



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

<https://github.com/steveyoungca/PurviewSession> - Resources

<https://aka.ms/AzurePurviewSession1> - Survey



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.

Slide Deck and Link Repository

<https://github.com/steveyoungca/PurviewSession>

steveyoungca / PurviewSession Public

Code Issues Pull requests Actions Projects Wiki Security Insights Settings

main 1 branch 0 tags Go to file Add file ▾ Code ▾

steveyoungca Initial files from workshop for share ... 2372a14 22 minutes ago 1 commit

Azure Purview Overview October 20... Initial files from workshop for share 22 minutes ago

Readme.md Initial files from workshop for share 22 minutes ago

Readme.md

Resources - References in Slides

Get started

The following links were provided in the Azure Purview Partner Accelerator Workshop, Oct 9 and 10 2021.

This site will be archived December 1, 2021.

- 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

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)

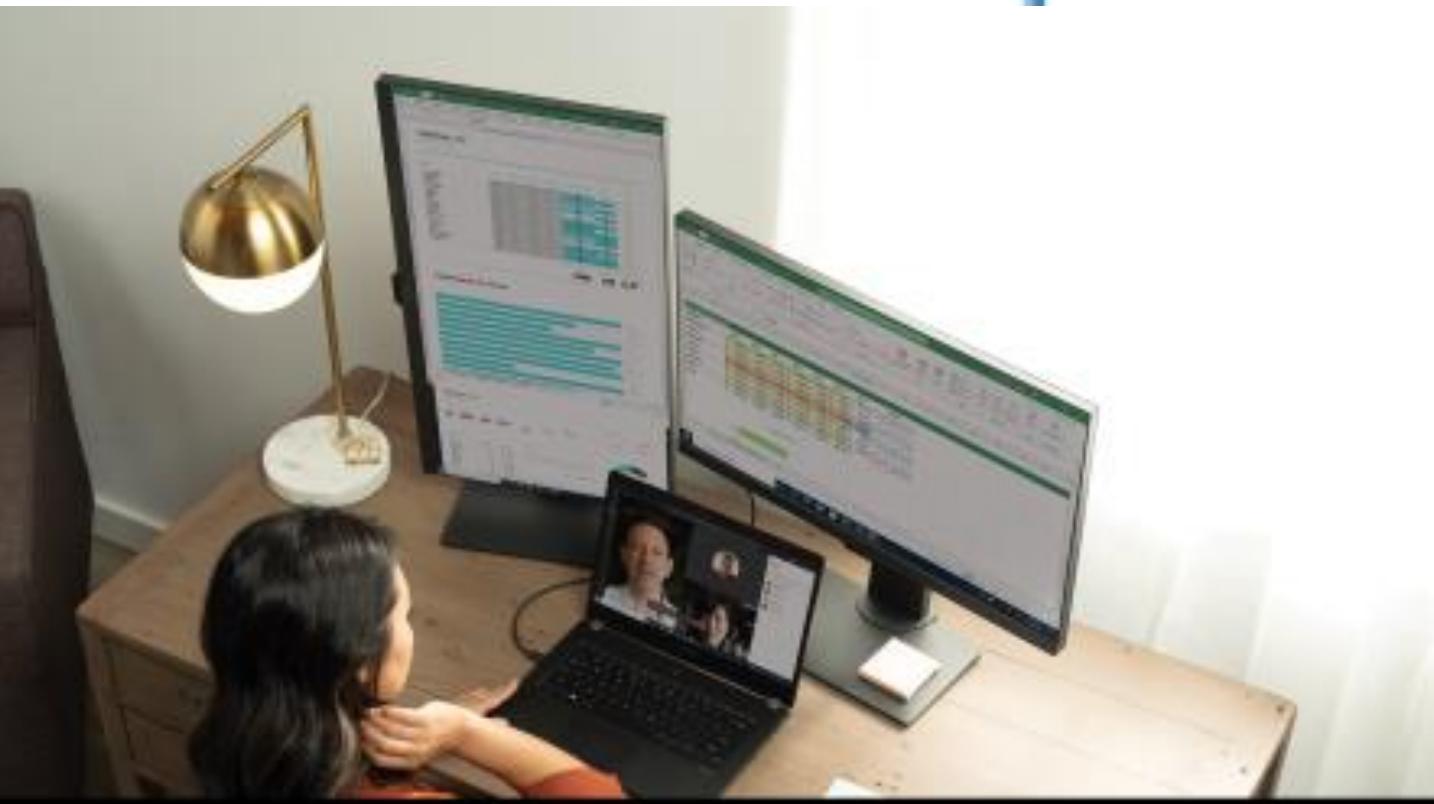
zeinam@microsoft.com

Agenda – Day 2

- Welcome and Introduction
- Setting the Stage - Data Governance Problem
- An introduction to Azure Purview
- Machine Learning Use Case
- Break
- Technical Presentations
 - General Technical Presentation
 - What is Collection
 - Connect Power BI to Azure Purview
- Break
- Technical Presentations
 - Programming the API (Overview)
- Q&A / Your Thoughts
- Additional Resources & Close Out



