



Azure Purview Partner Accelerator

Maximize the Value of Your Data in the Cloud

Steve Young - Data & AI CSA

Franck Mercier- CSA

Will Johnson- CSA - SR SPECIALIST

Zeinab Mokhtarian Koorabbasloo - Sr. Program Manager

November 9 & 10, 2021



LAND ACKNOWLEDGEMENT

Toronto, Mississauga Ontario

We acknowledge the land we are meeting on is the traditional territory of many nations including:

- **Mississaugas of the Credit**
- **Anishinaabeg** (Ah-nish-ee-Na-bey)
- **Chippewa** (Chip-ee-wah)
- **Haudenosaunee** (Ho-Da-No-Shau- Nee)
- **Huron Wendat peoples** (When-Dat)

It is now home to many diverse First Nations, Inuit and Métis (**May-tee**) peoples.

We also acknowledge that Toronto is covered by Treaty 13 with the Mississaugas of the Credit.

Virtual Workshop Code of Conduct

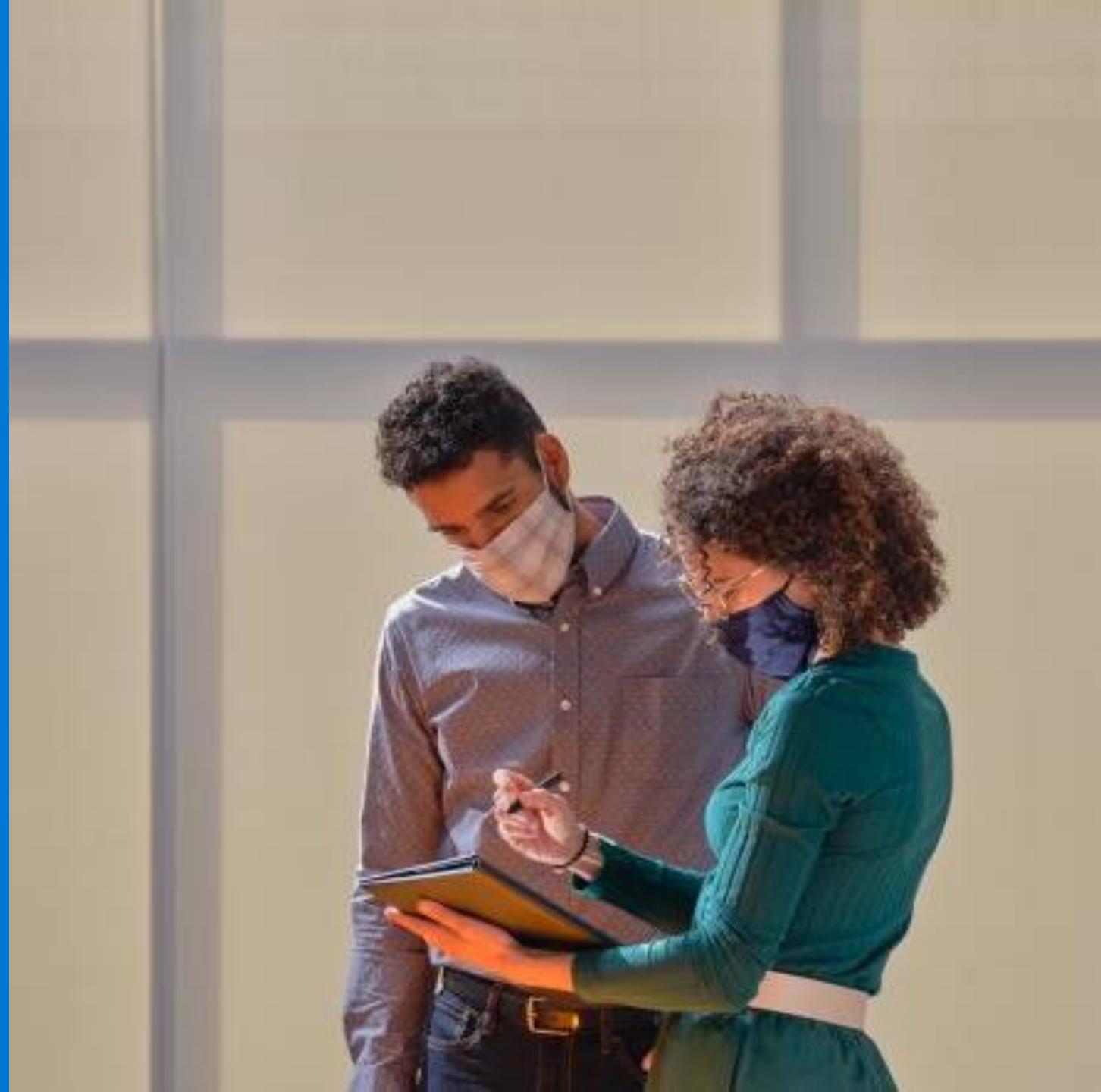
To ensure all participants have a safe and enjoyable learning experience, please be respectful to other participants and to the training facilitators.

Please do not use audio or the chat window for any of the following:

- Swearing or profanity
- Negative or hateful comments to other participants
- Loud or disruptive conduct
- Posting inappropriate comments or images
- Any conduct deemed inappropriate by the facilitators

If any participant does not follow this code of conduct, they may be removed from the training by the facilitator.

Introduction



Speakers

Steve Young - Data & AI CSA (Microsoft)

styoung@microsoft.com <http://5minutebi.com>

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

Will Johnson- CSA (Microsoft)

will.johnson@microsoft.com

Franck Mercier- CSA (Microsoft)

franmer@microsoft.com

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

Zeinab Mokhtarian Koorabbasloo

Sr. Program Manager (Microsoft)

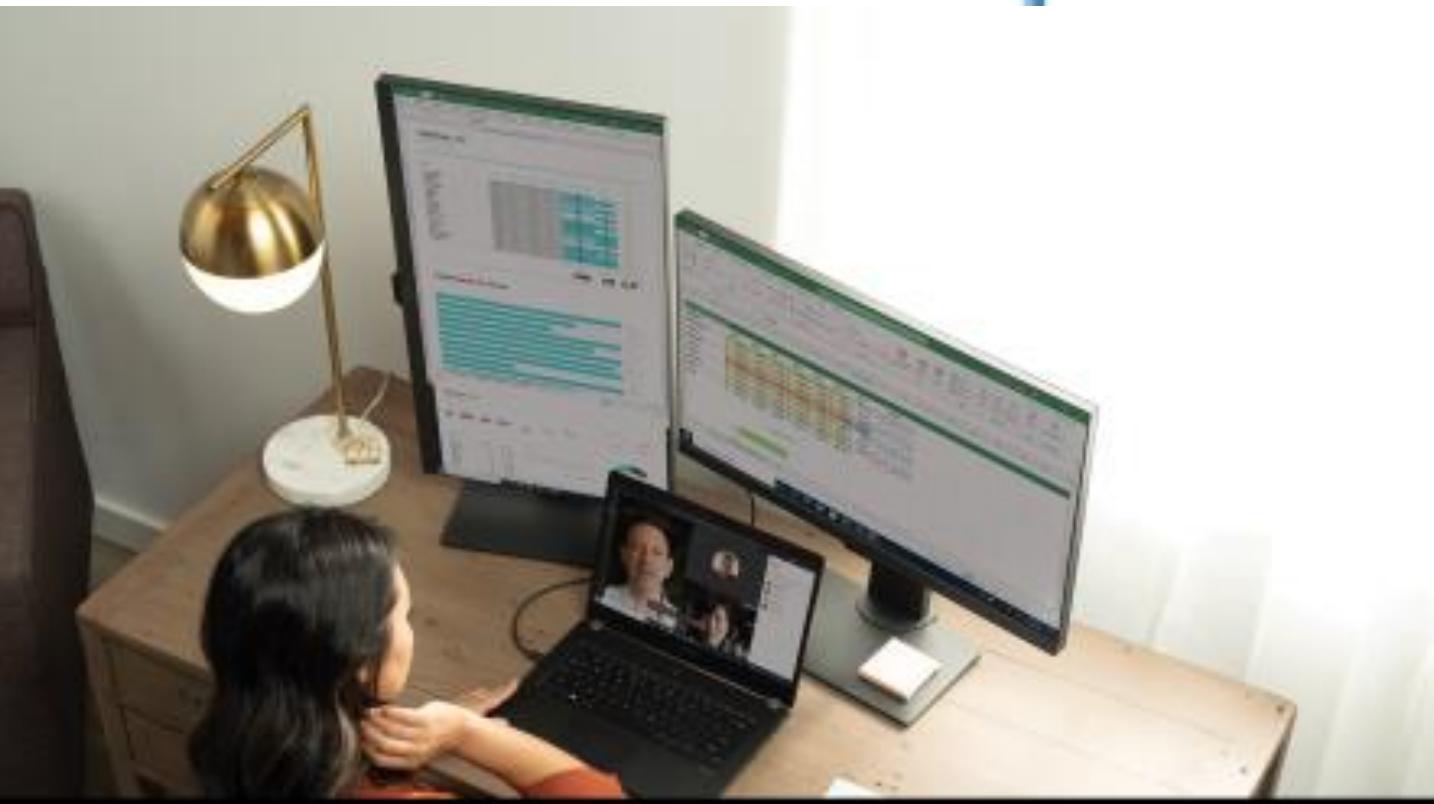
zeinam@microsoft.com

Agenda - Day 1

- Welcome and Introduction
 - Setting the Stage - Data Governance Problem
 - An introduction to Azure Purview
 - Partner Opportunity & Use Cases
 - Demonstration
-
- Q&A / Your Thoughts
 - Additional Resources & Close Out



Setting The Stage - The Data Governance Problem



Data Governance Problem – Business has changed

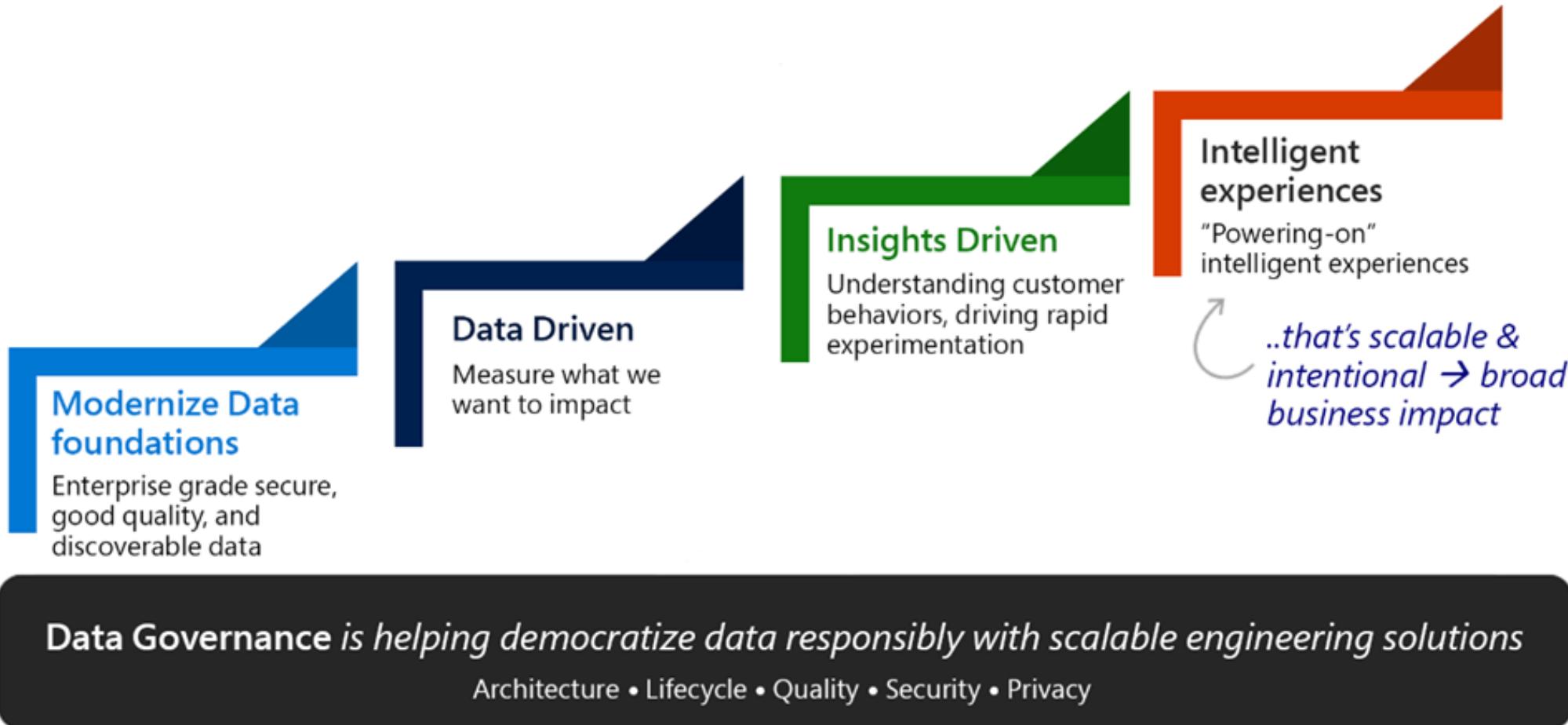


Figure 1. Data governance is the foundational pillar of the Microsoft Enterprise Data Strategy

Data Governance Problem – The Danger of Dark Data

Need:

