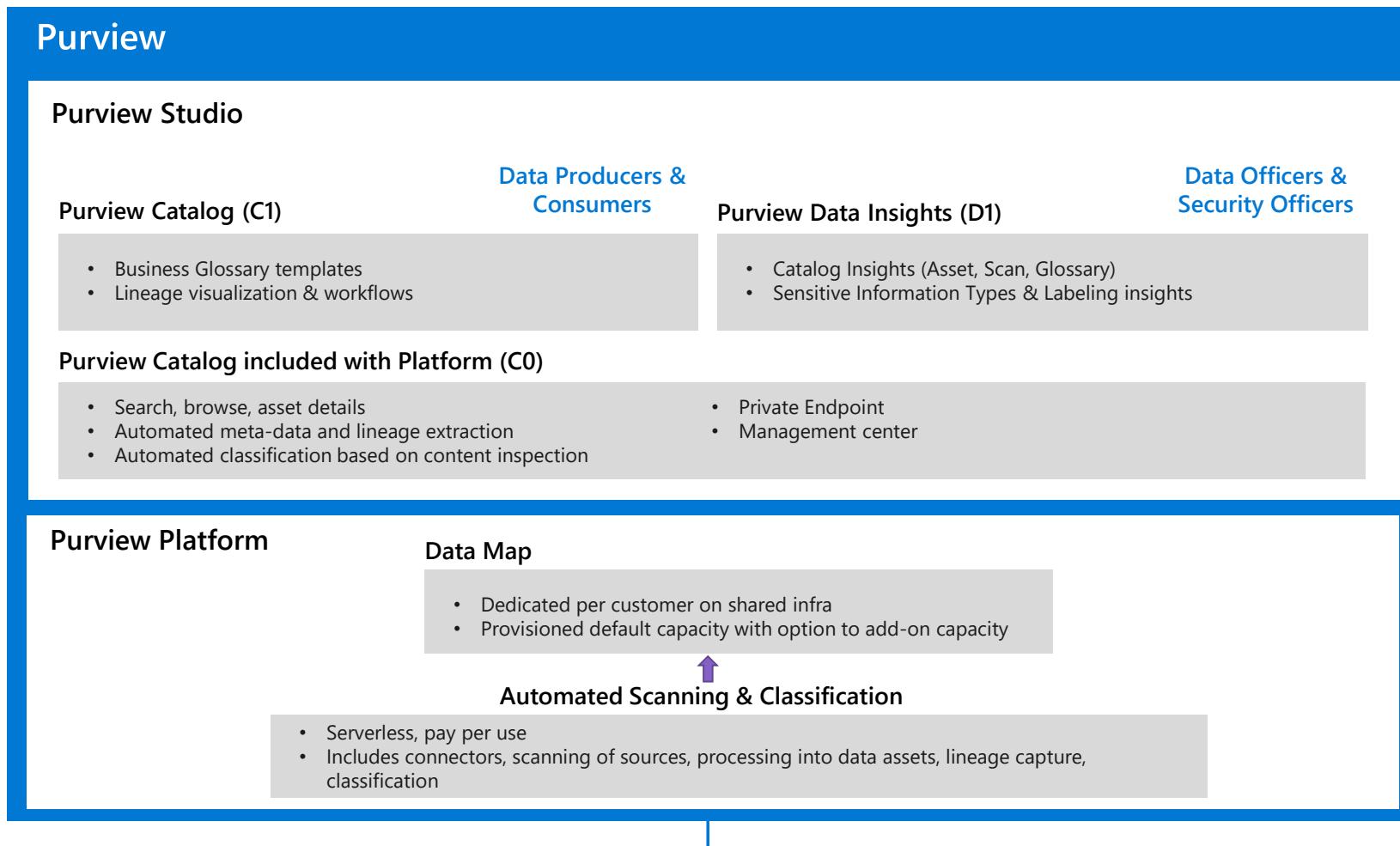
Microsoft Purview Data Governance in Azure Features

	Generally Available	In Preview
Purview Data Map		
Automated scanning of hybrid sources	●	
Multi Cloud Scanning for AWS S3	●	
Data Classification	●	
Apache Atlas API support	●	
Purview Data Catalog		
Search and Browse	●	
Business Glossary	●	
Data Lineage	●	
Purview Data Insights		
Assets and Scans Reports		●
Glossary reports		●
Classification and Labelling Reports		●
Asset-level drill down by sensitivity		●
Purview Integrations		
Purview in Azure Synapse workspaces	●	
Power BI Integration	●	
SQL Server Integration	●	
Azure SQL Integration	●	
Microsoft Purview Information Protection Sensitivity Labels support		●

Microsoft Purview: Unified Data Governance



Microsoft Purview Features at Public Preview



* Power BI, SQL Server on-prem, Azure Data Services including Synapse, Cosmos DB & Storage, Non-Microsoft systems including SAP ECC, SAP S4 HANA & Teradata, Multi-cloud systems including AWS S3



Microsoft Purview Studio (Unified Experience)

Purview Studio

A single, centralized place that provides unified experience for data producers, data consumers, data & security officers

The screenshot shows the Purview Studio interface for the organization "Adatum Corp" (adatumcorpretail). The top navigation bar includes Microsoft Azure, Purview, and Adatum Corp, along with user information for nayenama@microsoft.com and a Microsoft profile icon.

The main dashboard displays the following statistics:

- 33 sources
- 10,234,314 assets
- 114 glossary terms

Below the stats, there is a search catalog bar and three main navigation cards:

- Browse assets**: Explore assets by source type and collection.
- Manage glossary**: Create and edit terms with custom term templates.
- Knowledge center**: Discover learning materials and tutorials.

The central area features a 3D network diagram illustrating the interconnectedness of data assets, represented as blue cubes and cylinders connected by lines.

On the left, a sidebar lists recently accessed items and my items, showing the last update time for each. A link to view all recently accessed items is also provided.

On the right, there is a section titled "Links" with links to Purview overview, Getting started, and Documentation.

Name	Last update
starter3	12 days ago
customer.docx	12 days ago
demoreportingblob	12 days ago
democatalogadcadls	12 days ago
democatalogadcadls	12 days ago
Contoso Dev DebtServiceCoverageRatio	12 days ago

Purview Studio

Purview Studio divided into **Activity hubs**.

These organize the tasks needed for managing data governance activities

The screenshot shows the Microsoft Purview Studio interface. At the top, there's a navigation bar with 'Microsoft Azure', 'Purview', and 'Adatum Corp'. On the right side of the header is a user profile for 'nayenama@microsoft.com' and the 'MICROSOFT' logo. Below the header, there are five main sections, each representing an 'Activity hub':

- Home:** Represented by a house icon. Description: 'Quick actions, recently accessed items, owned items, search bar and documentations.'
- Sources:** Represented by a cylinder with an arrow icon. Description: 'Create collections, register data sources and set up scans.' Below it is an icon showing a monitor, a cylinder, and a wrench.
- Glossary:** Represented by a book icon. Description: 'Manage glossary terms, search glossary terms, manage term templates and custom attributes, Import and export terms using .CSV.'
- Insights:** Represented by a lightbulb icon. Description: 'Get Insights on your data.'
- Manage:** Represented by a gear icon. Description: 'Metadata management - classifications, resource sets, data sources, Integration run time, Alerts, Security, data factories and data share connections'

A red box highlights the 'Sources' hub icon. A red arrow points from the left edge of the slide towards the 'Sources' hub icon.

Purview Studio

Ease of access to get updates, to switch accounts, to get notifications and to provide feedback

The screenshot shows the Microsoft Azure Purview Studio interface for the Adatum Corp account. The top navigation bar includes 'Microsoft Azure', 'Purview', 'Adatum Corp', and a user profile for 'nayenama@microsoft.com'.

Key features highlighted with red boxes:

- Release Notes**: Located in the main content area below the account name.
- Switch Accounts**: Located in the main content area below the account name.
- Notifications**: Located in the main content area below the account name.
- Feedback**: Located in the main content area below the account name.
- Search Catalog**: A search bar at the top of the main content area.
- Recently accessed** and **My items**: Navigation tabs at the top of the main content area.
- Links**: A section on the right side containing links to 'Purview overview', 'Getting started', and 'Documentation'.

Below the main content area, there are three cards:

- Browse assets**: Explore assets by source type and collection.
- Manage glossary**: Create and edit terms with custom term templates.
- Knowledge center**: Discover learning materials and tutorials.

At the bottom left, there is a link to 'View all recently accessed'.

Microsoft Purview Activities

Purview Icons are enabled or disabled based on the role assigned to the user in Purview. All roles are assigned by Purview account collection admin in Purview Governance Portal

Appropriate Role	Annotate Glossary, Classification, and rules	Collections Access control, Lineage, connections	Data catalog Search, Browse, and Asset details	Data Estate Insights All reports	Data map Sources, Credentials, IR	Policies Data policies	Scan rule set Scan rule set	Shares Data sharing	Workflows Access to workflow authoring information
 Collection admin CollectionRead,CollectionWrite	✗	✓	✗	✗	✗	✗	✗	✗	✗
 Data curator Data Read, Data Write, Collection Read, Insight Read	✓	_READONLY	✓	_READONLY	✗	_READONLY	✗	✗	✗
 Data reader Data Read, Collection Read	_READONLY	_READONLY	_READONLY	✗	✗	✗	✗	✗	✗
 Data share contributor Share Read, Share Write, Source Read, Source Write, Collection Read	✗	_READONLY	✗	✗	✓	✗	✗	✓	✗
 Data Source admin Scan Read, Scan Write, Source Read, Source Write, Collection Read	✗	_READONLY	✗	✗	✓	*_READONLY	✗	✗	✗
 Insights reader Insight Read	✗	✗	✗	_READONLY	✗	✗	✗	✗	✗
 Policy author Policy Read, Policy Write, Source Read, Collection Read, Data Read	✗	_READONLY	_READONLY	✗	_READONLY	✓	_READONLY	✗	✗
 Workflow admin Workflow Read, Workflow Write, Workflow Bind	✗	✗	✗	✗	✗	✗	✗	✗	✓
 Account creator All access	✓	✓	✓	✓	✓	✓	✓	✓	✓
 Data source admin + Data curator Scan Read, Scan Write, Data Read, Data Write, Collection Read	✓	_READONLY	✓	_READONLY	✓	✗	✓	✗	✗
 Data source admin + Data reader Scan Read, Scan Write, Data Read, Collection Read	_READONLY	_READONLY	_READONLY	✗	✓	✗	✓	✗	✗

* The creator of Purview account will automatically have full access in Purview

i.e., Collection Admin + Data Source Admin + Data Curator

* Just having 'Data Source Admin' roles does not provide access to Purview Studio.

This role should always be associated with Data Curator or Data Reader



Purview Studio

Data Catalog

Data Catalog Page

All Data Catalog related activities

The screenshot shows the Microsoft Purview Data Catalog Page for the organization 'Adatum Corp' (adatumcorpretail). The page includes a top navigation bar with Microsoft Azure, Purview, and Adatum Corp, along with user information and a search bar. Below the navigation is a summary section with metrics: 33 sources, 10,234,314 assets, and 114 glossary terms. A red box highlights the 'Search catalog' input field. To the right is a 3D data visualization showing a stack of blue cylinders and cubes. Red arrows point from text labels to specific elements: 'Catalog metrics' points to the summary stats, 'Search bar to browse and search data catalog' points to the search input, and 'Data Catalog tiles for key activities' points to three cards: 'Browse assets', 'Manage glossary', and 'Knowledge center'. Another red box highlights the 'Recently accessed' and 'My items' tabs at the top of a list of recently accessed entities. A red arrow points from text to this list with the label 'Recently accessed entities and list of entities owned by logged-on data user.' On the right side, a 'Links' section contains three items: 'Purview overview', 'Getting started', and 'Documentation', each with a brief description. A red arrow points from text to this section with the label 'Links for Overview, getting started and purview documentation'.

Recently accessed entities and list of entities owned by logged-on data user.

Name	Last update
DimOrganization	3 hours ago
starter	12 days ago
customer.docx	12 days ago
demoreportingblob	12 days ago
democatalogadcadls	12 days ago
democatalogadcadls	12 days ago

→ View all recently accessed

Catalog metrics

33 sources | 10,234,314 assets | 114 glossary terms

Search bar to browse and search data catalog

Data Catalog tiles for key activities

Recently accessed **My items**

Links

- Purview overview
- Getting started
- Documentation

Links for Overview, getting started and purview documentation



Data Catalog **Browse & Search**

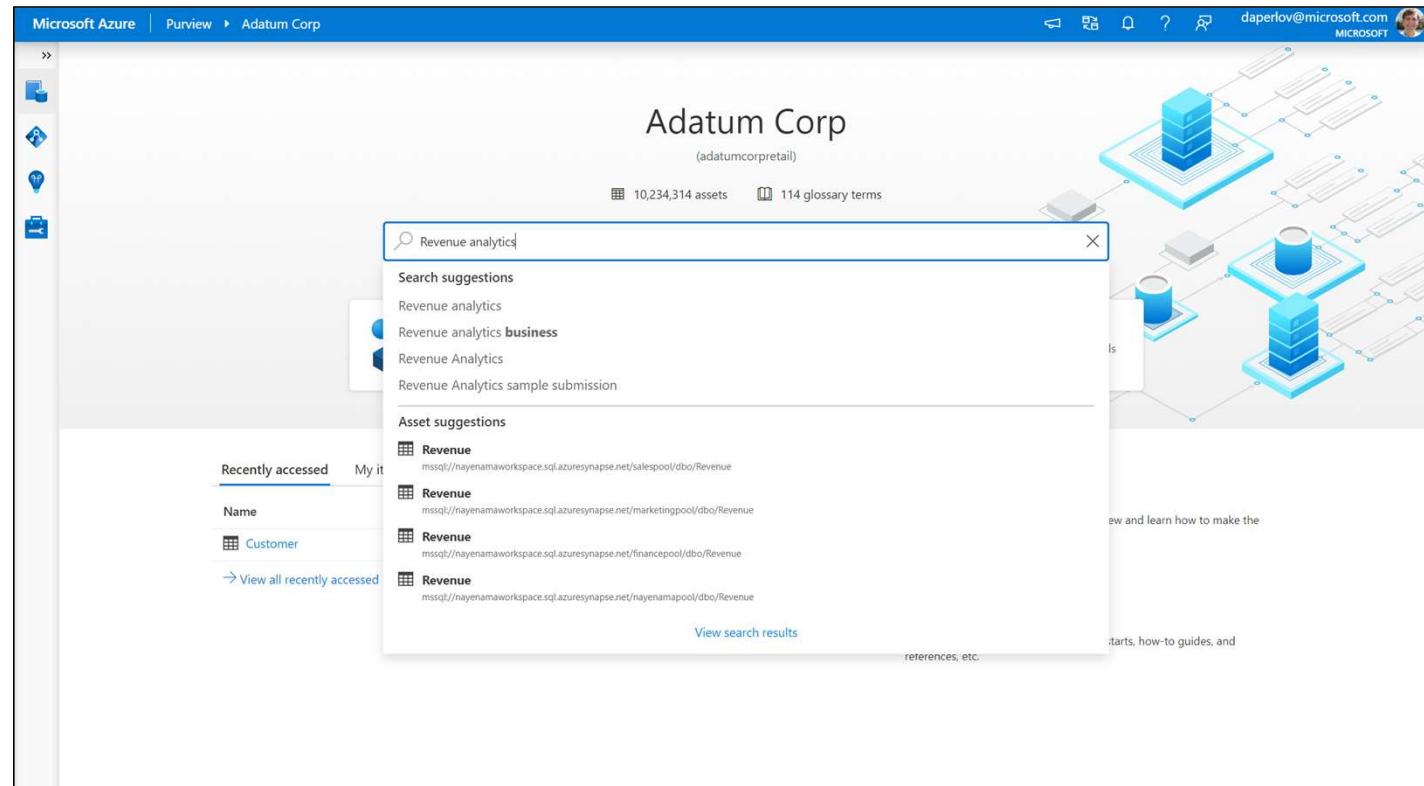
(Effortless discovery of trusted & accurate data)

Browse & Search

Leverage Purview's search relevance engine to quickly discover useful data assets through browse and search

Benefits

- Provide intelligent recommendations based on data relationships, business context, search history
- Easy data exploration by browsing by collection and source type
- Relevant search results without writing complex queries or applying advanced filters
- Leverage facets to find the exact data you are looking for



Browse & Search

Top search results include curated and relevant assets that match to a keyword

The screenshot shows the Microsoft Purview Data Catalog interface with the search term "customer north america" entered in the search bar. The results page displays 14702 results, with the first few items being "Customer" assets from various sources:

- Customer** (Azure SQL Table): mssql://demosqlserveroct2020.database.windows.net/demoOct2020/SalesLT/**Customer**
- Customer** (Amazon RDS Table (SQL)): mssql://demodatabase.ccjlzhcphjg.us-east-2.rds.amazonaws.com/AdventureWorks/SalesLT/**Customer**
- Customer** (Azure Blob Resource Set): https://sep2019demostoragev2.blob.core.windows.net/customerdata/**Customer**[N].csv
- Customer Profitability Sample** (Power BI Dataset): https://app.powerbi.com/groups/cced5bbf-f465-4785-b910-a6dbd1a0a86c/datasets/26e624e8-f186-4e9d-9303-b67449af0d3e
- Customer Profitability Sample** (Power BI Dataset): https://app.powerbi.com/groups/15c3ff9f-b998-494b-9467-4dbf7e175198/datasets/f70861f2-be96-4c55-9f8b-7059d91f6ff1
- Customer Profitability Sample** (Power BI Dashboard): https://app.powerbi.com/groups/15c3ff9f-b998-494b-9467-4dbf7e175198/dashboards/254bbd8e-f48e-41cf-97af-b330e0177a0c

The left sidebar includes filters for Source types (all), Instances (all), and a "Filter by keyword" input field. Other filter categories like Collection, Classification, Contact, and Label are also present.

Browse & Search

Use facets and filters based on asset metadata to narrow down your search results

The screenshot shows the Microsoft Purview Data Catalog interface. At the top, the navigation bar includes 'Microsoft Azure | Purview > Adatum Corp'. A search bar contains the query 'customer north america'. The main area displays 'Search results for customer north america' with 12 results found. The results list includes:

- Customer**: Azure Blob Resource Set, https://sep2019demostoragev2.blob.core.windows.net/customerdatalcustomer/Customer[N].csv
- DimCustomer**: Azure Synapse Dedicated SQL Table, mssql://nayenamaworkspace.sql.azuresynapse.net/salespool/dbo/DimCustomer
- DimCustomer**: Azure Synapse Dedicated SQL Table, mssql://nayenamaworkspace.sql.azuresynapse.net/marketingpool/dbo/DimCustomer
- DimCustomer**: Azure Synapse Dedicated SQL Table, mssql://nayenamaworkspace.sql.azuresynapse.net/adatumcorppool/dbo/DimCustomer
- DimCustomer**: Azure Synapse Dedicated SQL Table, mssql://nayenamaworkspace.sql.azuresynapse.net/financepool/dbo/DimCustomer
- DimCustomer**: Azure Synapse Dedicated SQL Table, mssql://nayenamaworkspace.sql.azuresynapse.net/nayenamapool/dbo/DimCustomer

The left sidebar features a navigation tree under 'Data catalog >' and several filter categories, each with a red border:

- Collection**:
 - AzureAndBINorthAmerica (14688)
 - Adatum Corp (12)
 - AmazonNorthAmerica (2)
- Classification**:
 - Email Address (5)
 - Person's Name (5)
 - U.S. Phone Number (1)
- Contact**: ...
- Label**: ...
- Glossary term**: ...

At the bottom, there are navigation links: '< Previous', 'Page 1 of 1', and 'Next >'. The top right corner shows the user's email (daperlov@microsoft.com) and profile picture.

Browse & Search

Browse by collection to find assets in a specific business unit

Microsoft Azure | Purview > Adatum Corp

Search assets

Data catalog >

Browse assets

Refresh

By collection By source type

Filter by keyword

Showing 9 collections

Name	Description	Assets	Collection admin
Adatum Corp	The root collection.	276	Naga Krishna Yenamandra, +5 more
AmericasDataCenters		-	Naga Krishna Yenamandra
AmazonNorthAmerica		19	Naga Krishna Yenamandra
AzureAndBINorthAmerica		10,234,019	Naga Krishna Yenamandra
GoogleNorthAmerica		-	Naga Krishna Yenamandra
NorthAmericaDataCenter		-	Naga Krishna Yenamandra
EuropeDataCenters		-	Naga Krishna Yenamandra
AzureEurope		-	Naga Krishna Yenamandra
NorthEuropeDataCenter		-	Naga Krishna Yenamandra

Browse & Search

Browse by collection to find assets in a specific business unit

The screenshot shows the Microsoft Purview Data Catalog interface, specifically the 'Browse assets' section for the 'Adatum Corp' collection. The left sidebar includes navigation links like 'Data catalog', 'Browse assets', 'Refresh', and filters for 'By collection' and 'By source type'. A 'View collection tree' link is also present. The main content area displays a list of assets under the 'AmazonNorthAmerica' collection, which is itself a sub-collection of 'AmericasDataCenters' under 'Adatum Corp'. The assets listed are:

- Address**: Amazon RDS Table (SQL) - mssql://demodatabase.ccjlzhcgphjg.us-east-2.rds.amazonaws.com/AdventureWorks/SalesLT/Address
- AdventureWorks**: Amazon RDS Database (SQL) - mssql://demodatabase.ccjlzhcgphjg.us-east-2.rds.amazonaws.com/AdventureWorks
- BuildVersion**: Amazon RDS Table (SQL) - mssql://demodatabase.ccjlzhcgphjg.us-east-2.rds.amazonaws.com/AdventureWorks/dbo/BuildVersion
- Customer**: Amazon RDS Table (SQL) - mssql://demodatabase.ccjlzhcgphjg.us-east-2.rds.amazonaws.com/AdventureWorks/SalesLT/Customer
- CustomerAddress**: Amazon RDS Table (SQL) - mssql://demodatabase.ccjlzhcgphjg.us-east-2.rds.amazonaws.com/AdventureWorks/SalesLT/CustomerAddress
- dbo**: Amazon RDS Table (SQL)

Below the asset list, there are navigation controls for 'Previous', 'Page 1 of 1', and 'Next'.

Browse & Search

Browse the data catalog based on the source type of a data asset

The screenshot shows the Microsoft Purview Data Catalog interface, specifically the 'Browse by source type' section. The top navigation bar includes 'Microsoft Azure', 'Purview', 'Adatum Corp', a search bar labeled 'Search assets', and a user profile for 'daperlov@microsoft.com MICROSOFT'. On the left, there's a sidebar with icons for 'Data catalog', 'Browse assets', 'Refresh', and two filtering options: 'By collection' and 'By source type' (which is selected). The main area displays a grid of data asset categories:

Asset Type	Count
Amazon RDS (SQL)	1 item
Azure Blob Storage	3 items
Azure Data Factory	1 item
Azure Data Lake Storage Gen2	1 item
Azure Database for MySQL	1 item
Azure Database for PostgreSQL	1 item
Azure Dedicated SQL Pool (formerly SQL DW)	2 items
Azure SQL Database	4 items
Azure SQL Server	6 items
Azure Synapse Analytics	1 item
Power BI	11 items
Azure Storage Account	4 items

Asset Overview

Discover technical, semantic and business information about an asset including applied classifications, associated glossary terms, sensitivity labels, and endorsement labels

The screenshot shows the Asset Overview page for a Power BI Dataset named "ProductSalesReport". The dataset is marked as "Certified". The page includes navigation links for "Overview", "Properties", "Schema", "Lineage", "Contacts", and "Related". It displays the last update time as "Updated on July 27, 2021 10:02 PM UTC by automated scan".

Asset description: Dataset with sales and customer details.

Classifications (4): Canada Social Insurance Number, Franmer Product, Franmer SIN, Person's Name.

Schema classifications (1): Franmer Product.

Fully qualified name: <https://app.powerbi.com/groups/657d261d-247b-4234-9c2d-515b0bbdd328/datasets/0a0ca805-3934-4265-be58-3afc833d8ac5>

Hierarchy:

- Shared
Power BI Capacity
- SalesDepartment
Power BI Workspace
- ProductSalesReport
Power BI Dataset

Glossary terms (7):

- Customer Product
- Product client
- Global Markets
- Products
- Sales
- Saleswithtax
- SSN

Asset Schema

See an assets schema and the metadata associated with each column

The screenshot shows the Purview Data Catalog interface for the "Customers" asset. The asset is identified as an "Azure Synapse Dedicated SQL Table". The "Schema" tab is selected, showing the following columns and their properties:

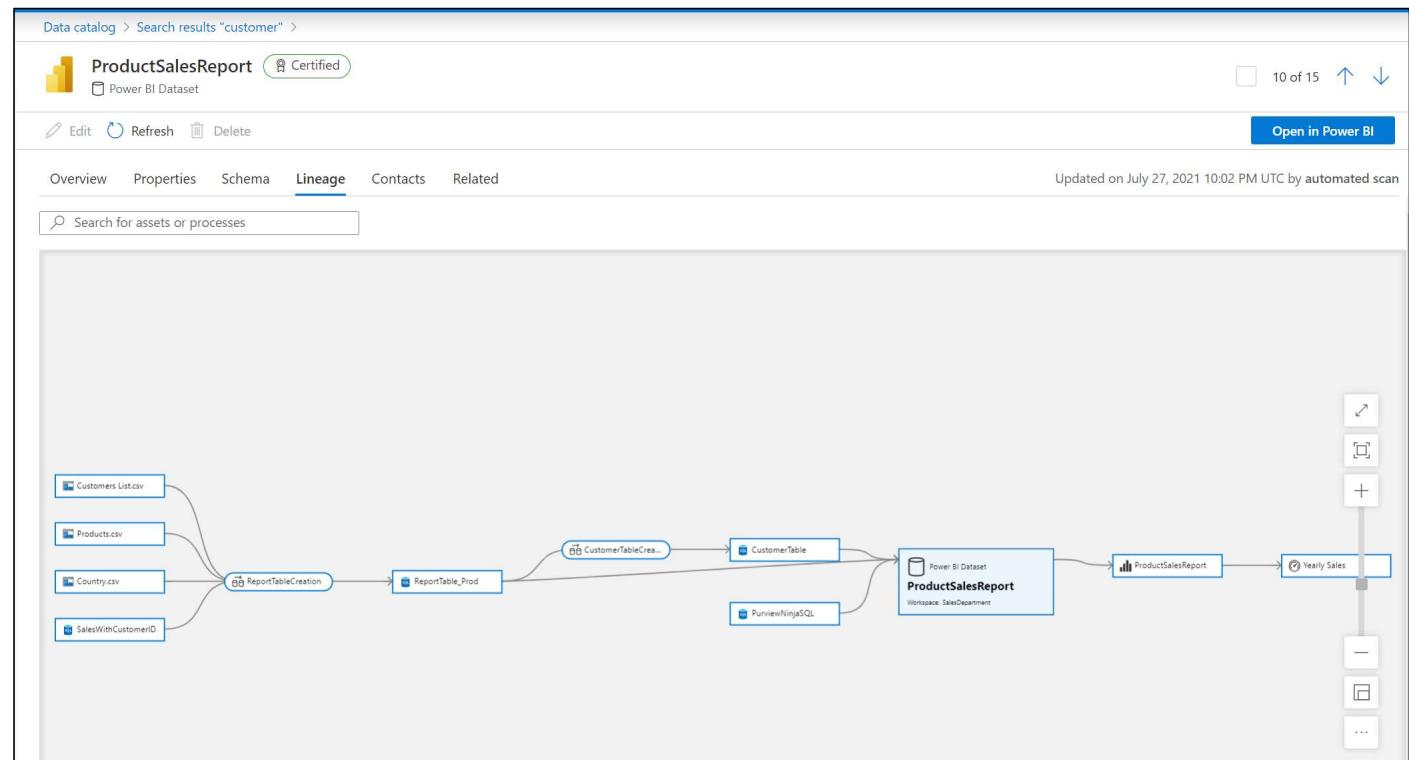
Column name	Classifications	Sensitivity label	Glossary terms	Data type	Asset description
AddressLine1				varchar	
AddressLine2				varchar	
City				varchar	
CreatedDate				date	
CustomerID				varchar	
FirstName	⚡ Person's Name			varchar	
LastName	⚡ Person's Name			varchar	
PhoneNumber	⚡ EU Phone Number ⚡ U.S. Phone Number			varchar	
State				char	
UpdatedDate				date	
ZipCode				varchar	

Asset Lineage

Trace lineage of data assets across the data estate

Benefits

- Ensure data provenance with a visual representation of owners, sources, transformation, and lifecycle
- Leverage support of Apache Atlas's open-source lineage APIs and built-in integrations with solutions such as Azure Data Factory, Azure Synapse Analytics, Azure Data Share and Power BI
- Analyze impact of changes to data and understand dependencies visually.
- Root cause analysis of failures by inspecting dependencies upstream and determine downstream impact



Asset Contacts

Identify experts and owners of the data asset

Data catalog > Search results "customer" >

 **ProductSalesReport** Certified
Power BI Dataset

[Edit](#) [Refresh](#) [Delete](#)

[Overview](#) [Properties](#) [Schema](#) [Lineage](#) [Contacts](#) [Related](#)

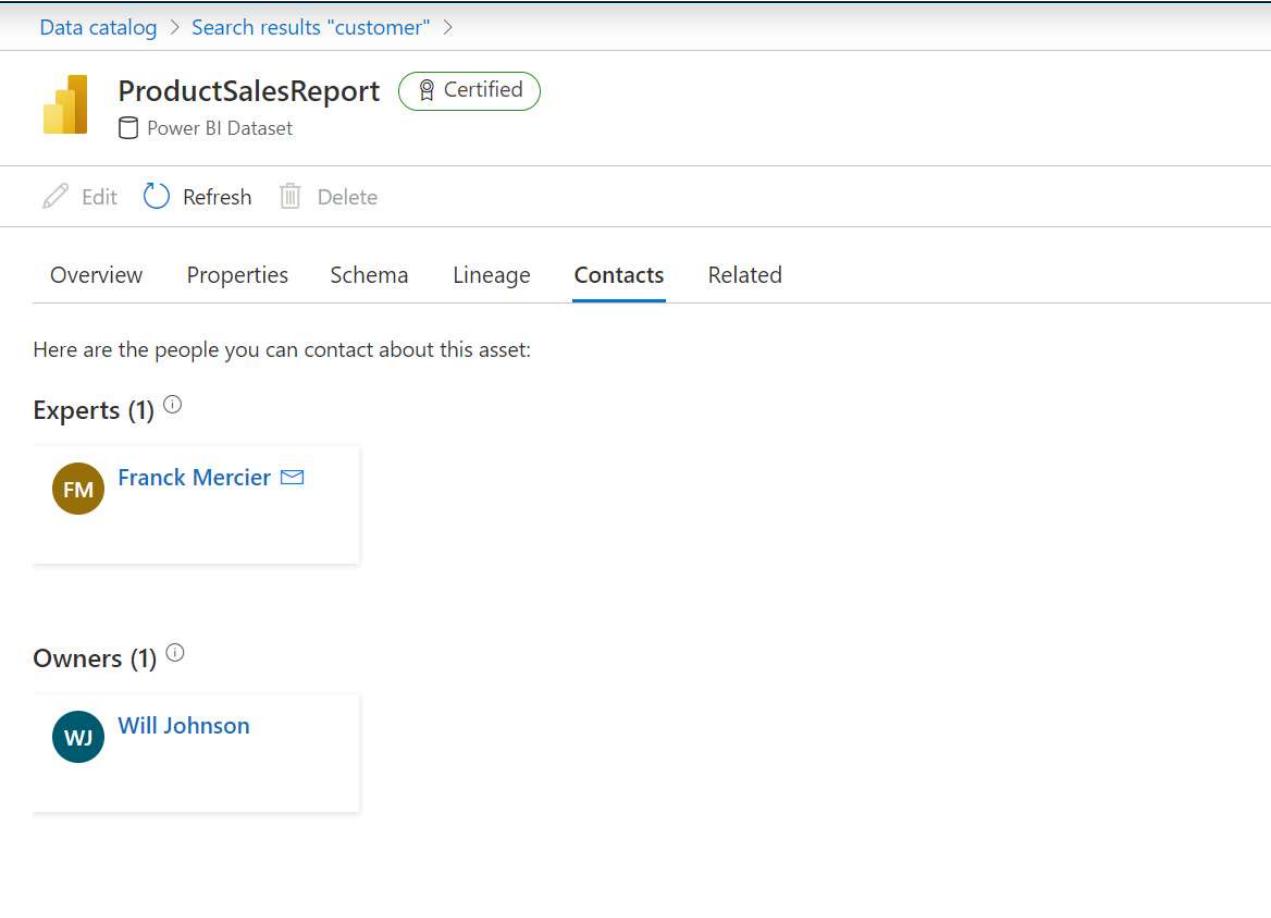
Here are the people you can contact about this asset:

Experts (1) ⓘ

 Franck Mercier [✉](#)

Owners (1) ⓘ

 Will Johnson





Purview Catalog **Business Glossary**

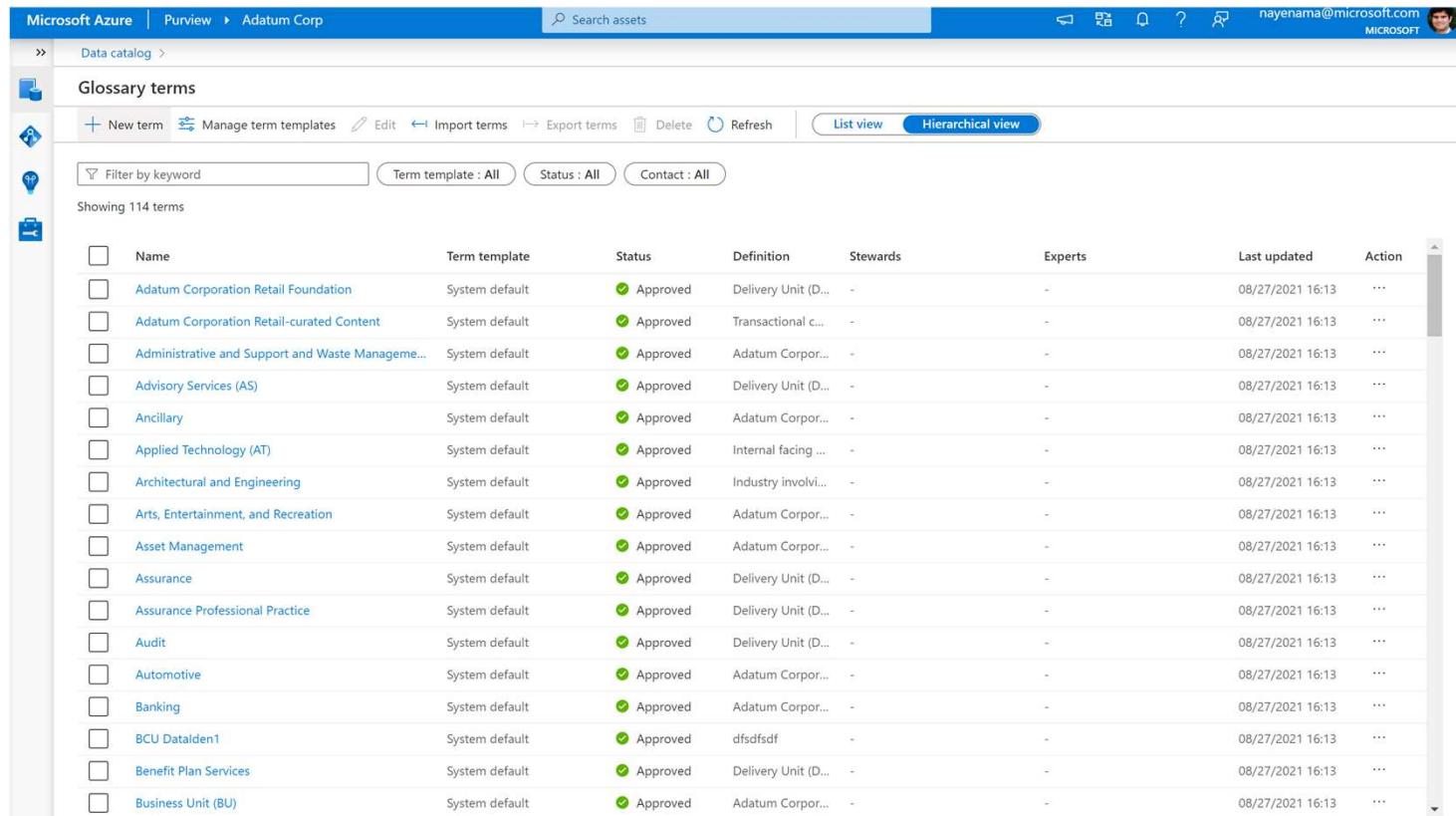
(Search & Browse your data estate from a business lens)

Business Glossary

Consistent and curated understanding of business terms and definitions

Benefits

- Understand business context associated with data in the organization
- Bulk Import glossary terms from existing data dictionaries easily
- Flexible business terms definition with custom attributes per business domain
- Browse & Search your data estate from a business lens
- Establish hier



The screenshot shows the Microsoft Purview Business Glossary interface. The top navigation bar includes 'Microsoft Azure', 'Purview', 'Adatum Corp', a search bar, and user information for 'nayenama@microsoft.com'.