Setting The Stage - The Data Governance Problem



Data Governance Problem – Business has changed

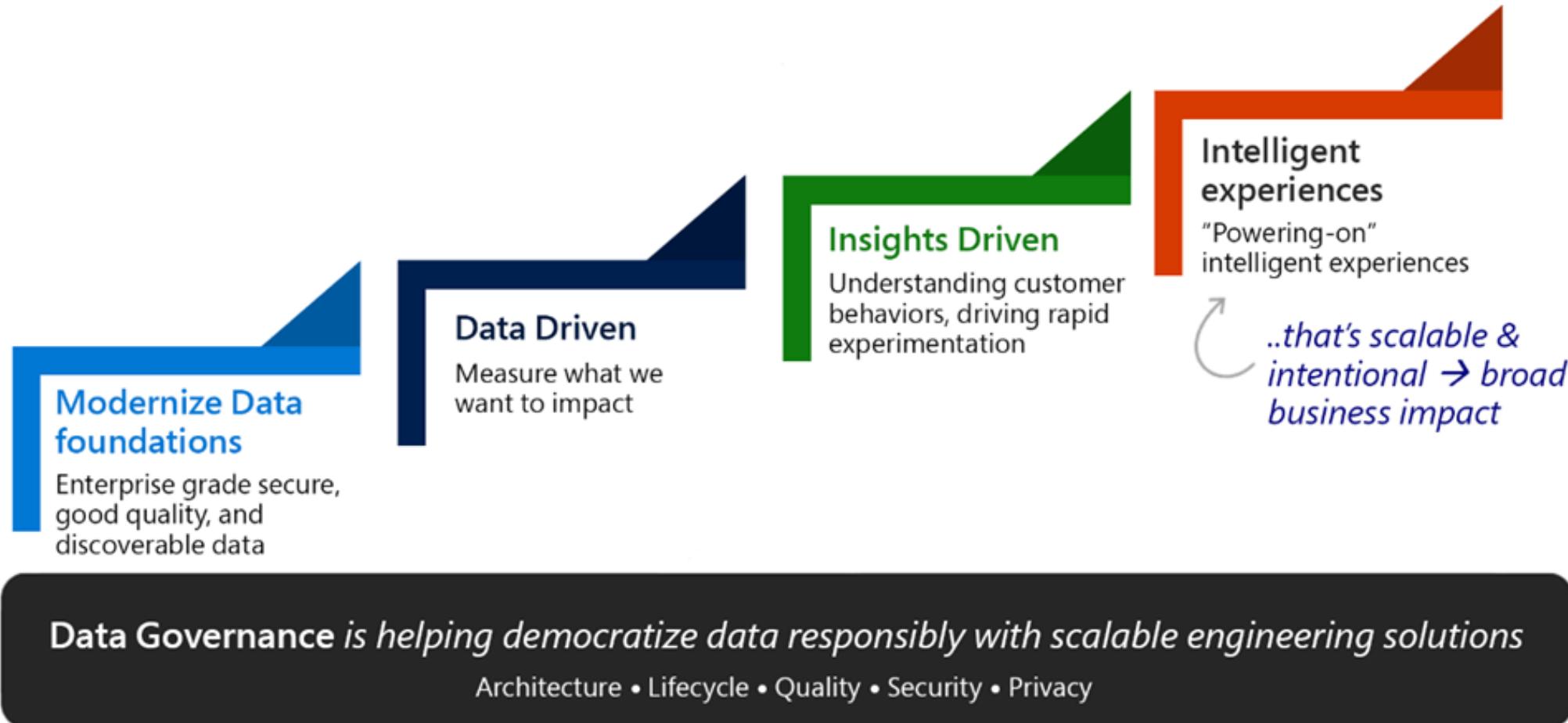


Figure 1. Data governance is the foundational pillar of the Microsoft Enterprise Data Strategy