Provide our Customers the

ability to **discover, catalog and secure** their data

Why do we Need Data Governance? – Not Just Finding

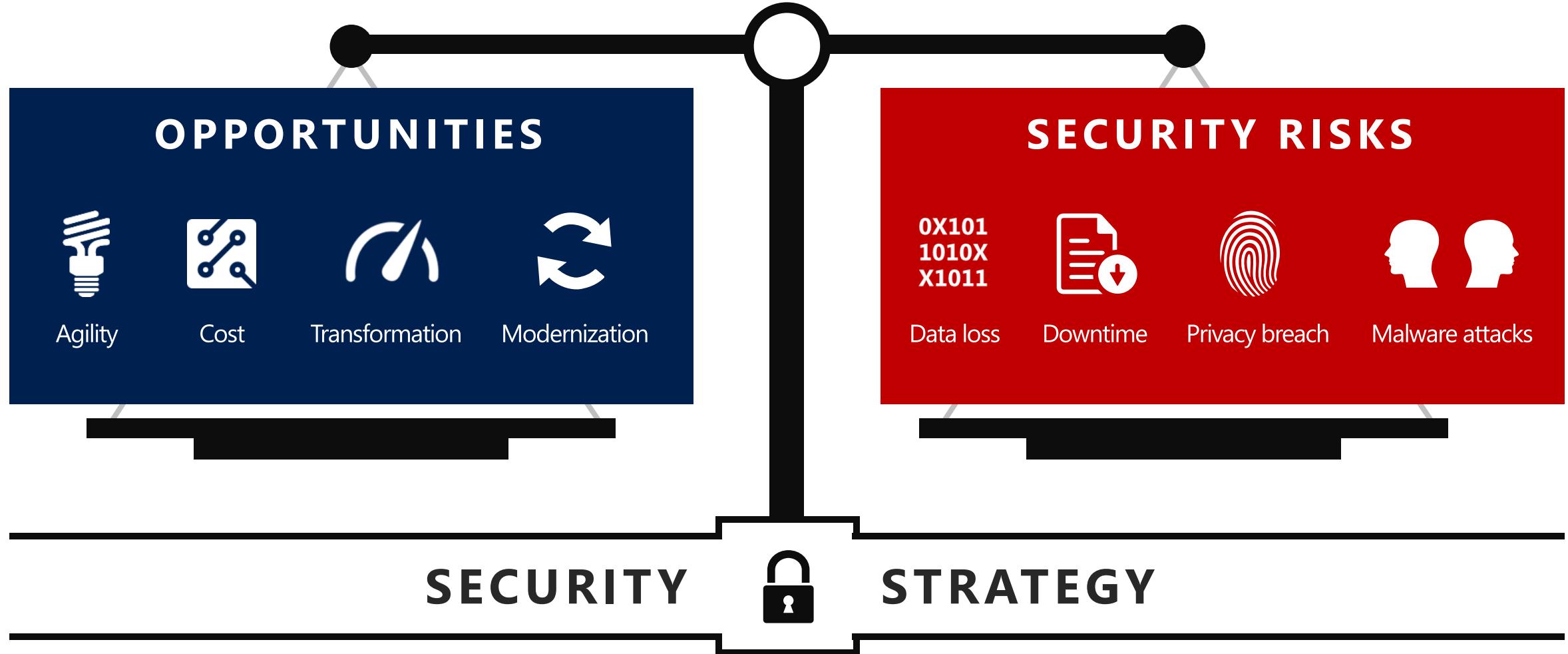
Data Breach Statistics: How big of a problem?

- An estimated 74% of companies have more than 1,000 stale sensitive files. ([Varonis](#))
- An estimated 41% of companies have more than 1,000 sensitive files including credit card numbers and health records left unprotected. ([Varonis](#))
- An estimated 21% of all files are not protected in any way. ([Varonis](#))

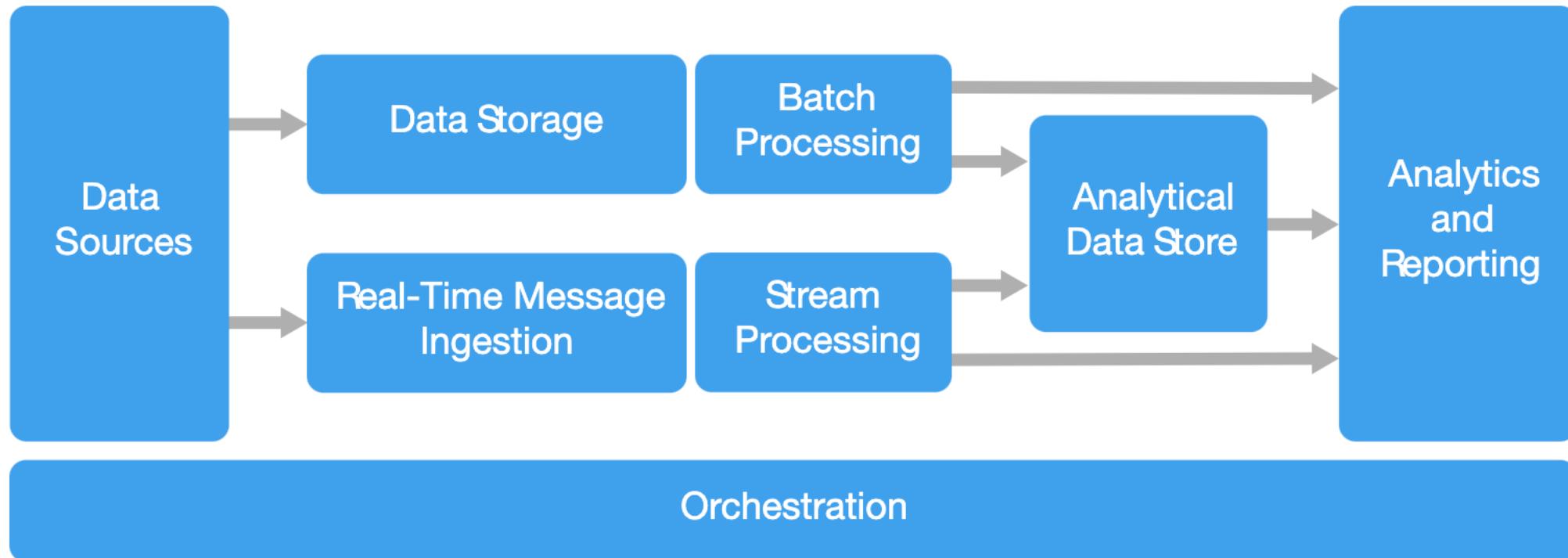
Not all Data Breaches have be hacking.

People can expose data that they shouldn't as they are unaware of its sensitivity.
In some cases, a data breach is just an Excel Export away.

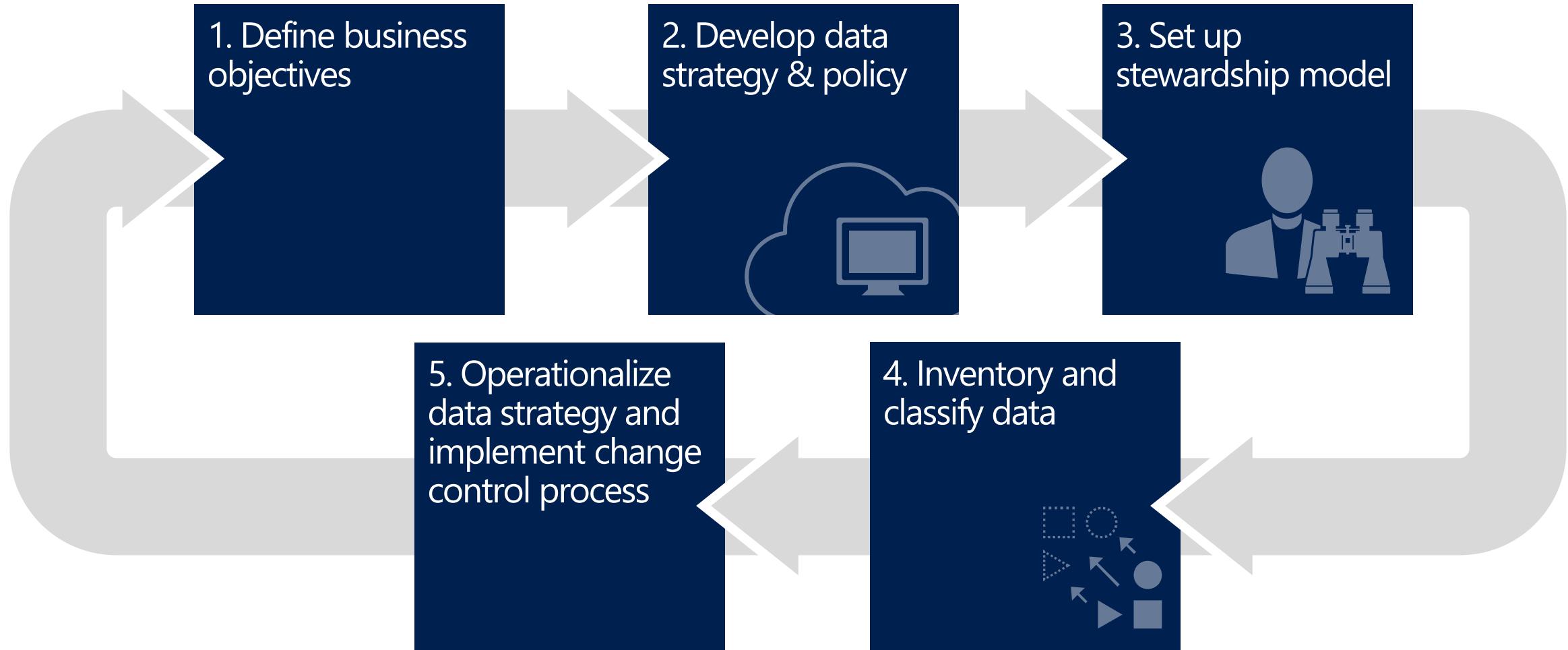
Balance Opportunities and Risks



Modern Data Warehouse



Data Governance is more than just an App



Azure Purview In Detail



Azure Purview Now in General Availability

Maximize the business value of Your data... no matter where it resides.