The main page title is 'Glossary terms'. Below it are buttons for 'New term', 'Manage term templates', 'Edit', 'Import terms', 'Export terms', 'Delete', and 'Refresh'. The 'Hierarchical view' button is highlighted.

Below these buttons is a filter bar with 'Filter by keyword', 'Term template : All', 'Status : All', and 'Contact : All'.

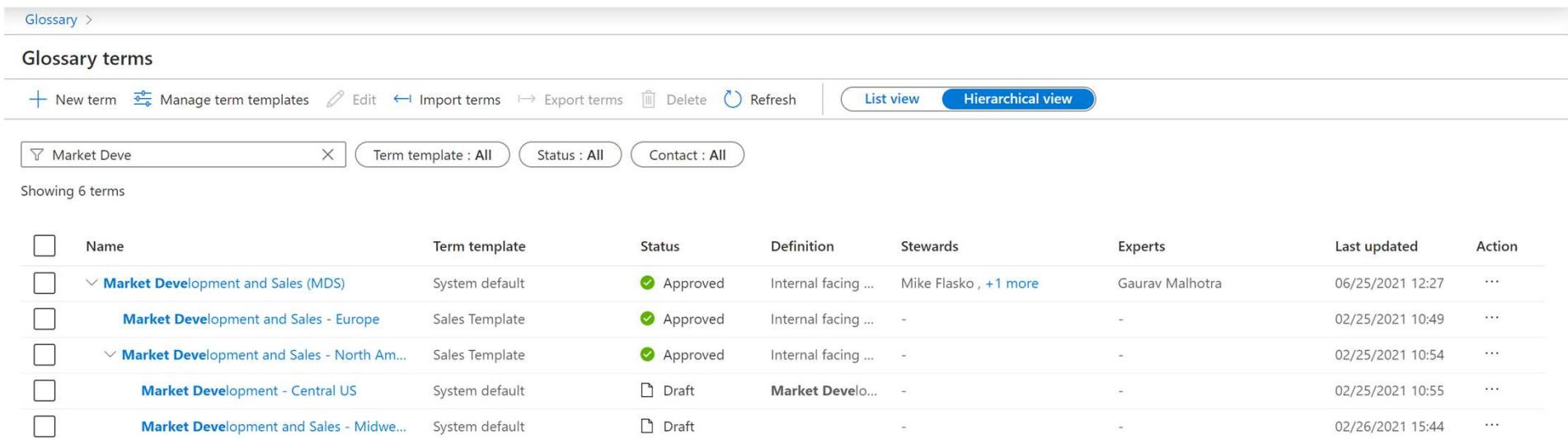
The table displays 114 terms. The columns are: Name, Term template, Status, Definition, Stewards, Experts, Last updated, and Action. Each row shows a term name, its template (e.g., System default), status (Approved), definition (e.g., Delivery Unit (D...)), and last updated date (08/27/2021 16:13). The 'Action' column contains three dots (...).

Name	Term template	Status	Definition	Stewards	Experts	Last updated	Action
Adatum Corporation Retail Foundation	System default	Approved	Delivery Unit (D...)	-	-	08/27/2021 16:13	...
Adatum Corporation Retail-curated Content	System default	Approved	Transactional c...	-	-	08/27/2021 16:13	...
Administrative and Support and Waste Manageme...	System default	Approved	Adatum Corpor...	-	-	08/27/2021 16:13	...
Advisory Services (AS)	System default	Approved	Delivery Unit (D...)	-	-	08/27/2021 16:13	...
Ancillary	System default	Approved	Adatum Corpor...	-	-	08/27/2021 16:13	...
Applied Technology (AT)	System default	Approved	Internal facing ...	-	-	08/27/2021 16:13	...
Architectural and Engineering	System default	Approved	Industry involvi...	-	-	08/27/2021 16:13	...
Arts, Entertainment, and Recreation	System default	Approved	Adatum Corpor...	-	-	08/27/2021 16:13	...
Asset Management	System default	Approved	Adatum Corpor...	-	-	08/27/2021 16:13	...
Assurance	System default	Approved	Delivery Unit (D...)	-	-	08/27/2021 16:13	...
Assurance Professional Practice	System default	Approved	Delivery Unit (D...)	-	-	08/27/2021 16:13	...
Audit	System default	Approved	Delivery Unit (D...)	-	-	08/27/2021 16:13	...
Automotive	System default	Approved	Adatum Corpor...	-	-	08/27/2021 16:13	...
Banking	System default	Approved	Adatum Corpor...	-	-	08/27/2021 16:13	...
BCU Datalden1	System default	Approved	dfsdfsdf	-	-	08/27/2021 16:13	...
Benefit Plan Services	System default	Approved	Delivery Unit (D...)	-	-	08/27/2021 16:13	...
Business Unit (BU)	System default	Approved	Adatum Corpor...	-	-	08/27/2021 16:13	...

Business Glossary

Establishing hierarchical business glossary to achieve the following:

- Define parent-child relationship between terms.
- Same term name can be created in different parents to contextualize them as per organization needs.

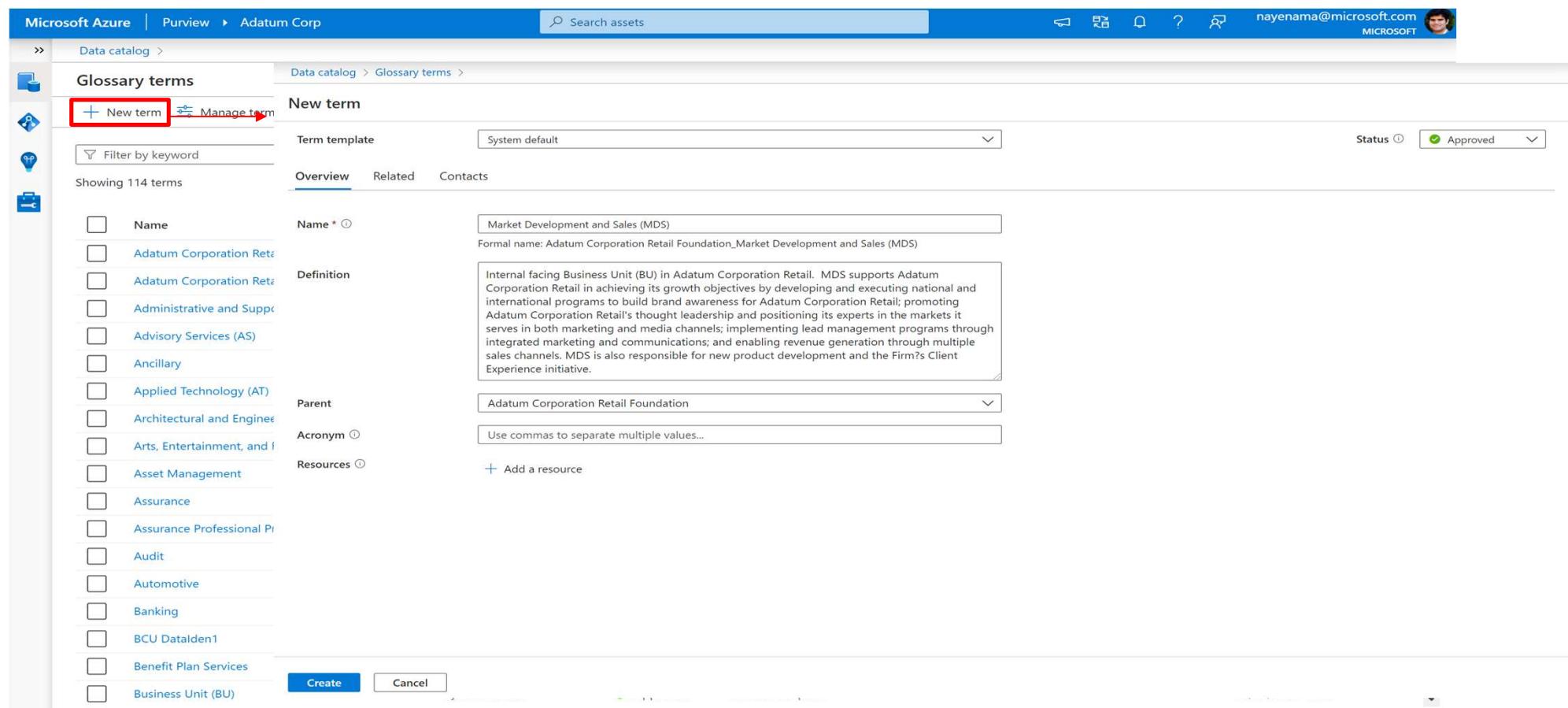


The screenshot shows the Purview Studio Business Glossary interface. At the top, there's a navigation bar with 'Glossary' and a back arrow. Below it is a section titled 'Glossary terms' with a subtitle 'Market Dev'. The main area contains a table with the following columns: Name, Term template, Status, Definition, Stewards, Experts, Last updated, and Action. There are six rows of data, each representing a term under the 'Market Development and Sales (MDS)' parent category. The first row is expanded to show its children: 'Market Development and Sales - Europe', 'Market Development and Sales - North Am...', 'Market Development - Central US', and 'Market Development and Sales - Midwe...'. The last three rows are collapsed, indicated by a minus sign icon. The table also includes a header row with checkboxes for selecting multiple items. At the bottom of the table, there are several buttons: 'New term', 'Manage term templates', 'Edit', 'Import terms', 'Export terms', 'Delete', 'Refresh', 'List view' (which is currently selected), and 'Hierarchical view'.

Name	Term template	Status	Definition	Stewards	Experts	Last updated	Action
Market Development and Sales (MDS)	System default	Approved	Internal facing ...	Mike Flasko, +1 more	Gaurav Malhotra	06/25/2021 12:27	...
Market Development and Sales - Europe	Sales Template	Approved	Internal facing ...	-	-	02/25/2021 10:49	...
Market Development and Sales - North Am...	Sales Template	Approved	Internal facing ...	-	-	02/25/2021 10:54	...
Market Development - Central US	System default	Draft	Market Develo...	-	-	02/25/2021 10:55	...
Market Development and Sales - Midwe...	System default	Draft	-	-	-	02/26/2021 15:44	...

Business Glossary

Vocabulary for business users attached to assets in the catalog



The screenshot shows the Microsoft Purview Business Glossary interface. The top navigation bar includes 'Microsoft Azure', 'Purview', 'Adatum Corp', a search bar ('Search assets'), and a user profile ('nayenama@microsoft.com MICROSOFT'). The left sidebar has icons for Data catalog, Glossary terms, Assets, and Reports. The main area is titled 'Glossary terms' and shows a list of 114 terms. A red box highlights the 'New term' button at the top left of the main content area. The 'New term' form is displayed, with fields for Name (set to 'Market Development and Sales (MDS)'), Definition (a detailed description of MDS), Parent (set to 'Adatum Corporation Retail Foundation'), and Acronym (left empty). The 'Create' button is at the bottom left of the form.

New term

Term template: System default

Status: Approved

Overview Related Contacts

Name * (Market Development and Sales (MDS))

Formal name: Adatum Corporation Retail Foundation_Market Development and Sales (MDS)

Definition: Internal facing Business Unit (BU) in Adatum Corporation Retail. MDS supports Adatum Corporation Retail in achieving its growth objectives by developing and executing national and international programs to build brand awareness for Adatum Corporation Retail; promoting Adatum Corporation Retail's thought leadership and positioning its experts in the markets it serves in both marketing and media channels; implementing lead management programs through integrated marketing and communications; and enabling revenue generation through multiple sales channels. MDS is also responsible for new product development and the Firm's Client Experience initiative.

Parent: Adatum Corporation Retail Foundation

Acronym:

Resources: Use commas to separate multiple values... + Add a resource

Create Cancel

Business Glossary

Understand business context associated with data in the organization

The screenshot shows the Microsoft Purview Business Glossary interface. At the top, there's a navigation bar with 'Microsoft Azure' and 'Purview' followed by 'Adatum Corp'. A search bar says 'Search assets'. On the right, there are icons for notifications, help, and a user profile for 'nayenama@microsoft.com MICROSOFT'.

The main content area shows a term card for 'Market Development and Sales (MDS)'. It has a blue header with the term name and a 'System default' status. Below the header are 'Edit', 'Delete', and 'Refresh' buttons. A 'Overview' tab is selected, with 'Related' and 'Contacts' tabs available.

Formal name: Market Development and Sales (MDS)

Definition: Internal facing Business Unit (BU) in Adatum Corporation Retail. MDS supports Adatum Corporation Retail in achieving its growth objectives by developing and executing national and international programs to build brand awareness for Adatum Corporation Retail; promoting Adatum Corporation Retail's thought leadership and positioning its experts in the markets it serves in both marketing and media channels; implementing lead management programs through integrated marketing and communications; and enabling revenue generation through multiple sales channels. MDS is also responsible for new product development and the Firm's Client Experience initiative.

Acronym: ⓘ No acronym for this term

Catalog assets: 6 assets associated with "Market Development and Sales (MDS)" [View assets](#)

Parent: No parent term for this term

Last updated: 08/27/2021 16:13 by Naga Krishna Yenamandra

Created on: 08/27/2021 16:13 by Naga Krishna Yenamandra

Resources: ⓘ No resources

Business Glossary

Bulk import business terms from existing dictionaries

The screenshot shows the Microsoft Purview Studio Business Glossary interface. On the left, there's a list of glossary terms for 'Adatum Corp' with various filters and actions like 'New term', 'Manage term templates', and 'Import terms'. A red box highlights the 'Import terms' button. On the right, a modal window titled 'Import terms' lists different term templates:

- System default**: System default term template has only the basic fields.
- Banking Template**: Banking Template [View detail](#)
- CountryDetails**: CountryDetails [View detail](#)
- Customer Template**: Customer Template [View detail](#)
- Human Resources Template**: Term Template for creating Human Resources glossary terms [View detail](#)
- MachineManufacturer**: MachineManufacturer [View detail](#)
- ManufacturingTemplate**: ManufacturingTemplate [View detail](#)

At the bottom of the modal are 'Continue' and 'Cancel' buttons.

Name	Term template	Status	Definition
Adatum Corporation Retail Foundation	System default	Approved	Delivery Unit (D...
Adatum Corporation Retail-curated Content	System default	Approved	Transactional c...
Administrative and Support and Waste Manageme...	System default	Approved	Adatum Corpor...
Advisory Services (AS)	System default	Approved	Delivery Unit (D...
Ancillary	System default	Approved	Adatum Corpor...
Applied Technology (AT)	System default	Approved	Internal facing ...
Architectural and Engineering	System default	Approved	Industry involvi...
Arts, Entertainment, and Recreation	System default	Approved	Adatum Corpor...
Asset Management	System default	Approved	Adatum Corpor...
Assurance	System default	Approved	Delivery Unit (D...
Assurance Professional Practice	System default	Approved	Delivery Unit (D...
Audit	System default	Approved	Delivery Unit (D...
Automotive	System default	Approved	Adatum Corpor...
Banking	System default	Approved	Adatum Corpor...
BCU Datalden1	System default	Approved	dfsdfsdf
Benefit Plan Services	System default	Approved	Delivery Unit (D...
Business Unit (BU)	System default	Approved	Adatum Corpor...

Business Glossary

Flexible business terms definition per business domain

The screenshot shows the Microsoft Purview Studio Business Glossary interface. On the left, there's a sidebar with icons for Data catalog, Glossary terms, Assets, and Settings. The main area displays a list of 'Glossary terms' with columns for Name, Term template, Status, and Definition. A red box highlights the 'Manage term templates' button in the top navigation bar. On the right, a modal window titled 'Manage term templates' is open, showing a list of term templates like Banking Template, CountryDetails, Customer Template, etc., with columns for Attribute name, Field type, and Description. An 'OK' button is at the bottom of the modal.

Name	Term template	Status	Definition
Adatum Corporation Retail Foundation	System default	Approved	Delivery Unit (D...
Adatum Corporation Retail-curated Content	System default	Approved	Transactional c...
Administrative and Support and Waste Manageme...	System default	Approved	Adatum Corpor...
Advisory Services (AS)	System default	Approved	Delivery Unit (D...
Ancillary	System default	Approved	Adatum Corpor...
Applied Technology (AT)	System default	Approved	Internal facing ...
Architectural and Engineering	System default	Approved	Industry involvi...
Arts, Entertainment, and Recreation	System default	Approved	Adatum Corpor...
Asset Management	System default	Approved	Adatum Corpor...
Assurance	System default	Approved	Delivery Unit (D...
Assurance Professional Practice	System default	Approved	Delivery Unit (D...
Audit	System default	Approved	Delivery Unit (D...
Automotive	System default	Approved	Adatum Corpor...
Banking	System default	Approved	Adatum Corpor...
BCU Datalden1	System default	Approved	dfsdfsdf
Benefit Plan Services	System default	Approved	Delivery Unit (D...
Business Unit (BU)	System default	Approved	Adatum Corpor...

Manage term templates

System Custom

+ New term template Refresh

You can create new term template with custom attributes and manage them here.

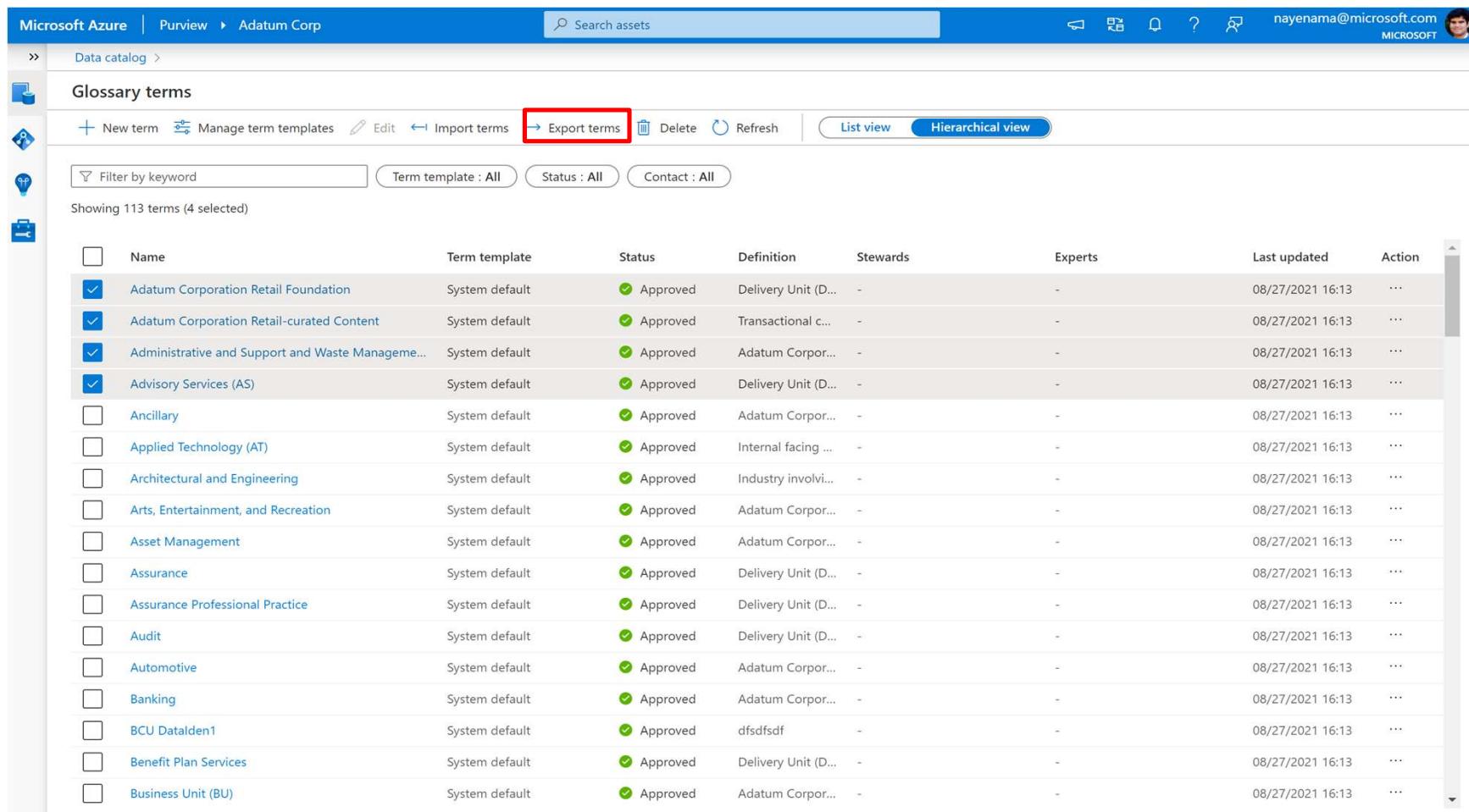
Filter by keyword All term templates

Attribute name	Field type	Description
Banking Template	Text	Bank Location *
CountryDetails	Single choice	Country
Customer Template	Text	Tier
Human Resources Template	Single choice	Data Privacy *
MachineManufacturer	Multiple choice	MachineType
ManufacturingTemplate	Single choice	Code *
Sales Template	Text	Code & Values

OK

Business Glossary

Download terms into .CSV file using export functionality



The screenshot shows the Microsoft Purview Business Glossary interface. At the top, there's a navigation bar with 'Microsoft Azure', 'Purview', 'Adatum Corp', a search bar labeled 'Search assets', and a user profile for 'nayenama@microsoft.com'. Below the navigation is a toolbar with icons for 'New term', 'Manage term templates', 'Edit', 'Import terms', 'Export terms' (which is highlighted with a red box), 'Delete', and 'Refresh'. There are also tabs for 'List view' (which is selected) and 'Hierarchical view'. A filter bar below the toolbar includes 'Filter by keyword', 'Term template : All', 'Status : All', and 'Contact : All'. The main area displays a table titled 'Glossary terms' showing 113 terms. The columns are: Name, Term template, Status, Definition, Stewards, Experts, Last updated, and Action. Several terms are selected, indicated by checked checkboxes in the first column. The 'Status' column shows 'Approved' with green checkmarks. The 'Last updated' column shows the date '08/27/2021 16:13' for all rows.

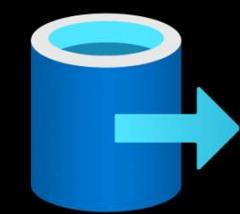
Action	Last updated	Experts	Stewards	Definition	Status	Term template	Name
...	08/27/2021 16:13	-	-	Delivery Unit (D...)	Approved	System default	Adatum Corporation Retail Foundation
...	08/27/2021 16:13	-	-	Transactional c...	Approved	System default	Adatum Corporation Retail-curated Content
...	08/27/2021 16:13	-	-	Adatum Corpor...	Approved	System default	Administrative and Support and Waste Manageme...
...	08/27/2021 16:13	-	-	Delivery Unit (D...)	Approved	System default	Advisory Services (AS)
...	08/27/2021 16:13	-	-	Adatum Corpor...	Approved	System default	Ancillary
...	08/27/2021 16:13	-	-	Internal facing ...	Approved	System default	Applied Technology (AT)
...	08/27/2021 16:13	-	-	Industry involvi...	Approved	System default	Architectural and Engineering
...	08/27/2021 16:13	-	-	Adatum Corpor...	Approved	System default	Arts, Entertainment, and Recreation
...	08/27/2021 16:13	-	-	Adatum Corpor...	Approved	System default	Asset Management
...	08/27/2021 16:13	-	-	Delivery Unit (D...)	Approved	System default	Assurance
...	08/27/2021 16:13	-	-	Delivery Unit (D...)	Approved	System default	Assurance Professional Practice
...	08/27/2021 16:13	-	-	Delivery Unit (D...)	Approved	System default	Audit
...	08/27/2021 16:13	-	-	Adatum Corpor...	Approved	System default	Automotive
...	08/27/2021 16:13	-	-	Adatum Corpor...	Approved	System default	Banking
...	08/27/2021 16:13	-	-	dfsdfsdf	Approved	System default	BCU Datalden1
...	08/27/2021 16:13	-	-	Delivery Unit (D...)	Approved	System default	Benefit Plan Services
...	08/27/2021 16:13	-	-	Adatum Corpor...	Approved	System default	Business Unit (BU)

Business Glossary

Change glossary UX visualization between list and hierarchy views

The screenshot shows the Microsoft Purview Business Glossary interface. At the top, there is a navigation bar with 'Microsoft Azure' and 'Purview' tabs, followed by 'Data catalog' and 'Adatum Corp'. A search bar says 'Search assets'. On the right, there is a user profile for 'nayenama@microsoft.com' and a 'MICROSOFT' logo. Below the navigation is a toolbar with icons for 'New term', 'Manage term templates', 'Edit', 'Import terms', 'Export terms', 'Delete', and 'Refresh'. Two buttons are highlighted with a red box: 'List view' (which is blue and has a white border) and 'Hierarchical view' (which is grey). Below the toolbar are filter options: 'Filter by keyword', 'Term template : All', 'Status : All', and 'Contact : All'. The main area shows 'Showing 113 terms' and a grid of terms. The first term listed is 'Adatum Corporation Retail Foundation'. It has a 'System default' checkbox, a description 'Delivery Unit (DU) under the Firmwide (FW) Business Unit focused on charitable contributions and volunteerism.', and status information 'Approved' (green checkmark) and 'Updated 12 days ago'. The second term is 'Adatum Corporation Retail-curated Content', which is also a 'System default'. Its description is 'Transactional content managed internally by Adatum Corporation Retail BUS for the purpose of research, creating industry-based benchmarks and testing as part of client engagements, PoCs, research publications and pilots'. It also has 'Approved' status and was updated 12 days ago. The third term is 'Administrative and Support and Waste Management and Remediation Services', which is a 'System default'. Its description is 'Adatum Corporation Retail industry sub segment involving Business Support Services; Employment Services; Facilities Support Services; Investigation and Security Services; Office Administrative Services; Other Support Services; Remediation and Other Waste Management Services; Services to Buildings and Dwellings; Travel Arrangement and Reservation Services; Waste Collection; and Waste Treatment and Disposal.' It has 'Approved' status and was updated 12 days ago. The fourth term listed is 'Advisory Services (AS)'. Each term has a checkbox to its right.

Term	Description	Status	Last Updated
Adatum Corporation Retail Foundation	Delivery Unit (DU) under the Firmwide (FW) Business Unit focused on charitable contributions and volunteerism.	Approved	Updated 12 days ago
Adatum Corporation Retail-curated Content	Transactional content managed internally by Adatum Corporation Retail BUS for the purpose of research, creating industry-based benchmarks and testing as part of client engagements, PoCs, research publications and pilots	Approved	Updated 12 days ago
Administrative and Support and Waste Management and Remediation Services	Adatum Corporation Retail industry sub segment involving Business Support Services; Employment Services; Facilities Support Services; Investigation and Security Services; Office Administrative Services; Other Support Services; Remediation and Other Waste Management Services; Services to Buildings and Dwellings; Travel Arrangement and Reservation Services; Waste Collection; and Waste Treatment and Disposal.	Approved	Updated 12 days ago
Advisory Services (AS)			



Purview DataMap **Sources**

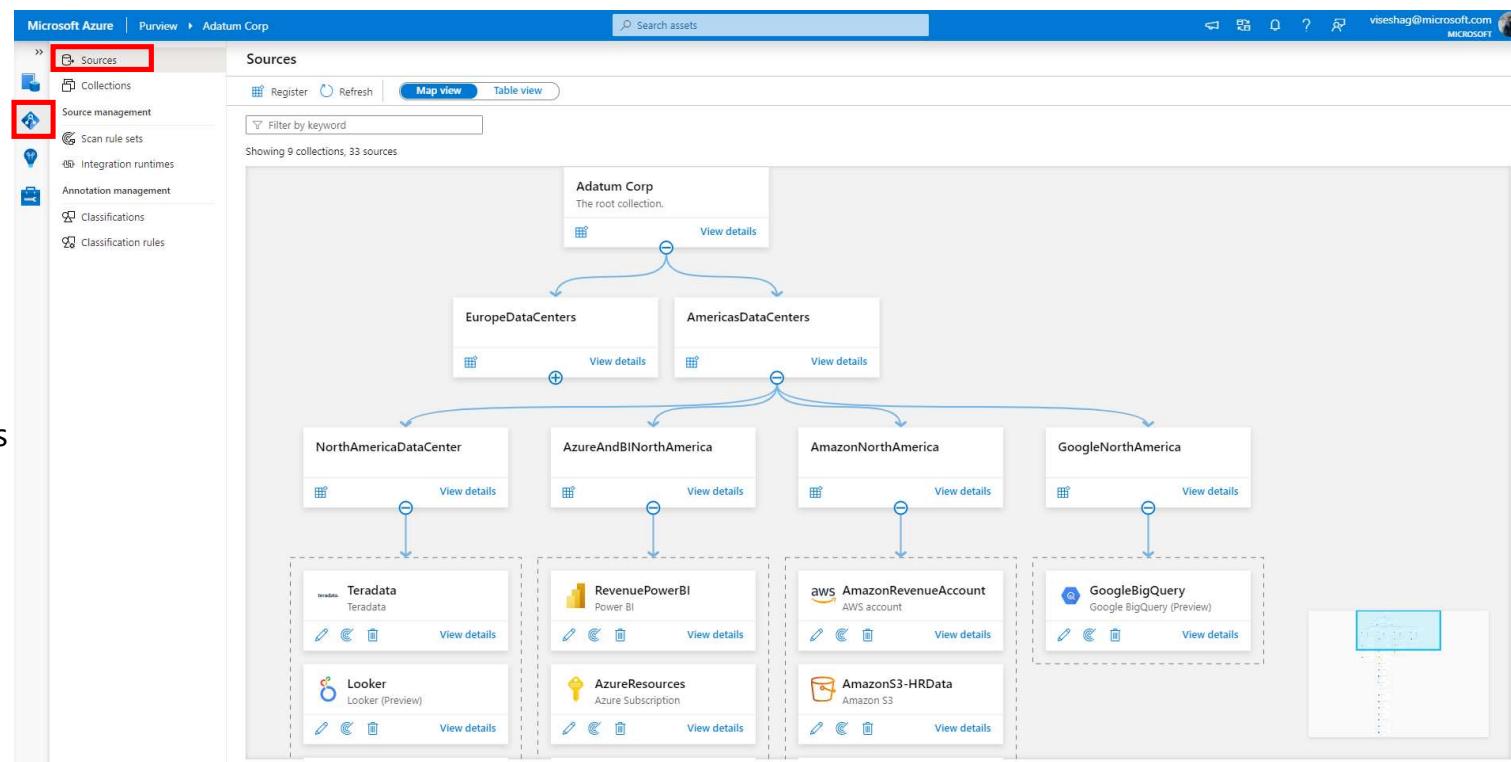
(Automated Scanning, Classification, Open APIs to populate the DataMap)

Sources

Map your data to manage an enriched metadata map of operational and transactional data no matter where it lives

Benefits

- Automated scanning of on-prem, multicloud, SaaS data
- Discover sensitive data stored in Azure, AWS, Google cloud and other services
- Discover Azure data sources, PowerBI, SQL better. Leverage turnkey integrations with Power BI, SQL (on-prem, azure, MI) and key Azure Data Services such as Azure Synapse, Cosmos DB, ADLS.
- Manage metadata and scale understanding of data with automated, fully managed, serverless metadata management capability
- Leverage Apache Atlas Open APIs to programmatically publish metadata and lineage from a wide range open-source data systems



Sources

Automated scanning & classification of on-prem, multicloud and SaaS data

The screenshot shows the Microsoft Purview DataMap Sources interface. On the left, a navigation sidebar includes links for Sources, Collections, Source management, Scan rule sets, Integration runtimes, Annotation management, Classifications, and Classification rules. The main area displays a hierarchical tree of data sources under the root collection 'Adatum Corp'. The tree structure is as follows:

- Adatum Corp (The root collection)
 - EuropeDataCenters
 - NorthAmericaDataCenter
 - AzureAndBINorthAmerica
 - AmericasDataCenters
 - AmazonNorthAmerica
 - Teradata
 - RevenuePowerBI
 - aws AmazonRevenueAccount
 - AmazonS3-HRData
 - AzureResources

Each node in the tree has a 'View details' button. Below the tree, there are three cards: 'Teradata', 'RevenuePowerBI', and 'aws AmazonRevenueAccount'. Each card has a 'View details' button and edit icons.