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

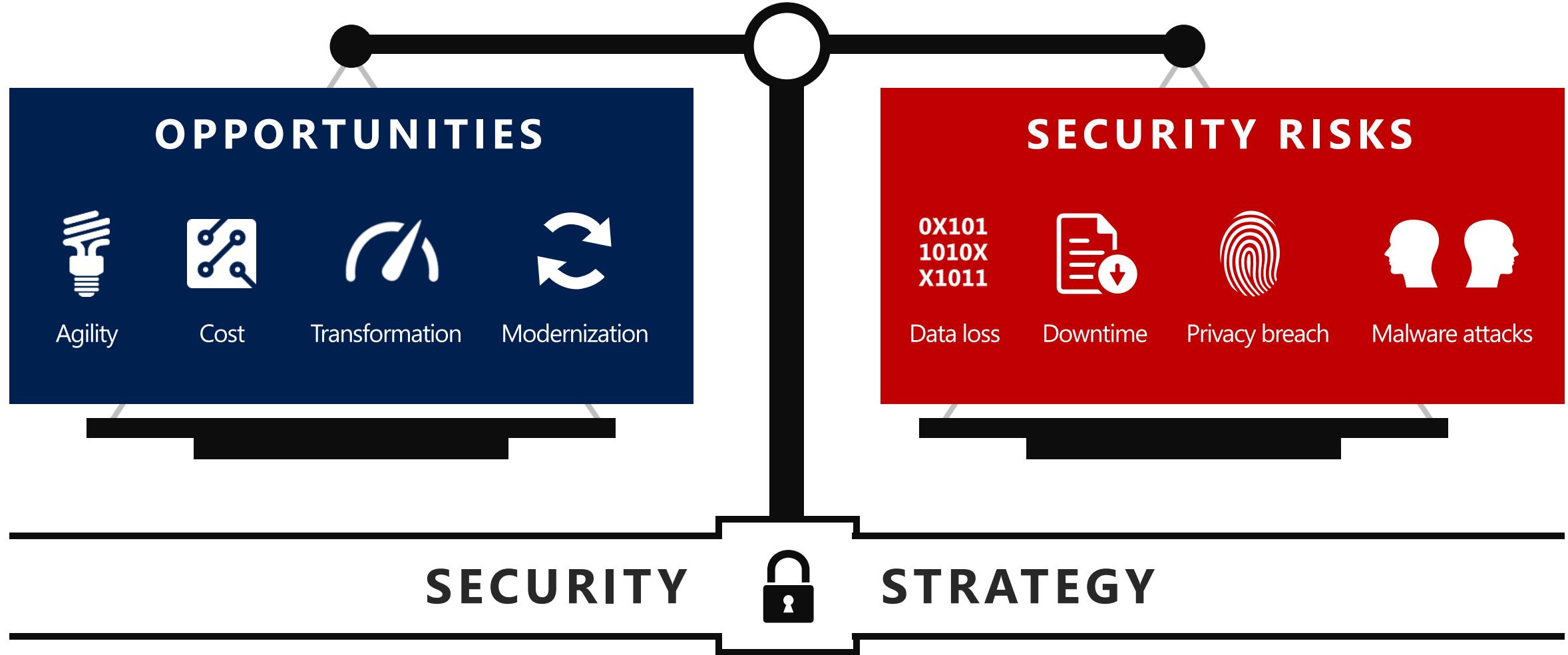
Data Breach Statistics: How big of a problem?

- 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](#))

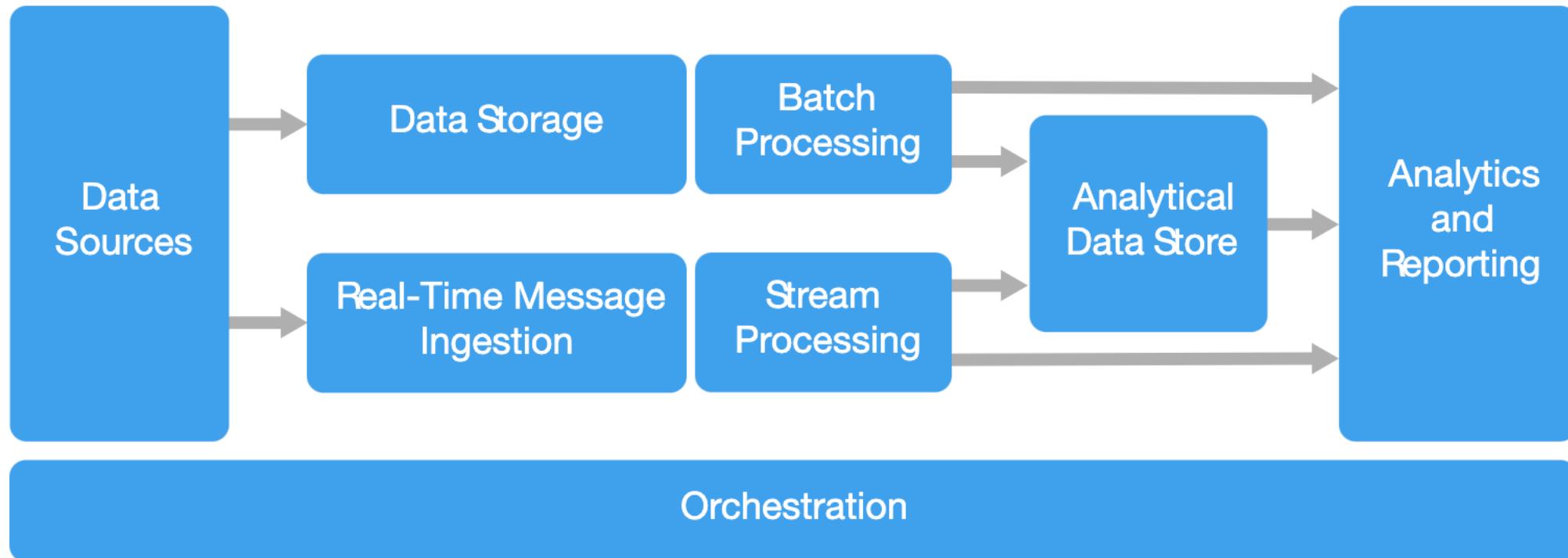
Not all Data Breaches have be hacking.