Azure Purview

57B+ **2,300+**

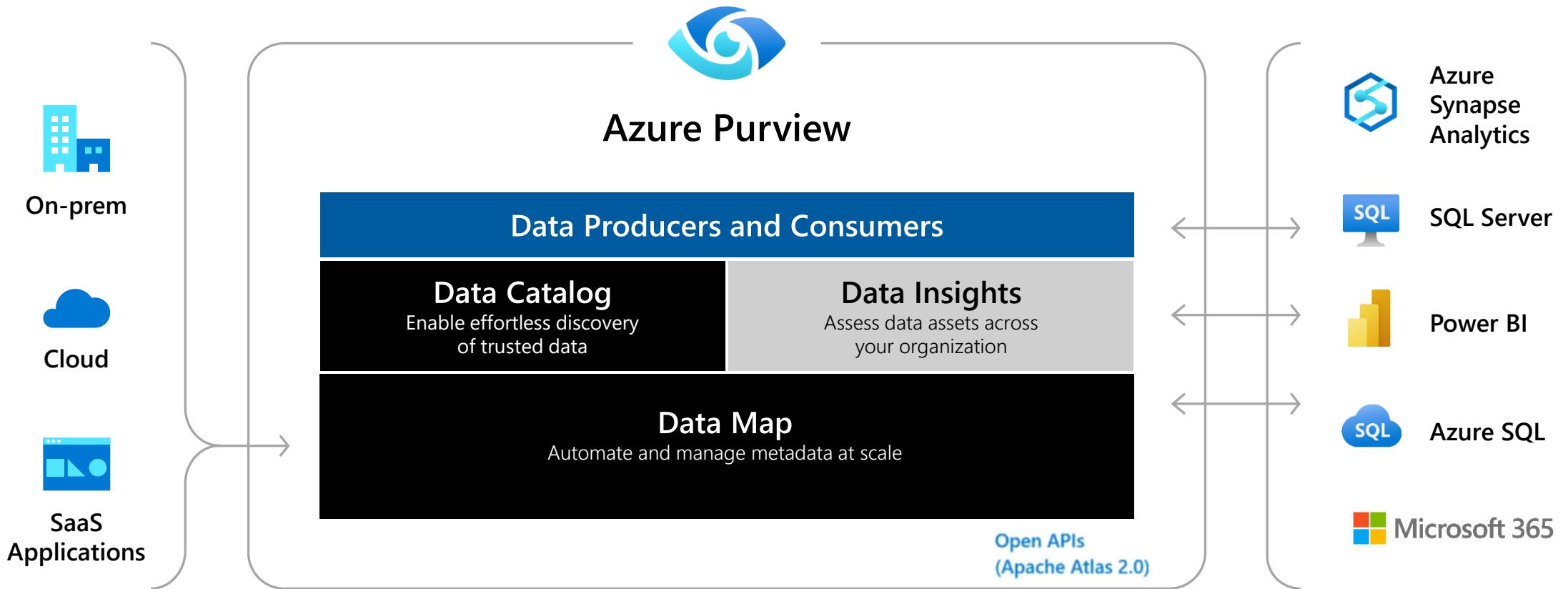
Data assets
discovered

Purview
customers



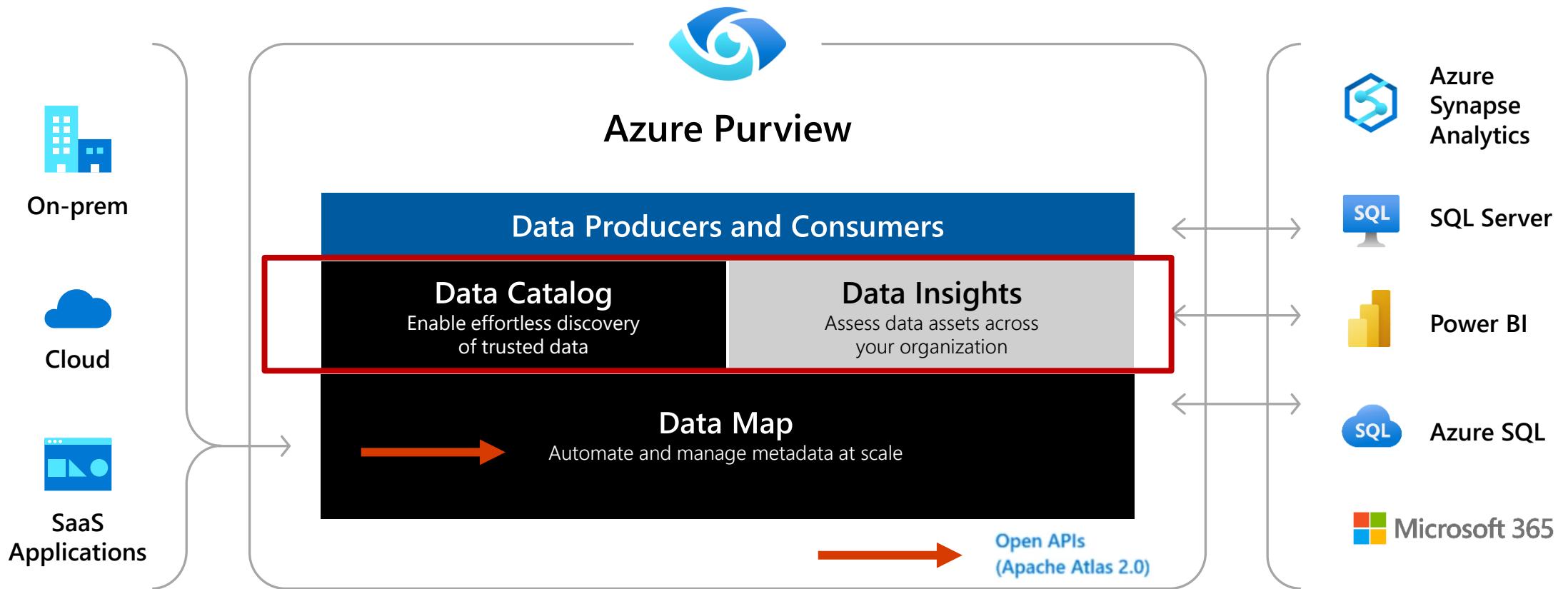
Unified Data Governance with Azure Purview

Generally Available
Preview



Unified Data Governance with Azure Purview

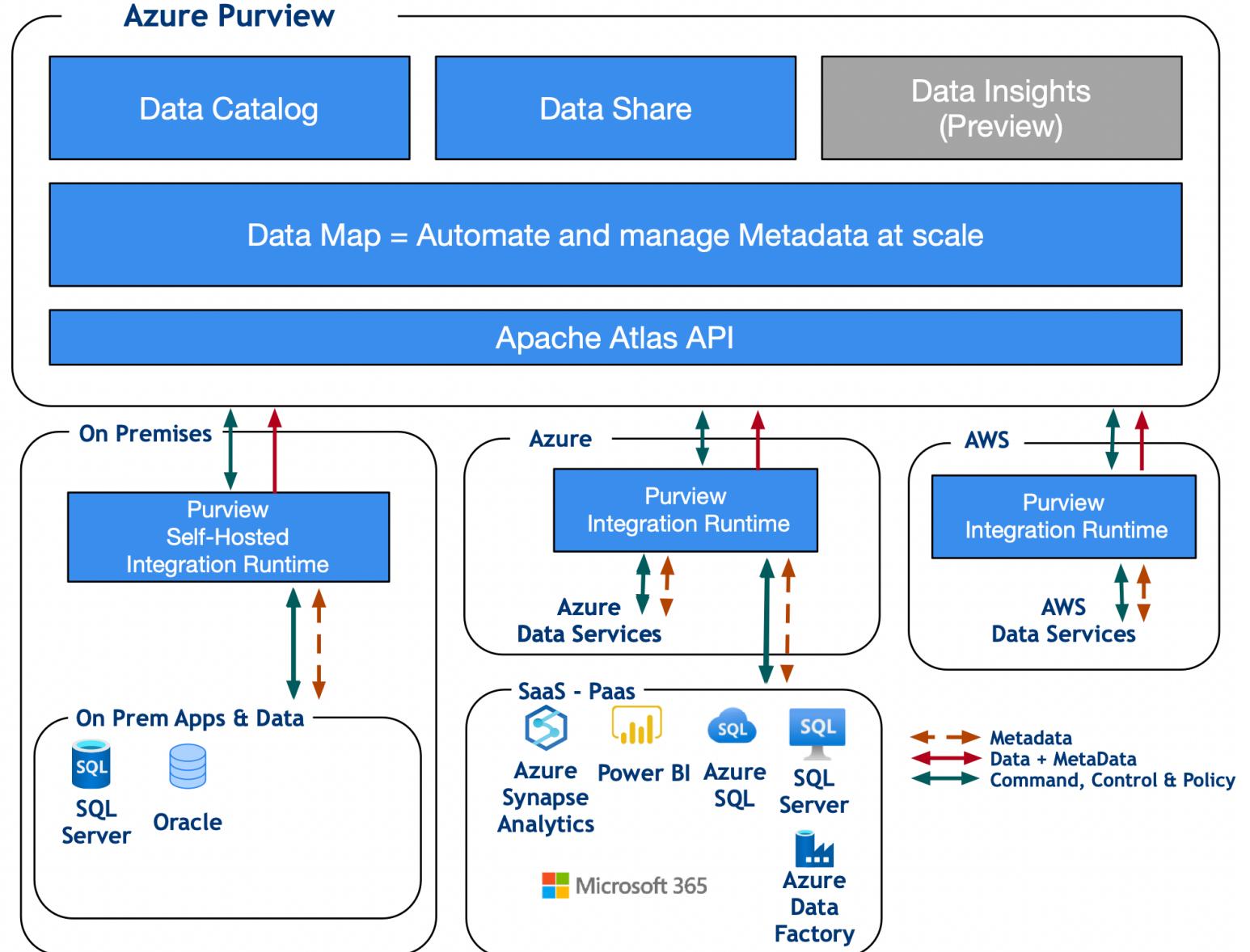
Generally Available
Preview



Azure Purview Data Overview

IMPORTANT!!

The data stays in the source,
only the metadata
is brought over.



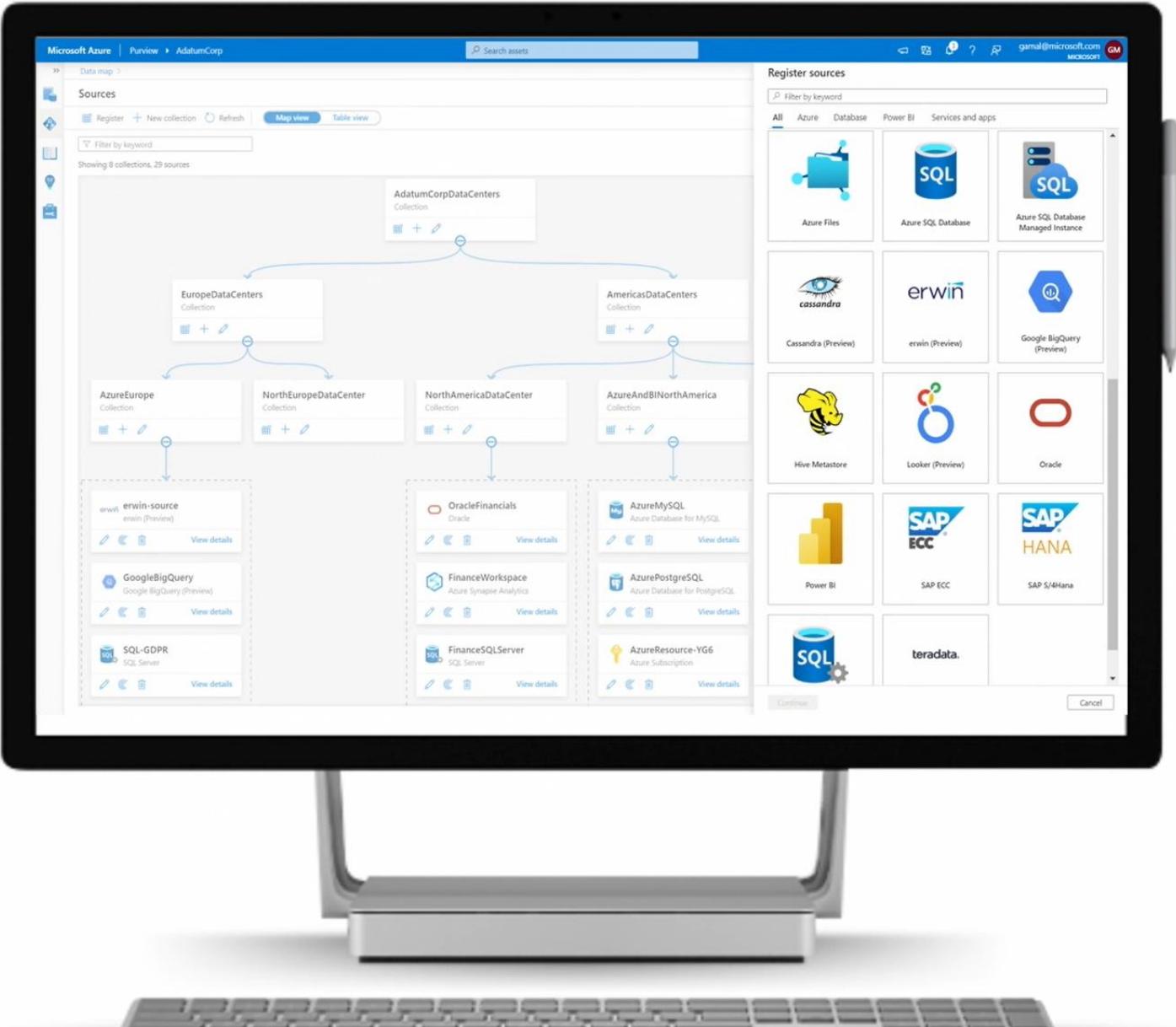
Automated Data Discovery using Purview Data Scan

Supports hybrid data landscapes

Automated data scanning, classification & lineage extraction

200+ built-in data classifiers

Support for 35 data sources and growing



Data Scanning

Azure Purview

Public Preview

Coming Soon

Google Cloud

erwin

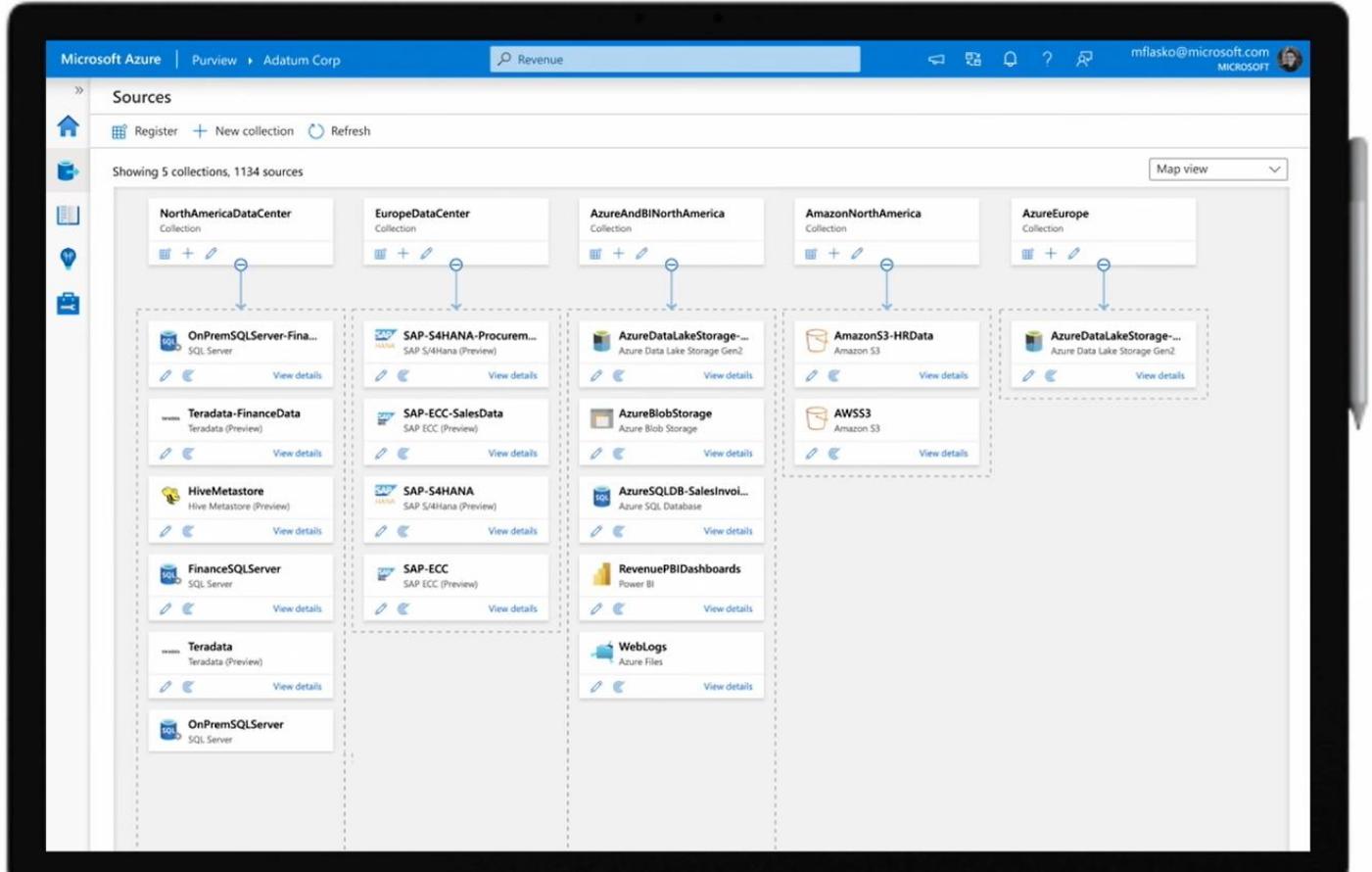


Make data meaningful using Purview Data Map

A graph describing the data assets and their relationships accross your data estate