To the right of the tree, a 'Register sources' dialog is open. It features a search bar labeled 'Filter by keyword' and a category selector with options: All, Azure, Database, Power BI, and Services and apps. The 'Services and apps' tab is selected. A grid of service icons is shown, including:

- Azure (MULTIPLE)
- AWS account (MULTIPLE)
- Azure Synapse Analytics (MULTIPLE)
- Amazon S3
- Azure Blob Storage
- Azure Cosmos DB (SQL API)
- Azure Data Explorer (Kusto)
- Azure Data Lake Storage Gen1
- Azure Data Lake Storage Gen2
- Azure Database for MySQL
- Azure Database for PostgreSQL
- Azure Dedicated SQL Pool

At the bottom of the dialog are 'Continue' and 'Cancel' buttons.

Scan Sources

Set up scans to enrich business value of data with technical, business and semantic metadata

The screenshot shows the Microsoft Purview DataMap Sources interface. On the left, a sidebar navigation includes 'Sources', 'Collections', 'Source management', 'Scan rule sets', 'Integration runtimes', 'Annotation management', 'Classifications', and 'Classification rules'. The main area is titled 'Sources' and displays a 'Map view' of 9 collections and 33 sources. Three main categories are shown: 'NorthAmericaDataCenter', 'AzureAndBINorthAmerica', and 'AmazonNorthAmerica'. Under 'NorthAmericaDataCenter', there are four items: 'Teradata' (with a red box around its 'Edit' icon), 'Looker (Preview)', 'erwin (Preview)', and 'FinanceSQLServer' (with a red box around its 'Edit' icon). Under 'AzureAndBINorthAmerica', there are five items: 'RevenuePowerBI', 'AzureResources', 'AzureSynapseAnalytics', 'PowerBI', and 'PowerBI-Finance'. Under 'AmazonNorthAmerica', there are four items: 'AmazonRevenueAccount', 'AmazonS3-HRData', 'AmazonSql', and 'PowerBI-Finance'. A search bar at the top right says 'Search assets'. On the far right, a 'Scan "FinanceSQLServer"' configuration panel is open, prompting for 'Name' (Scan-prod), 'Connect via integration runtime' (OnPremSHIR (Running)), 'Server endpoint' (10.0.1.7), 'Database name' (empty), 'Credential' (AdatumCorpCredential-SQL), and 'Select a collection' (Adatum Corp > AmericasDataCenters > NorthAmericaDataCenter). It also includes a note: 'All assets scanned will be included in the collection you select.' Buttons for 'Continue', 'Test connection', and 'Cancel' are at the bottom.

Scan Sources

Scope your scan to select the entire storage account or few containers, files

The screenshot shows the Microsoft Purview DataMap Sources interface. On the left, there's a navigation sidebar with options like Sources, Collections, Source management, Scan rule sets, Integration runtimes, Annotation management, Classifications, and Classification rules. The main area displays a list of sources under 'Sources' with buttons for Register, Refresh, Map view (which is selected), and Table view. A search bar at the top right says 'Search assets'. On the right, a 'Scope your scan' dialog is open, showing a tree view of Azure Data Lake Storage with many checkboxes checked. The dialog includes a 'Refresh' button, a 'Search' input field, and 'Continue' and 'Cancel' buttons at the bottom.

Scope your scan

Refresh

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

Search

- AzureDataLakeStorage
 - adfstagedcommandtempdata
 - productiondata
 - 2000
 - 2001
 - 2002
 - 2003
 - 2004
 - 2005
 - 2006
 - 2007
 - 2008
 - 2009
 - 2010
 - 2011
 - 2012
 - 2013
 - 2014
 - 2015
 - 2016
 - 2017

Continue Back Cancel

Scan Sources

Select files types for scanning. Define custom file types

The screenshot shows the Microsoft Purview DataMap interface for defining sources. On the left, a sidebar lists various management options like Collections, Source management, Scan rule sets, and Integration runtimes. The main area displays a list of sources, including AzureSQLDB-SalesInvo..., FinancialBlob, BlobStorage, AzureDataLakeStorage, AzureMySQL, and AzurePostgreSQL. To the right, a 'Select file types' panel is open, featuring a grid of file formats with checkboxes. A red arrow points to the 'Custom File Type' button at the top of this panel.

Microsoft Azure | Purview > Adatum Corp

Sources

Search assets

Custom File Type →

Register Refresh Map view Table view

Filter by keyword

Showing 9 collections, 33 sources

AzureSQLDB-SalesInvo... Azure SQL Database

FinancialBlob Azure Blob Storage

BlobStorage Azure Blob Storage

AzureDataLakeStorage Azure Data Lake Storage Gen2

AzureMySQL Azure Database for MySQL

AzurePostgreSQL Azure Database for PostgreSQL

Select file types

+ New file type

File types for schema extraction and classification.

CSV	JSON	PSV	SSV
TSV	TXT	XML	PARQUET
AVRO	ORC	Document file types	GZIP

Continue Back Cancel

Scan Sources

Discover sensitive data using 200+ built in classifiers, define your own custom classifiers

The screenshot shows the Microsoft Purview DataMap interface for scanning sources. On the left, there's a navigation bar with 'Sources' selected, and a list of collections: NorthAmericaDataCenter, EuropeDataCenter, and AzureAndBINorthAmerica. The AzureAndBINorthAmerica collection is currently active, displaying various data sources like OnPremSQLServer-Fina..., SAP-S4HANA-Procurement, and AzureDataLakeStorage.

In the top right, a search bar contains the text 'Revenue'. Below it is a 'Select classification rules' section with the sub-instruction: 'Choose classification rules that will run on the dataset.'

The classification rules are divided into two sections:

- System rules** (highlighted with a red box):
 - Government
 - Financial
 - Personal
 - Security
 - Miscellaneous
- Custom rules** (highlighted with a red box):
 - Custom classification rules
 - TransactionID
 - SupplierNumber
 - StoreUid
 - ProductCode
 - ProductCategory
 - PONumber
 - InvoiceNumber
 - EmployeeID
 - CouponCode

At the bottom right, there are 'Create', 'Back', and 'Cancel' buttons, along with a status message: '105 classification rules selected'.

Scan Sources

Run the scan one-time or on a schedule

The screenshot shows the Microsoft Purview DataMap Sources interface. On the left, a sidebar navigation includes Sources, Collections, Source management, Scan rule sets, Integration runtimes, Annotation management, Classifications, and Classification rules. The main area displays a list of 33 sources under 'Sources' (Map view selected). The sources listed are:

- AzureSqlManagedInsta... (Azure SQL Database Managed Instance)
- AzureCosmos (Azure Cosmos DB (SQL API))
- AzureDedicatedSQLPool (Azure Dedicated SQL Pool (formerly SQL DW))
- ReportingBlob (Azure Blob Storage)
- AzureSQLDB-SalesInvo... (Azure SQL Database)
- FinancialBlob (Azure Blob Storage)

To the right, a 'Set a scan trigger' dialog is open. It includes fields for 'Recurring' (selected) or 'Once', 'Time zone' (UTC), 'Recurrence' (Every 1 Month(s)), 'Select day of the month to scan' (a grid from 1 to 31), 'Schedule scan time (UTC)' (hh:mm:ss AM), 'Start recurrence at (UTC)' (2021-09-09 10:19:00 PM), and a checkbox for 'Specify recurrence end date (UTC)'. Navigation buttons at the bottom are Continue, Back, and Cancel.

DataMap— Integration runtimes

Integration runtimes are the compute infrastructure used by scanners to provide automated data scanning, lineage and classification capabilities across different network and Multi-Cloud environments including on-prem environments

Benefits

- Offers Self-Hosted Integration Runtime
- Self-Hosted Integration Runtime— use compute resources in on-premises machine or a VM inside private network

Integration runtime setup

Integration Runtime is the native compute used to execute scans. Choose the network environment of the integration runtime which will connect to the data source and run scans.

Self-Hosted

Use this for running scans in an on-premises / private network

Name	Type	Status
OnPremSHIR	Self-Hosted	Running

Continue Cancel

DataMap – Custom Classification rules

Define your own data pattern along with thresholds to reduce false positives

The screenshot shows the Microsoft Azure Purview DataMap interface. The left sidebar has a tree view with nodes like Sources, Collections, Source management, Scan rule sets, Integration runtimes, Annotation management, Classifications, and Classification rules. The Classification rules node is highlighted with a red box. The main area is titled 'Classification rules' and shows a table with a header row: 'Name' and 'State'. The table lists numerous classification rules, each with a green checkmark in the 'State' column, indicating they are enabled. The rules include: ABA Routing Number, Age of an individual, Argentina National Identity (DNI) Number, Australia Bank Account Number, Australia Business Number, Australia Company Number, Australia Driver's License Number, Australia Medical Account Number, Australia Passport Number, Australia Tax File Number, Austria Driver's License Number, Austria Identity Card, Austria Passport Number, Austria Social Security Number, Austria Tax Identification Number, Austria Value Added Tax, Belgium Driver's License Number, Belgium National Number, Belgium Passport Number, Belgium Value Added Tax Number, and Brazil Individual Taxpayer Identification Number (CPF).

Name	State
ABA Routing Number	Enabled
Age of an individual	Enabled
Argentina National Identity (DNI) Number	Enabled
Australia Bank Account Number	Enabled
Australia Business Number	Enabled
Australia Company Number	Enabled
Australia Driver's License Number	Enabled
Australia Medical Account Number	Enabled
Australia Passport Number	Enabled
Australia Tax File Number	Enabled
Austria Driver's License Number	Enabled
Austria Identity Card	Enabled
Austria Passport Number	Enabled
Austria Social Security Number	Enabled
Austria Tax Identification Number	Enabled
Austria Value Added Tax	Enabled
Belgium Driver's License Number	Enabled
Belgium National Number	Enabled
Belgium Passport Number	Enabled
Belgium Value Added Tax Number	Enabled
Brazil Individual Taxpayer Identification Number (CPF)	Enabled

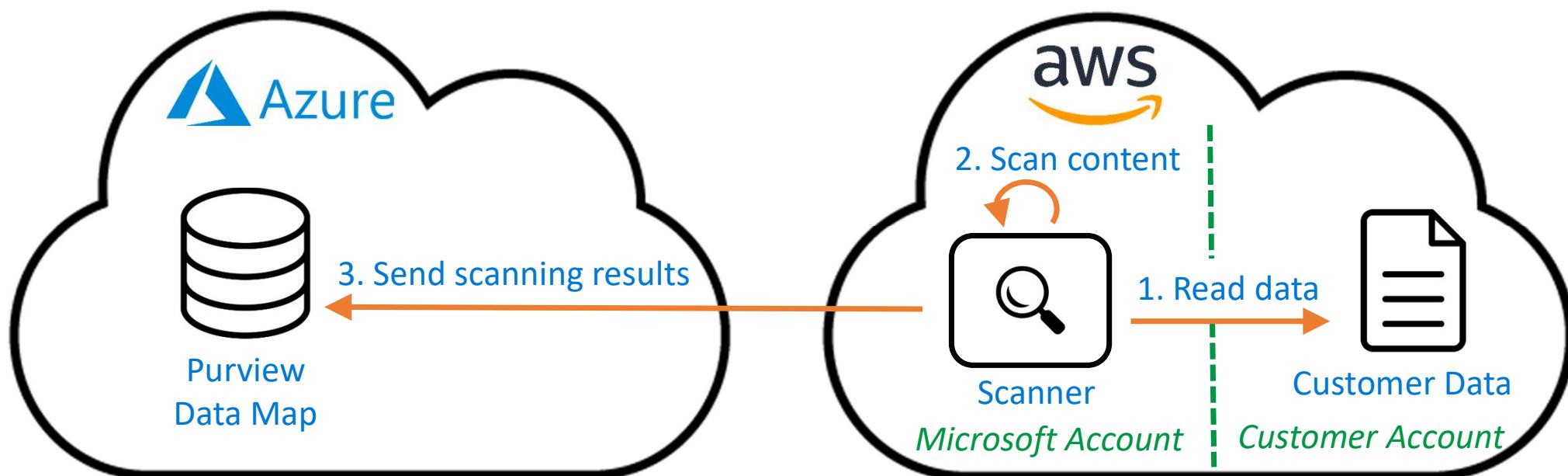
Scanning Data in Remote Cloud

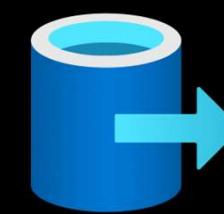
Easy configuration

- Fully managed solution
- No need to deploy and maintain agent in the customer's account
- Requires a dedicated role in AWS IAM to allow Purview scanner read the data

Data Privacy

- Sensitive data is scanned locally in the remote cloud
- Only scanning results are sent to Azure





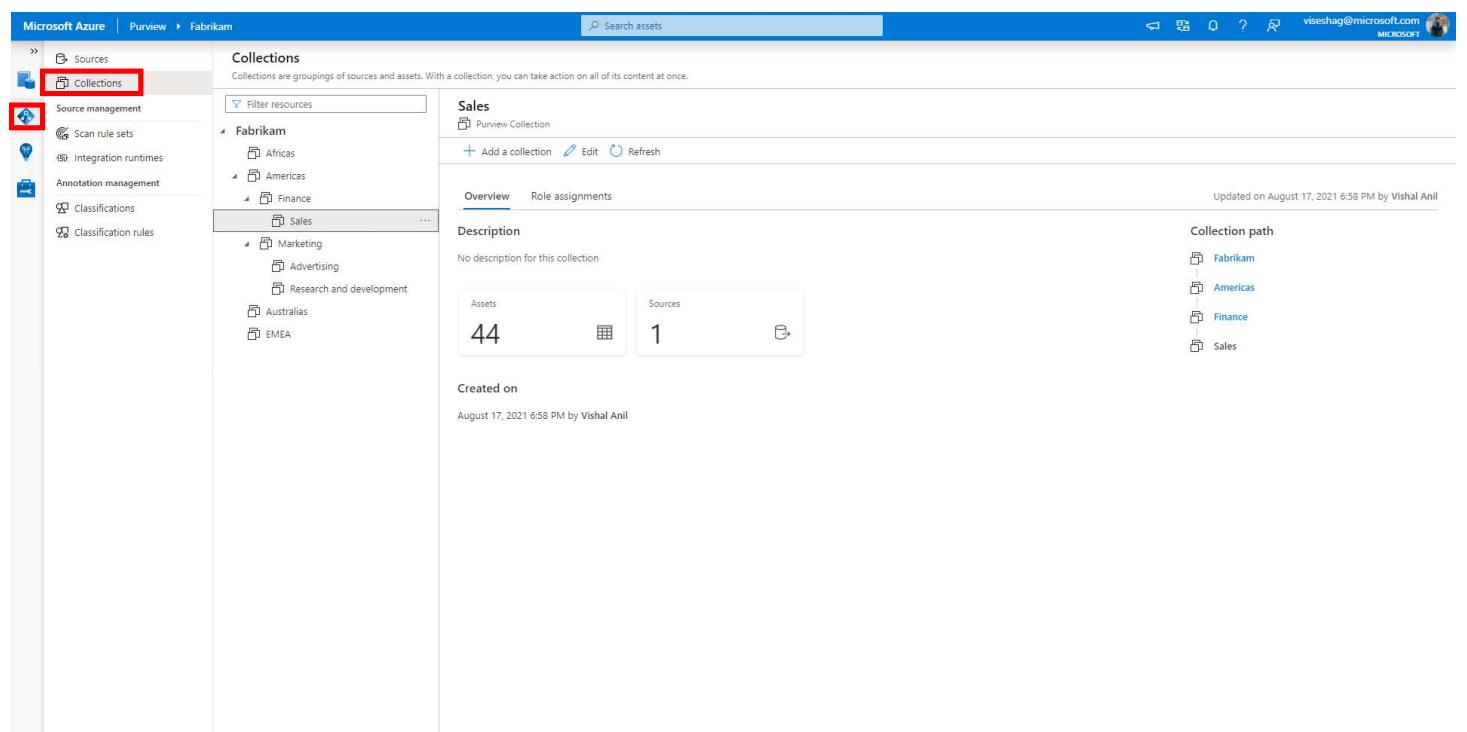
Purview DataMap Collections

Collections

Organize assets and sources based on your business' requirements, and restrict access to them using Purview Collections

Benefits

- A collection is a tool Purview uses to group assets, sources, and other artifacts into a hierarchy
- Provide fine-grained access control
- Discover assets by collection



The screenshot shows the Microsoft Purview Collections interface. The left sidebar has a 'Collections' icon highlighted with a red box. The main area displays a tree view of collections under 'Fabrikam'. The 'Sales' collection is expanded, showing 'Marketing', 'Advertising', 'Research and development', 'Australia', and 'EMEA'. The 'Assets' section shows 44 items, and the 'Sources' section shows 1 item. A 'Collection path' sidebar on the right shows the hierarchy: Fabrikam > Americas > Finance > Sales.

Collections

Collections are groupings of sources and assets. With a collection, you can take action on all of its content at once.

Sales

Purview Collection

+ Add a collection Edit Refresh

Overview **Role assignments**

No description for this collection

Assets 44 View

Sources 1 View

Created on

August 17, 2021 6:58 PM by Vishal Anil

Collection path

```

graph TD
    Root[Fabrikam] --> Americas[Americas]
    Americas --> Finance[Finance]
    Finance --> Sales[Sales]
  
```

Collections – assigning roles

Assign Purview roles at the collection level to users, groups and service principals

The screenshot shows the Microsoft Purview Collections interface. The left sidebar navigation includes Sources, Collections (selected), Source management, Scan rule sets, Integration runtimes, Annotation management, Classifications, and Classification rules. The main area displays the 'Collections' page for the 'Fabrikam' organization. A search bar at the top right says 'Search assets'. Below it, a 'Filter resources' button and a 'Sales' Purview Collection are listed. A red box highlights the 'Sales' collection. Below the collection name are buttons for '+ Add a collection', 'Edit', and 'Refresh'. The page is divided into sections: 'Overview' (disabled) and 'Role assignments' (selected). A red box highlights the 'Role assignments' tab. Under 'Role assignments', there are tabs for 'Collection admins', 'Data source admins', 'Data curators', and 'Data readers'. A 'Restrict inherited permissions' toggle switch is present. The 'Collection admins' section lists four users: Daniel Perlovsky (Inherited), Hoc Phan, Vishal Anil, and Vishal Anil (Inherited). The 'Data source admins' section lists one user: Vishal Anil (Inherited). The 'Data curators' and 'Data readers' sections are currently empty. A note states: 'Collection admins can manage the collection, its details, and add subcollections. They can also add data curators, data readers, and other Purview roles to a collection scope. Collection admins that are part of a parent collection can't be removed.' The page footer indicates it was updated on August 17, 2021, at 6:58 PM by Vishal Anil.

Name	Type
Daniel Perlovsky (Inherited) Daniel.Perlovsky@microsoft.com	User
Hoc Phan Hoc.Phan@microsoft.com	User
Vishal Anil Vishal.Seshagirirao@microsoft.com	User
Vishal Anil (Inherited) Vishal.Seshagirirao@microsoft.com	User
Data source admins	
Vishal Anil (Inherited) Vishal.Seshagirirao@microsoft.com	User
Data curators	

Collections – restricting access

Lock down certain collections of assets by restricting inherited permissions

The screenshot shows the Microsoft Purview Collections interface. On the left, there's a navigation sidebar with options like Sources, Collections, Source management, Scan rule sets, Integration runtimes, Annotation management, Classifications, and Classification rules. The 'Collections' option is selected. The main area displays a 'Sales' Purview Collection under the 'Fabrikam' hierarchy. A 'Role assignments' tab is selected, showing a table of users with their names, emails, and types (User). At the top of this table, there's a 'Restrict inherited permissions' toggle switch, which is currently turned on (indicated by a red box). Below the table, there are sections for 'Collection admins', 'Data source admins', and 'Data curators', each with a list of users.

Name	Type
Daniel Perlovsky (Inherited) Daniel.Perlovsky@microsoft.com	User
Hoc Phan Hoc.Phan@microsoft.com	User
Vishal Anil Vishal.Seshagirirao@microsoft.com	User
Vishal Anil (Inherited) Vishal.Seshagirirao@microsoft.com	User
Data source admins	
Data curators	



Microsoft Purview Data Sharing

Data Sharing

(Share data within and between organizations)

Data Sharing

Share data within and between organizations using Microsoft Purview Data Sharing

Benefits

- Easily share 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

The screenshot shows the Microsoft Purview Governance Portal interface for Data Sharing. The top navigation bar includes 'Overview', 'Sent shares' (highlighted with a red box), 'Received shares', and a central search bar. The main content area is titled 'Data share (preview)' and features a large button labeled '+ Create a share' with the sub-instruction 'Share data internally or with other organizations.' Below this are three cards: 'What is Data share?' (explaining it's a simple and secure way to share data between SOURCE and TARGET), 'How to send shares' (instructions to create a share, choose data, and send invitations), and 'How to receive shares' (instructions to accept shared data and decide where to access it). On the right side, there's a decorative graphic of data storage and sharing icons.

Share Data

Data provider creates new share

The screenshot shows the Microsoft Purview Governance Portal interface for creating a new share. On the left, there's a sidebar with icons for Overview, Sent shares (highlighted with a red box), Received shares, and other governance categories. The main area is titled 'Sent shares' and lists several existing shares. A 'New share' button is highlighted with a red box. To the right, a 'New share' dialog box is open, prompting for a 'Share name' (with a placeholder 'Enter a name') and a 'Share type'. The 'In-place share' option is selected, described as 'Provide access to data at its source'. There are fields for 'Description' (placeholder 'Enter a description') and 'Select a collection' (dropdown menu showing '(Root) FabrikamPurview'). At the bottom of the dialog are 'Continue' and 'Cancel' buttons.

Microsoft Purview | Governance Portal > FabrikamPurview

Search catalog

Overview

Sent shares

Received shares

New share

Share name *

Enter a name

Share type *

In-place share

Provide access to data at its source.

Description

Select a collection *

(Root) FabrikamPurview

Continue

Cancel

Share Data

Add assets from Blob Storage or ADLS Gen2 data sources

The screenshot shows the 'New share' wizard in the Microsoft Purview Governance Portal. The current step is 'Assets'. It displays a list of available assets from various data sources:

Share name	Created by	Share type
ProductData41	jife@microsoft.com Contoso Marketplace	In-place sha
ProductData40	jife@microsoft.com Contoso Marketplace	In-place sha
ForecastData	jife@microsoft.com Contoso Marketplace	In-place sha
InventoryData	jife@microsoft.com Contoso Marketplace	In-place sha
ProductData	jife@microsoft.com Contoso Marketplace	In-place sha

A red box highlights the '+ Add assets' button at the bottom of the list.

The screenshot shows the 'Add assets' dialog in the Microsoft Purview Governance Portal. It lists assets under the path 'jife@previewsstorage > products'. Two assets are selected: 'catalog' and 'sales'. A red box highlights the selected assets.

Asset Path	Selected
jife@previewsstorage > products	2 asset(s) selected
catalog	<input checked="" type="checkbox"/>
sales	<input checked="" type="checkbox"/>
creative.json	<input type="checkbox"/>
products.parquet	<input type="checkbox"/>

At the bottom, it says 'Showing 1 to 5 of 5 assets'.

Share Data

Add Share recipients and expiration date if desired

The screenshot shows the Microsoft Purview Governance Portal Data Sharing interface. On the left, there's a sidebar with icons for Overview, Sent shares (selected), Received shares, and other governance categories. The main area has two sections: 'Sent shares' on the left and 'New share' on the right.

Sent shares: A table listing six existing shares. Each row includes a checkbox, the share name, the created by user (with their initials in a teal circle), and the share type (In-place share). The share names are: demo-cc, demo, Demo-Thales, Demo-DS-Fhe, and ARIBAN-DS.

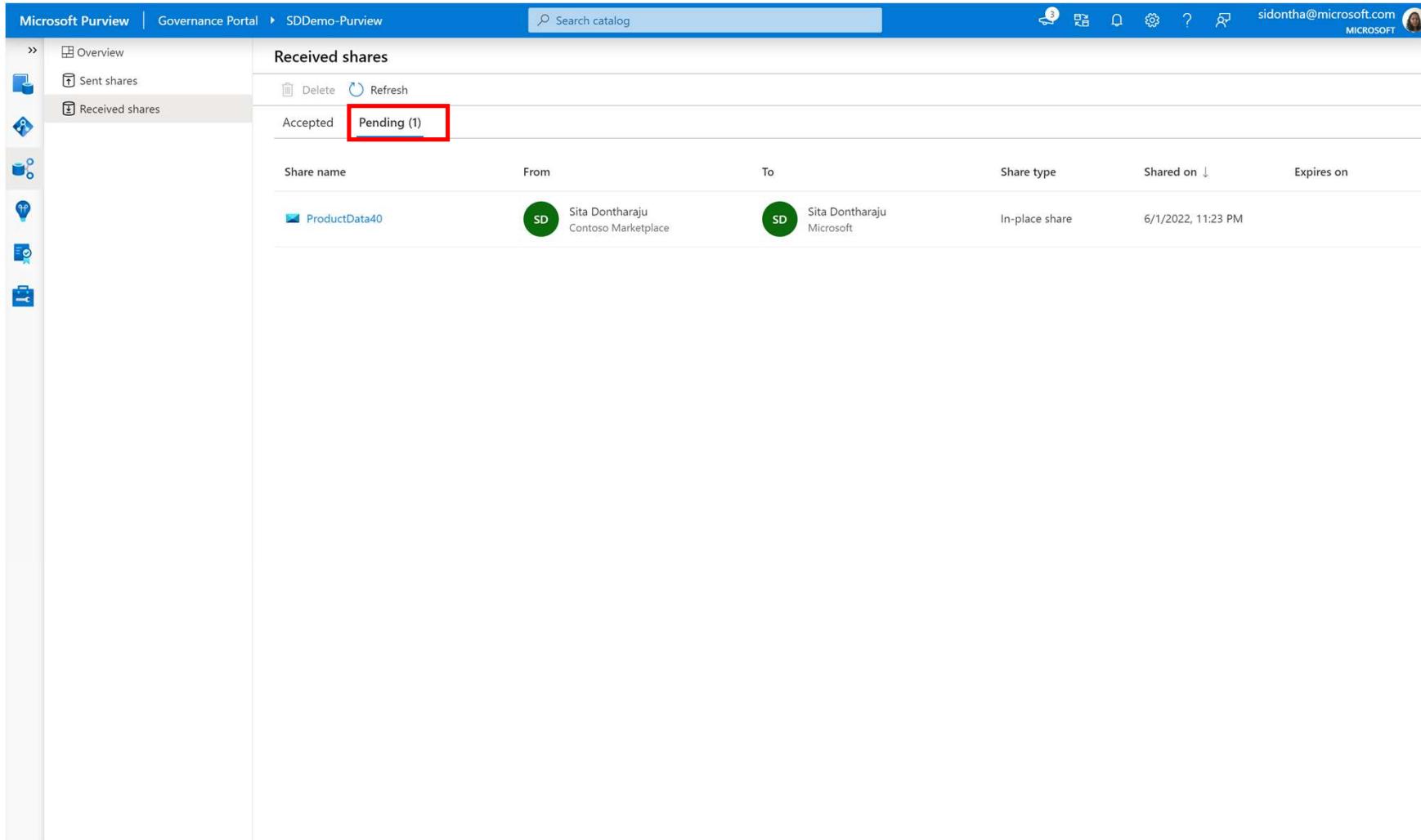
	Share name	Created by	Share type
<input type="checkbox"/>	demo-cc	FH Farhad Heybati Microsoft	In-place sha
<input type="checkbox"/>	demo	FH Farhad Heybati Microsoft	In-place sha
<input type="checkbox"/>	Demo-Thales	FH Farhad Heybati Microsoft	In-place sha
<input type="checkbox"/>	Demo-DS-Fhe	FH Farhad Heybati Microsoft	In-place sha
<input type="checkbox"/>	ARIBAN-DS	AB Arindam Bandyopadhyay Microsoft	In-place sha

New share: A form for creating a new share. It has fields for 'Email' (containing sidontha@microsoft.com) and 'Expiration date' (set to 2022-06-08 at 10:39:22 PM). There are three buttons at the top of this section: 'Add recipient' (highlighted with a red box), 'Delete all', and 'Update expiration for all'.

At the bottom of the page are three buttons: 'Create and Share' (in blue), 'Back', and 'Cancel'.

Receive Share

Data consumer receives email invitation, sees notification in Microsoft Purview governance portal



The screenshot shows the Microsoft Purview Governance Portal interface. The top navigation bar includes links for Overview, Sent shares, and Received shares. The main content area is titled "Received shares" and displays a table of pending shares. The table has columns for Share name, From, To, Share type, Shared on, and Expires on. One row is shown, detailing a share named "ProductData40" from "Sita Dontharaju Contoso Marketplace" to "Sita Dontharaju Microsoft". The share type is "In-place share" and it was shared on 6/1/2022, 11:23 PM. A red box highlights the "Pending (1)" button in the filter section above the table.

Share name	From	To	Share type	Shared on	Expires on
ProductData40	 Sita Dontharaju Contoso Marketplace	 Sita Dontharaju Microsoft	In-place share	6/1/2022, 11:23 PM	

Receive Share

Data consumer accepts and configures received share, maps assets to target Storage account.

The screenshot shows the Microsoft Purview Governance Portal interface. On the left, there's a vertical sidebar with various icons. The main navigation bar at the top includes 'Microsoft Purview', 'Governance Portal', and 'SDDemo-Purview'. A search bar labeled 'Search catalog' is also present. The main content area has a header 'Received shares' with tabs for 'Accepted' (selected) and 'Pending (1)'. Below this, a table lists a single share named 'ProductData40' from 'Sita Dontharaju' (Contoso Marketplace) to 'Sita Dontha' (Microsoft). The right side of the screen displays a 'Pending share "ProductData40"' configuration dialog. It shows the 'From' field as 'Sita Dontharaju' and 'Company' as 'Contoso Marketplace'. The 'Description' and 'Expiration date' fields are empty. Below this, there are two input fields: 'Received share name *' containing 'ProductData40' and 'Select a collection *' with '(Root) SDDemo-Purview' selected. At the bottom of the dialog are three buttons: 'Accept and configure' (highlighted with a red box), 'Reject', and 'Cancel'.

Receive Share

Data consumer views status of the mapping, navigates to the target storage account to work with received data.