Fine-grained access controls

Open platform



>

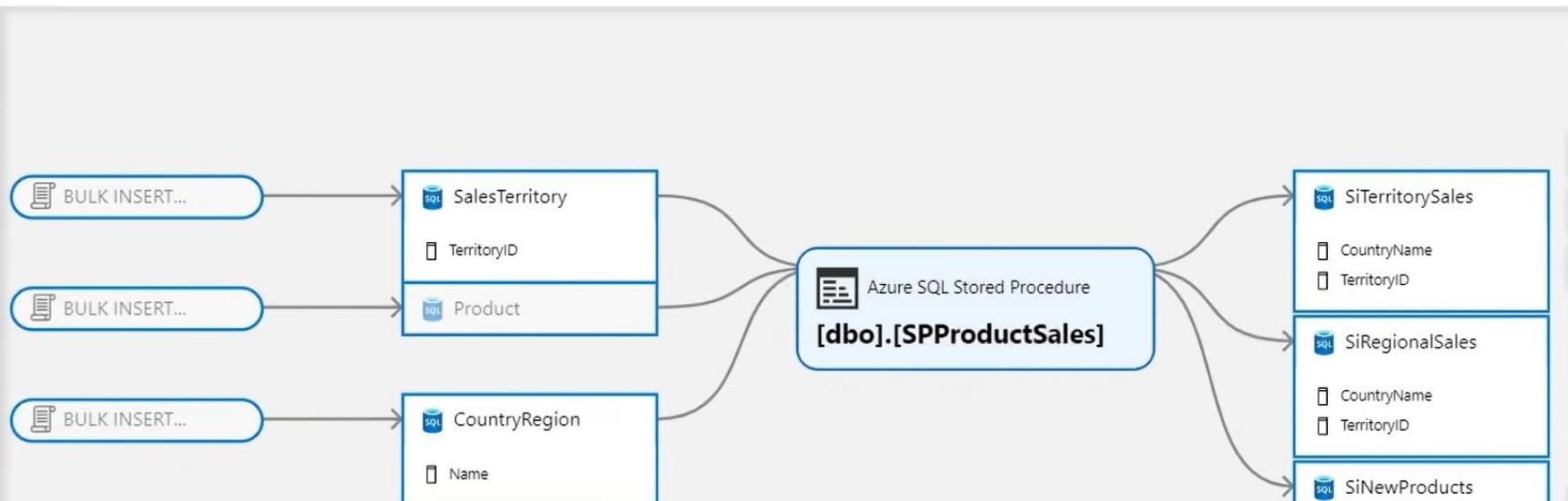
Data catalog >

RevByCustomer

Azure Dedicated SQL Pool Table

[Edit](#) [Refresh](#) [Delete](#)[Overview](#) [Properties](#) [Lineage](#) [Contacts](#) [Related](#)

Updated on September 10, 2021 7:44 PM UTC

 Search for assets or processes[Input datasets](#) [Columns](#)[Search and select columns](#) Show only assets with all 2 columns Columns CountryRegion ModifiedDate Name SalesTerritory CountryRegionCode Name TerritoryID Product ListPrice ModifiedDate

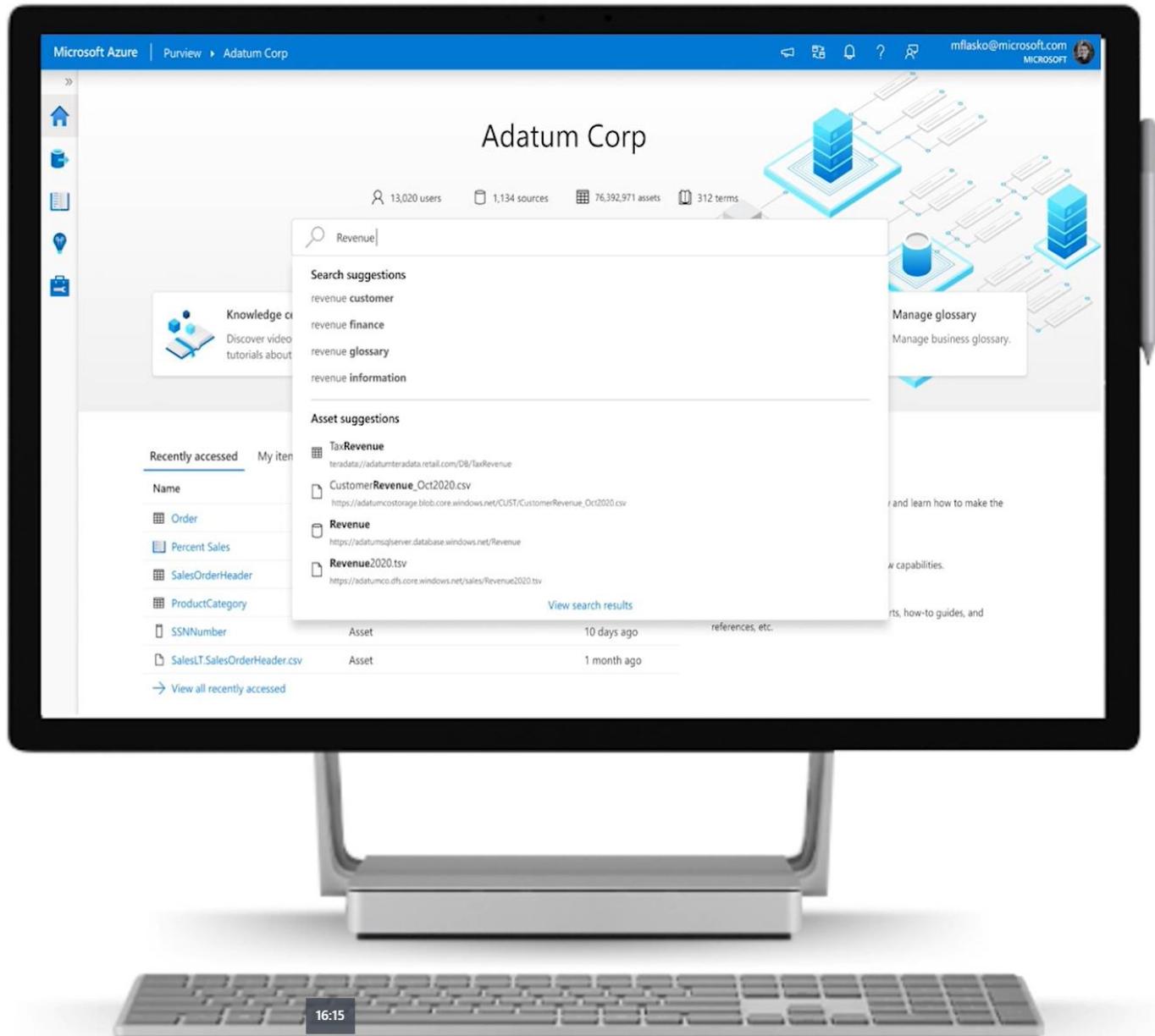
25:04

Find Trusted Data with Purview Data Catalog

Search, Browse & Curate

Intelligent Recommendations

Lineage Visualization



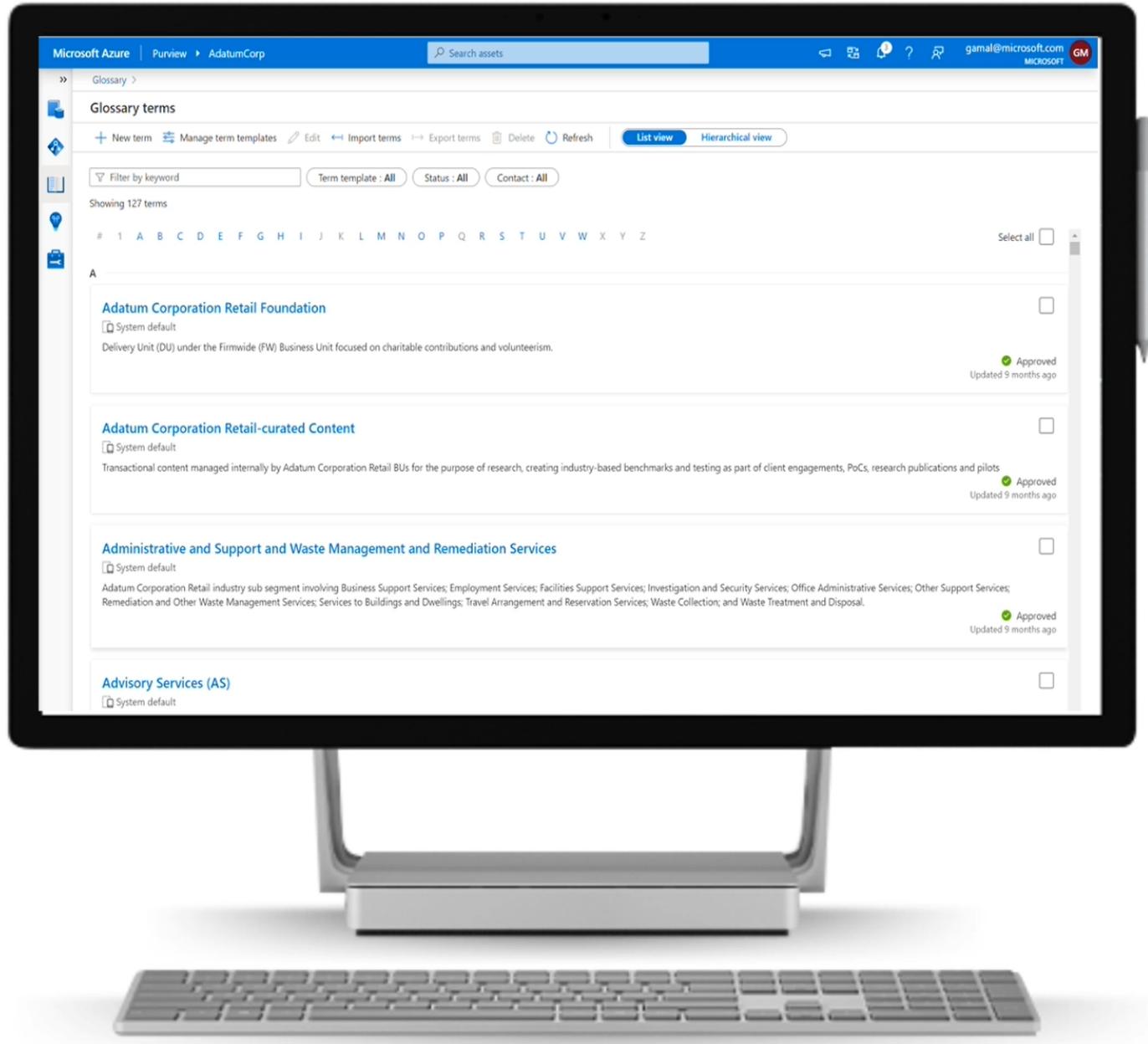
Make data relevant with business context

Azure Purview Business Glossary

Search by Business Term

Hierarchy Support

Integrated Approval Workflows



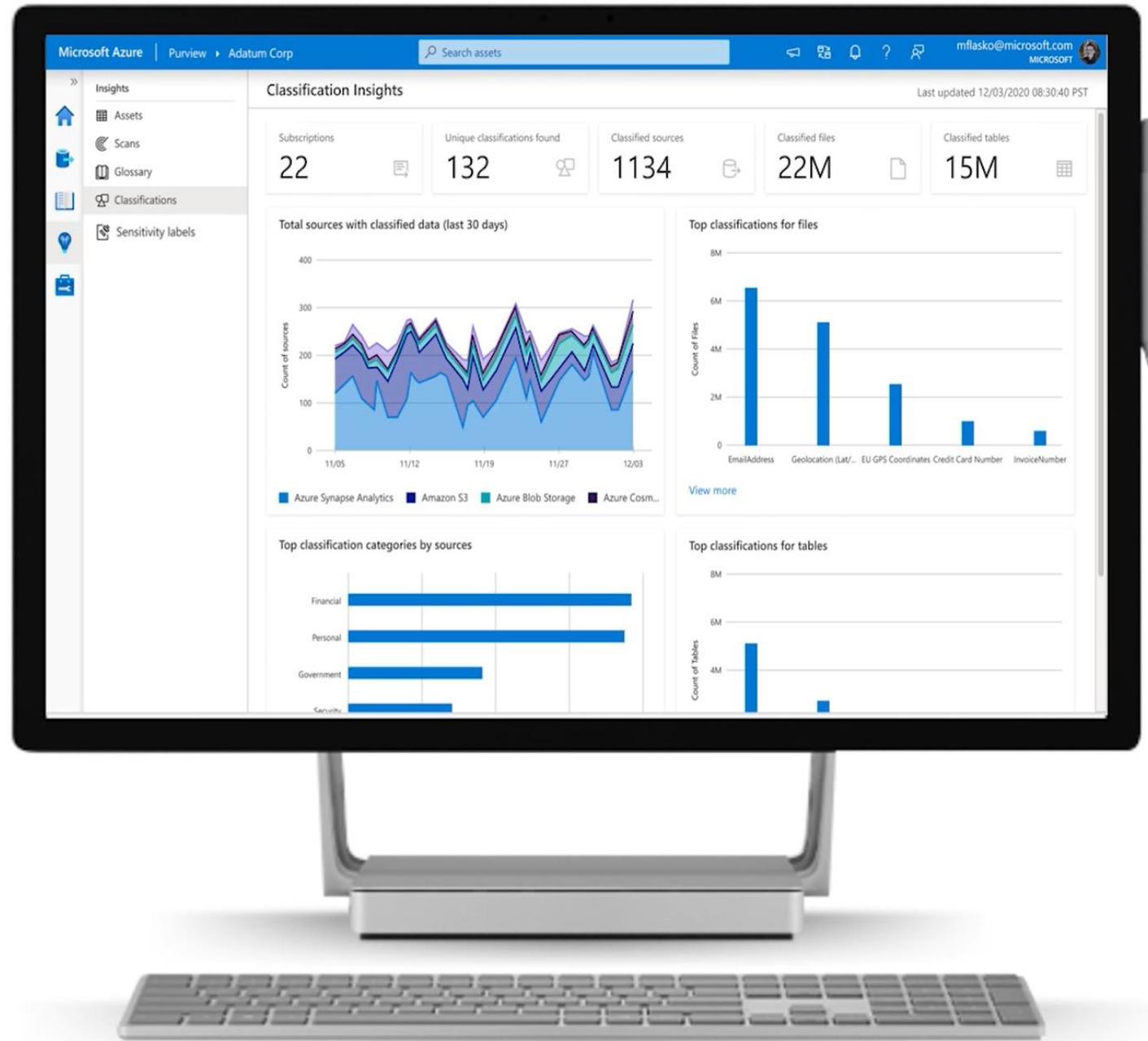
Get a bird's-eye view of your data landscape with Purview Data Insights

Data Asset distribution

Sensitive Data

Data Scanning Coverage

Business Glossary Utilization



Roadmap & Pricing

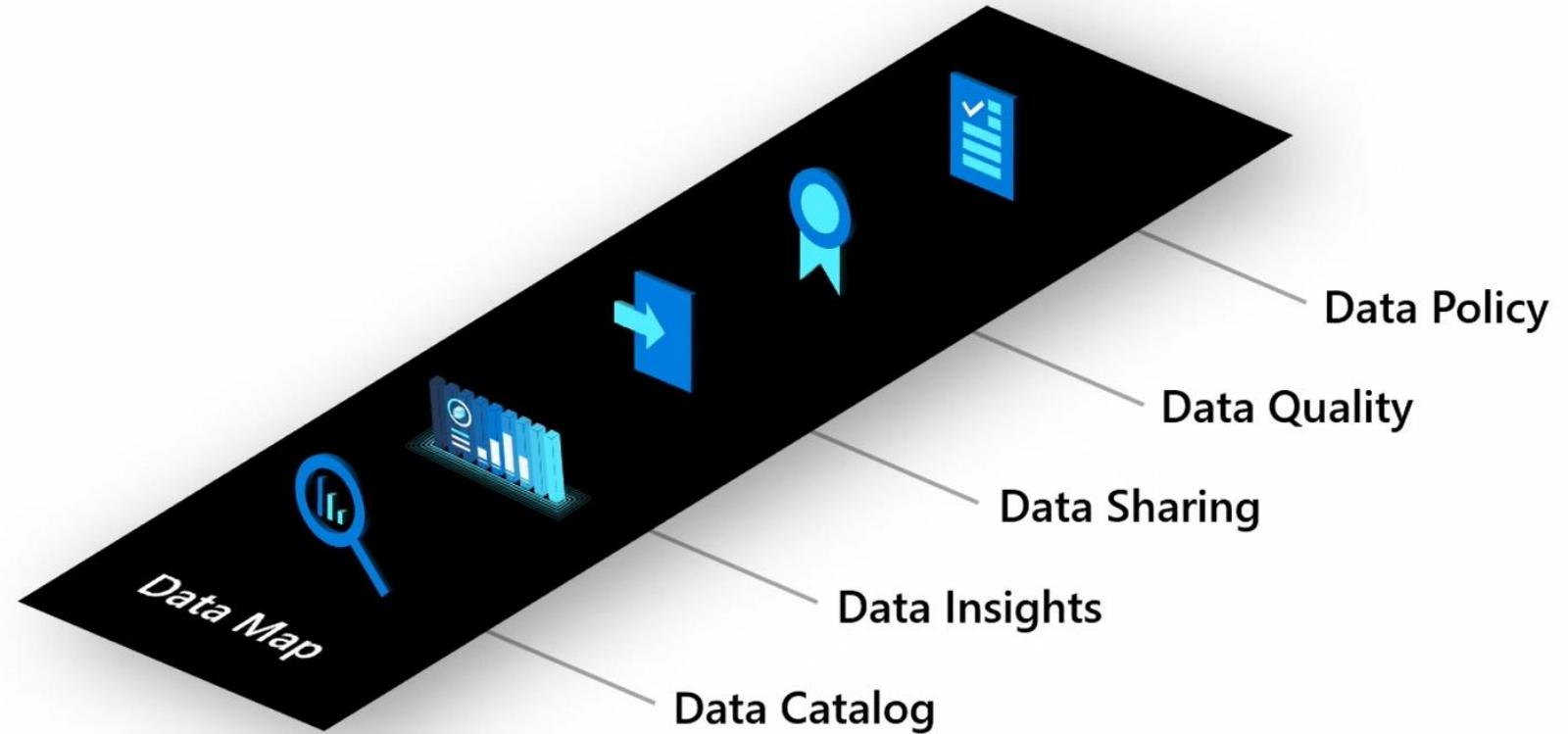


Azure Purview Features Roadmap

	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 Information Protection Sensitivity Labels support		●

Unified Data Governance with Azure Purview

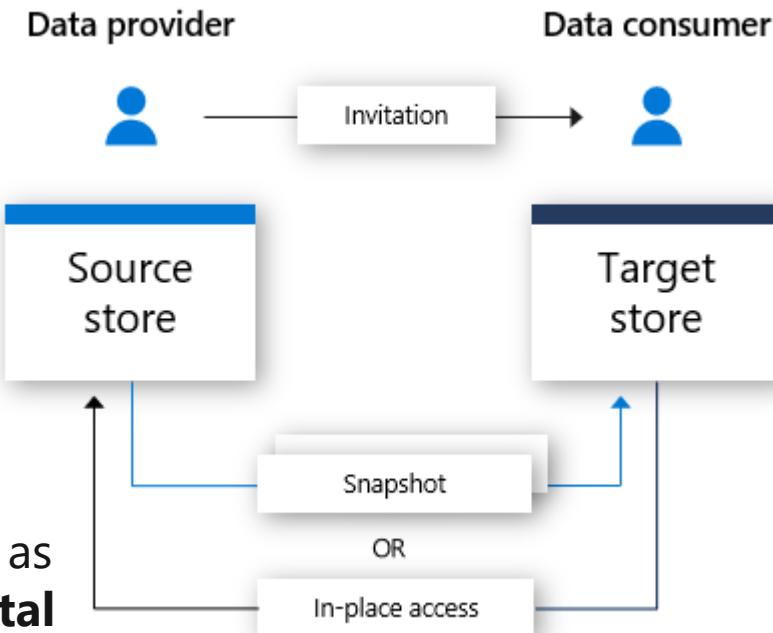
Roadmap



Azure Data Share - GA

Key capabilities

- Share data from the list of [supported data stores](#) with customers and partners outside of your organization
- Keep track of who you have shared your data with
- Choice of snapshot or in-place sharing
- Allow your customers to pull the latest version of your data as needed, or allow them **to automatically receive incremental changes** to your data at an interval defined by you

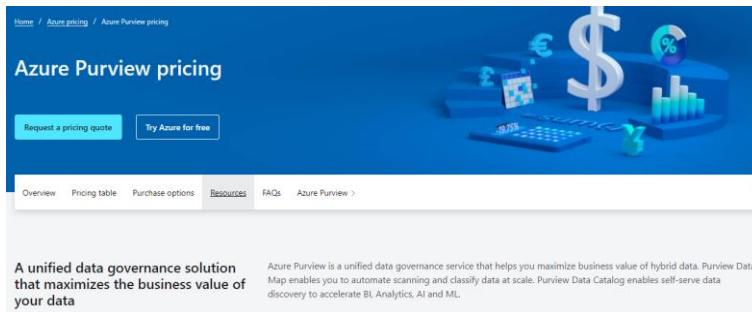


Source:

[What is Azure Data Share? | Microsoft Docs](#)

Azure Purview Pricing

- [Pricing – Azure Purview – Pricing Calculator](#)
- [Pricing - Azure Purview – An Overview](#)



The screenshot shows the Azure Purview pricing page. At the top, there are two buttons: "Request a pricing quote" and "Try Azure for free". Below these are navigation links: "Overview", "Pricing table", "Purchase options", "Resources" (which is selected), "FAQs", and "Azure Purview >". A large section titled "A unified data governance solution that maximizes the business value of your data" contains a sub-section about Purview Data Map. The main content area is titled "Azure Purview Data Map" and describes its function as the foundation for managing data governance at cloud scale. It mentions Data Map stores metadata, lineage, classifications, and other annotations associated with data assets.

Azure Purview Data Map

The Azure Purview Data Map is the foundation for managing data governance at cloud scale. Data Map stores metadata, lineage, classifications, and other annotations associated with data assets and serves the requests for this information through Apache Atlas APIs or through the applications built on Data Map including Data Catalog. A search request for the data assets tagged as sensitive or a rendering of a lineage graph in the Data Catalog are examples of a request served up by the Data Map.

Data Map can be automatically populated at cloud scale and kept up to date with the scanning of data sources registered with the Purview account. The scans can be configured to also classify data based on metadata and content inspection. Data Map comes with built-in features (e.g., Resource Sets) to optimize storage of data assets.

Elastic Data Map

By default, a Purview account is provisioned with a Data Map of at least 1 Capacity Unit. 1 Capacity Unit supports requests of up to 25 data map operations per second and includes storage of up to 2 GB of metadata about data assets.

The Data Map can scale capacity elastically based on the request load. Request load is measured in terms of data map operations per second. As a cost control measure, a Data Map is configured by default to elastically scale within the [elasticity window](#).

A data map operation is a create, read, update, or delete of an entity in the Data Map. Examples of an entity include a data asset or a lineage relationship between two data assets. A search request may require multiple operations depending on the assets returned and complexity of the request. Storage size of an entity may vary in size depending on the type of entity and annotations associated with the entity.

Data Map requires an additional Capacity Unit for every 2 GB of metadata storage required. For example, a Data Map with 2 GB of metadata storage is billed at 1 Capacity Unit per hour. If the addition of new data assets increases the size to 2.1 GB, the Data Map is billed at 2 Capacity Unit per hour.

Capacity Unit	Price
	\$0.411 per 1 Capacity Unit Hour

Pricing calculator

Configure and estimate the costs for Azure products

Products Example Scenarios Saved Estimates FAQs

Select a product to include it in your estimate.