The screenshot shows the Microsoft Purview Governance Portal interface. The top navigation bar includes 'Microsoft Purview | Governance Portal > Contoso'. The left sidebar has icons for Overview, Sent shares, and Received shares, with 'Received shares' being the active tab. The main content area is titled 'Received shares > blob'. It shows an 'In-place share' named 'blob'. Below this are buttons for 'Map Assets', 'Delete', and 'Refresh'. The 'Assets' tab is selected in the navigation bar. A table lists the asset details:

Name ↑	Source type	Status	Target type	Target
namjapaneastproblob	Azure Blob Storage	Mapped	Azure Blob Storage	japaneastconblob/japaneast/test



Purview Insights

Data Insights

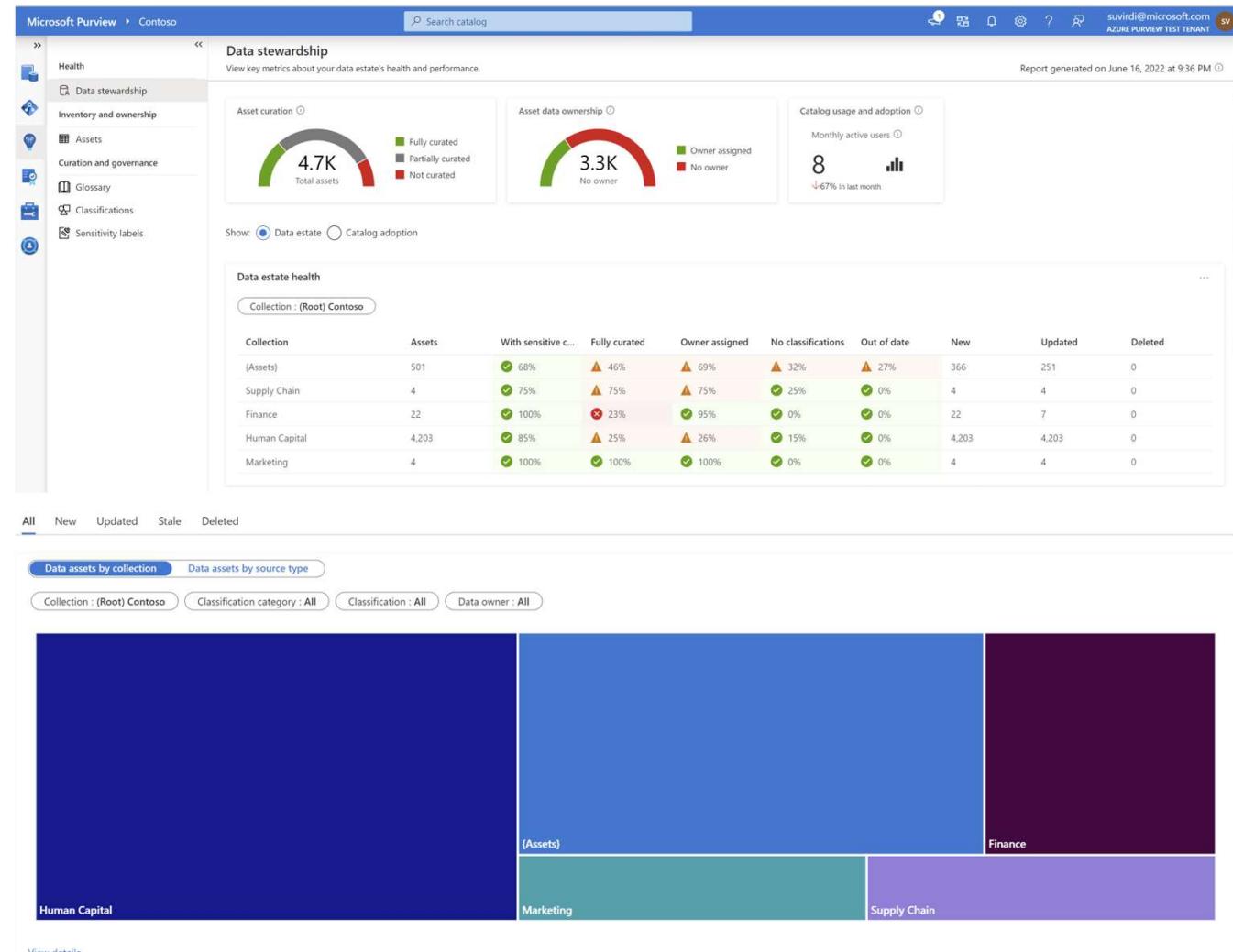
(Bird's eye view of your Data Estate)

Data Estate A0 Insights

Bird's eye view of data landscape

Benefits

- Intended to help users such as Chief Data Officers quickly understand their data estate at large and gain key insights such as where sensitive data resides
- Asset Insights to see where all data resides across Collections and range of source types. Glossary insights to understand changes made to business terms and how much coverage glossary has over your data map
- Sensitive data Insights to simplify compliance risk assessment across operational and transactional data sources. Assess risk and derive audit trails of data qualified by sensitivity and business relevance



Diapositive 111

A0 [Mention was removed] for final approval. cc: [Mention was removed]

Author; 2022-06-21T21:36:14.718

Data stewardship

A view of the data estate health and insights into the catalog's adoption

Health

- Data stewardship** (selected)
- Inventory and ownership
- Assets
- Curation and governance
- Glossary
- Classifications
- Sensitivity labels

Data stewardship
View key metrics about your data estate's health and performance.

Report generated on June 16, 2022 at 9:36 PM ⓘ

Asset curation ⓘ

4.7K
Total assets

Fully curated
Partially curated
Not curated

Asset data ownership ⓘ

3.3K
No owner

Owner assigned
No owner

Catalog usage and adoption ⓘ

Monthly active users ⓘ

8

↓67% In last month

Show: Data estate Catalog adoption

Data estate health

Collection : (Root) Contoso

Collection	Assets	With sensitive c...	Fully curated	Owner assigned	No classifications	Out of date	New	Updated	Deleted
{Assets}	501	✓ 68%	⚠ 46%	⚠ 69%	⚠ 32%	⚠ 27%	366	251	0
Supply Chain	4	✓ 75%	⚠ 75%	⚠ 75%	✓ 25%	✓ 0%	4	4	0
Finance	22	✓ 100%	✗ 23%	✓ 95%	✓ 0%	✓ 0%	22	7	0
Human Capital	4,203	✓ 85%	⚠ 25%	⚠ 26%	✓ 15%	✓ 0%	4,203	4,203	0
Marketing	4	✓ 100%	✓ 100%	✓ 100%	✓ 0%	✓ 0%	4	4	0

Data stewardship

A view of the data estate health and insights into the catalog's adoption

[Data stewardship](#)

View key metrics about your data estate's health and performance.

Report generated on June 16, 2022 at 9:36 PM ⓘ

Show: Data estate Catalog adoption

Active users by feature category

Date range: Daily Weekly Monthly

Date	All	Search and browse	Asset curation
May 18	4	3	1
May 19	0	0	0
May 20	2	0	1
May 21	0	0	0
May 22	0	0	0
May 23	1	0	1
May 24	0	1	1
May 25	0	1	1
May 26	0	0	0
May 27	0	0	0
May 28	0	0	0
May 29	0	0	0
May 30	0	0	0
May 31	0	2	0
Jun 1	0	0	0
Jun 2	0	0	0
Jun 3	0	0	0
Jun 4	0	0	0
Jun 5	0	0	0
Jun 6	0	0	0
Jun 7	1	1	1
Jun 8	0	0	0
Jun 9	0	0	0
Jun 10	0	0	0
Jun 11	0	0	0
Jun 12	0	0	0
Jun 13	1	1	0
Jun 14	0	0	0
Jun 15	2	0	1

Number of users

May 18 May 20 May 22 May 24 May 26 May 28 May 30 Jun 1 Jun 3 Jun 5 Jun 7 Jun 9 Jun 11 Jun 13 Jun 15

— All — Search and browse — Asset curation

Most viewed assets Last 30 days

Asset name	Curation	Views
customclassifications_renamed_9.csv	Fully curated	55
supportedclassifications_renamed_2.tsv	Partially curated	50
blobdatascansrscus	Partially curated	36
1mbdata-lessthan-Sunique-records	Fully curated	32
WithoutColumnName.csv	Fully curated	30

Top searched keywords Last 30 days

Show: Searches with results Searches with no results

Keyword	Search volume
*	6
mssql://10.1.0.4/mssqlserver/sinkonpremdb/sink/adc_lineagecopy_source	2

Assets

Understand distribution of data assets across a range of source types, classification types, collections and more

[Assets \(preview\)](#)

View key metrics about your data estate's assets.

Report generated on June 17, 2022 at 8:52 AM

Total assets 1.1M Adatum Corp

Asset classification 1.1M No classification

Asset data ownership 1.1M No owner

New assets (Last 30 days) 23 New

Deleted (Last 30 days) 2 Existing

All New Deleted

Data assets by collection Data assets by source type

Collection : (Root) Adatum Corp Classification category : All Classification : All Data owner : All

(Assets) AmericasDataCenters

View details

File-based source types

Top file extensions

By count By size Source type : All

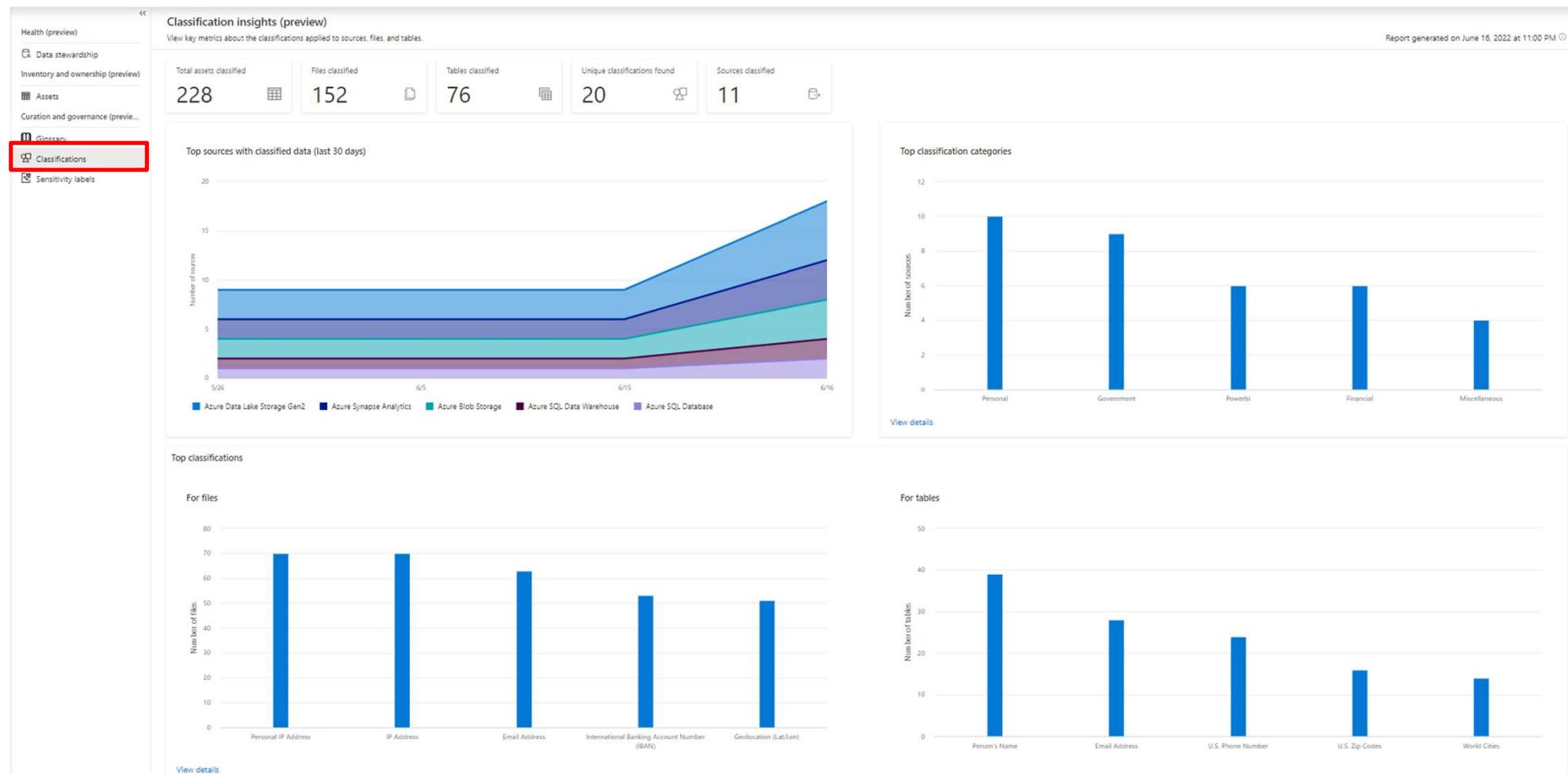
File Extension	Count
JSON	~50
CSV	~50
SSV	~30
TSV	~30

Files not stored in resource sets

Storage Location	Count
Azure Blob Storage	~80

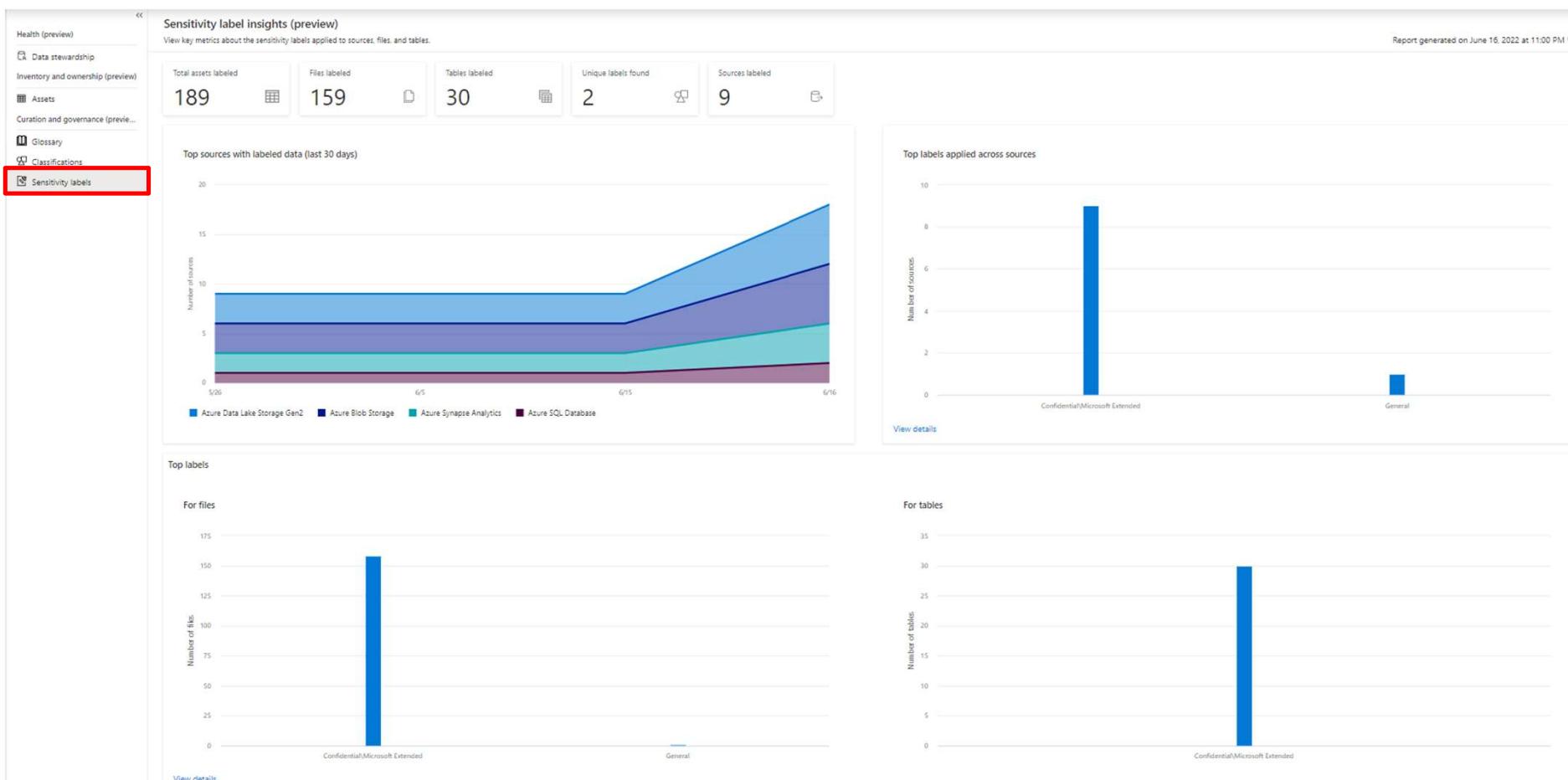
Classifications

Understand what sensitive data exists across the data estate from various lens



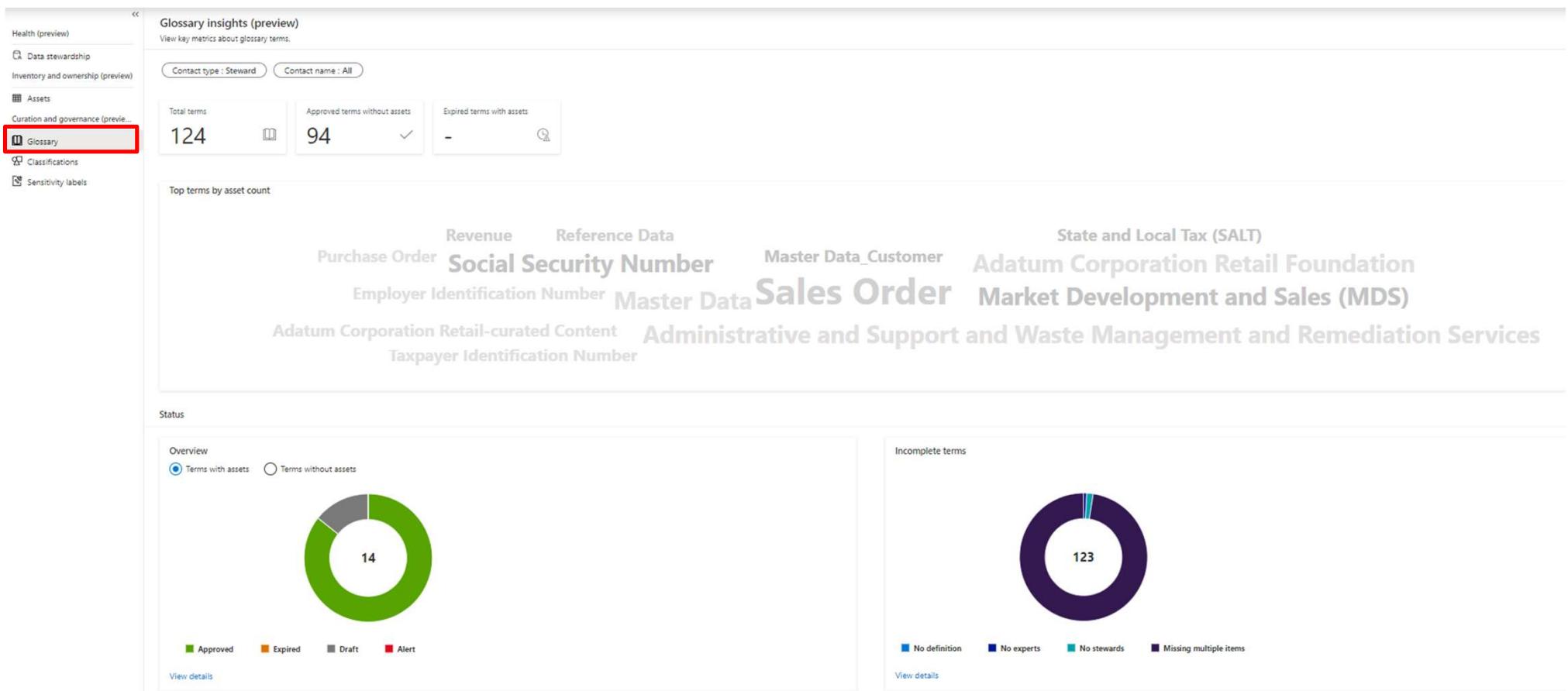
Sensitivity Labels

Understand what sensitivity labels have been applied across the data estate



Glossary

Understand changes made to business terms and assess how much coverage glossary has over the data map





Purview Studio Management Center

Management Center

Provides ability to manage account information, scan rulesets, custom classifiers, integration runtimes & external connections

The screenshot shows the Microsoft Purview Management Center interface. At the top, there's a navigation bar with 'Microsoft Azure' and 'Purview' followed by a breadcrumb trail 'Adatum Corp'. On the right side of the header is a search bar labeled 'Search assets'. The main content area has a left sidebar under 'General' with icons for Home, Account information (selected), Data sources, Scan rule sets, Integration runtimes, Metrics, Metadata management, Classifications (highlighted with a red box), Classification rules, External connections, Data Factories, and Data Shares. The main panel is titled 'Account information' and contains an informational message about participating in the Purview Public Preview. Below that is an 'Overview' section with fields for 'Account name' (adatumcorpretail) and 'Friendly name *' (Adatum Corp). A 'Save' button is located to the right of the friendly name field.

Management Center – Account Information

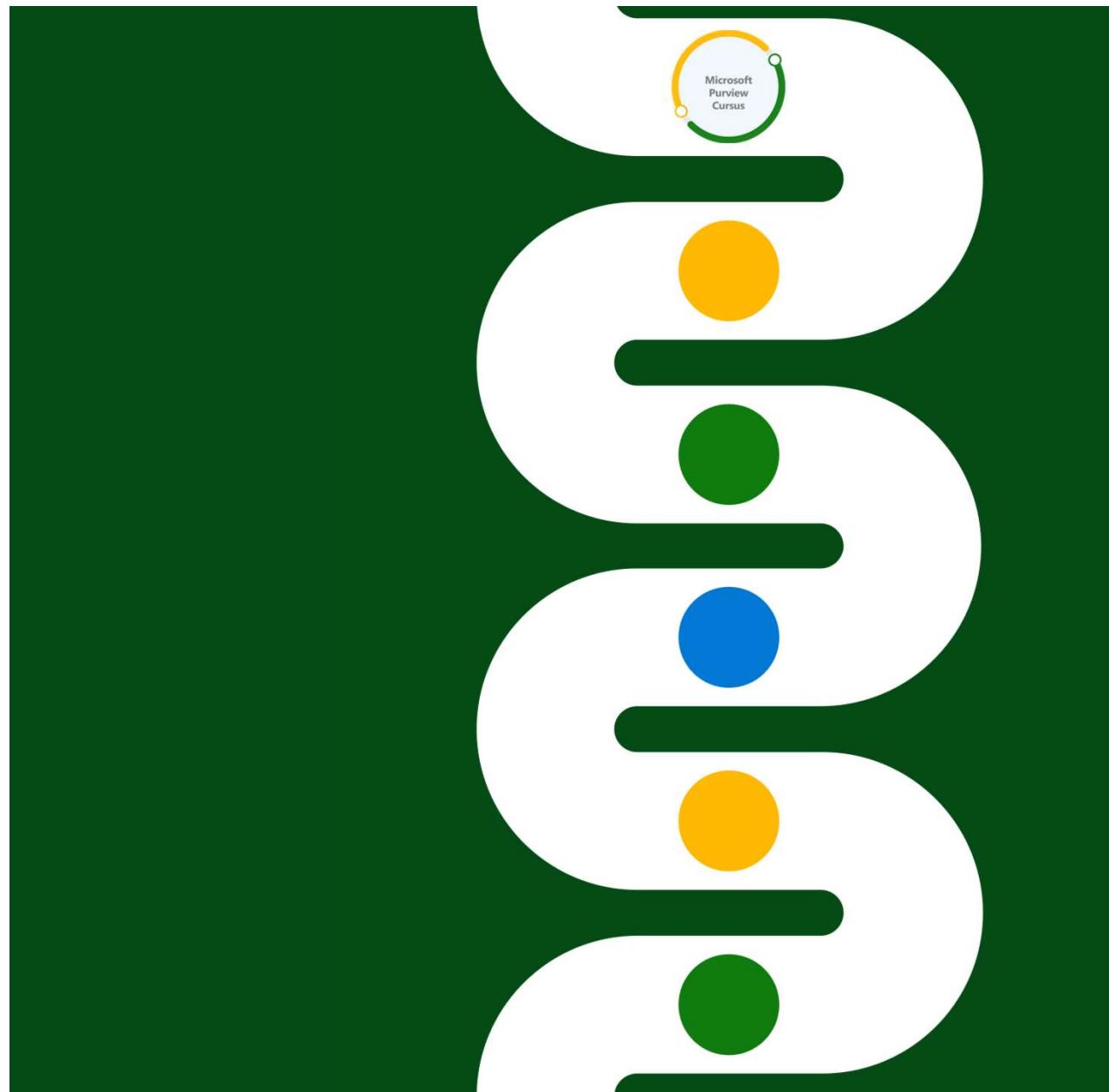
See Purview Account information and set a friendly name for Purview account

The screenshot shows the Microsoft Purview Management Center interface. The left sidebar has a 'General' category with several items: Overview (selected and highlighted with a red box), Metrics, Lineage connections, Data Factory, Data Share, Security and access, and Credentials. The main content area is titled 'Overview' and displays the following information:

- Account name:** adatumcorpretail
- Friendly name:** Adatum Corp (in a text input field)
- Save:** A button to save the changes.

The top navigation bar includes links for Microsoft Azure, Purview, and the specific account 'Adatum Corp'. It also features a search bar labeled 'Search assets', and icons for help, refresh, and user profile.

Microsoft Purview Data Governance in Azure REST API



Purview / Atlas API



TYPE SYSTEM
AN EXAMPLE: AZURE SQL TABLE

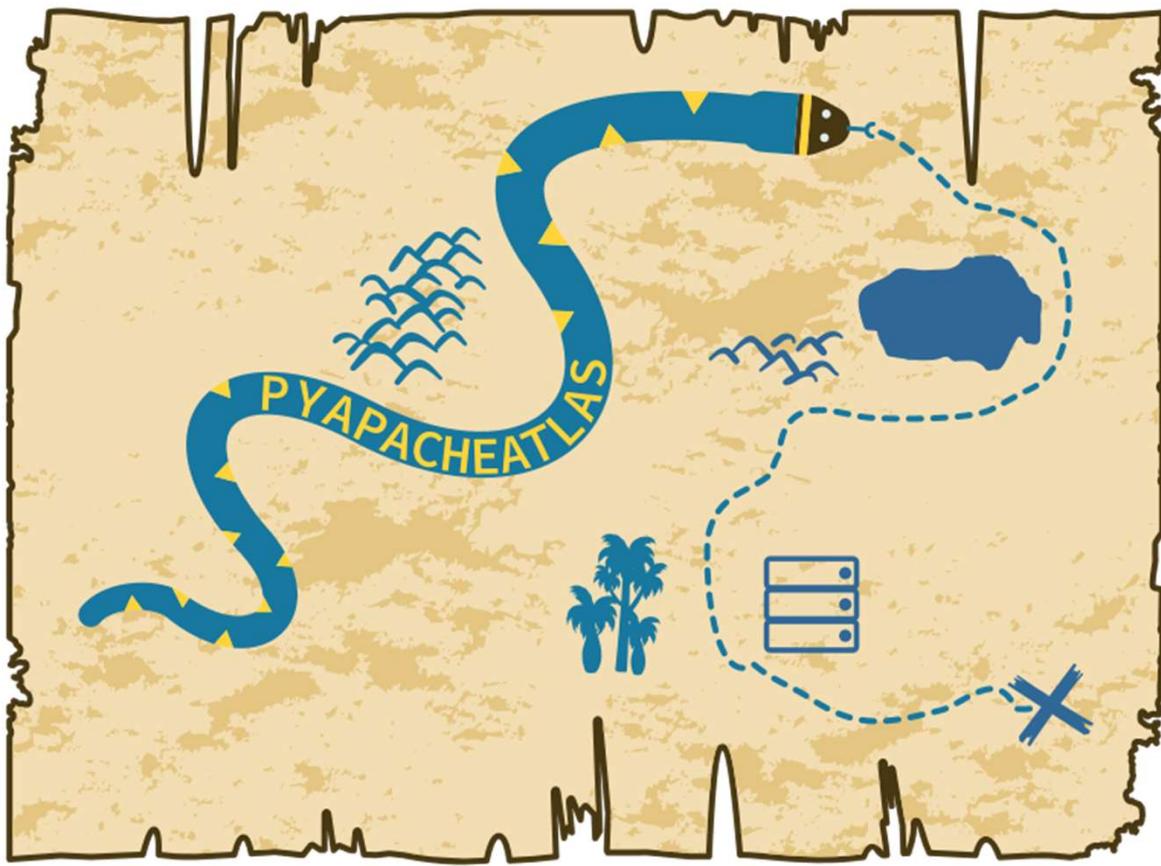


CUSTOM LINEAGE AND
RELATIONSHIPS



TIPS AND TRICKS:
RESOURCE SETS AND POWER BI

PyApacheAtlas



<https://github.com/wjohnson/pyapacheatlas>

<https://pypi.org/project/pyapacheatlas/>

<https://wjohnson.github.io/pyapacheatlas-docs/latest/>

Microsoft Purview Data Governance in Azure

- Microsoft Purview Data Governance in Azure is a unified data governance service that helps you manage and govern your on-premises, multi-cloud, and software-as-a-service (SaaS) data.
- The key component in Purview is 'Data Map'. Purview's data map provides foundation for data map for data discovery and effective data governance.
 - Data Map is auto-populated by Purview's built-in scanning infrastructure.
 - Business users can curate Data Map using interactive UI.
 - Developers can interact with data map programmatically using open-source Apache Atlas APIs.
 - Purview provides Atlas REST APIs. Atlas APIs provide a powerful and flexible mechanism handling both push and pull scenarios. Information can be published to Purview using Atlas APIs for bootstrapping or to push latest updates from another system into Purview. The information available in Purview can also be read using Atlas APIs and then synced back to existing products.

Apache Atlas

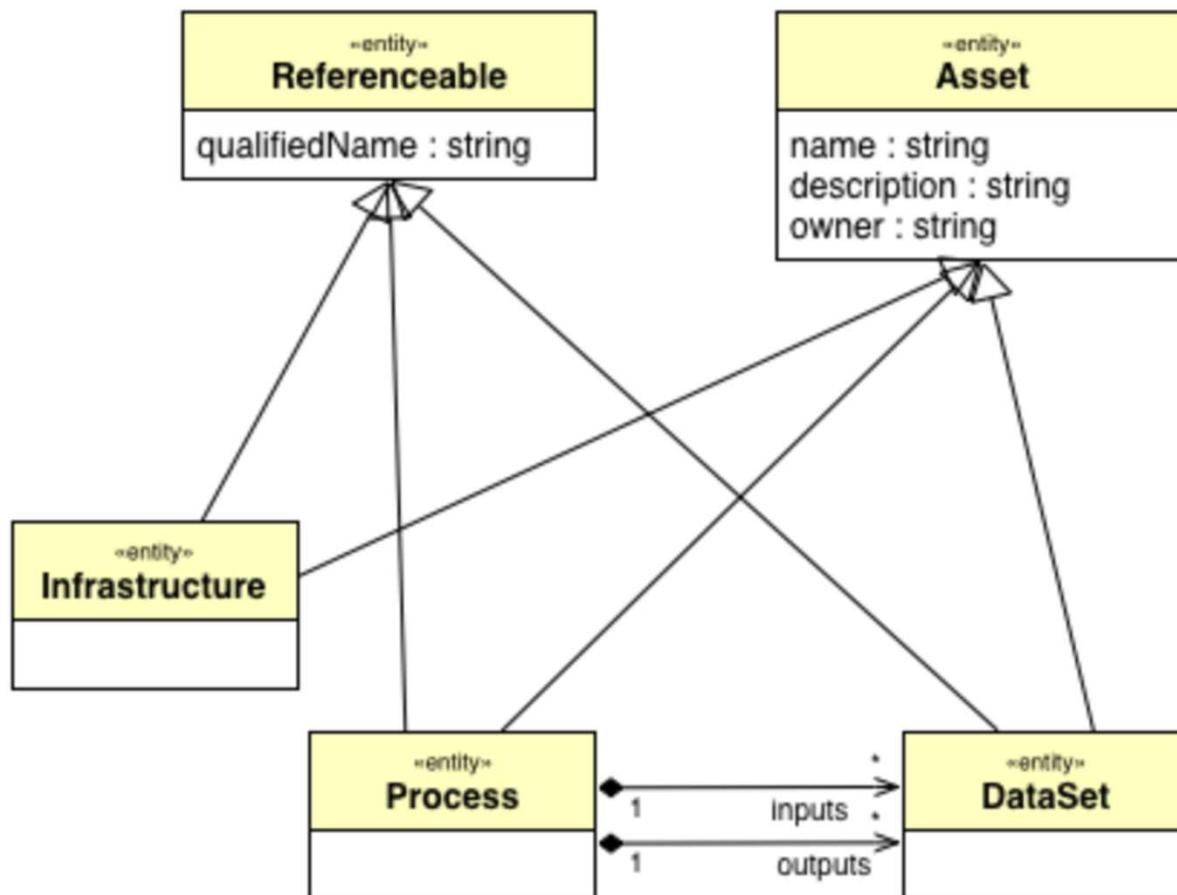
- [Atlas](#) is a scalable and extensible set of core foundational governance service.
- Apache Atlas provides open metadata management and governance capabilities for organizations to build a catalog of their data assets, classify and govern these assets and provide collaboration capabilities
- Apache Atlas is a Type System. It allows users to define and manage objects in Atlas.
 - Definitions are called Types. It represents how a metadata is stored and accessed in Atlas.
 - The actual object (files, tables etc.) or in other words Instances of Type are called entities.