Purview



Azure Purview

A unified data governance solution that maximizes the business value of your data

Your Estimate

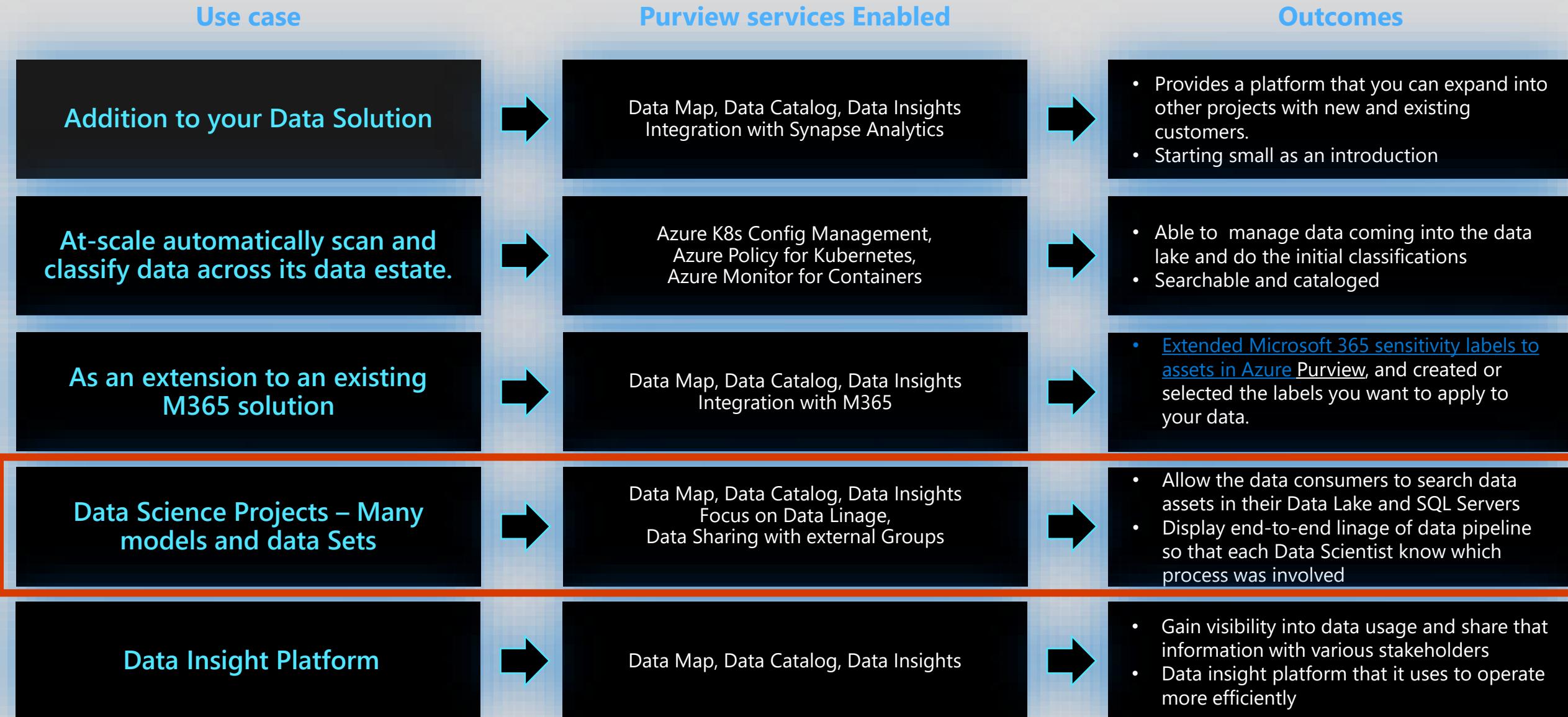
Your Estimate

Select a product, choose an example scenario, or open a saved estimate to view monthly costs.

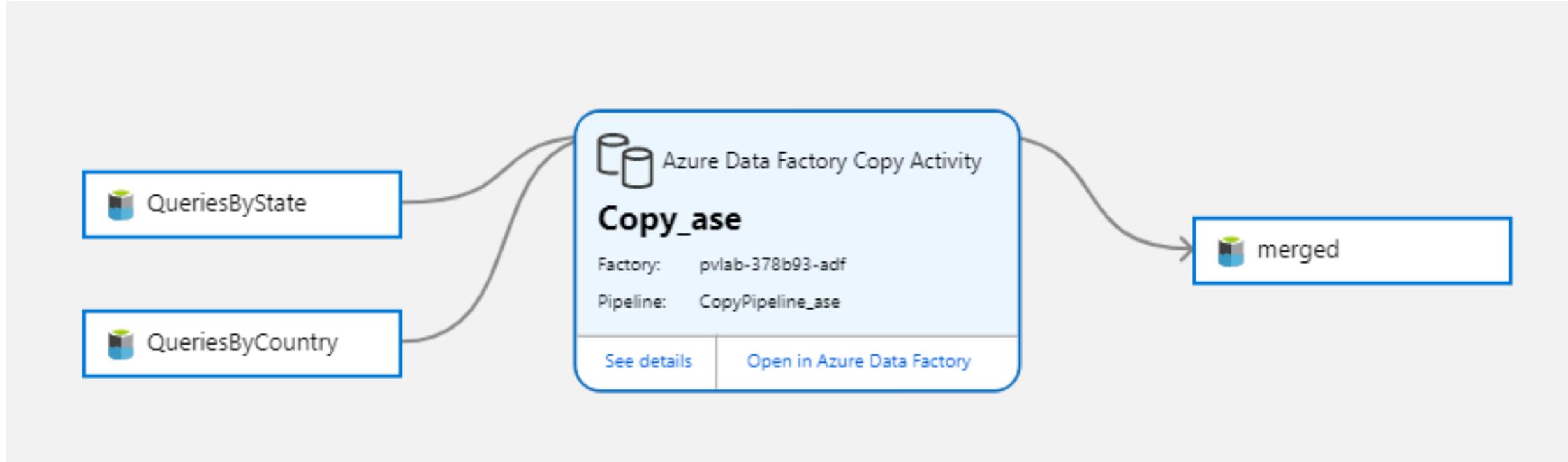
Partner Opportunities and Use Cases



Azure Purview Partner Opportunities



Use Cases – Theme is Lineage



Source:

<https://docs.microsoft.com/en-us/azure/purview/catalog-lineage-user-guide>
Data lineage in Azure Purview - Azure Purview | Microsoft Docs

Use Case #1 - Machine Learning Lineage

MICROSOFT SOLUTIONS

Purview Machine Learning Lineage Solution Accelerator

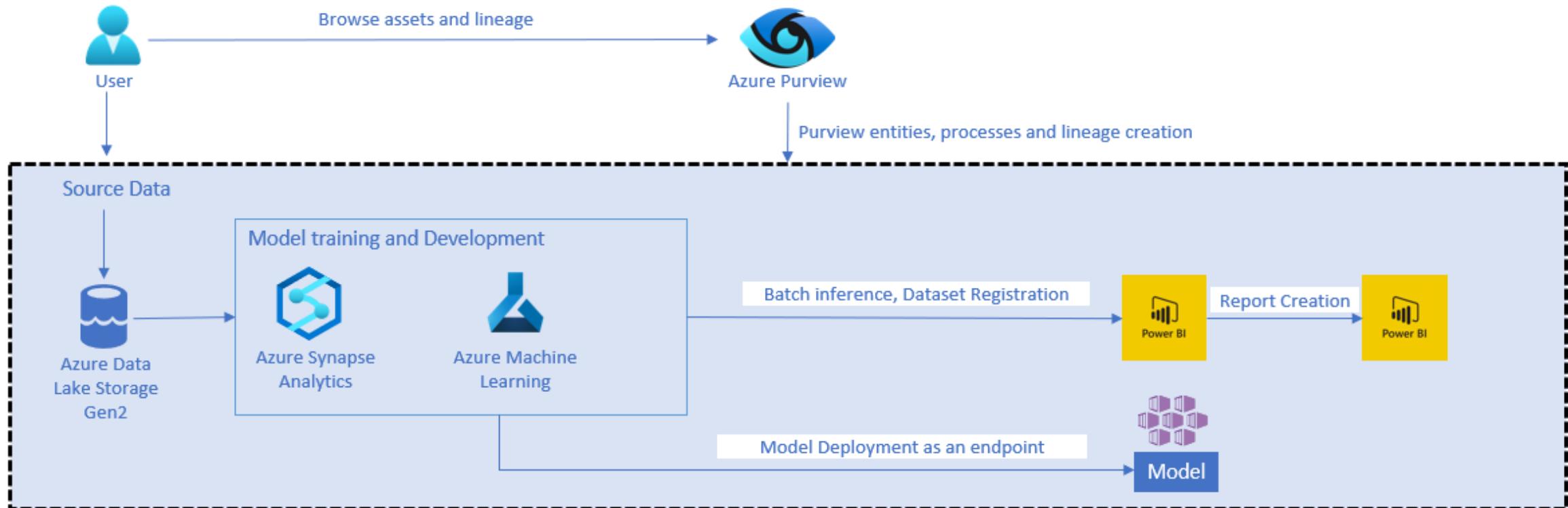
EARLY ACCESS ENGINEERING



Source:

[Purview-Machine-Learning-Lineage-Solution-Accelerator: Solution accelerator to help build Machine Learning Lineage \(github.com\)](https://github.com/microsoft/Purview-Machine-Learning-Lineage-Solution-Accelerator)

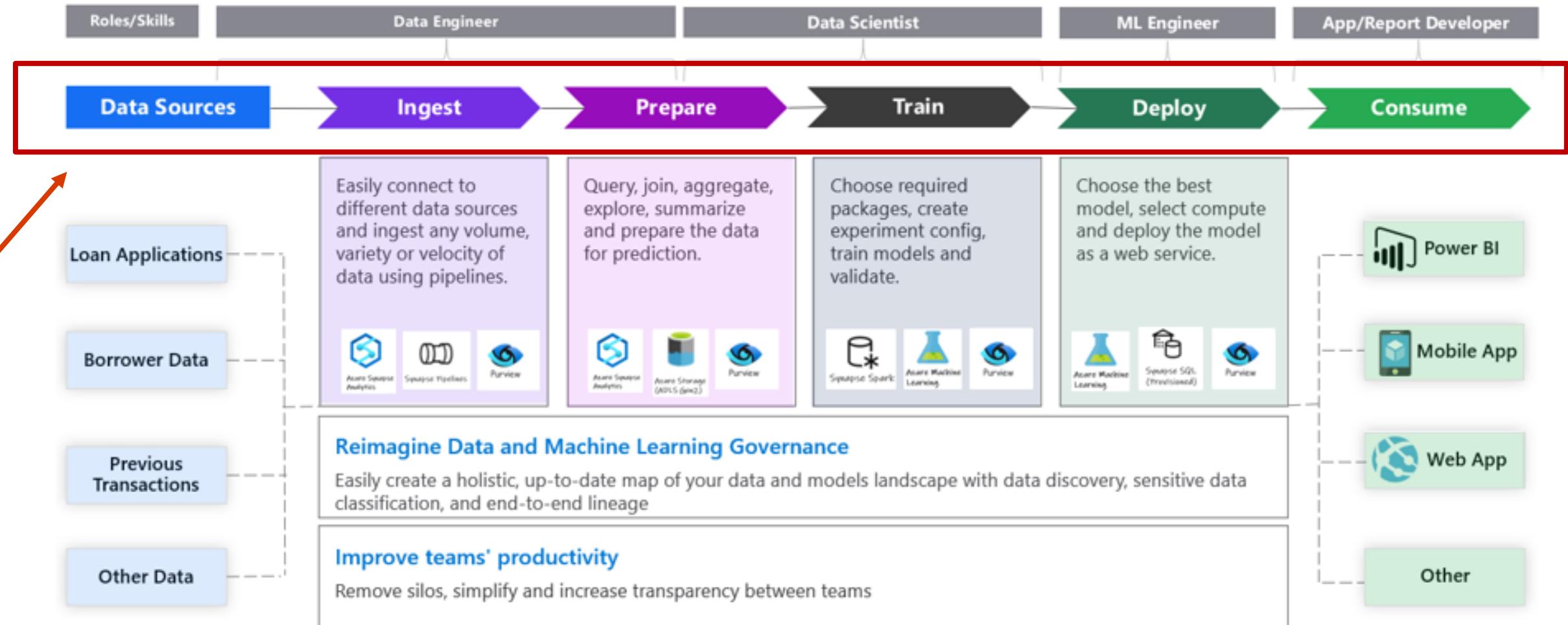
Use Case - Machine Learning Lineage



Source:

[Purview-Machine-Learning-Lineage-Solution-Accelerator: Solution accelerator to help build Machine Learning Lineage \(github.com\)](https://github.com/Azure/Purview-Machine-Learning-Lineage-Solution-Accelerator)

Use Case - Sample Credit Risk Prediction ML Process

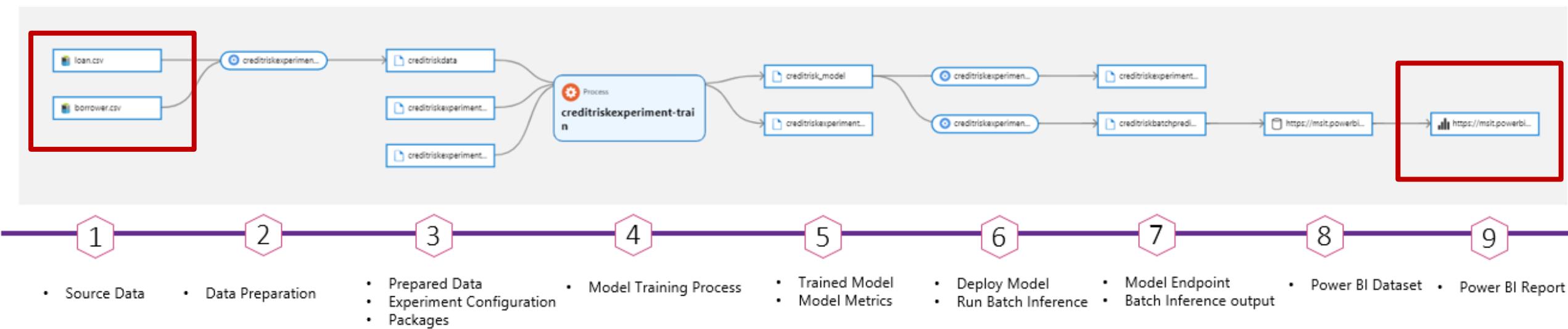


Source:

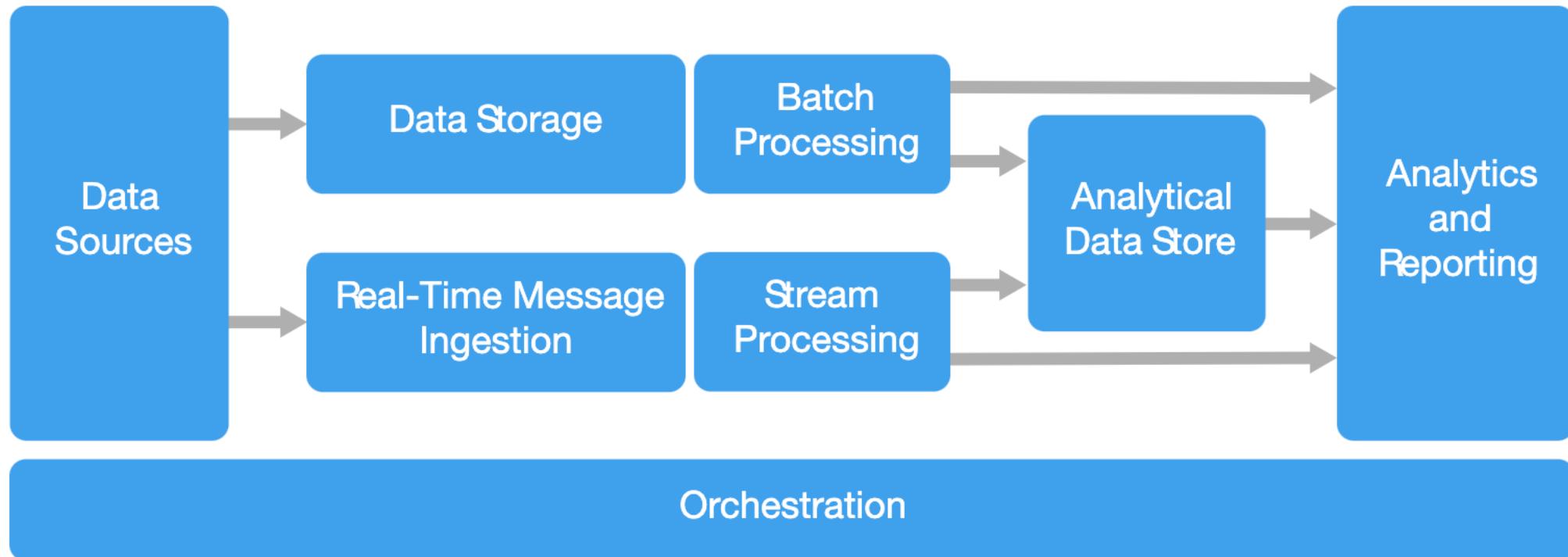
[Purview-Machine-Learning-Lineage-Solution-Accelerator: Solution accelerator to help build Machine Learning Lineage](#)

Machine Learning Use Case - Purview ML Process Lineage

Machine Learning pipeline Lineage in Purview Studio

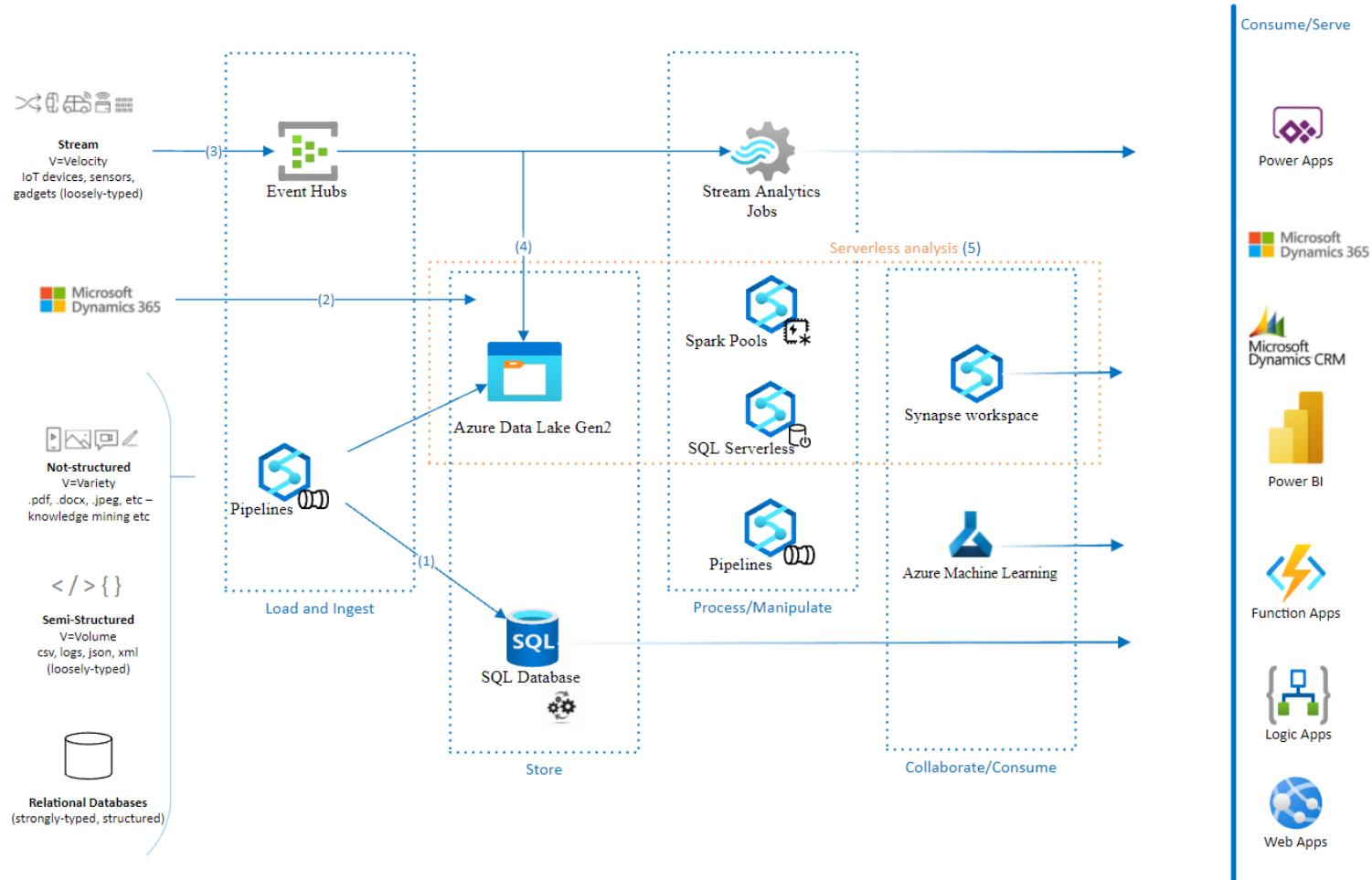


Use Case #2 – Classic Analytic Data Solution



Use Case #2a – Classic Analytic Data Solution (SMB)

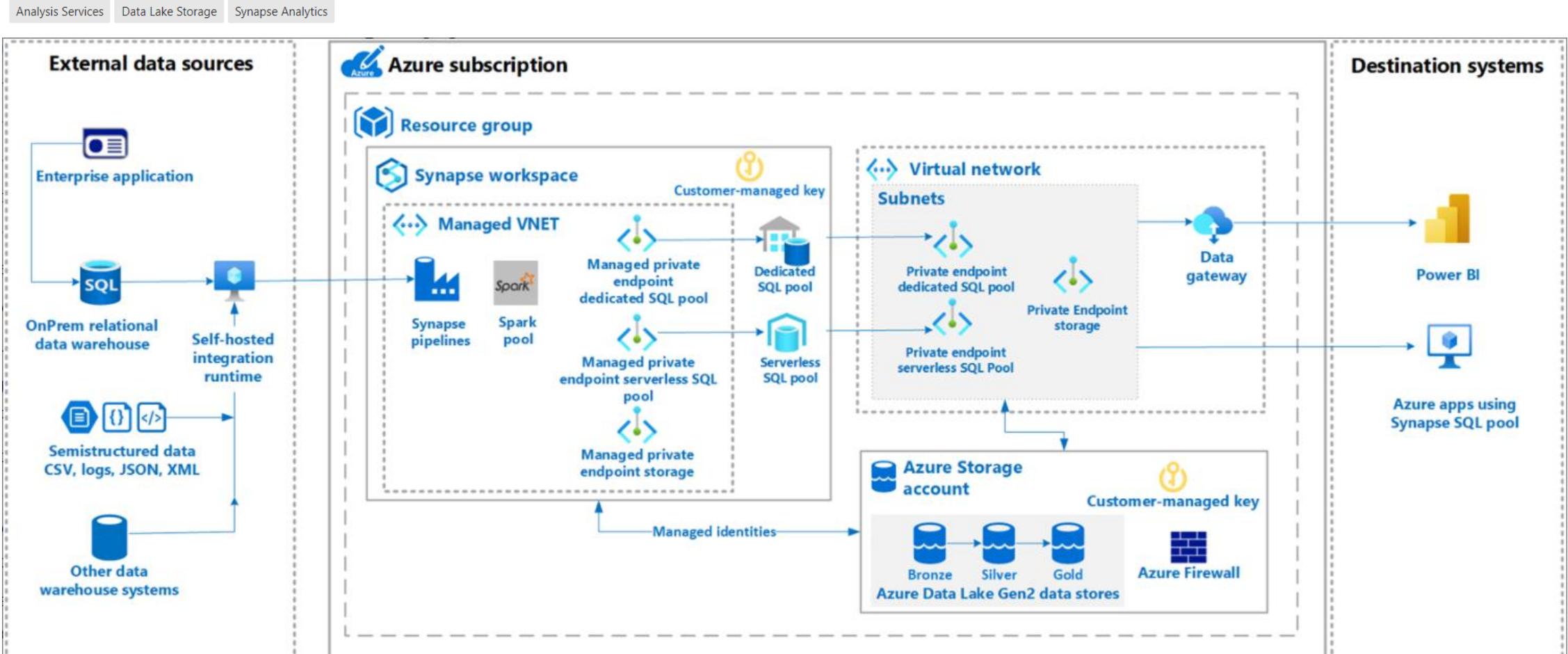
Data Lake SQL Database Synapse Analytics Dynamics 365 Microsoft Power Platform



Source:

[Modern data warehouse for small and medium business - Azure Example Scenarios | Microsoft Docs](#)

Use Case #2b – Classic Analytic Data Solution (ENT)



Source:

[Big data analytics with enterprise-grade security using Azure Synapse - Azure Solution Ideas | Microsoft Docs](#)

Use Case #2 – Classic Analytic Data Solution

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

The screenshot shows the Microsoft Azure Purview Data Catalog interface. The top navigation bar includes 'Microsoft Azure', 'Purview', 'Adatum Corp', a search bar containing 'customer north america', and user information for 'daperlov@microsoft.com'.

The main area displays 'Search results for customer north america'. The left sidebar features a 'Collection' section with 'AzureAndBINGNorthAmerica' (14688), 'Adatum Corp' (12), and 'AmazonNorthAmerica' (2). Below this are sections for 'Classification' (Email Address, Person's Name, U.S. Phone Number), 'Contact', 'Label', and 'Glossary term'.

The search results list 12 out of 12 results, sorted by Relevance. The results include:

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

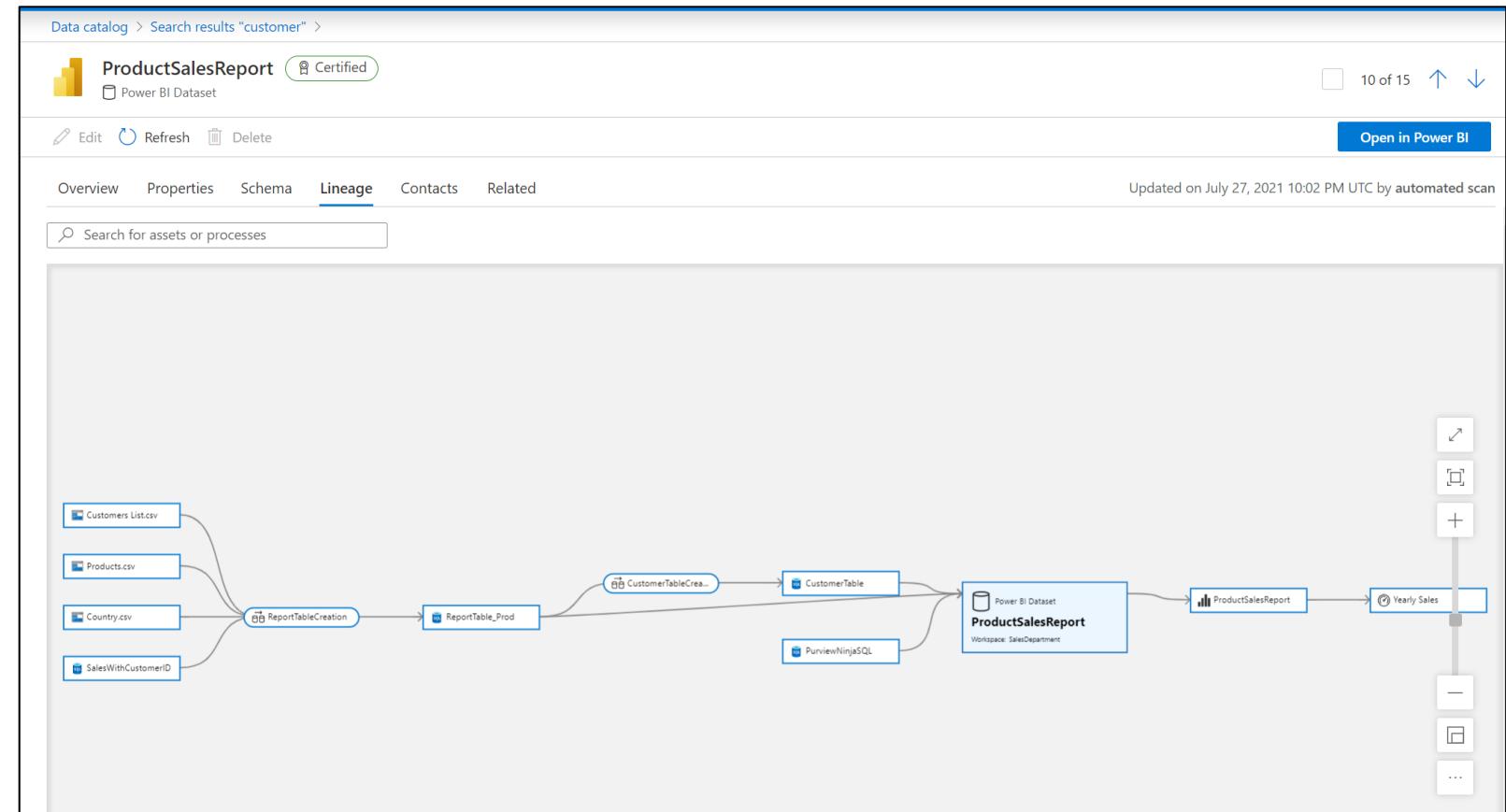
Pagination controls at the bottom show '< Previous', 'Page 1 of 1', and 'Next >'.