Type System

Atlas Base Types

- Referenceable:
 - Represents entities.
 - Has attribute called "qualified Name":
- Asset:
 - Extends Referenceable
 - Has attributes like name, description and owner. Name is a required attribute the others are optional.
- DataSet:
 - Extends Asset. E.g., tables, files etc.
- Process:
 - Extends Asset.
 - Represents any data transformation operation.

Base Model



Apache Type System

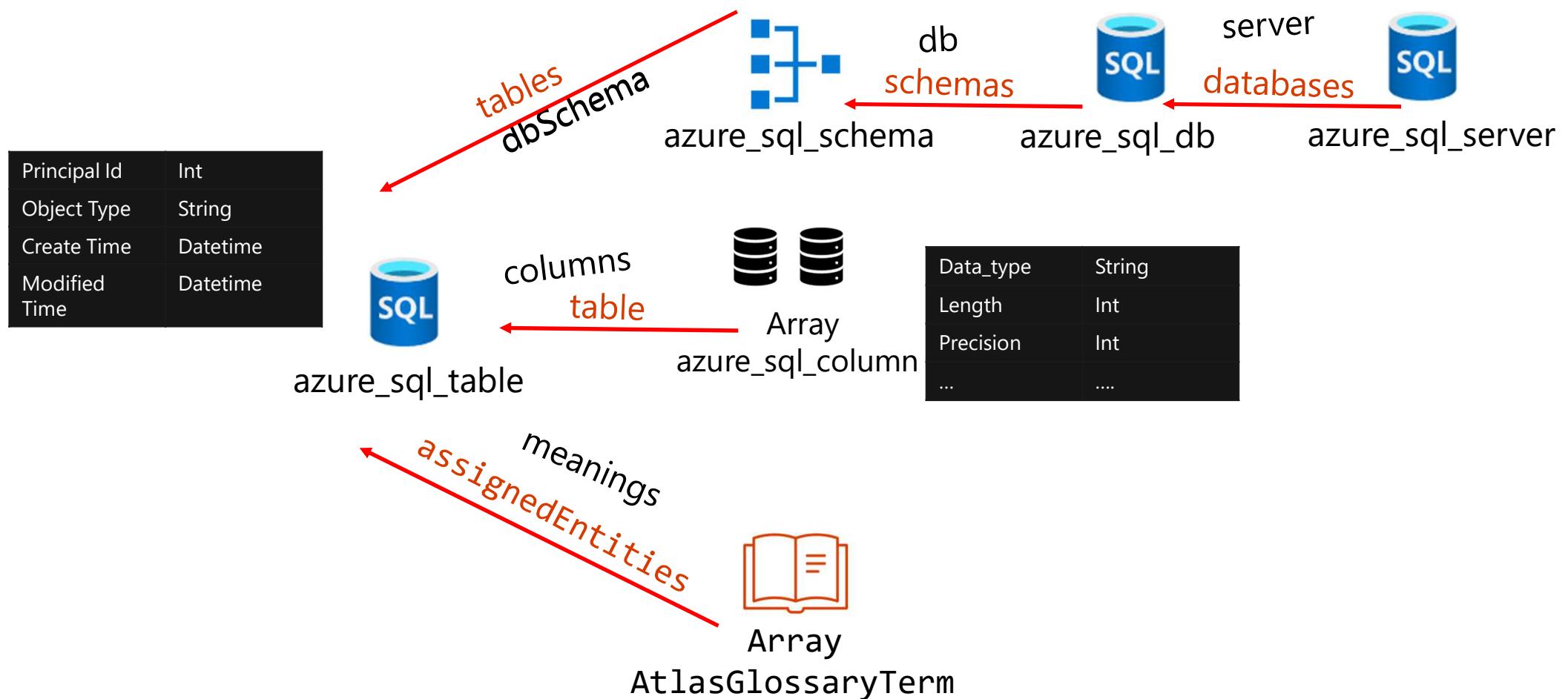
- **Entity** – Name and list of attributes (properties) for the metadata.
 - Can Inherit from other entity types
 - Attributes defines the data type and cardinality i.e., providing rules around the actual value.
 - GUID – Unique identifier generated by Atlas
- **Relationship** – Relationship type defines relationship between two entities.
- **Classification** – Used to classify entities
- **Struct** – Restrict data values
 - Does not support inheritance
 - Cannot exist independently
- **Enum** – Restrict values

Data Curation using Atlas APIs

- Types API
- Glossary APIs
- Entity APIs
- Discovery APIs
- Relationship APIs

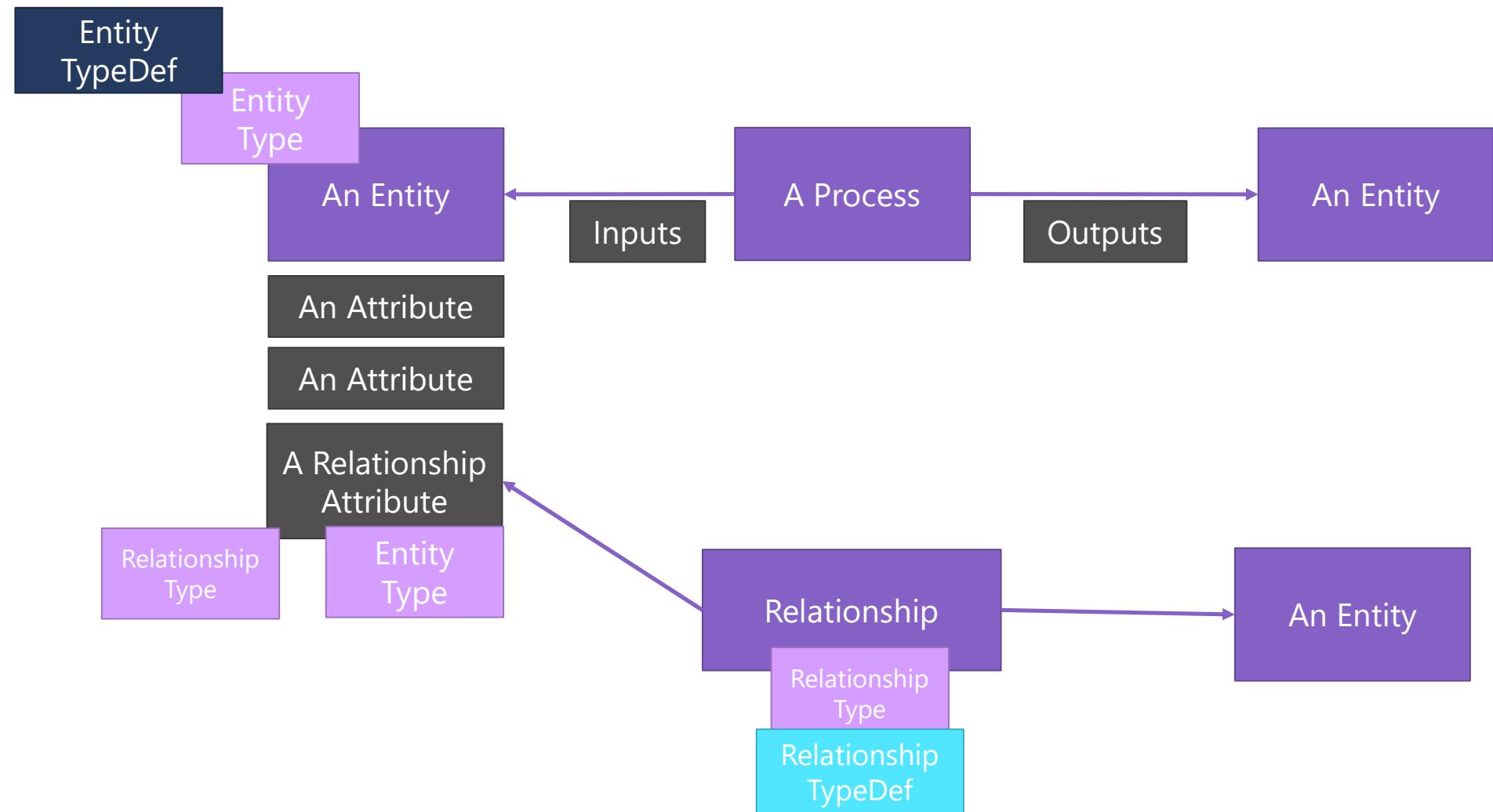
Type System Expanded

An Instance of Azure SQL Table and its Relationships



Type System Expanded

Most Common Patterns of Types and Entities



Azure SQL Table: A Type Definition

```
"category": "ENTITY",
"guid": "5f94b8b9-0430-4210-ade2-7b6f7e2d2db4",
"name": "azure_sql_table",
"description": "azure_sql_table",
"serviceType": "Azure SQL Database",
"options": {
    "schemaElementsAttribute": "columns"
},
"superTypes": ["DataSet"],
"subTypes": [],

"attributeDefs": [
    {"name": "principalId", ...},
    {"name": "objectType", ...},
    {"name": "createTime", ...},
    {"name": "modifiedTime", ...}
],
"relationshipAttributeDefs": [
    {
        "name": "dbSchema",
        "typeName": "azure_sql_schema",
        "isOptional": false,
        "relationshipTypeName": "azure_sql_schema_tables"
    },
    {
        "name": "columns",
        "typeName": "array<azure_sql_column>",
        "cardinality": "SET",
        "constraints": [{"type": "ownedRef"}],
        "relationshipTypeName": "azure_sql_table_columns"
    }
]
```

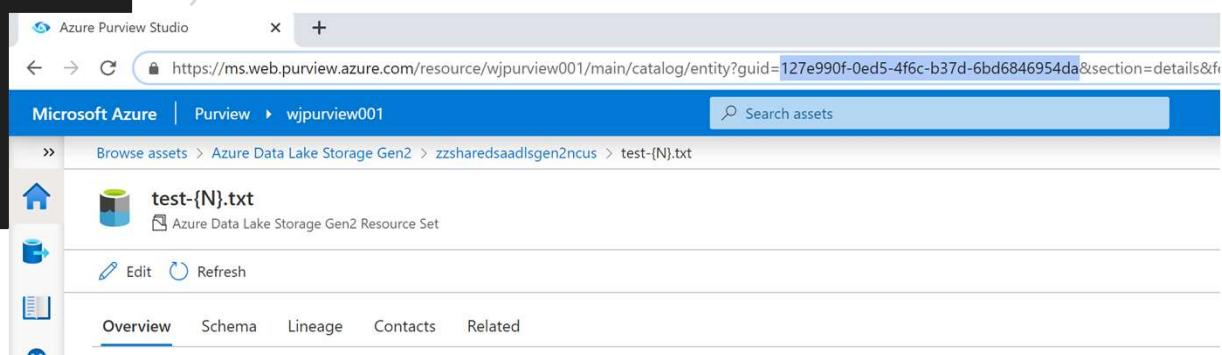
Exploring Types

```
1  {
2    "enumDefs": [ ...  
566  ],
567  "structDefs": [ ...  
1030  ],
1031  "classificationDefs": [ ...  
3121  ],
3122  "relationshipDefs": [ ...  
1279  ],
1280  "entityDefs": [  
1281    {  
1282      "category": "ENTITY",  
1283      "guid": "73afdd7c-dd88-4351-b81f-103a43ce8163",  
1284      "createdBy": "admin",  
1285      "updatedBy": "admin",  
1286      "createTime": 1604729949698,  
1287      "updateTime": 1604729949698,  
1288      "version": 1,  
1289      "name": "azure_mariadb_view_column",  
1290      "description": "azure_mariadb_view_column",  
1291      "typeVersion": "1.0",  
      "...  
    }  
  ]  
}
```

```
# Get all types  
client.get_all_typedefs()
```

```
# Get a single type  
client.get_typedef(  
  TypeCategory.ENTITY,  
  name="azure_sql_table"  
)
```

```
# Get one entity to understand its type  
entity = client.get_entity("abc-123-jkl")  
print(  
  json.dumps(entity["entities"], indent=2)  
)
```



Creating a Custom Type



```
from pyapacheatlas.core import AtlasAttributeDef, EntityTypeDef

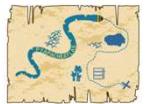
custom_type = EntityTypeDef(
    name="MyCustomType",
    superTypes=["DataSet"], # Default in PyApacheAtlas
)

custom_process = EntityTypeDef(
    name="MyCustomProcess",
    superTypes=["Process"],
)

custom_type.addAttributeDef(
    AtlasAttributeDef(name="attribute1", isOptional=False, typeName="int"),
    AtlasAttributeDef(name="attribute2", typeName="string")
)
client.upload_typedefs(
    entityDefs=[custom_type, custom_process],
    force_update=True
)
```

PUT/POST /api/atlas/v2/types/typedefs

Creating an Entity



```
from pyapacheatlas.core import AtlasEntity
from pyapacheatlas.core.util import GuidTracker

gt = GuidTracker()
ae = AtlasEntity(
    name="myentity", # Analogous to the display name / searchable name
    typeName="MyCustomType", # The type of entity you're uploading
    qualified_name="Custom://myentity", # Must be a unique name across Purview
    guid=gt.get_guid(), # Get a unique negative number for this upload
    # Add required or optional attributes
    attributes = {"attribute1":123, "attribute2":"abc"}
)
client.upload_entities(batch = [ae])
```

POST /api/atlas/v2/entity/bulk

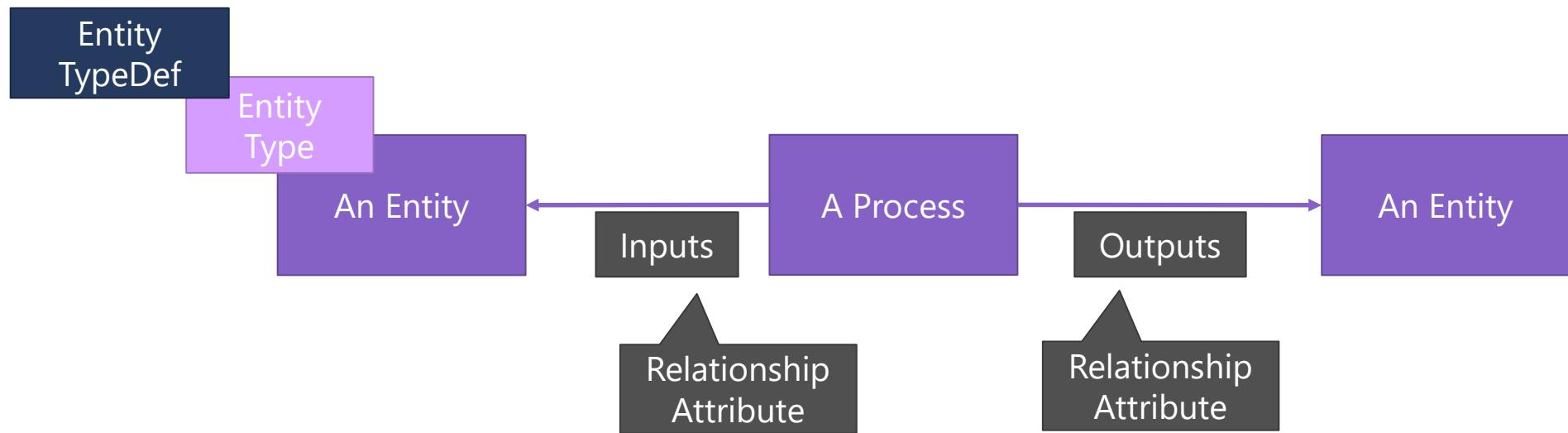
```
client.get_entity(guid = [...])
```

GET /api/atlas/v2/entity/bulk

Lineage and Relationships

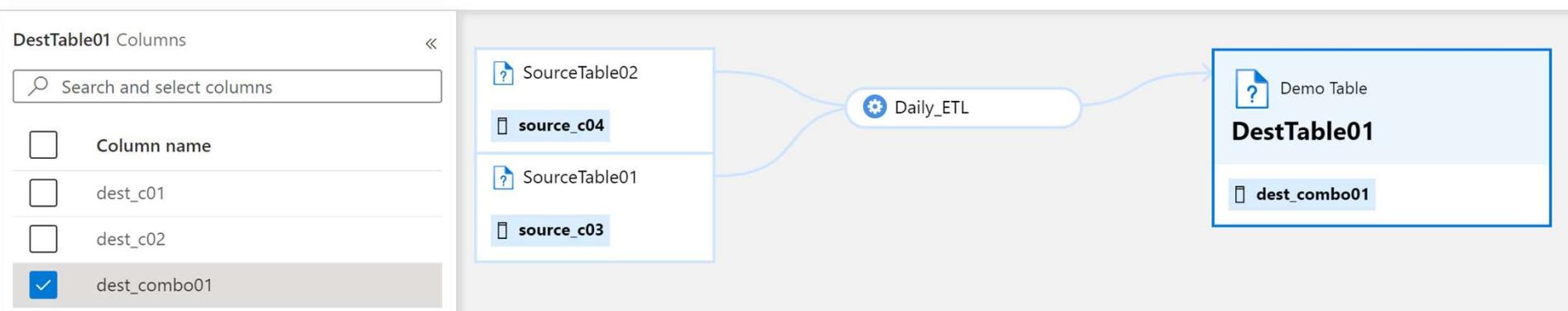
Type System Expanded

Lineage Overview



Lineage in Purview UI

Overview Schema **Lineage** Contacts Related

 Search for assets or processes

Custom Lineage

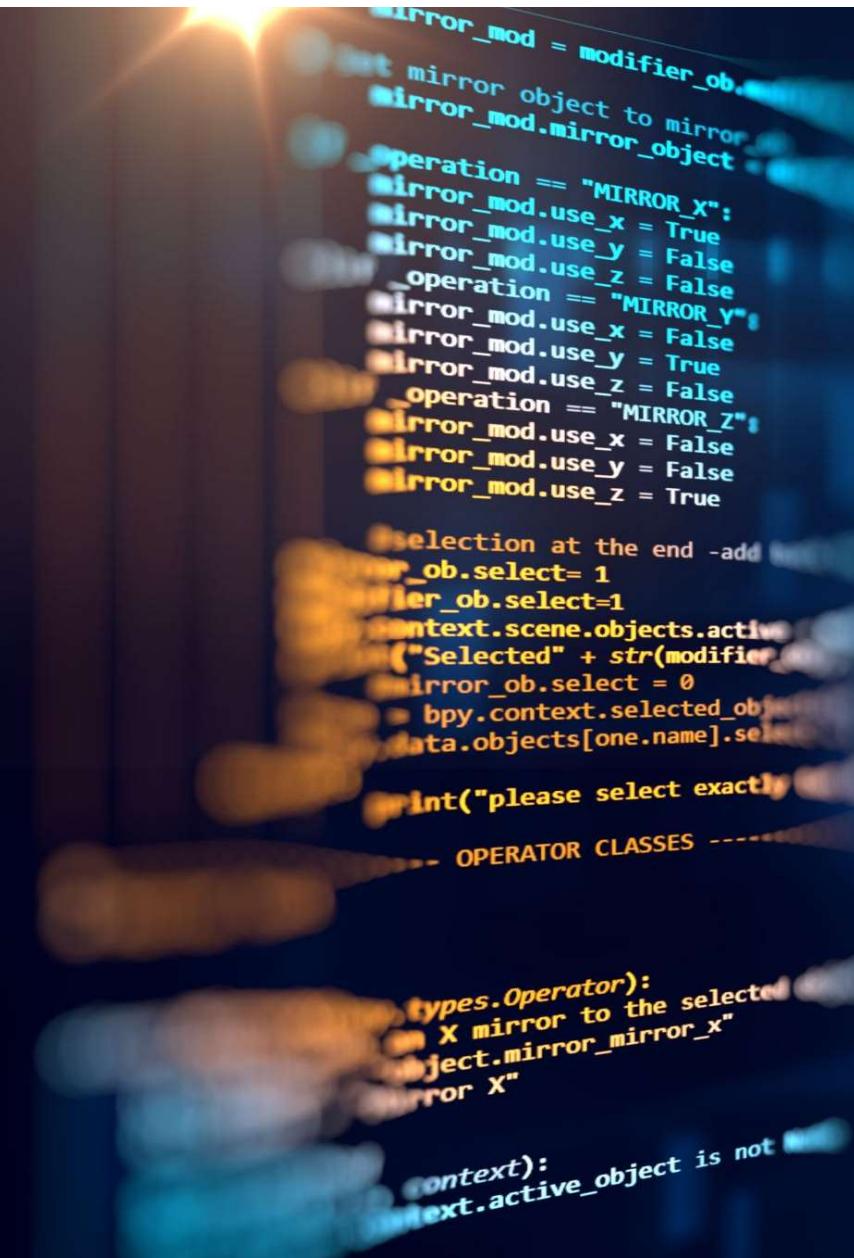


```
input01 = AtlasEntity(  
    name="input01",  
    typeName="DataSet",  
    qualified_name="pyapacheatlas://demoinput01",  
    guid=-100  
)  
output01 = AtlasEntity(  
    name="output01",  
    typeName="DataSet",  
    qualified_name="pyapacheatlas://demooutput01",  
    guid=-101  
)  
  
# The Atlas Process is the lineage component that links the two  
# entities together. The inputs and outputs need to be the "header"  
# version of the atlas entities, so specify minimum = True to  
# return just guid, qualifiedName, and typeName.  
process = AtlasProcess(  
    name="sample process",  
    typeName="Process",  
    qualified_name="pyapacheatlas://democustomprocess",  
    inputs=[input01],  
    outputs=[output01],  
    guid=-102  
)  
  
# Batch upload the entities!  
client.upload_entities(  
    batch=[output01, input01, process]  
)  
  
# If you know the guids, you can pass them as guid dicts  
# No need to instantiate "dummy" entities for updates.  
inputs = [{"guid": "aeb02811-1f89-4d3f-98b9-5e7bda050422"}],  
outputs= [{"guid": "xyz-123-4d3f-98b9-5e7bda050422"}]
```

https://github.com/wjohnson/pyapacheatlas/blob/master/samples/CRUD/create_entity_and_lineage.py

Demo Custom Entities

[Purview - Custom Process](#)



```
mirror_mod = modifier_obj.modifiers["Mirror"]
# set mirror object to mirror
# mirror_mod.mirror_object = mirror_obj
if mirror_mod.mirror_object == None:
    print("no mirror object selected")
else:
    if mirror_mod.operation == "MIRROR_X":
        mirror_mod.use_x = True
        mirror_mod.use_y = False
        mirror_mod.use_z = False
    elif mirror_mod.operation == "MIRROR_Y":
        mirror_mod.use_x = False
        mirror_mod.use_y = True
        mirror_mod.use_z = False
    elif mirror_mod.operation == "MIRROR_Z":
        mirror_mod.use_x = False
        mirror_mod.use_y = False
        mirror_mod.use_z = True

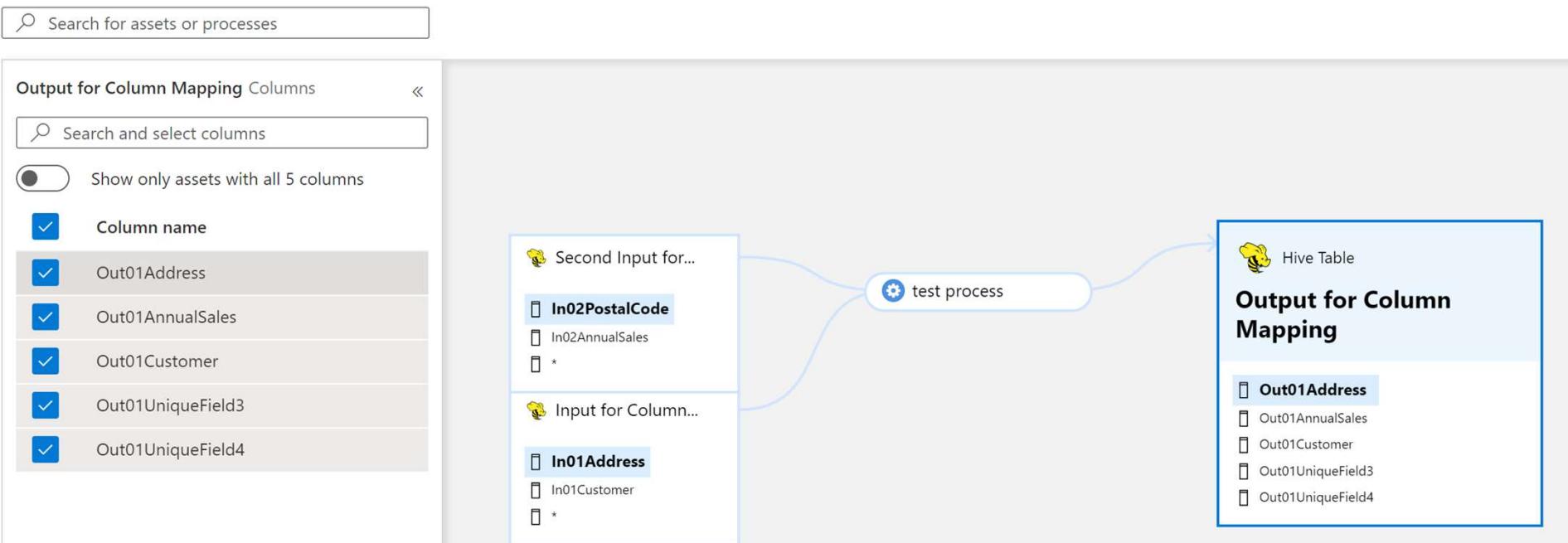
#selection at the end -add
#mirror_mod.select= 1
#mirror_mod.select=1
context.scene.objects.active = bpy.context.selected_objects[-1]
print("Selected" + str(modifier_obj))
mirror_ob.select = 0
bpy.context.selected_objects[0].select = 1
data.objects[one.name].select = 1
print("please select exactly one object")

- OPERATOR CLASSES -
# Mirror Operator
class MirrorOperator(bpy.types.Operator):
    bl_idname = "object.mirror"
    bl_label = "X mirror to the selected object.mirror_mirror_x"
    bl_options = {'REGISTER', 'UNDO'}
    bl_description = "X mirror to the selected object.mirror_mirror_x"

    def execute(self, context):
        if context.active_object == None:
            print("no active object selected")
        else:
            if context.active_object.type != "MESH":
```

Column Mapping in Lineage UI

Overview Schema **Lineage** Contacts Related



Column Mapping



```
column_mapping = [
    # This object defines the column lineage between input01 and output01
    {"ColumnMapping": [
        {"Source": "In01Address", "Sink": "Out01Address"},           ColumnMappings: List of dicts with source/sink column names
        {"Source": "In01Customer", "Sink": "Out01Customer"}],
     "DatasetMapping": {
        "Source": colMapInput01.qualifiedName, "Sink": colMapOutput01.qualifiedName}          DatasetMapping: Dict with source/sink qualified name.
    },
    # This object defines the column lineage between input02 and output01
    {"ColumnMapping": [
        {"Source": "In02AnnualSales", "Sink": "Out01AnnualSales"},           ColumnMappings: List of dicts with source/sink column names
        {"Source": "In02PostalCode", "Sink": "Out01Address"}],
     "DatasetMapping": {"Source": colMapInput02.qualifiedName, "Sink": colMapOutput01.qualifiedName}          DatasetMapping: Dict with source/sink qualified name.
    }
]

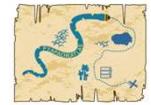
# Create the process with the stringified column mapping json.
process = AtlasProcess(
    name="test process",                                         Needs a Process Type with columnMapping attribute definition
    typeName="ProcessWithColumnMapping",
    qualified_name="pyapacheatlas://colMapOutputProcessDemo",
    inputs=[colMapInput01, colMapInput02],
    outputs=[colMapOutput01],
    guid=gt.get_guid(),
    attributes={"columnMapping": json.dumps(column_mapping)}          Attribute's value is 'stringified' json
)
```

Defining Relationship Types



```
rd = RelationshipTypeDef(  
    name="azure_sql_table_columns",  
    attributeDefs=[],  
    relationshipCategory="COMPOSITION", # Means the child can't exist without the parent  
endDef1={ # End Def 1 describes what the parent will have as an attribute  
    "type": "azure_sql_table",  
    "name": "columns",  
    "isContainer": True,  
    "cardinality": "SET",  
    "isLegacyAttribute": False  
},  
endDef2={# End Def 2 describes what the children will have as an attribute  
    "type": "azure_sql_column",  
    "name": "table",  
    "isContainer": False,  
    "cardinality": "SINGLE",  
    "isLegacyAttribute": False  
}  
client.upload_typedefs(relationshipDefs=[rd])
```

PUT/POST /api/atlas/v2/types/typedefs



Assigning Terms: Using Relationships

```
relationship = {  
    "typeName": "AtlasGlossarySemanticAssignment",  
    "attributes": {},  
    "end1": {  
        "guid": TermId,  
    },  
    "end2": {  
        "typeName": "YourAssetsType",  
        "uniqueAttributes":{  
            "qualifiedName": "YourAssetsQualifiedName"}  
    }  
}  
  
try:  
    # Attempt the upload, but the relationship may already exist  
    # So simply skip log it and skip past it.  
    results = client.upload_relationship(relationship)  
except AtlasException as e:  
    print(f"\tException for {TermId} and was not uploaded: {e}")
```

POST /api/atlas/v2/relationship

https://github.com/wjohnson/pyapacheatlas/blob/master/samples/CRUD/update_entities_with_term.py

Working with Schemas



```
# Creating the entities that will be used in uploads.  
# One table will be added  
table = AtlasEntity("rel10","hive_table", "tests://rel10", guid=-1)  
# Four columns will be added  
c1 = AtlasEntity("rel10#01", "hive_column", "tests://rel10#c", guid=-2, attributes={"type":"str"})  
c2 = AtlasEntity("rel10#02", "hive_column", "tests://rel02#c", guid=-3, attributes={"type":"str"})  
c3 = AtlasEntity("rel10#03", "hive_column", "tests://rel03#c", guid=-4, attributes={"type":"str"})  
c4 = AtlasEntity("rel10#04", "hive_column", "tests://rel04#c", guid=-5, attributes={"type":"str"})  
  
# Add relationships to the columns from the table overwriting existing columns  
# Good if you want to overwrite existing schema or creating a brand new table  
# and Schema.  
columns_to_add = [ c1, c2, c3 ]  
# Use a list comprehension to convert them into dictionaries when adding a list  
table.addRelationship(columns=[c.to_json(minimum=True) for c in columns_to_add])  
  
# OR Add a table relationship to a column. This lets you essentially APPEND  
# a column to a table's schema.  
c4.addRelationship(table = table)  
  
client.upload_entities([table, c1, c2, c3, c4])
```

Excel Interface

Using Excel to Upload Entities / Lineage



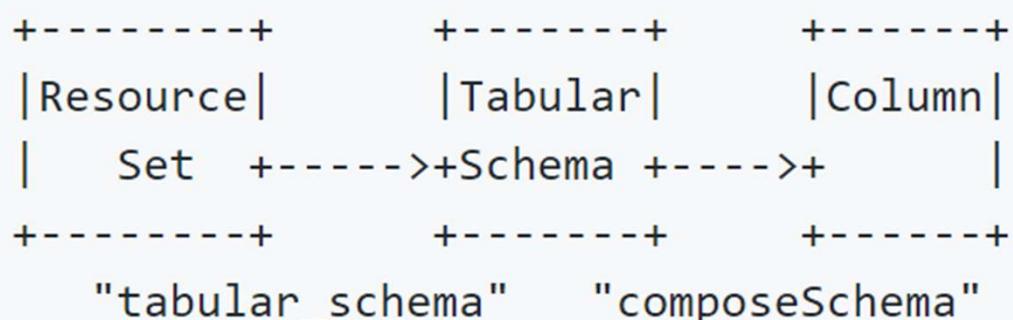
```
bulk_entities = excel_reader.parse_update_lineage(file_path)
bulk_results = client.upload_entities(bulk_entities)
lineage_processes = excel_reader.parse_update_lineage(file_path)
lineage_results = client.upload_entities(lineage_processes)
```

Header	Value
Target typeName	hive_table
Target qualifiedName	pyapacheatlas://demo_update_lineage_output
Source typeName	hive_table
Source qualifiedName	pyapacheatlas://demo_update_lineage_input
Process name	demo_some_custom_hive_query
Process qualifiedName	pyapacheatlas://process_update_lineage_custom_hive_quer
Process typeName	Process

typeName	name	qualifiedName	classifications	[Relationship] table	type
DataSet	exampledataset	pyapacheatlas://dataset			
hive_table	hivetable01	pyapacheatlas://hivetable01			
hive_column	columnA	pyapacheatlas://hivetable01#colA	pyapacheatlas://hivetable01		string
hive_column	columnB	pyapacheatlas://hivetable01#colB	pyapacheatlas://hivetable01		long
hive_column	columnC	pyapacheatlas://hivetable01#colC	pyapacheatlas://hivetable01		int

Tips and Tricks

Tips and Tricks: Tabular Schema and Resource Sets



Updating Resource Set Columns:

1. Find Resource Set
2. RelationshipAttribute: Tabular Schema
3. Get Tabular Schema Entity
4. RelationshipAttribute composeSchema
 1. Update individual Column entities
 2. Add new Column entities as new Relationships

Tips and Tricks: PBI Lineage

```
LINEAGE_TO_CHANGE_TYPE = "powerbi_dataset_process"
dummy_existing_process = AtlasProcess(
    name=LINEAGE_TO_CHANGE_NAME,
    typeName=LINEAGE_TO_CHANGE_TYPE,
    qualified_name=LINEAGE_TO_CHANGE_QUALIFIED_NAME,
    inputs=None, # Set to None so no update will occur
    outputs=None, # We will update this with set_outputs below
    guid=-104
)

real_existing_process = client.get_entity(
    typeName=LINEAGE_TO_CHANGE_TYPE,
    qualifiedName=LINEAGE_TO_CHANGE_QUALIFIED_NAME
)["entities"][0]

# Create one more output to be added.
one_more_input = AtlasEntity( ... )
# Add the existing and new output to the dummy process
dummy_existing_process.inputs = ([one_more_input.to_json(minimum=True)] +
    real_existing_process["entities"]["inputs"] )
complex_results = client.upload_entities(
    batch=[dummy_existing_process, one_more_input]
)
```

<https://github.com/wjohnson/pyapacheatlas/wiki/Azure-Purview-Tips#update-powerbi-lineage>

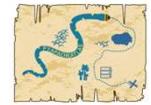
Frequently Asked API Questions

Frequently Asked API Questions

- **Charges for the API**
 - Underlying Event Hub with 1 Standard Tier (~\$USD 30/month)
 - No additional charges for working with REST API
- **Throttling**
 - Underlying Event Hub architecture (1 MB/s per TU) ([docs](#))

Appendix

Setting up PyApacheAtlas



- Requires Python 3.6 or higher
- Install using...
`pip install pyapacheatlas`
- Follow the instructions in the [Purview REST API tutorial](#).
 - Create a Service Principal. Collect Tenant Id, Client Id, and Client Secret
 - Assign the Service Principal the "Purview Data Curator" role
- In the sample Python scripts the tenant id, client id and secret, and the purview service name are saved as environment variables but could be hard coded (in the quotes).
- Use the PurviewClient and ServicePrincipalAuthentication classes.

```
from pyapacheatlas.auth import ServicePrincipalAuthentication
from pyapacheatlas.core.client import PurviewClient

oauth = ServicePrincipalAuthentication(
    tenant_id=os.environ.get("TENANT_ID", ""),
    client_id=os.environ.get("CLIENT_ID", ""),
    client_secret=os.environ.get("CLIENT_SECRET", ""))
)
client = PurviewClient(
    account_name = os.environ.get("PURVIEW_NAME", ""),
    authentication=oauth
)
```

Where to Learn More

- [PyApacheAtlas Github](#)
- [Purview REST API tutorial](#)
 - Includes the most up-to-date link to the Swagger API.
- [Apache Atlas documentation](#)
 - Caution: Not all the endpoints are implemented in the same way or at all as in Purview.
 - Defer to the Purview Swagger for most accurate information.
- **Additional Tools**
 - [PurviewCLI Package](#) supports command line tooling and scans API.

Azure Purview CLI

purviewcli

Azure Purview

Interface



Purview Studio

Classifications | Glossary | Insights | Lineage | Search | Sensitivity Labels | Sources



purviewcli

Platform (REST API Endpoints)



Apache Atlas 2.0

catalog

Entity | Glossary | Lineage | Relationship | Types | Search



scan

Collections | Sources | Scans | Scan Rule Sets



Data Insights

guardian

Assets | Scans | Glossary | Classification
Sensitivity Labels | File Extensions

Azure Purview High-Level Concepts

What is purviewcli?

`purviewcli` is a Python package that provides a cross-platform command line interface (CLI) to Azure Purview. Suitable for ad-hoc tasks and quick exploratory operations.

Installation

```
pip install purviewcli
```

Python Package Index (PyPi) <https://pypi.org/project/purviewcli/>

Environment Variables

Environment Variable	Description
PURVIEW_NAME (★)	Azure Purview account name.
AZURE_CLIENT_ID	The Azure AD tenant(directory) ID.
AZURE_TENANT_ID	Client(application) ID of an App Registration.
AZURE_CLIENT_SECRET	Client secret generated for the App Registration.
AZURE_CLIENT_CERTIFICATE_PATH	A path to certificate and private key pair.
AZURE_USERNAME	The username(upn) of an Azure AD user account.
AZURE_PASSWORD	The password of the Azure AD user account.

★ PURVIEW_NAME is mandatory, all other environment variables are optional.

Authentication

The `purviewcli` package leverages `DefaultAzureCredential` from [azure-identity](#). `purviewcli` will attempt to authenticate via the following mechanisms in this order, stopping when one succeeds.

Mechanism	Description
Environment	If account information has been specified via Environment Variables, <code>purviewcli</code> will use it to authenticate.
Managed Identity	If deployed to an Azure host with Managed Identity enabled, <code>purviewcli</code> will authenticate with it.
Visual Studio Code	If a user has signed in to the Visual Studio Code Azure Account extension, <code>purviewcli</code> will authenticate as that user.
Azure CLI	If a user has signed in via the Azure CLI <code>az login</code> command, <code>purviewcli</code> will authenticate as that user.

Authorization

The identity executing Azure Purview CLI commands will require the following role assignments:

Role	Description
Purview Data Curator	Has access to the Purview portal and can read all content in Azure Purview except for scan bindings, can edit information about assets, can edit classification definitions and glossary terms, and can apply classifications and glossary terms to assets..
Purview Data Source Administrator	Does not have access to the Purview Portal (the user needs to also be in the Data Reader or Data Curator roles) and can manage all aspects of scanning data into Azure Purview but does not have read or write access to content in Azure Purview beyond those related to scanning.

Commands

Commands	Description
pv entity	Entities are a collection of attributes that model or represent an asset.
pv glossary	A glossary provides vocabulary for business users.
pv lineage	Understand the data's origin and where it moves over time.
pv relationship	Relationships describe connections between two entities.
pv types	A Type is a definition of how a particular object type is stored.
pv scan	Azure Purview scan.
pv insight	Azure Purview insights.
pv search	Azure Purview advanced search.

```
pv --help
```

Usage

```
pv command sub-command --parameter1='value' --parameter2='value'
```

- All parameters are required by default.
- Parameters enclosed with square brackets "[]" are optional.
- Mutually exclusive parameters are enclosed with parens "()" and separated with a pipe "|".
- The "= <val>" indicates parameters which require an input (e.g. --parameter=<val>). Input can be specified after a space (e.g. --parameter 'value') or equal "=" sign (e.g. --parameter='value').
- Parameters that do not require an input are boolean operators and are **False** by default and **True** if present (e.g. --ignoreRelationships).
- The ellipsis "..." indicates parameters that are allowed to repeat (e.g. --guid='12345' --guid='23451' --guid='34512')

Example

An example of how you can get started with `purviewcli` within a Jupyter Notebook.

1. Install `purviewcli`
2. Set Environment Variables
3. Execute Command

```
In [21]: # 1. Install purviewcli package
%pip install --upgrade purviewcli

In [20]: # 2. Set up environment variables
import os
os.environ['PURVIEW_NAME']="purview_account_name"
os.environ['AZURE_CLIENT_ID']="client_id"
os.environ['AZURE_TENANT_ID']="tenant_id"
os.environ['AZURE_CLIENT_SECRET']="client_secret"

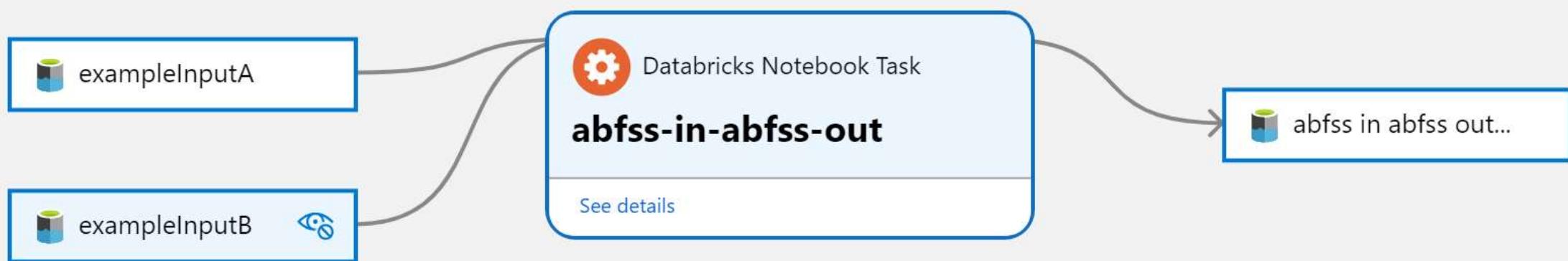
In [22]: # 3. Execute command
!pv glossary read
[{"id": "b5544147-c1c5-b835-d3e93e7eaae2", "name": "Glossary", "qualifiedName": "Glossary", "description": "This is a long description.", "language": "Spanish", "lastModifiedTS": "4", "longDescription": "This is a long description.", "terms": [{"termGuid": "6eb3cf63-1ed6-4c64-934f-e38ce2c2ffec", "displayText": "MyTerm", "relationGuid": "03752bbc-57f0-47d0-a685-3620bb73a80b"}]}
```

Notebook Samples:

<https://github.com/tayganr/purviewcli/tree/master/samples>



Azure Databricks to Purview Lineage Connector Solution Accelerator



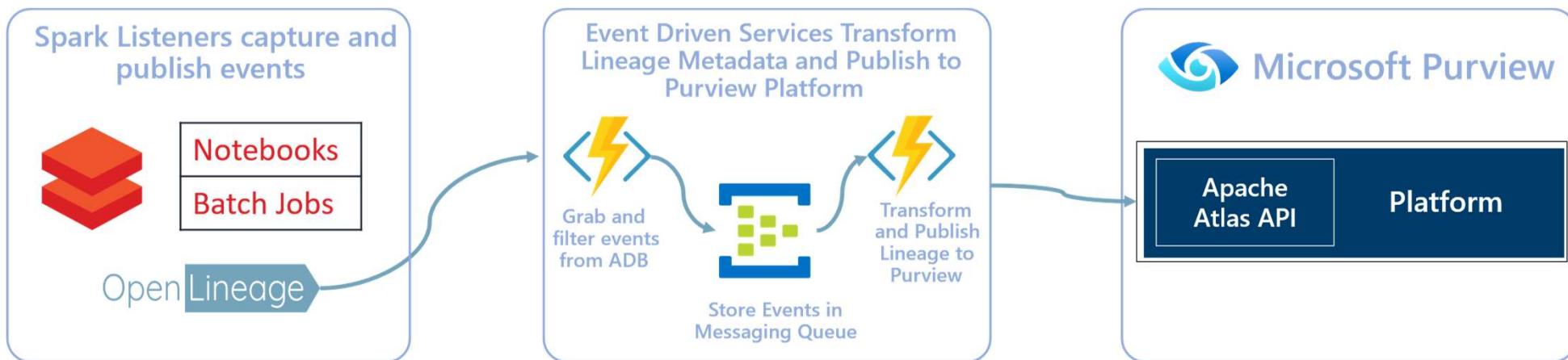
NOTE: In addition to [this solution accelerator](#), Microsoft Purview is creating native models for Azure Databricks (e.g.: Notebooks, jobs, job tasks...) to integrate with Catalog experiences. With native models in Microsoft Purview for Azure Databricks, customers will get enriched experiences such as search, browse, hierarchies, collection etc. If you choose to use this solution accelerator in a Microsoft Purview account before the native models are released, these enriched experiences are not backward compatible. Current models will be released shortly, please reach out to your Microsoft account representative for timeline related questions on the native model support for Azure Databricks in Microsoft Purview.



Azure Databricks to Purview Lineage Connector Solution Accelerator



Gathering lineage data is performed in the following steps:



1. Azure Databricks clusters are configured to initialize the OpenLineage Spark Listener with an endpoint to receive data.
2. Spark operations will output data in a standard OpenLineage format to the endpoint configured in the cluster.
3. Endpoint provided by an Azure Function app that will filter incoming data and pass it to an Azure EventHub.
4. Events are captured by a second Function app to transform the data into a format compatible with Atlas and Purview.
5. Lineage data is synchronized with existing Purview metadata and uploaded to Purview using standard Apache Atlas APIs.

[Azure Databricks to Purview Lineage Connector - Code Samples | Microsoft Docs](#)

DEMO

There are two deployment options for this solution accelerator:

- **Demo Deployment**

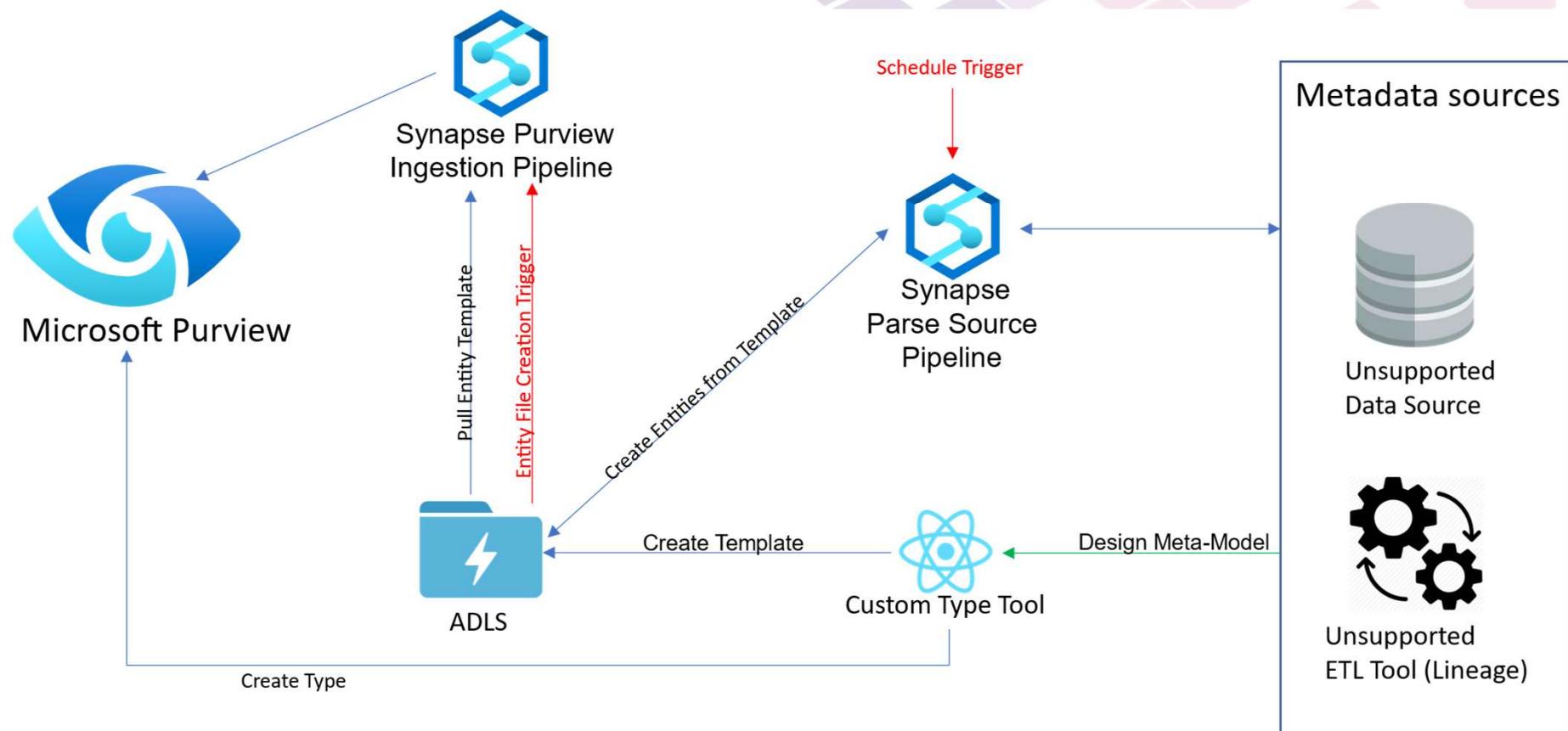
No additional prerequisites are necessary as the demo environment will be setup for you, including Azure Databricks, Purview, ADLS, and example data sources and notebooks.

- **Connector Only Deployment**

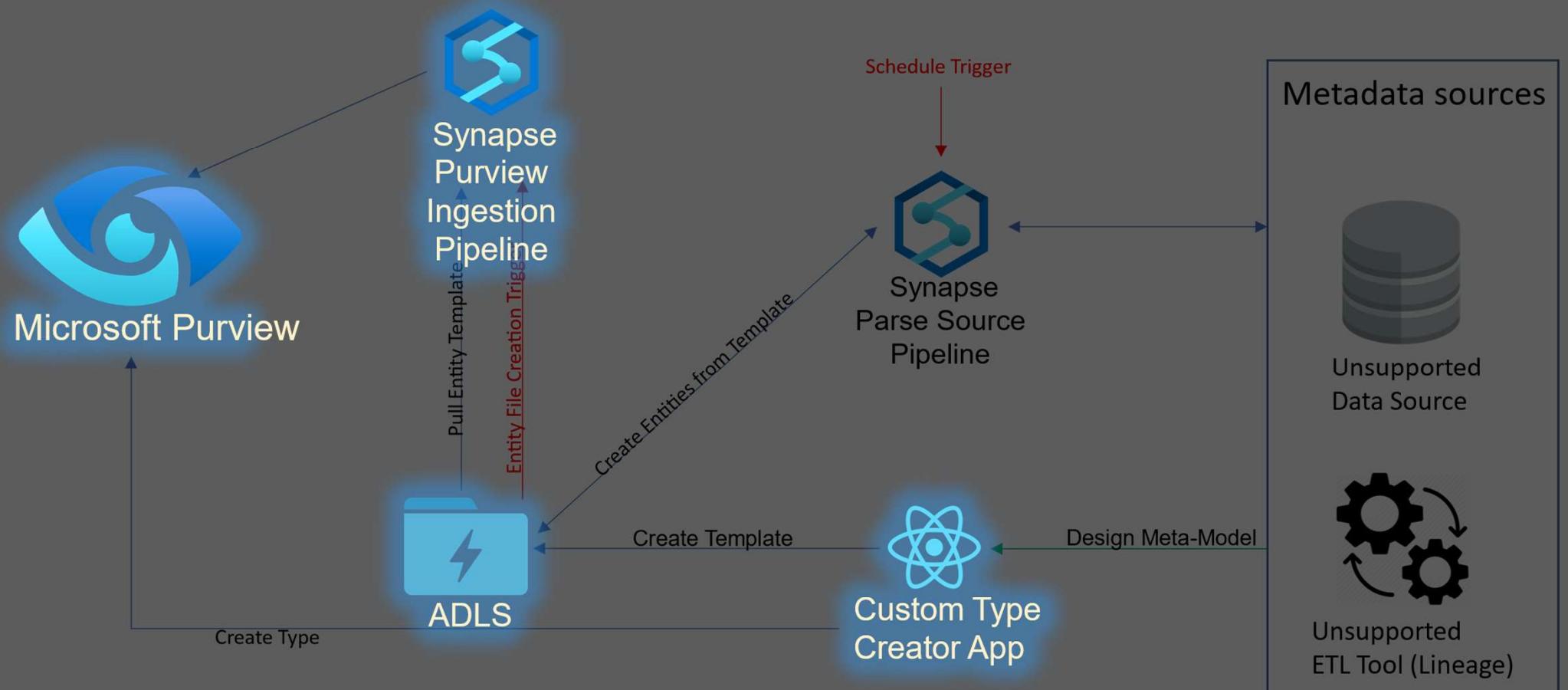
If installed as a working connector, Azure Databricks, data sources, and Microsoft Purview are assumed to be setup and running.

Purview Custom Connector Solution Accelerator

EARLY ACCESS ENGINEERING

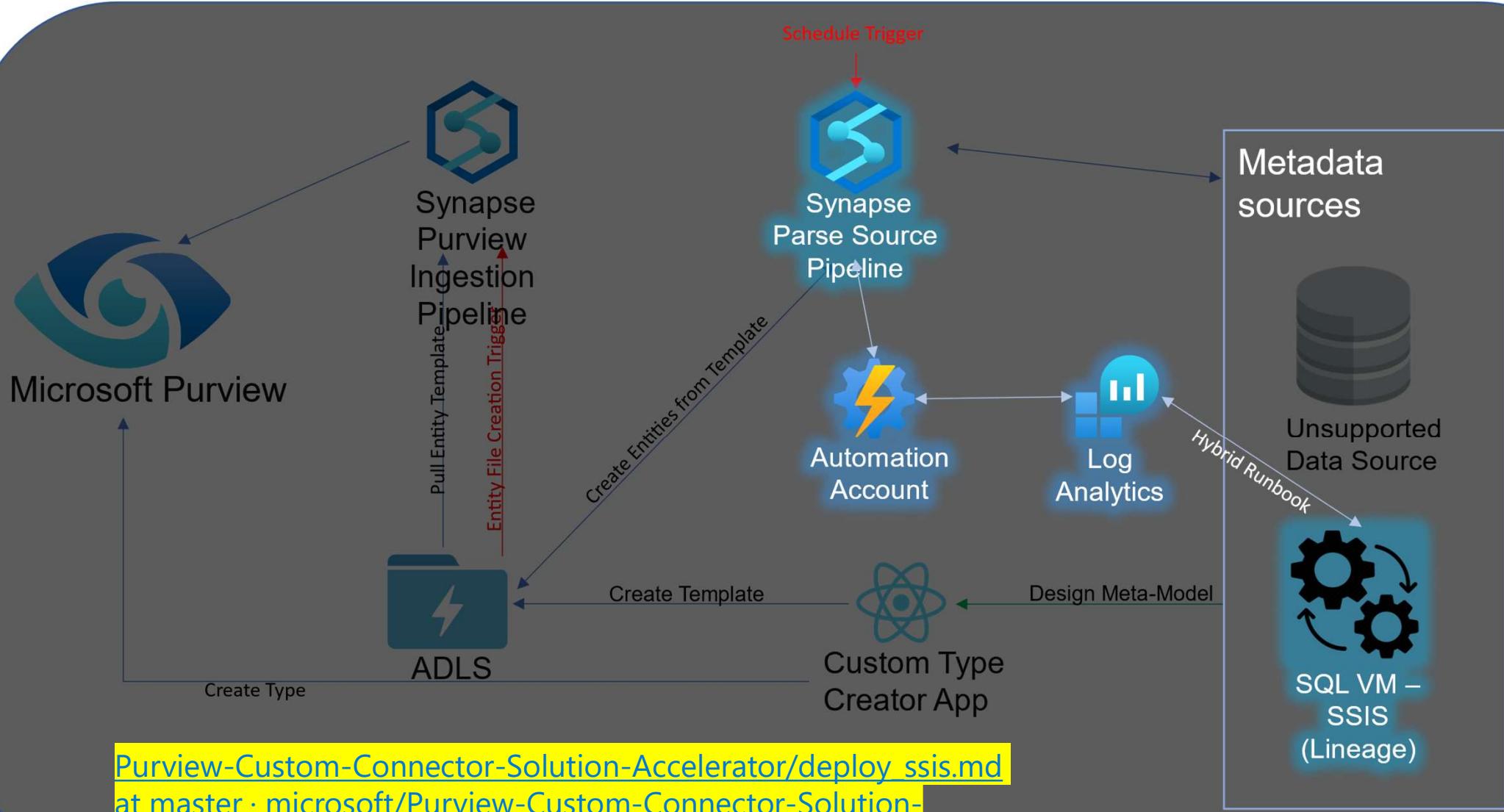


Demo - Deploy Tag Database Source Example



[Purview-Custom-Connector-Solution-Accelerator/deploy_tag_db.md at master · microsoft/Purview-Custom-Connector-Solution-Accelerator \(github.com\)](https://github.com/microsoft/Purview-Custom-Connector-Solution-Accelerator/blob/master/microsoft/Purview-Custom-Connector-Solution-Accelerator.github.com)

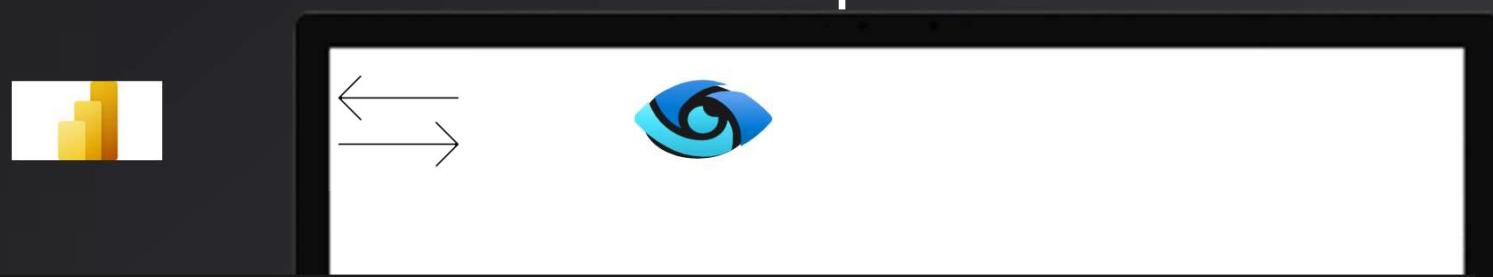
DEMO - Deploy ETL Lineage Example (SSIS)



Microsoft Purview Data Governance in Azure and enterprise services integration

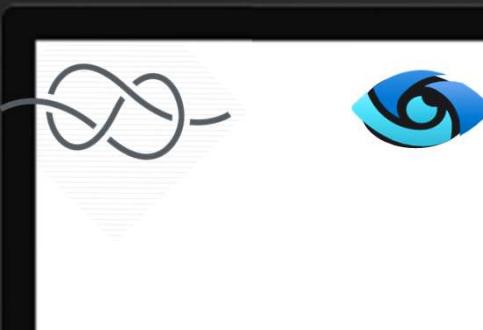


Purview & Power BI : Maximize business value of enterprise data



Area	Details
1. Registration	Register PowerBI <-> Purview Tech Metadata - Artifacts, Tables, Column Name & Types Business Metadata - Description, configured by, modified by, modified date time Classification & Labelling <ul style="list-style-type: none">• Sensitivity Insights• Auto labels in PBI• Microsoft Information Protection (MIP) Labels Integration (Purview->Synapse/SQL Server->PBI ->Office) - Purview labels automatically synced to Synapse and Azure SQL DB
2. Metadata Integration	Lineage <ul style="list-style-type: none">• Datasets• Reports• Dashboards

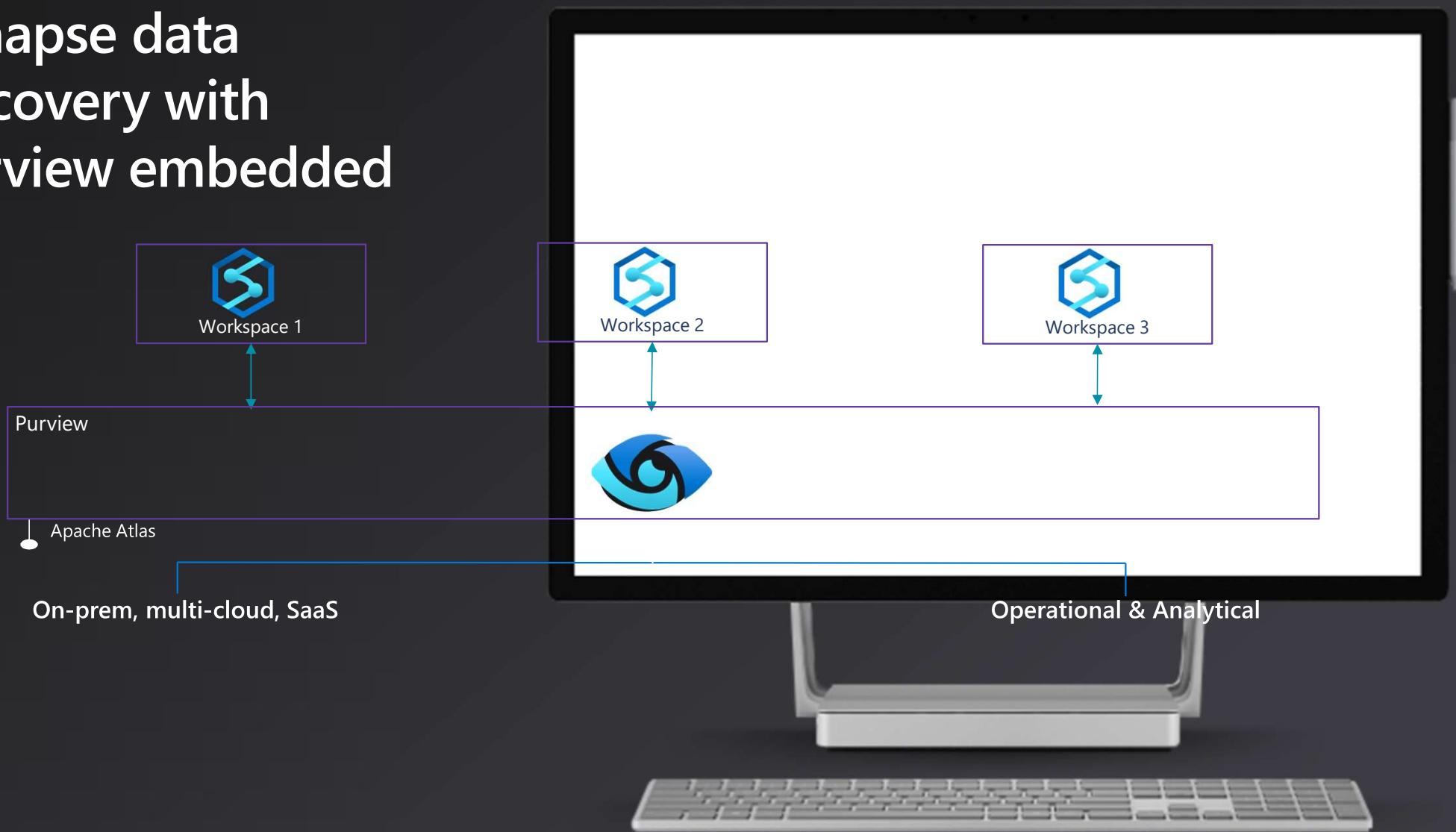
Purview & Synapse: Unified Analytics and Data Governance



Area	Details
1. Registration	Associate Synapse with Purview
2. Metadata Integration	Tech Metadata, Classification and Labelling
3. Data Discovery & Search	Discover data across Synapse workspaces Discover data across the entire data estate
4. Embedded experience	Contextual gesture and time to analytics



Synapse data discovery with Purview embedded



Purview Integrations: Public Preview

Synapse + Purview: Better Together



Register

Register Synapse workspaces with Purview



Browse & Search

Search datasets in Synapse workspaces from linked Purview catalog



Enhanced metadata, lineage and annotation

View enhanced metadata, lineage from Purview for Synapse tables and add annotations

Register Synapse Workspaces

The screenshot shows the Microsoft Purview Sources interface. The top navigation bar includes 'Microsoft Azure', 'Purview', 'ContosoBabylon', a search bar 'Search assets', and a user profile 'smallAlias@contoso.com'.

The left sidebar contains icons for Home, Sources, Audit Log, Metrics, Alerts, and Security.

The main area displays 'Sources' under 'Group A' and 'Group B'.

Group A: Contains 7 sources. One source, 'New-Synapse-Source' (Azure Synapse Analytics), is highlighted with a red border.

Group B: Contains 8 sources. Two sources, 'Cosmos DB test' (Cosmos DB) and 'SQL Server1' (SQL Server), are highlighted with red borders.

Each source card provides details such as 'Scan applied:', 'Policy applied:', 'Asset scanned:', 'Asset classified:', and 'Asset enforced:'.

Buttons for 'Register', 'New group', 'Refresh', 'Filter by keyword', 'Group: All', and 'Add filter' are visible at the top of the list.

Browse & Search: Synapse (powered by Purview)

The screenshot displays two views of the Microsoft Azure Synapse Analytics interface, illustrating the integration of Purview for search and discovery.

Left View: Shows the 'Data' workspace. A search bar at the top right contains the query 'retail sales'. Below it, the 'Search suggestions' section lists items like 'retail sales model', 'retail saleslt product', and 'retail salesmasterdata lifetimedata...'. The 'Asset suggestions' section shows database objects such as 'RetailSalesTable', 'retail-region-codes', 'Sales', and 'SalesTeam'. A 'View search results' button is located at the bottom of this sidebar.

Right View: Shows the search results for 'retail sales'. The results are categorized under 'Sources: Azure Cosmos DB', 'Type: all', and 'Instances: all'. The results are sorted by Relevance and include:

- StoreRetailSales**: Azure Cosmos DB Database, https://sales-cosmosdb.documents.azure.com/dbs/StoreRetailSales. Last updated: 8-12-2019 10:32AM. Column Matches: -
- RetailSalesActivity**: Azure Cosmos DB SQL API Collection, https://sales-cosmosdb.documents.azure.com/dbs/SalesRegions/colls/RetailSalesActivity. Last updated: 8-12-2019 10:32AM. Column Matches: -
- SalesCatalogAnalysis**: Azure Cosmos DB Database, https://sales-cosmosdb.documents.azure.com/dbs/SalesCatalogAnalysis. Last updated: 8-12-2019 10:32AM. Column Matches: -
- sales-cosmosdb**: Azure Cosmos DB Account, https://sales-cosmosdb.documents.azure.com/. Last updated: 8-12-2019 10:32AM. Column Matches: -
- SalesRegions**: Azure Cosmos DB Database, https://sales-cosmosdb.documents.azure.com/dbs/SalesRegions. Last updated: 8-12-2019 10:32AM. Column Matches: -

Enhanced Metadata, Lineage & Annotation

The screenshot shows the Microsoft Azure Synapse Analytics Data workspace interface. The left sidebar navigation includes icons for Home, Databases, Tables, External tables, SQL Analytics od (SQL on-demand), default (Spark), Views, and Pipelines. The main content area displays the 'Data' workspace for the 'contosoworkspace'. A search bar at the top right shows the query 'retail sales'. The results list 'StoreRetailSales' under 'Search results for "retail sales"'. The asset details page for 'StoreRetailSales' (GENERAL) shows the following sections:

- Description:** No description for this asset.
- Classifications:** Personal.Address, Personal.Email, Personal.Phone
- Properties:**

lastModifiedTime	Mon Jan 27 2020 10:23:45 GMT-0800 (Pacific Standard Time)
replicatedTo	
replicatedFrom	
createTime	Fri Nov 15 2019 10:23:45 GMT-0800 (Pacific Standard Time)
qualifiedName	https://sales-cosmosdb.documents.azure.com/dbs/StoreRetailSales
name	StoreRetailSales
- Hierarchy:** sales-cosmosdb.documents.azure... (Azure Cosmos DB Account) → StoreRetailSales (Azure Cosmos DB Database)
- Glossary terms:** No glossary terms for this asset.