Use Case #2 – Classic Analytic Data Solution

Trace lineage of data assets across the data estate

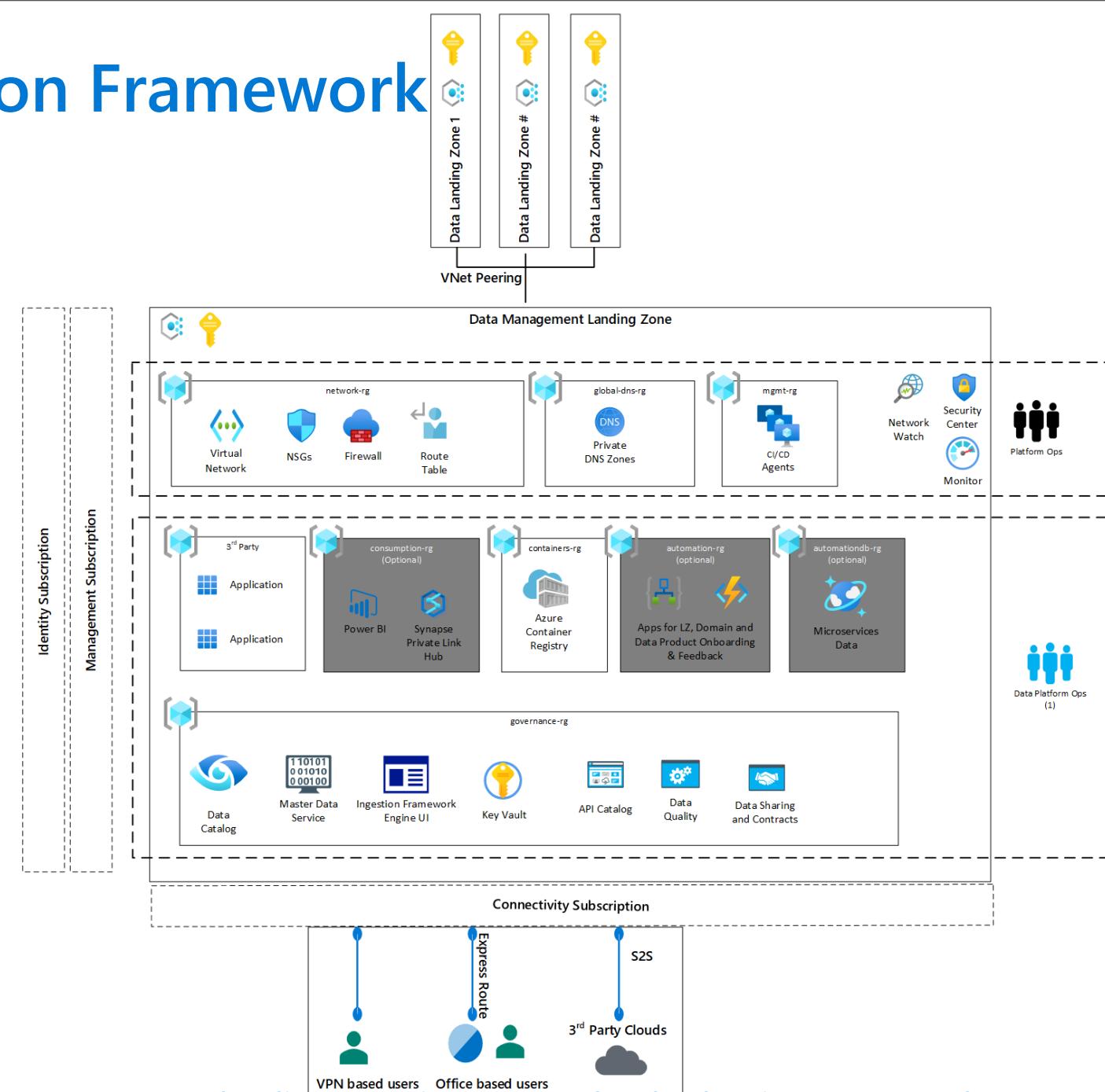
Benefits

- 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



Use Case #3 – Cloud Adoption Framework

Overview of the Azure enterprise Scale for analytics and AI data management landing zone



Overview of the enterprise-scale for analytics and AI data management landing zone in Azure - Cloud Adoption Framework

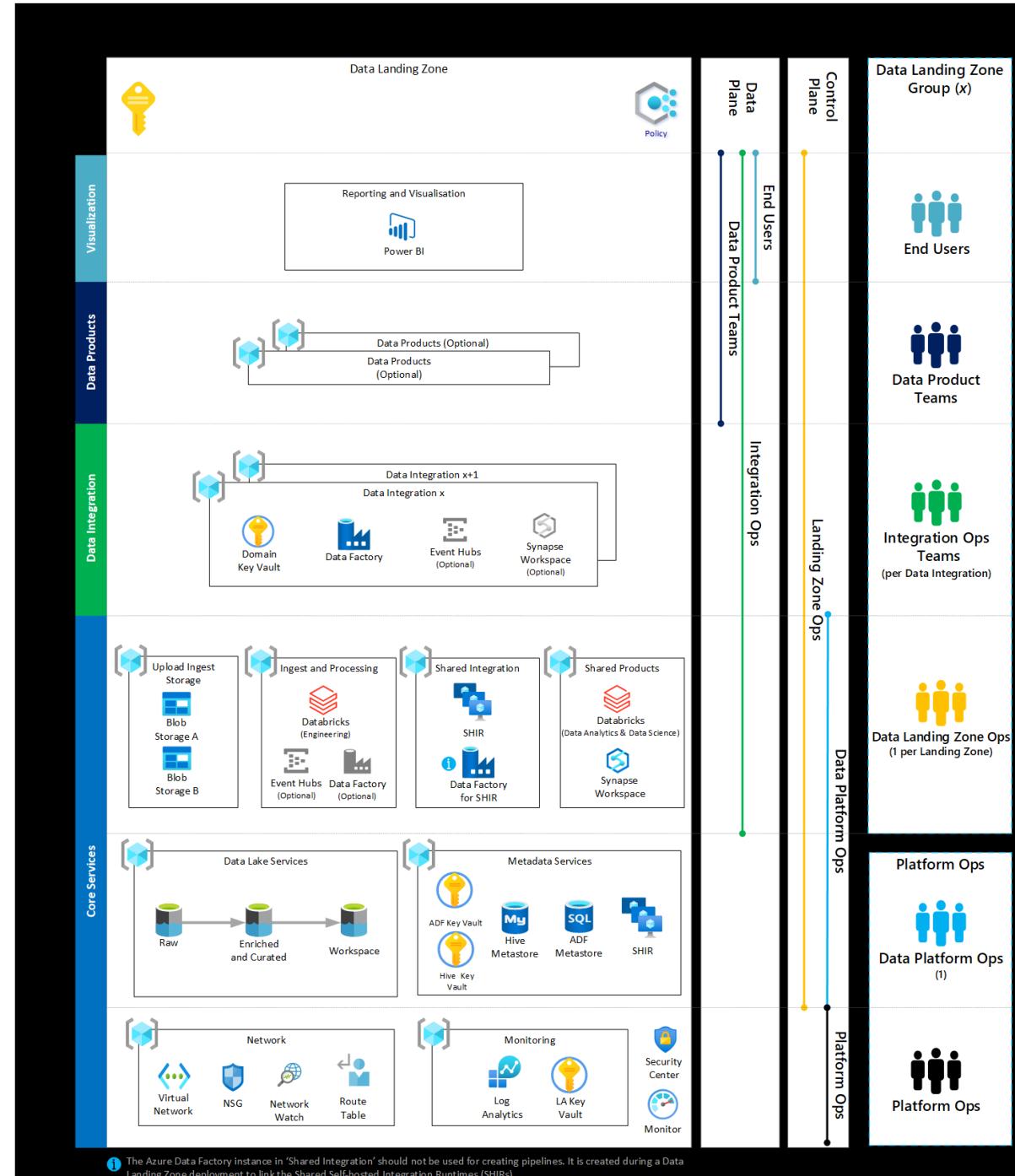
Use Case #3 – Cloud Adoption

For the CAF - We recommend you have a catalog service provisioned to define the metadata of the datasets stored across the data landing zones.

The catalog registers and maintains the data information in a centralized place and makes it available for the organization. It ensures that enterprises avoid duplicate datasets caused by redundant data ingestion by different project teams.

Enterprise-scale for analytics and AI relies on Azure Purview to serve as:

- System of registration
- Discovery for enterprise data sources
- Data classification engine
- Policy store
- API for registering and reading data information.
- Compliance dashboard hub



Franck Mercier

Demo Purview Studio



Resources & Close Out



Resources - References in Slides

- [Labeling in Azure Purview - Azure Purview | Microsoft Docs](#)
- [Understanding resource sets - Azure Purview | Microsoft Docs](#)
- [Frequently asked questions \(FAQ\) - Azure Purview | Microsoft Docs](#)
- [Why You Need a Data Governance Process Now!! - 5MinuteBI](#)
- [Learning Guide: Introduction to Azure Purview \(GA Update\) - 5MinuteBI](#)
- [Creating a modern data governance strategy to accelerate digital transformation \(microsoft.com\)](#)
- [Network isolation and Security: Use private endpoints for your Azure Purview account - Microsoft Tech Community](#)
-

Links to Use Case References

- [Data Catalog lineage user guide - Azure Purview | Microsoft Docs](#)
- [Purview-Machine-Learning-Lineage-Solution-Accelerator: Solution accelerator to help build Machine Learning Lineage \(github.com\)](#)
- [Modern data warehouse for small and medium business - Azure Example Scenarios | Microsoft Docs](#)
- [Big data analytics with enterprise-grade security using Azure Synapse - Azure Solution Ideas | Microsoft Docs](#)
- [Overview of the enterprise-scale for analytics and AI data management landing zone in Azure - Cloud Adoption Framework | Microsoft Docs](#)

Resources - References in Slides

Links to Use Case References

- [Data Catalog lineage user guide - Azure Purview | Microsoft Docs](#)
- [Purview-Machine-Learning-Lineage-Solution-Accelerator: Solution accelerator to help build Machine Learning Lineage \(github.com\)](#)
- [Modern data warehouse for small and medium business - Azure Example Scenarios | Microsoft Docs](#)
- [Big data analytics with enterprise-grade security using Azure Synapse - Azure Solution Ideas | Microsoft Docs](#)
- [Overview of the enterprise-scale for analytics and AI data management landing zone in Azure - Cloud Adoption Framework | Microsoft Docs](#)
- [Map your data estate with Azure Purview - Microsoft Tech Community](#)
- [Overview of the enterprise-scale for analytics and AI data management landing zone in Azure - Cloud Adoption Framework](#)
- [Overview of the enterprise-scale for analytics and AI architecture data landing zone in Azure - Cloud Adoption Framework](#)
- [What is Azure Data Share? | Microsoft Docs](#)

- [Understand access and permissions - Azure Purview | Microsoft Docs](#)
- [Data Catalog lineage user guide - Azure Purview | Microsoft Docs](#)
- [Azure SQL and Azure Purview work better together - Microsoft Tech Community – Stored Procedure lineage in preview](#)

Resources

Get started

- [Get started quickly & easily with Azure Purview!](#)
- [Microsoft Learn: Introduction to Azure Purview](#)
- [Azure Purview documentation](#)
- [Data lineage in Azure Purview - Azure Purview | Microsoft Docs](#)

Communities

- [Join our Azure Data Community](#)
- [Join our Azure Purview Tech Community](#)
- [Azure Purview - Microsoft Tech Community - Blog](#)

Customer stories

- [Grundfos - Danish pump manufacturer develops sustainable water solution with unified data governance from Azure Purview](#)
- [Heathrow Airport - Heathrow boosts operational efficiency and improves decision making with Azure Purview](#)
- [illimity Bank - illimity optimizes data governance and streamlines compliance with Azure Purview](#)

Resources - Azure Purview Pricing

- [Pricing – Azure Purview – Pricing Calculator](#)
- [Pricing - Azure Purview](#)
- [Pricing overview – How Azure pricing works](#)

Purview Overview

+

Some New Info!!