Power BI + Purview: Better Together



Discover

Discover Power BI assets (tenant, workspace level) in Purview Catalog



Browse & Search

Find Power BI assets (dashboards, reports, datasets etc.) in Purview catalog



Enhanced metadata, lineage, sensitivity labels & endorsements

View enhanced metadata, lineage, sensitivity labels, endorsement labels in Purview for Power BI assets

Discover

Purview > gababu0508

Search assets

Customer360 model HIGHLY CONFIDENTIAL CERTIFIED

Power BI Dataset

Edit Refresh Pin Open in Power BI

Overview Schema Lineage Contacts Related

Column name	Column level classifications	Data type	Description
type_details	⚡ Financial.US ABA Number	varchar	
payment		varchar	
location	⚡ Financial.US Bank Number	varchar	
export_type		varchar	
date		timestamp	
volume		int	Export notes
unit		varchar	
notes		varchar	

Browse & Search

The screenshot shows the Microsoft Purview Data Catalog interface. The left sidebar has a breadcrumb trail: Purview > gababu0508 > Customer. The main area displays a search results page for the term "customer".

Search Bar: customer

Quick filters:

- Asset type: Select an asset type
- Classification: Select a classification
- Contact: Select a contact

Search suggestions:

- customer **customer**
- customer **n**
- customer**data customer**

Asset suggestions:

- Customers**
https://app.powerbi.com/groups/063687dc-2c52-4986-a173-09981287331...
- customer[N].csv**
https://sep2019demostoragev2.blob.core.windows.net/customerdata/custo...
- Customer Profitability Sample**
https://app.powerbi.com/groups/063687dc-2c52-4986-a173-09981287331...
- Customers**
https://app.powerbi.com/groups/cced5bb4-f465-4785-b910-a6dbd1a0a86c...

View search results

Last updated: 05/21/2020 00:45:50 UTC by DataScan

Hierarchy:

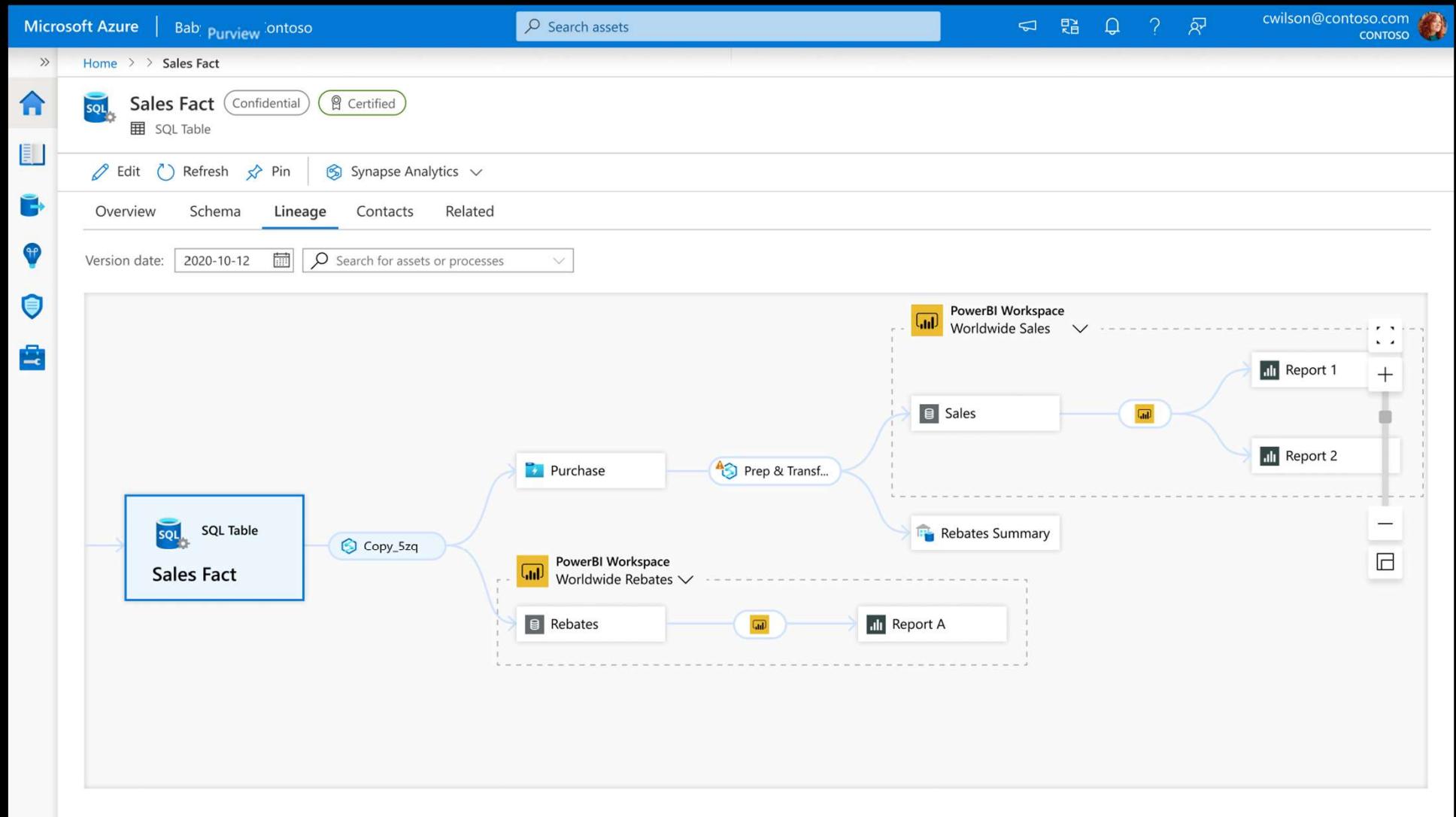
- Customer360 model**
Power BI Dataset

Glossary terms:
No glossary terms for this asset.

Properties:

qualifiedName	f722ad8c3d93/datasets/56071215-7721-41dd-ab53-3a1163244a75
description	
name	Customer360 model
configuredBy	yaronc@firsttomarket.onmicrosoft.com

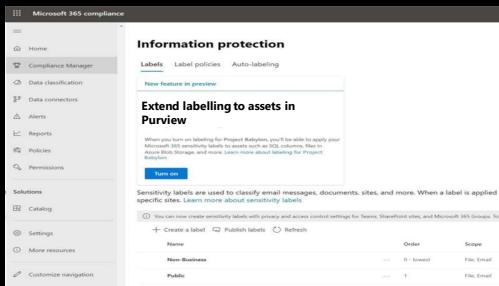
Enhanced Metadata, Lineage & Labels



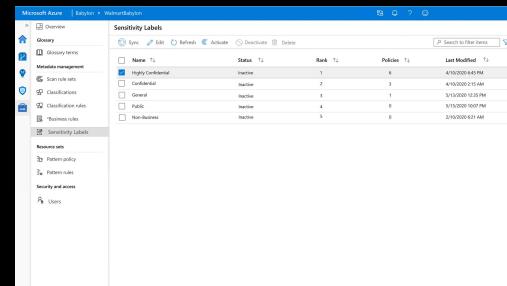
Microsoft 365 + Purview: Better Together

Scenario

Consistent classification and Labels across M365, Azure, SQL Svr, Power BI



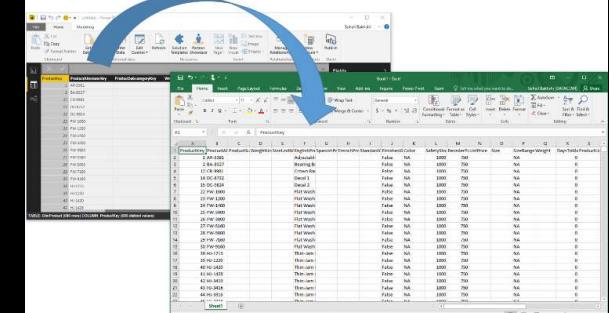
Discover Purview from M365
Compliance Center & get started with
classifiers and labels



Classification & Labels applied across breadth of Data Estate by Purview

Auto labelling

Labels automatically cascade via lineage



Information worker exports data from Power BI & saves to Excel with same sensitivity label

M365 -> Azure -> PBI -> M365

Microsoft Purview | Data Governance in Azure & Information Protection integration

The collage consists of five screenshots arranged in a grid-like fashion:

- Top Left:** Microsoft 365 compliance - Edit sensitivity label. Shows the 'Scope' step, which includes options for 'Files & emails', 'Groups & sites', and 'Azure Purview assets (preview)'. It also displays a list of sensitive info types such as Australia Driver's License Number, Canada Driver's License Number, EU Driver's License Number, France Driver's License Number, German Driver's License Number, Italy Driver's License Number, Japan Driver's License Number, New Zealand Driver License Number, U.K. Driver's License Number, U.S. Driver's License Number, Australia Bank Account Number, Canada Bank Account Number, International Banking Account Number (IBAN), Israel Bank Account Number, Japan Bank Account Number, New Zealand bank account number, and U.S. Bank Account Number.
- Top Right:** Microsoft 365 compliance - Edit sensitivity label. Shows the 'Scope' step, specifically the 'Files & emails' section. It includes configuration for encryption and content marking settings, and a note about auto-labeling conditions. It also lists 'Groups & sites' and 'Azure Purview assets (preview)'.
- Middle Left:** Microsoft 365 compliance - Information protection. Shows the 'Label policies' tab, which allows users to create sensitivity labels with privacy and access control settings for Teams, SharePoint sites, and Microsoft 365 Groups. It includes sections for 'Create a label', 'Publish labels', and 'Refresh'.
- Middle Right:** Microsoft Azure - Purview - SalesForceExpenseCard.xlsx. This screenshot shows a detailed view of a file asset in the Purview Management Center. It includes tabs for 'Overview', 'Lineage', 'Contacts', and 'Related'. The 'Overview' tab displays properties like 'Last updated' (08/25/2020 18:19:40 UTC by IPADemobulyan), 'Classification' (Credit Card Number, EU Debit Card Number), and 'Properties' (modifiedTime, replicatedTo, etc.). The 'Lineage' tab shows the data flow from various sources like 'Sales Force Expense Cards' through 'Azure Data Lake Storage Gen2 File' to 'Azure Data Lake Storage Gen2 File' and 'Azure Data Lake Storage Gen2 File'.
- Bottom Right:** Microsoft Azure - Purview - SalesForceExpenseCard.xlsx. This screenshot provides a more detailed view of the lineage diagram, showing the complex relationships between different data sources and storage locations.

Microsoft Purview | Data Governance in Azure & Information Protection integration



Supported data sources

Sensitivity labels are supported in Azure Purview for the following data sources:

Data type	Sources
Automatic labeling for files	<ul style="list-style-type: none">- Azure Blob Storage- Azure Files- Azure Data Lake Storage Gen 1 and Gen 2- Amazon S3
Automatic labeling for schematized data assets	<ul style="list-style-type: none">- 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

Microsoft Purview extend labeling capabilities to Cloud resources whose Azure, but also with hybrid and multi-cloud support

Microsoft Purview

Data Governance in Azure
& Information Protection integration
+ Microsoft Defender for Cloud

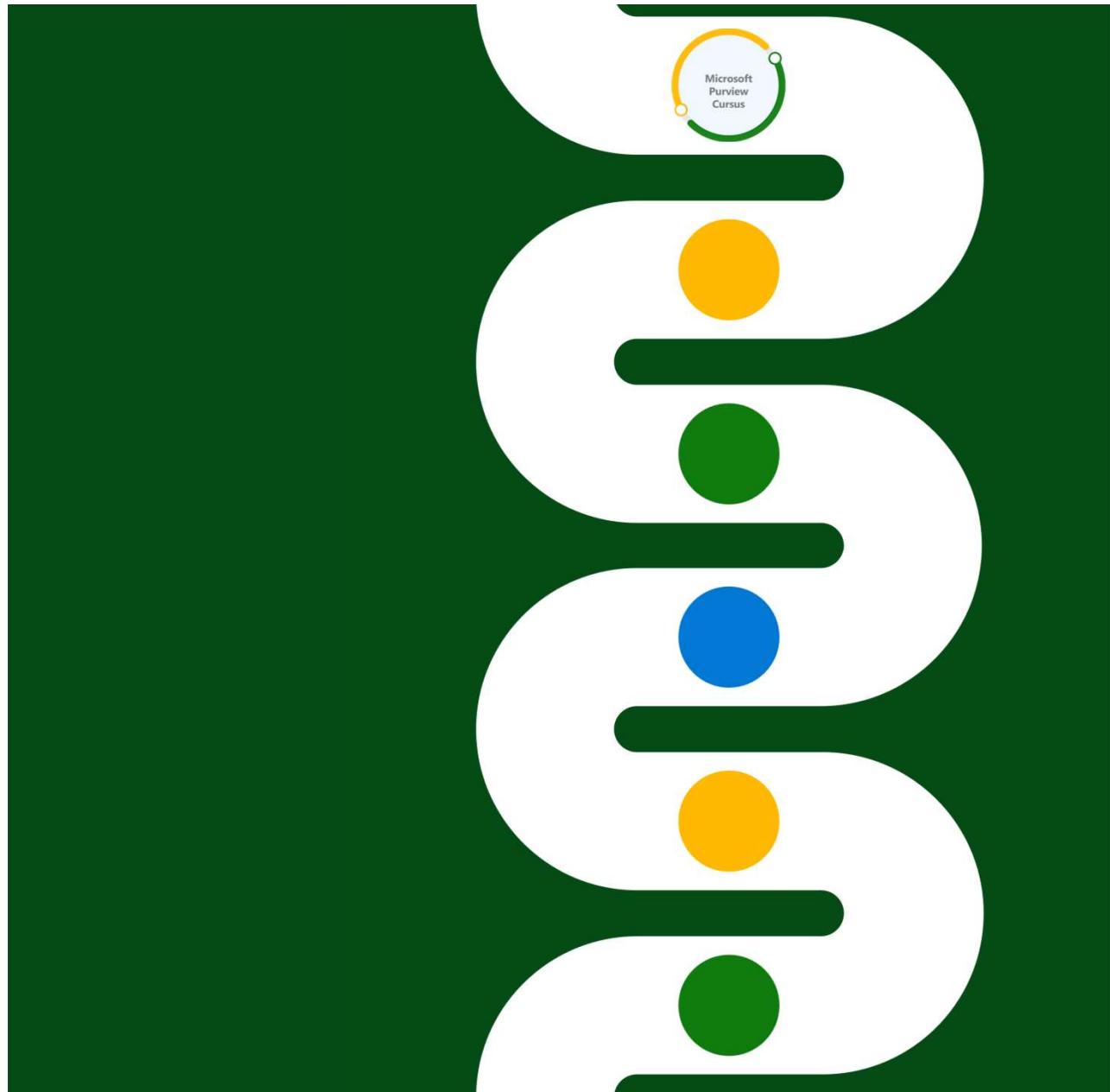
Identify sensitive data in cloud resources

- Extend visibility from cloud infrastructure resources into the data layer
- Leverage an entirely new way to prioritize security policies and the investigation of alerts
- Filter recommendations and resources by data sensitivity
- Easily view the number of assets that contain sensitive information across your environment

The screenshot shows the Microsoft Defender for Cloud Inventory page. At the top, there are filters for Subscriptions, Resource Groups, Resource types, and Data sensitivity, all set to All. A specific filter for 'Data sensitivity' is highlighted with a blue box, showing results for Confidential (4) and Confidential/All Employee (4). Below the filters, four summary sections are displayed: Total Resources (27), Unhealthy Resources (24), Unmonitored Resources (0), and Unregistered subscriptions (0). The main area lists 27 resources, each with a small icon, a name, and a sensitivity label. A blue box highlights a row for 'traceytest' which is labeled 'HighlyConfidential'. Other rows include 'testcust1' (General + 1 more), 'ip4ae2e2-test' (Confidential/Microsoft Extended + 1 more), 'joselwai-test' (General), 'aipdem1' (Secret Label/Accounting + 3 more), 'aipclptest' (Secret Label/Accounting + 4 more), 'aipclptest-test' (Secret Label/Accounting + 3 more), 'aipclptest1' (Public), 'aashishkar01' (General), 'aipolptest1adsgen2' (Secret Label/Accounting), 'datrigan-cb' (HighlyConfidential), 'datrigan-cb' (HighlyConfidential), and 'aipclptest1' (HighlyConfidential).

Microsoft Purview Data Governance in Azure deployment

Security and Network isolation



Securing your Purview account



Access control and
roles



Set up your
credentials



Network isolation

Access control and roles



Alice
Data reader



Robert
Data curator



Claudia
Data source admin

Who should be assigned to what role?

User Scenario	Appropriate Role(s)
I just need to find assets, I don't want to edit anything	Purview Data Reader Role
I need to edit information about assets, put classifications on them, associate them with glossary entries, etc.	Purview Data Curator Role
I need to edit the glossary or set up new classification definitions	Purview Data Curator Role
My application's Service Principal needs to push data to Azure Purview	Purview Data Curator Role
I need to set up scans via the Purview Studio	Purview Data Source Administrator Role plus at least one of Purview Data Reader Role or Purview Data Curator Role
I need to enable a Service Principal or other programmatic identity to set up and monitor scans in Azure Purview without allowing the programmatic identity to access the catalog's information	Purview Data Source Administrator Role
I need to put users into roles in Azure Purview	Owner or User Access Administrator

Set up your credentials



Set up your key
vault connection



Set up your
credentials



Register your
source

Network isolation

Private Endpoint

Account + portal private endpoint

Capability

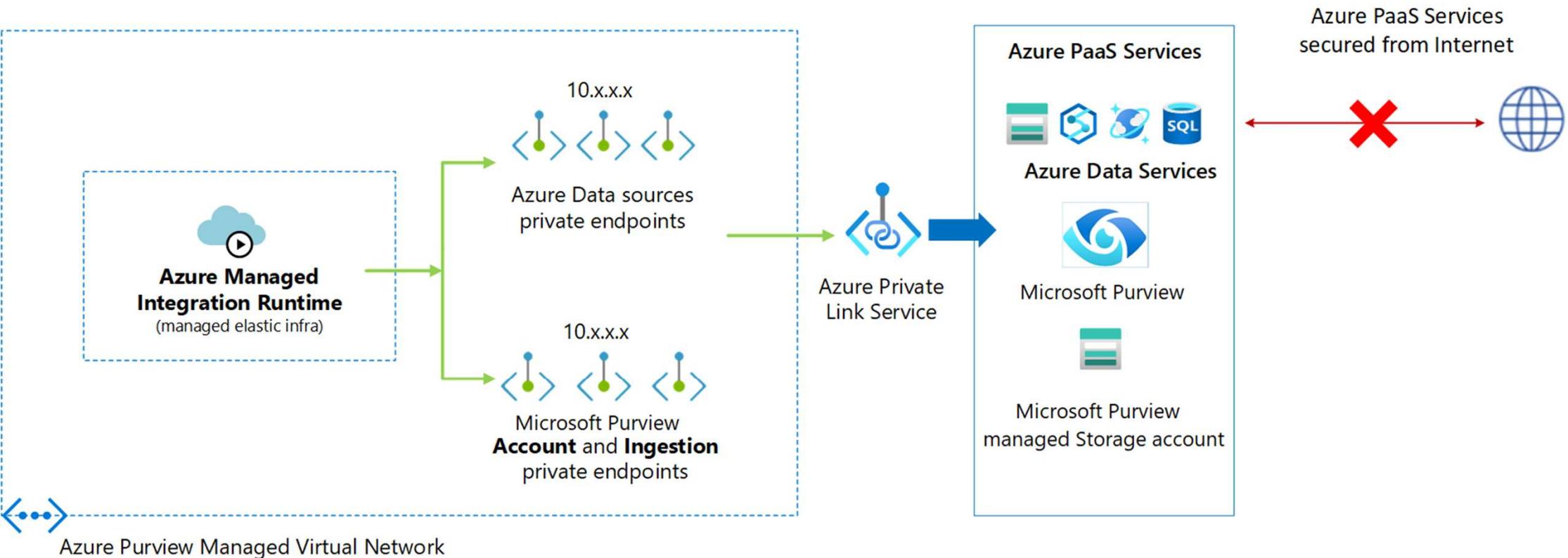
Access to Purview account will only be via private networks. This will include UX, scan and Atlas APIs

Ingestion private endpoint

Ability to scan on-premises and Azure data sources behind a vnet using SHIR ensuring E2E network isolation



Managed Virtual Network



A Managed Virtual Network in Microsoft Purview is a virtual network that is deployed and managed by Azure inside the same region as Microsoft Purview account to allow scanning Azure data sources inside a managed network, without having to deploy and manage any self-hosted integration runtime virtual machines by the customer in Azure.

Managed private endpoints

The screenshot shows the Azure Purview portal interface. On the left, there's a navigation sidebar under the 'General' category with options like Overview, Metrics, Lineage connections, Data Factory, Data Share, Security and access, Credentials, and Managed private endpoints. The 'Managed private endpoints' option is highlighted with a grey background. The main content area has a blue header bar with a search bar ('Search catalog') and notification icons. Below the header, the title 'Managed private endpoints' is displayed, followed by a row of buttons: '+ New', 'Delete', and 'Refresh'. A table lists five managed private endpoints, each with a checkbox, name, provisioning state (all succeeded), approval state (all approved), and linked resource ID. The table columns are: Name, Provisioning state, Approval state, and Linked resource ID.

	Name	Provisioning state	Approval state	Linked resource ID
<input type="checkbox"/>	Contoso-Purview-account	Succeeded	Approved	/subscriptions/225be6fe-ec1c
<input type="checkbox"/>	Contoso-Purview-managedBlob	Succeeded	Approved	/subscriptions/225be6fe-ec1c
<input type="checkbox"/>	Contoso-Purview-managedQue	Succeeded	Approved	/subscriptions/225be6fe-ec1c
<input type="checkbox"/>	AzureDataLakeStorage-contoso	Succeeded	Approved	/subscriptions/225be6fe-ec1c

Private endpoint uses a private IP address in the Managed Virtual Network to effectively bring the service into it. Private endpoints are mapped to a specific resource in Azure and not the entire service. Customers can limit connectivity to a specific resource approved by their organization. Learn more about [private links and private endpoints](#).

Managed private endpoints

contosoadls01 | Networking ...

Storage account | Directory: Microsoft

Search (Ctrl+ /) <> Firewalls and virtual networks Private endpoint connections

+ Private endpoint Approve Reject Remove Refresh

Filter by name... All connection states

<input checked="" type="checkbox"/> Connection name	Connection state	Private endpoint	Description
<input checked="" type="checkbox"/> contosoadls01.f269dd16-0805-4fda-... Pending	Pending	purview-f088c294-920d-4087-b1fa-...	Requested by DataFactory:purview-f088c294-920d-4087-b1fa-...

If the owner approves the connection, the private link is established. Otherwise, the private link won't be established. In either case, the Managed private endpoint will be updated with the status of the connection.

Deploy Managed VNet Runtimes

Prerequisites

- 1.A Microsoft Purview account deployed in one of the [supported regions](#).
- 2.From Microsoft Purview roles, you must be a data curator at root collection level in your Microsoft Purview account.
- 3.From Azure RBAC roles, you must be contributor on the Microsoft Purview account and data source to approve private links.

Create Private endpoint

Search catalog

Integration runtimes

+ New Edit Delete Refresh

Name	Type
Contoso-Managed-IR1	Managed

Notifications

Dismiss all

Approval required
Managed private endpoint (Contoso-Purview-account) requires approval.
[Manage approvals in Azure portal](#)

Approval required
Managed private endpoint (Contoso-Purview-managedQueue) requires approval.
[Manage approvals in Azure portal](#)

Approval required
Managed private endpoint (Contoso-Purview-managedBlob) requires approval.
[Manage approvals in Azure portal](#)

In progress
Enabling interactive authoring capability for Contoso-Managed-IR1.
a few seconds ago

Close

Approve Private endpoint

Contoso-Purview | Networking

Dashboard > Contoso-Purview | Networking

Search (Ctrl+F)

Firewall Private endpoint connections Ingestion private endpoint connections

+ Private endpoint ✓ Approve ✖ Reject Remove Refresh

Filter by name...	All connection states
<input checked="" type="checkbox"/> Connection name	Connection state
<input type="checkbox"/> Contoso-Purview/purview-eb416b2c-e6e7-4c8f-9ce8-f37381...	Approved

Private endpoint Description

Contoso-Purview/purview-eb416b2c-e6e7-4c8f-9ce8-f37381... Requested by DataFactory:purview-eb416b2c-e6e7-4c8f-9ce8-f37381...

In Azure portal, from your Microsoft Purview account resource window, approve the managed private endpoint.

scanwesteuropedhnkw | Networking

Dashboard > scanwesteuropedhnkw | Networking

Search (Ctrl+F)

Firewalls and virtual networks Private endpoint connections Custom domain

+ Private endpoint ✓ Approve ✖ Reject Remove Refresh

Filter by name...	All connection states
<input checked="" type="checkbox"/> Connection name	Connection state
<input checked="" type="checkbox"/> scanwesteuropedhnkw.076fdb69-2392-4...	Pending
<input checked="" type="checkbox"/> scanwesteuropedhnkw.8013c47a-2e34-4...	Pending

Private endpoint Description

purview-eb416b2c-e6e7-4c8f-9ce8-f37381... Requested by DataFactory:purview-eb416b2c-e6e7-4c8f-9ce8-f37381...

From Managed storage account page approve the managed private endpoints for blob and queue services

Deploy managed private endpoints for data sources

✓ Create Private endpoint

The screenshot shows the 'Managed private endpoints' section of the Azure Governance Portal. A red box highlights the '+ New' button. Below it, a list of existing managed private endpoints is shown, including 'Contoso-Purview-account', 'Contoso-Purview-managedBlob', and 'Contoso-Purview-managedQue'. A red box highlights the 'Azure Data Lake Storage Gen2' icon. At the bottom, a red box highlights the 'Continue' button.

✓ Approve Private endpoint

The screenshot shows the 'Managed' section of the Azure Governance Portal. A red box highlights the 'AzureDataLakeStorage-contosoadls01' entry in the list. To the right, a detailed view of the endpoint shows its provisioning state as 'Succeeded'. A red box highlights the 'Manage approvals in the Azure portal' link. A yellow box contains the following text:

From the list of managed private endpoints, select on the newly created managed private endpoint for your data source and then select on **Manage approvals in the Azure portal**, to approve the private endpoint in Azure portal

The screenshot shows the 'Managed private endpoints' section of the Azure Governance Portal. A red box highlights the 'New' button. Below it, a list of existing managed private endpoints is shown, including 'Contoso-Purview-account', 'Contoso-Purview-managedBlob', and 'Contoso-Purview-managedQue'. A red box highlights the 'Create' button at the bottom left. A yellow box contains the following text:

After creation, a private endpoint request will be generated that must get approved by an owner of the data source.

The screenshot shows the 'Networking' blade for the storage account 'contosoadls01'. A red box highlights the 'Private endpoint connections' tab. A yellow box contains the following text:

Successfully approved private endpoint connections
Successfully approved 1 private endpoint connections

Connection name	Connection state	Description
contosoadls01.f269dd16-0805-4fda-... purview-f088c294-920d-4087-b1fa-c9a62f9088e4	Approved	Requested by DataFactory:purview-f088c294-920d-4087-b1fa-c9a62f9088e4

Configure and verify DNS Name Resolution for Microsoft Purview private endpoints

Deployment options

Use any of the following options to set up internal name resolution when using private endpoints for your Microsoft Purview account:

01

[Deploy new Azure Private DNS Zones](#) in your Azure environment part of private endpoint deployment.
(Default option)

02

[Use existing Azure Private DNS Zones](#). Use this option if you using a private endpoint in a hub-and-spoke model from a different subscription or even within the same subscription.

03

[Use your own DNS Servers](#) if you do not use DNS forwarders and instead you manage A records directly in your on-premises DNS servers.

Configure and verify DNS Name Resolution for Microsoft Purview private endpoints

Option 1 - Deploy new Azure Private DNS Zones

The following table shows an example of Azure Private DNS zones and DNS A Records that are deployed as part of configuration of private endpoint for a Microsoft Purview account if you enable *Private DNS integration* during the deployment:

Private endpoint	Private endpoint associated to	DNS Zone (new)	A Record (example)
Account	Microsoft Purview	privatelink.purview.azure.com	Contoso-Purview
Portal	Microsoft Purview	privatelink.purviewstudio.azure.com	Web
Ingestion	Microsoft Purview managed Storage Account - Blob	privatelink.blob.core.windows.net	scaneastusabcd1234
Ingestion	Microsoft Purview managed Storage Account - Queue	privatelink.queue.core.windows.net	scaneastusabcd1234
Ingestion	Microsoft Purview managed Storage Account - Event Hub	privatelink.servicebus.windows.net	atlas-12345678-1234-1234-abcd-123456789abc

[Configure DNS Name Resolution for private endpoints - Microsoft Purview | Microsoft Docs](#)

Configure and verify DNS Name Resolution for Microsoft Purview private endpoints

Option 1 - Deploy new Azure Private DNS Zones

The following table shows an example of Azure Private DNS zones and DNS A Records that are deployed as part of configuration of private endpoint for a Microsoft Purview account if you enable *Private DNS integration* during the deployment:

Private endpoint	Private endpoint associated to	DNS Zone (new)	A Record (example)
Account	Microsoft Purview	privatelink.purview.azure.com	Contoso-Purview
Portal	Microsoft Purview	privatelink.purviewstudio.azure.com	Web
Ingestion	Microsoft Purview managed Storage Account - Blob	privatelink.blob.core.windows.net	scaneastusabcd1234
Ingestion	Microsoft Purview managed Storage Account - Queue	privatelink.queue.core.windows.net	scaneastusabcd1234
Ingestion	Microsoft Purview managed Storage Account - Event Hub	privatelink.servicebus.windows.net	atlas-12345678-1234-1234-abcd-123456789abc

[Configure DNS Name Resolution for private endpoints - Microsoft Purview | Microsoft Docs](#)

if a Microsoft Purview account name is 'Contoso-Purview', **when it is resolved from outside the virtual network** that hosts the private endpoint, it will be:

Name	Type	Value
Contoso-Purview.purview.azure.com	CNAME	Contoso-Purview.privatelink.purview.azure.com
Contoso-Purview.purview.azure.com	CNAME	<Microsoft Purview public endpoint>
<Microsoft Purview public endpoint>	A	<Microsoft Purview public IP address>
Web.purview.azure.com	CNAME	<Microsoft Purview governance portal public endpoint>

The DNS resource records for Contoso-Purview, **when resolved in the virtual network** hosting the private endpoint, will be:

Name	Type	Value
Contoso-Purview.purview.azure.com	CNAME	Contoso-Purview.privatelink.purview.azure.com
Contoso-Purview.purview.azure.com	A	<Microsoft Purview account private endpoint IP address>
Web.purview.azure.com	CNAME	<Microsoft Purview portal private endpoint IP address>

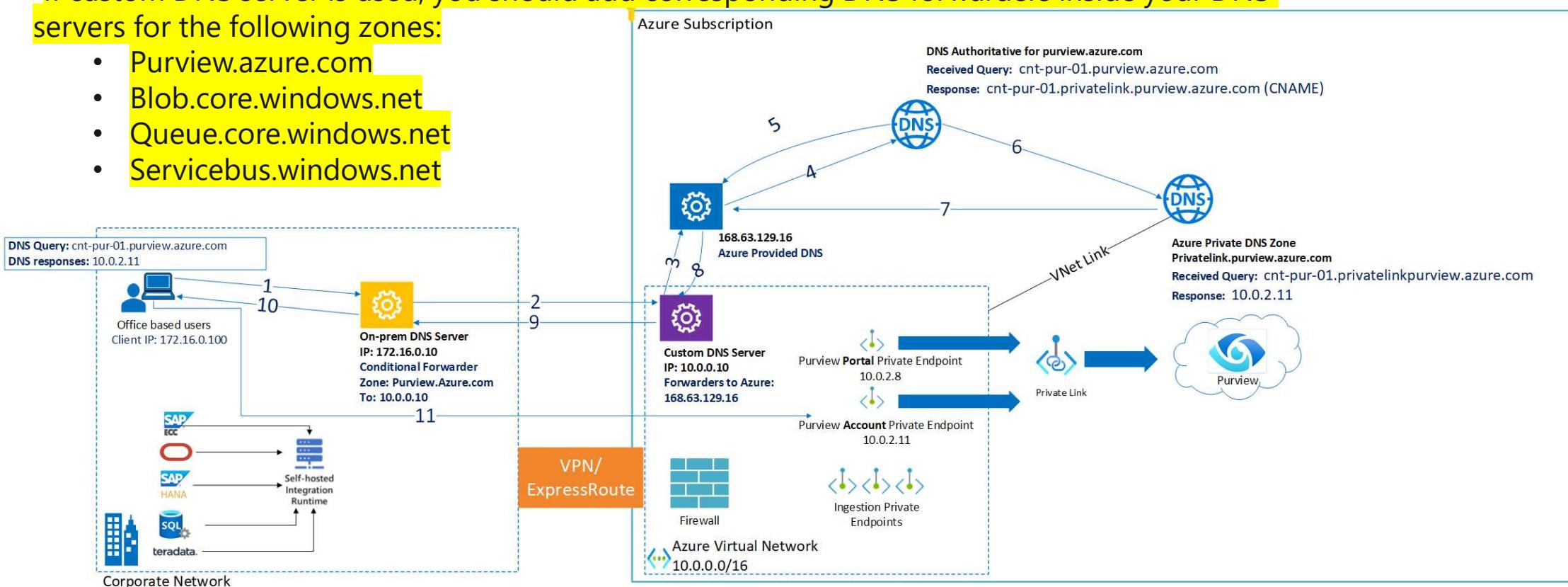
Configure and verify DNS Name Resolution for Microsoft Purview private endpoints

Option 2 - Use existing Azure Private DNS Zones

During the deployment of Microsoft Purview private endpoints, you can choose *Private DNS integration* using existing Azure Private DNS zones.

- If custom DNS server is used, you should add corresponding DNS forwarders inside your DNS servers for the following zones:

- Purview.azure.com
- Blob.core.windows.net
- Queue.core.windows.net
- Servicebus.windows.net



Configure and verify DNS Name Resolution for Microsoft Purview private endpoints

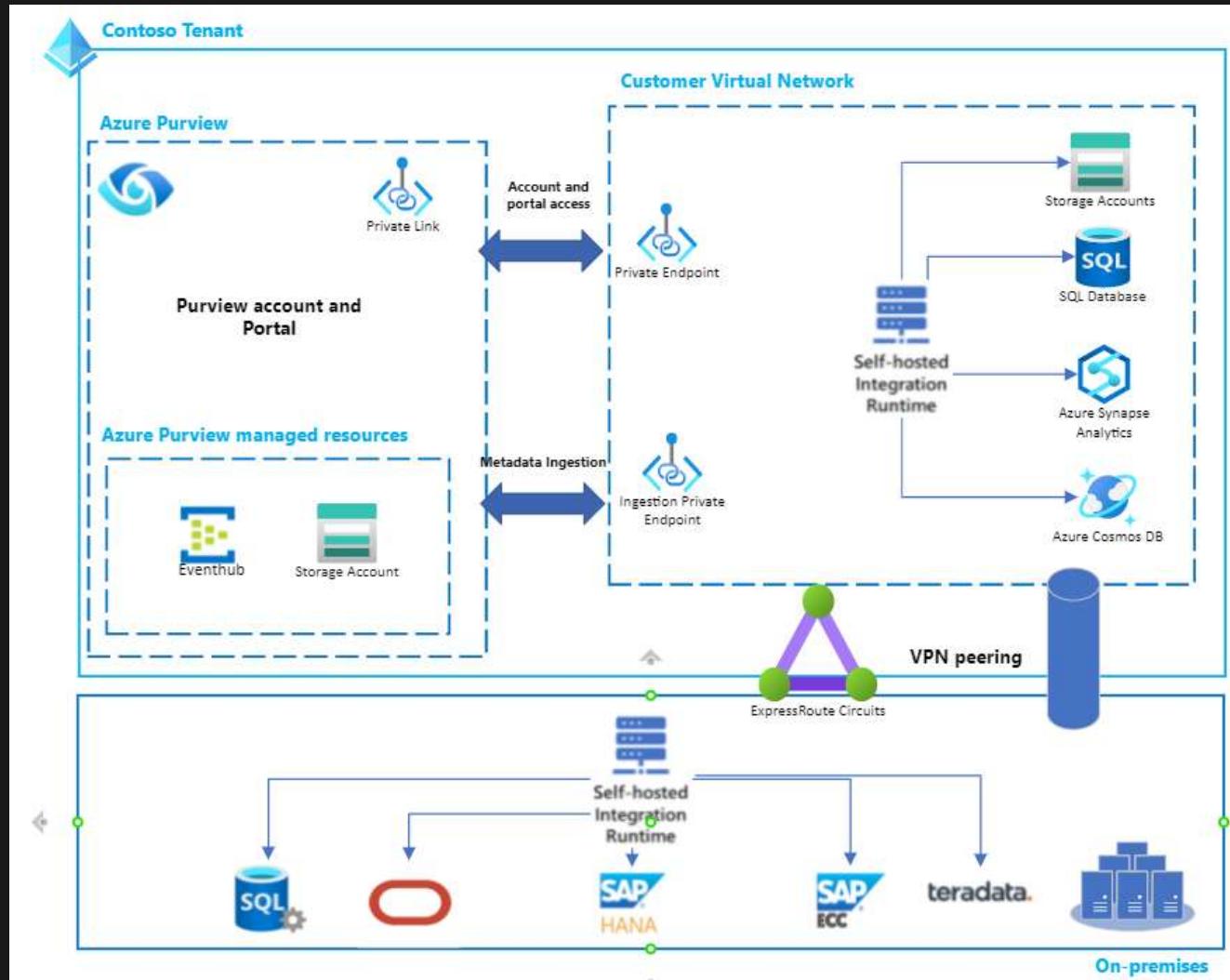
Option 3 - Use your own DNS Servers

If you do not use DNS forwarders and instead you manage A records directly in your on-premises DNS servers to resolve the endpoints through their private IP addresses, you might need to create the following A records in your DNS servers.

Name	Type	Value
web.purview.azure.com	A	<portal private endpoint IP address of Microsoft Purview>
scaneastusabcd1234.blob.core.windows.net	A	<blob-ingestion private endpoint IP address of Microsoft Purview>
scaneastusabcd1234.queue.core.windows.net	A	<queue-ingestion private endpoint IP address of Microsoft Purview>
atlas-12345678-1234-1234-abcd-123456789abc.servicebus.windows.net	A	<namespace-ingestion private endpoint IP address of Microsoft Purview>
Contoso-Purview.Purview.azure.com	A	<account private endpoint IP address of Microsoft Purview>
Contoso-Purview.scan.Purview.azure.com	A	<account private endpoint IP address of Microsoft Purview>
Contoso-Purview.catalog.Purview.azure.com	A	<account private endpoint IP address of Microsoft Purview>
Contoso-Purview.proxy.purview.azure.com	A	<account private endpoint IP address of Microsoft Purview>
Contoso-Purview.guardian.purview.azure.com	A	<account private endpoint IP address of Microsoft Purview>
gateway.purview.azure.com	A	<account private endpoint IP address of Microsoft Purview>

manifest.prod.ext.web.purview.azure.com	A	<portal private endpoint IP address of Microsoft Purview>
cdn.prod.ext.web.purview.azure.com	A	<portal private endpoint IP address of Microsoft Purview>
hub.prod.ext.web.purview.azure.com	A	<portal private endpoint IP address of Microsoft Purview>
catalog.prod.ext.web.purview.azure.com	A	<portal private endpoint IP address of Microsoft Purview>
cseo.prod.ext.web.purview.azure.com	A	<portal private endpoint IP address of Microsoft Purview>
datascan.prod.ext.web.purview.azure.com	A	<portal private endpoint IP address of Microsoft Purview>
datashare.prod.ext.web.purview.azure.com	A	<portal private endpoint IP address of Microsoft Purview>
datasource.prod.ext.web.purview.azure.com	A	<portal private endpoint IP address of Microsoft Purview>
policy.prod.ext.web.purview.azure.com	A	<portal private endpoint IP address of Microsoft Purview>
sensitivity.prod.ext.web.purview.azure.com	A	<portal private endpoint IP address of Microsoft Purview>

Network isolation and security



Deployment and Checklist

- Purview deployment challenges
- Purview deployment one-pager checklist
- Purview detailed deployment checklist
- Deployment readiness check and configuration scripts

Customer challenges when deploying Purview

- **Lack of access** to Purview, Azure AD or Azure Services by single team may slow down the deployment process.
- **Lack of knowledge** around the product by the data owners.
- Purview deployment may get delayed due **lack of prioritization** from other teams.
- **Scan frequency** requirements may vary across different teams in the organization.
- Lack of clarity around **costs** related to Purview and scans.
- Lack of visibility to **M365 Compliance Center** by security or data team to view Sensitivity Labels.
- Lack of **M365 E5 licenses**.
- Lack of **automation** and **scalability** when to onboard their data sources automatically.
- **Key management** to provide access to Purview to perform scans.

Purview Deployment Prerequisites

To help you prepare your environment, we have deployed powershell scripts so, you can evaluate your existing environment and identify missing configurations so Purview can scan data sources.

- ✓ Ideal for PoCs
- ✓ Lists all Purview requirements at high level.

<https://github.com/zeinab-mk/Purview-Deployment-Checklist>

Purview Readiness one-pager checklist

Purview Readiness Checklist is a checklist of high-level steps to guide you to plan and deploy Purview as your data governance solution. The guide is divided into four phases:

1. **Readiness** – Learn the pre-requisite tools and approaches important to all adoption efforts.
2. **Build Foundation** – Deploy Purview Accounts to establish your unified data governance model.
3. **Register Data Sources** – Setup first data landing zone (**start small and expand**) and onboard initial group of data sources.
4. **Curate and consume data** – Enable a unified Data Governance solution for data consumers using Purview for all data landing zones.

Readiness Checklist - Learn the pre-requisite tools and approaches important to all adoption efforts

1 Identify sources of data across your organization to digital state of your data across organization

- Who are the Data Owners?
- Who does have write access to the data?
- Where is Data? in the Cloud? on-premises?
- What Data Source types do you have in the organization? (e.g. Azure SQL DB, Amazon S3, on-premises SQL Servers)

3 Manage budget, costs and licensing

Currently, Microsoft offers [Purview](#) in a pay-as-you-go model. Understand how billing works, and see how you can control costs. Know the additional costs; to extend Microsoft 365 Sensitivity Labels to Purview, you need M365 E5 licenses.

4 Plan for communication, readiness and awareness across organization

Setup Authentication model for Purview

- Locate data sources. If they are in Azure, what subscriptions are in scope.
- Identify Subscription Owners in the data source subscriptions.
- Define what [authentication methods](#) to be used for each data source type to allow Purview to connect to data sources.
- If data sources reside in IaaS or on-premises VMs, deploy Microsoft Integration Runtime.
- Define Azure Key Vaults requirements to keep required keys and secrets for data governance
- Create required users for authentication to data sources and store secrets in Azure Key Vaults.

1
2

2 Define Roles and Responsibilities to build and manage a unified data governance solution

- Data Curators: Manage Classifications and Glossary Terms.
- Data Readers: Require read-only access to Search and Insights Reports.
- Data Source Administrators + Curators: Register and Scan Data Sources, Manage Catalog.
- Data Source Administrators + Readers: Register and Scan Data Sources.
- Supporting Roles for building an end-to-end data governance solution using Purview:
 - ✓ Azure Subscription Administrators to manage Azure resources, Policies, RBAC, Resource Providers.
 - ✓ Azure AD Administrators to manage identity and application registration.
 - ✓ NetOps to prepare the network to support Purview connectivity requirements.
 - ✓ SecOps to manage Azure Key Vaults and secrets.
 - ✓ M365 Administrators (Microsoft M365 Information Protection if M365 Sensitivity Labels will be used)
 - ✓ Data Source Owners to map them for roles in Purview.

3

4

5
6

6 Prepare Network and connectivity

- Define if you need to deploy Purview Account using Azure Private Endpoint.
- Check if Service Endpoint is enabled on Azure data sources. Allow AzureServices to bypass.
- Verify if data sources are deployed using Private Endpoints. Check Purview support for [Private Link](#).
- Configure required NSG rules to allow Purview access to VMs and Azure SQL Managed Instances.

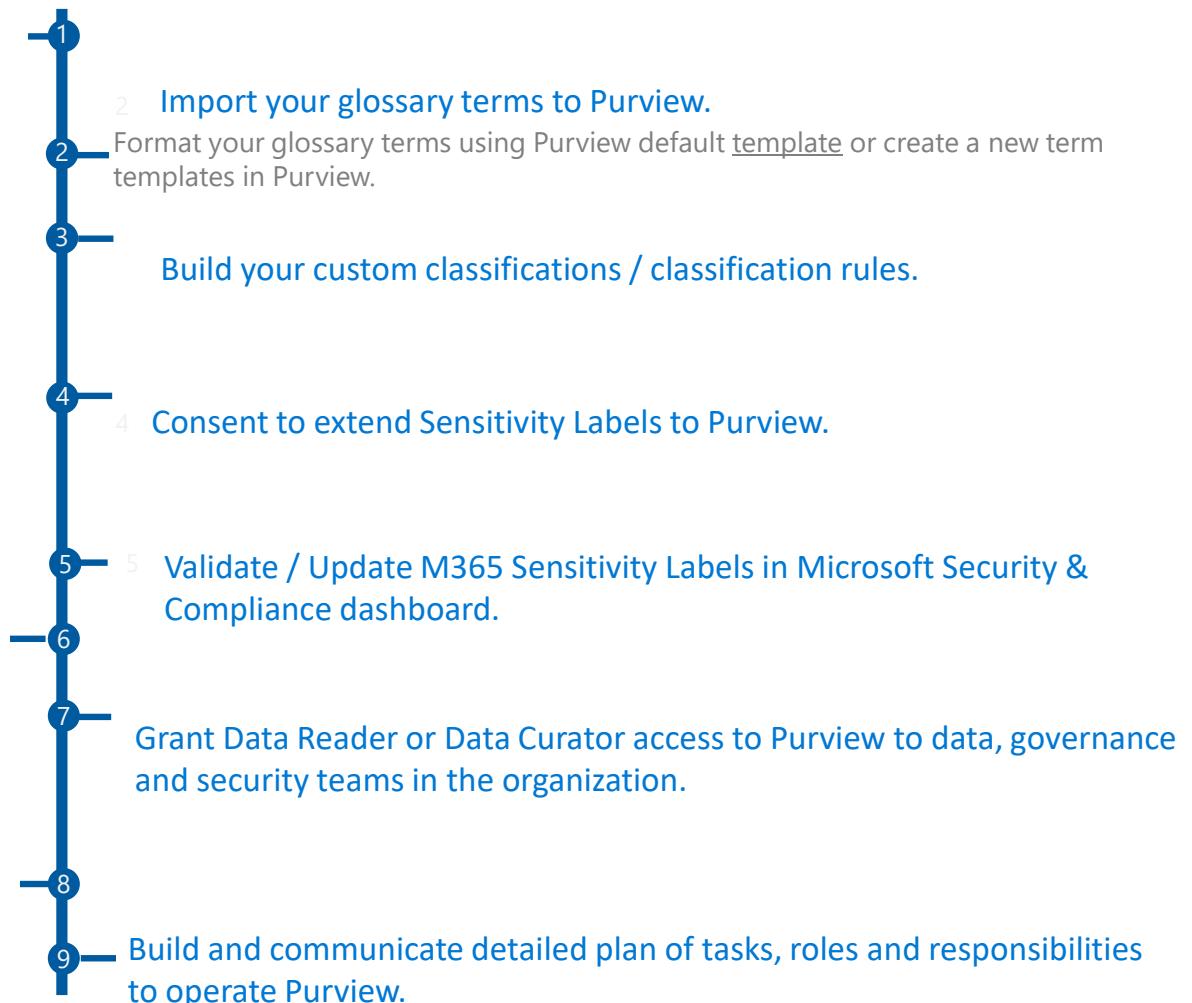
Build Foundation - Deploy Purview Accounts to establish your unified data governance model

1 Prepare your Azure Subscriptions and deploy Purview

- Define what Azure Subscriptions will be used to deploy Purview.
- Register required Azure Resource Providers in Purview subscription.
- Review and update Azure Policy assigned to Purview subscriptions to allow deployment of Azure Storage, EventHub Namespace and Purview Accounts.
- Deploy Purview Accounts for production and non-production environments.
- If Purview MSI used: assign required RBAC roles to Purview MSI.
- If Key Vault used, assign the required access to key vault secrets to Purview Account's MSI.
- Create Credentials and map to Key Vault's secrets in Purview.
- Register Application and configure service principals if SP is used.
- Generate and Register Integration Runtime for data sources inside IaaS or on-premises VMs.

Create Auto-labeling Rules in M365.

Perform initial Security review.



Register Data Sources - Setup first landing zone and onboard initial group of data sources

1 Define Purview Collection architecture.



Create Collection hierarchy inside Purview Account

2 Register pilot data sources, test connectivity, perform initial scans.

3 Verify classifications in assets and adjust custom classification rules.

4 Verify Assets with assigned labels.

5 Assign Data Experts and Data Owners to assets in Purview.

6 Assign Glossary Terms to data assets.

7 Setup automatic scan rules

8 Validate and adjust budget

9 Review security, business and operational requirements, identify gaps and adjust configuration.

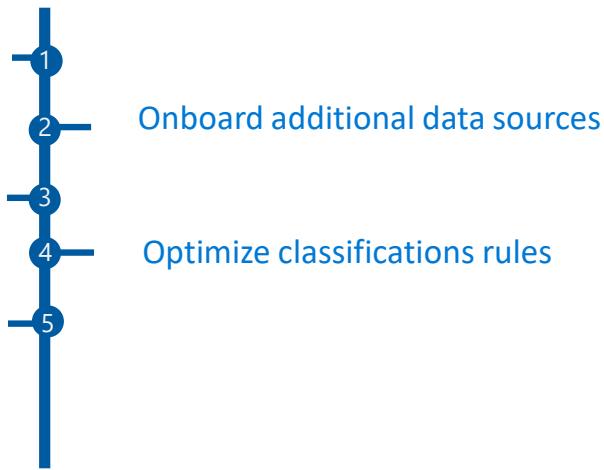
10 Adjust and share roles and responsibilities plan.

Curate and consume data – Enable a unified Data Governance solution for data consumers using Purview

1 Adjust scan rules

Optimize glossary terms

Adjust access levels



Purview Automated Readiness Checklist Scripts

Script	Use cases	Prerequisites	Required Permissions
Azure-Purview-automated-readiness-checklist.ps1	<ul style="list-style-type: none"> Possibility to run the script on a Top-level Management Group or a Subscription as data source scope Validates if Azure Purview Managed Identity has access to data sources on both network access and authorization Missing configurations are listed as output report 	<ul style="list-style-type: none"> An Azure Purview Account Powershell Az Module Powershell AzureAD Module Powershell Az.Synapse Module SQLCMD 	<ul style="list-style-type: none"> Contributor on Azure Purview Resource Group Reader at data source Subscription or Management Group Get/List Key Vault secret Azure AD Global Reader
Azure-Purview-automated-readiness-checklist-csv-Input.ps1	<ul style="list-style-type: none"> Possibility to run the script over data sources in multiple subscriptions using a csv list as input Validates if Azure Purview Managed Identity has access to data sources on both network access and authorization Missing configurations are listed as output report Reduced prompts when running the script 	<ul style="list-style-type: none"> An Azure Purview Account Powershell Az Module Powershell AzureAD Module Powershell Az.Synapse Module SQLCMD 	<ul style="list-style-type: none"> Contributor on Azure Purview Resource Group Reader at data source Subscription or Management Group Get/List Key Vault secret Azure AD Global Reader

Purview Automated MSI Configuration

Script	Use cases	Prerequisites	Required Permissions
Azure-Purview-MSI-Configuration.ps1	<ul style="list-style-type: none"> Possibility to run the script on a Top-level Management Group or a Subscription as data source scope Validates and configures required access for Azure Purview Managed Identity on both network and authorization Detected settings and progress in the script will generate an output report. Admin will receive several prompts to take action before configuring any settings 	<ul style="list-style-type: none"> An Azure Purview Account Powershell Az Module Powershell AzureAD Module Powershell Az.Synapse Module SQLCMD 	<ul style="list-style-type: none"> Contributor on Azure Purview Resource Group Owner at data source Subscription or Management Group Get/List Key Vault secret Azure AD Global Reader
Azure-Purview-MSI-Configuration-csv-Input.ps1	<ul style="list-style-type: none"> Possibility to run the script over data sources in multiple subscriptions using a csv list as input Validates and configures required access for Azure Purview Managed Identity on both network and authorization Admin is not required to provide any confirmation to setup the configurations. Reduced prompts when running the script 	<ul style="list-style-type: none"> An Azure Purview Account Powershell Az Module Powershell AzureAD Module Powershell Az.Synapse Module SQLCMD 	<ul style="list-style-type: none"> Contributor on Azure Purview Resource Group Reader at data source Subscription or Management Group Get/List Key Vault secret Azure AD Global Reader

Scripts current capabilities

- Automated readiness check of Azure data sources.
- Automated configuration of missing required settings for scanning data sources.
- Supported data sources:
 - All, Azure Blob Storage, ADLS Gen2, ADLS Gen1, Azure SQL, Azure SQL MI, Azure Synapse Analytics Dedicated Pool
- Supported scope:
 - Single Subscription
 - Multiple subscriptions based on a csv file
 - Top-level MG (includes all subscription in all 3 level down MGs)

Purview automated readiness checklist

- PowerShell-based script
- Select a Subscription or Management Group as data scope
- Guided execution (Admin input required during execution)
- No changes on data sources
- One Key Vault + Secret is deployed/used to use for Azure SQL DB, Azure SQL MI or Synapse Dedicated Pool
- Readiness reports generated based on findings

Purview automated readiness checklist csv input

- PowerShell-based script
- Subscription ids inside csv file e.g. c:\temp\subscriptions.csv
- Requires one key vault + two secrets on each subscription to authenticate to Azure SQL DB, Azure SQL MI or Synapse Dedicated Pool
- Reduced admin prompts during execution
- No changes on data sources
- Readiness reports generated based on findings

Purview MSI Configuration

- PowerShell-based script
- Select a Subscription or Management Group as data scope
- Guided execution (Admin input and confirmation is required during execution)
- One Key Vault + Secret is deployed/used to use for Azure SQL DB, Azure SQL MI or Synapse Dedicated Pool
- Data sources missing configurations are reported and remediated during execution
- Readiness reports generated based on findings and updates

Purview MSI Configuration csv input

- PowerShell-based script
- Subscription ids inside csv file e.g. c:\temp\subscriptions.csv
- Requires one key vault + two secrets on each subscription to authenticate to Azure SQL DB, Azure SQL MI or Synapse Dedicated Pool
- Reduced admin prompts during execution
- Data sources findings are remediated during execution
- Readiness reports generated based on findings and updates

Scripts Known issues

- On-prem / IaaS SQL VMs not included
- Admin client machine IP will not be whitelisted into Azure resources firewall intentionally
- Synapse serverless pool not included
- Synapse dedicated pool is also validated as Azure SQL Servers

Microsoft Purview Data Governance in Azure Roadmap



Roadmap



Data Steward



Organize and curate

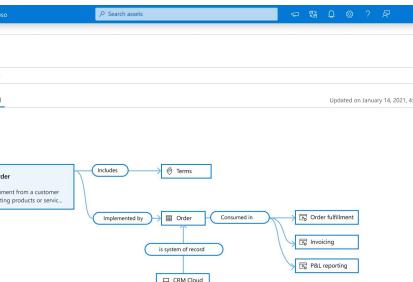
- Organize data into data domains.
- Create a business-friendly view of the data estate.
- 3 asset types

Private Preview

Establish ownership

- Establish business ownership and accountability for data use using business process assets.

Private Preview



Discover & understand

- Explore the data map by data domain
- Drill into and explore each domain
- 10+ asset types

Public Preview



Data Officer

CDO Insights & scale plays

Show data ownership and curation by:

- Organization
- Line of business
- Department

GA

Jun-Aug 2022

Oct-Dec 2022

2023

Roadmap: Discover all your data

Azure (15)	Database & DW (18)	File Storage (8)	File Formats (7)	NoSQL (2)	Services & Apps (21)	Stream (2)
Blob Storage	SQL Server	PostgreSQL	Amazon S3	AVRO	Cassandra	Informatica Power Center
Cosmos DB – SQL API	Teradata (pp)	SAP HANA	File System	Delimited Text	MongoDB	Azure Synapse (pp)
Cosmos DB – MongoDB API	Amazon RDS	SAP BW HANA	FTP	XML	Azure Data Factory	Workday
ADLS Gen1	Amazon Redshift	Oracle DW	Google Cloud Storage	JSON	Azure Data Share	Qlik Sense
ADLS Gen2	DB2	Oracle MYSQL	HDFS	ORC	SAP ECC (pp)	Dell Boomi
Data Explorer	Snowflake		SFTP	Parquet	SAP S/4 HANA (pp)	Dremio
Database for MariaDB	Google BigQuery		Sqoop	Document (Doc, PDF, XLS etc.)	Azure ML	Mulesoft
Database for MySQL	Oracle DW		Sharepoint		Salesforce	
Database for PostgreSQL	HBase				Erwin	
File Storage	Hive				E/R Studio	
SQL Database	Informix				Tableau	
SQL Managed Instance	Microsoft Access				Talend	
Azure Synapse Analytics (formerly SQL DW)	Netezza		Already Supported		Dynamics 365	
Search Index	Oracle		Support Planned (CY 2021)		SAP Business Objects	
Table Storage					SAP MDM	



Microsoft Azure Solutions Business

Thank you



Microsoft Azure Solutions Business

Tutorial & Demo environment

<https://docs.microsoft.com/fr-fr/azure/purview/tutorial-scan-data>

[tayganr/purviewdemo: Purview Demo Generator \(github.com\)](https://github.com/tayganr/purviewdemo)

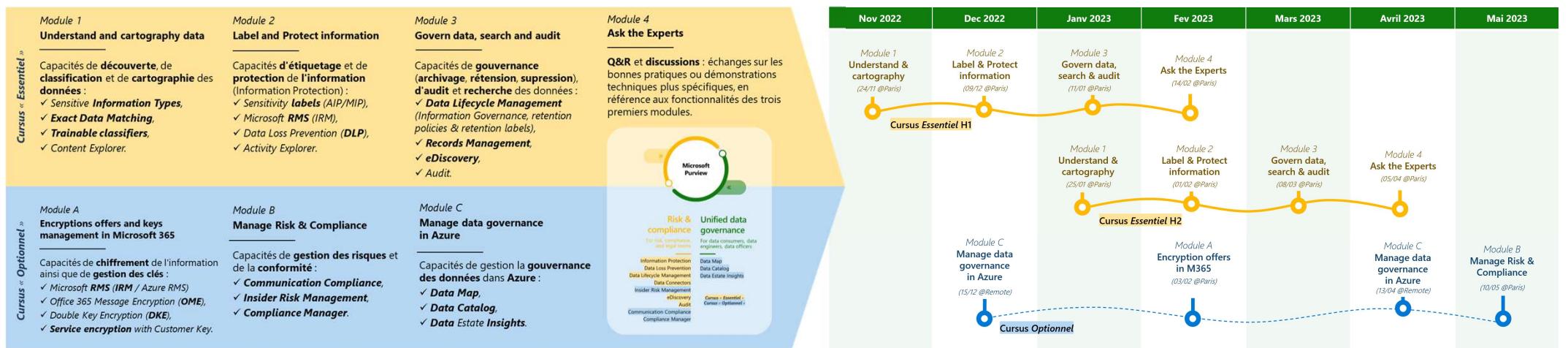
→ <https://aka.ms/MicrosoftPurviewCursus>

Microsoft Purview Cursus

Inscriptions aux modules

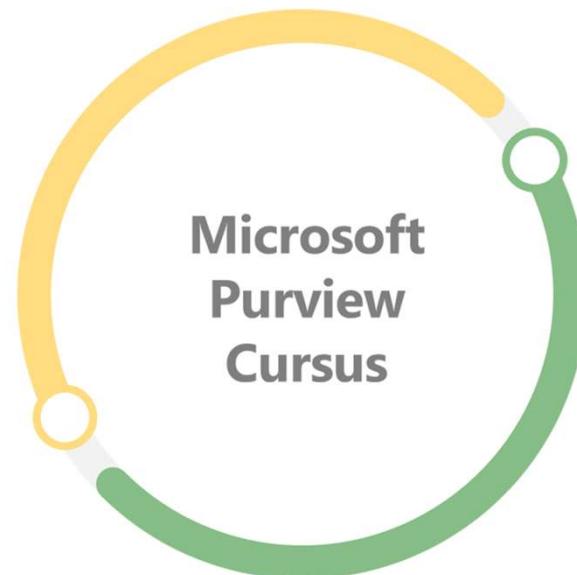
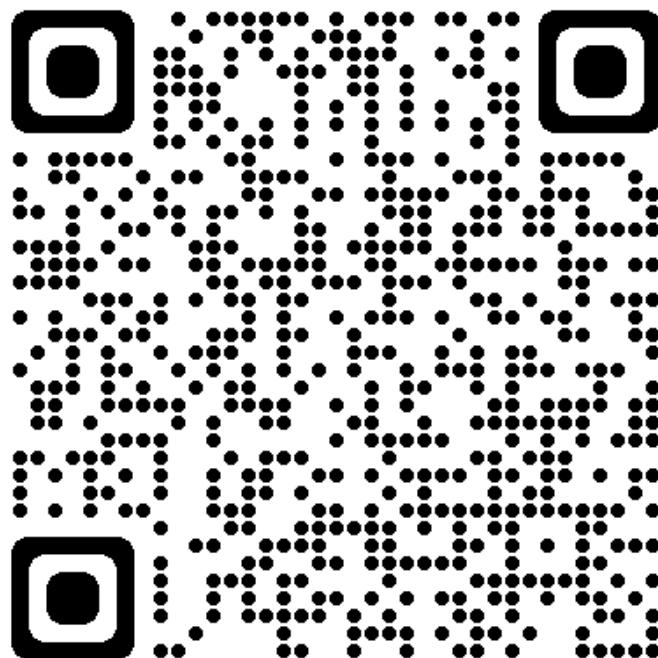
* Rejoignez-nous également à notre Centre de Conférence, afin d'assister physiquement aux modules et rencontrer nos spécialistes, en vous inscrivant en complément sur : <https://aka.ms/MicrosoftPurviewCursus/OnSite>

Cursus	Session	Titre de l'évènement	Date	Horaire en ligne	Inscription en ligne *	Horaire sur site	Inscription sur site *	Lieu
Essentiel	H1	Microsoft Purview Cursus - Understand and cartography data (module 1/4)	24/11/2022	9h30 - 12h30	Lien	9h - 13h	Lien	Paris
Essentiel	H1	Microsoft Purview Cursus - Label and Protect information (module 2/4)	09/12/2022	9h30 - 12h30	Lien	9h - 13h	Lien	Paris
Essentiel	H1	Microsoft Purview Cursus - Govern data, search and audit (module 3/4)	11/01/2023	9h30 - 12h30	Lien	9h - 13h	Lien	Paris
Essentiel	H1	Microsoft Purview Cursus – Ask the Experts (module 4/4)	14/02/2023	9h30 - 12h30	Lien	9h - 13h	Lien	Paris
Essentiel	H2	Microsoft Purview Cursus - Understand and cartography data (module 1/4)	25/01/2023	9h30 - 12h30	Lien	9h - 13h	Lien	Paris
Essentiel	H2	Microsoft Purview Cursus - Label and Protect information (module 2/4)	01/02/2023	9h30 - 12h30	Lien	9h - 13h	Lien	Paris
Essentiel	H2	Microsoft Purview Cursus - Govern data, search and audit (module 3/4)	08/03/2023	9h30 - 12h30	Lien	9h - 13h	Lien	Paris
Essentiel	H2	Microsoft Purview Cursus – Ask the Experts (module 4/4)	05/04/2023	9h30 - 12h30	Lien	9h - 13h	Lien	Paris
Optionnel	A	Microsoft Purview Cursus - Encryptions offers and keys management in Microsoft 365	03/02/2023	9h30 - 12h30	Lien	9h - 13h	Lien	Paris
Optionnel	B	Microsoft Purview Cursus - Manage Risk & Compliance	10/05/2023	9h30 - 12h30	Lien	9h - 13h	Lien	Paris
Optionnel	C	Microsoft Purview Cursus - Manage data governance in Azure	15/12 & 13/04	10h00 - 12h00	Lien 1 et 2	N/A	N/A	Remote



Microsoft Purview Cursus feedbacks

Share your feedback about the Microsoft Purview Cursus & this module!



➡ <https://aka.ms/MicrosoftPurviewCursus/ShareFeedback>

➡ <https://aka.ms/MicrosoftPurviewCursus>



Thank you.



Samuel Gaston-Raoul
Cloud Solution Architect

 aka.ms/Samuel



Thierry Matusiak
Information Security Architect

 aka.ms/Thierry



Philippe Dudognon
Cloud Solution Architect





Saida Laga
Information Security Architect





Frederic Gisbert
Cloud Solution Architect





Guillaume Lagache
Information Protection & Data Risk Specialist





Ali Bouhaddou
Cloud Solution Architect





Nabil Senoussaoui
Information Protection Specialist





Hélène Saulet
Partner Technology Strategist





Julien Lejard
Cloud Solution Architect





Jessica Grenat
Partner Marketing Advisor





Gauthier Lawny
Cyber Security Consultant





Charles Tostain
Cybersecurity Global Black Belt





Cyril Beurel
FastTrack Architect