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.

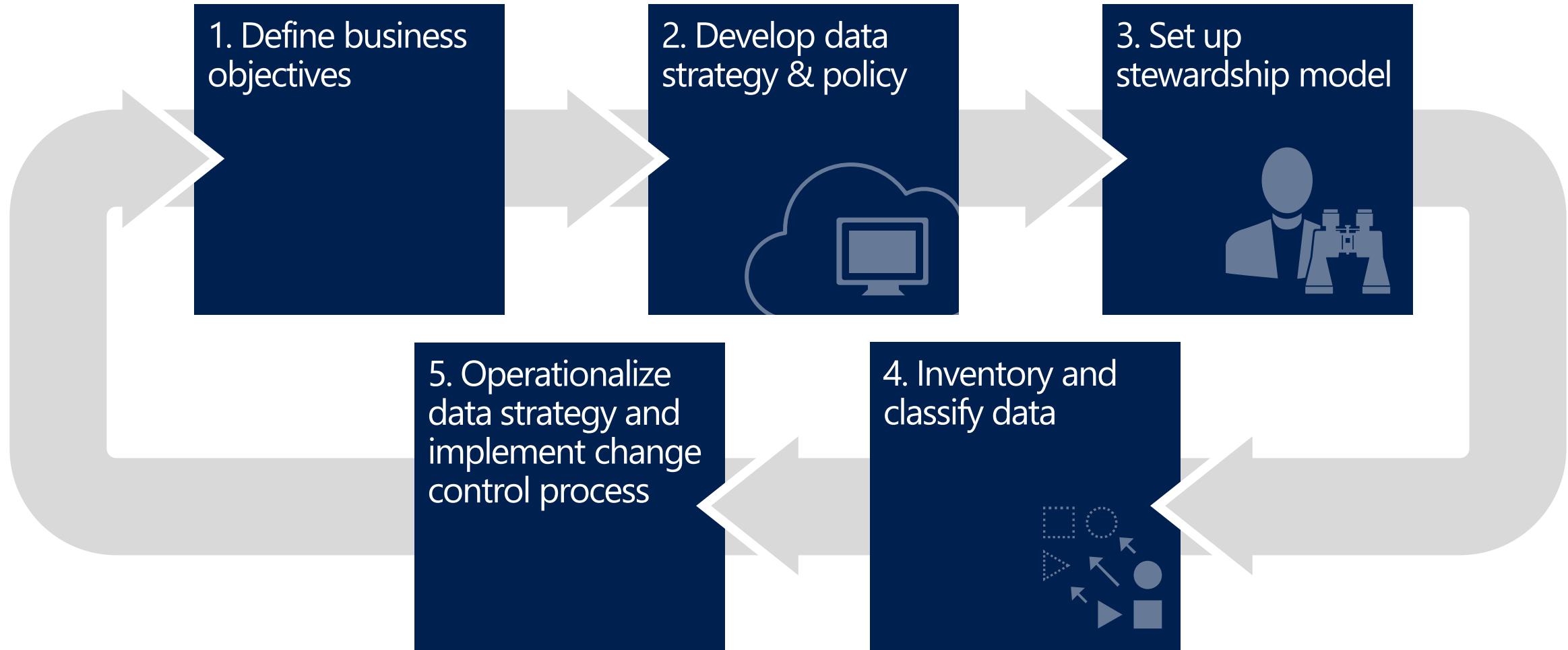
Balance Opportunities and Risks



Modern Data Warehouse



Data Governance is more than just an App

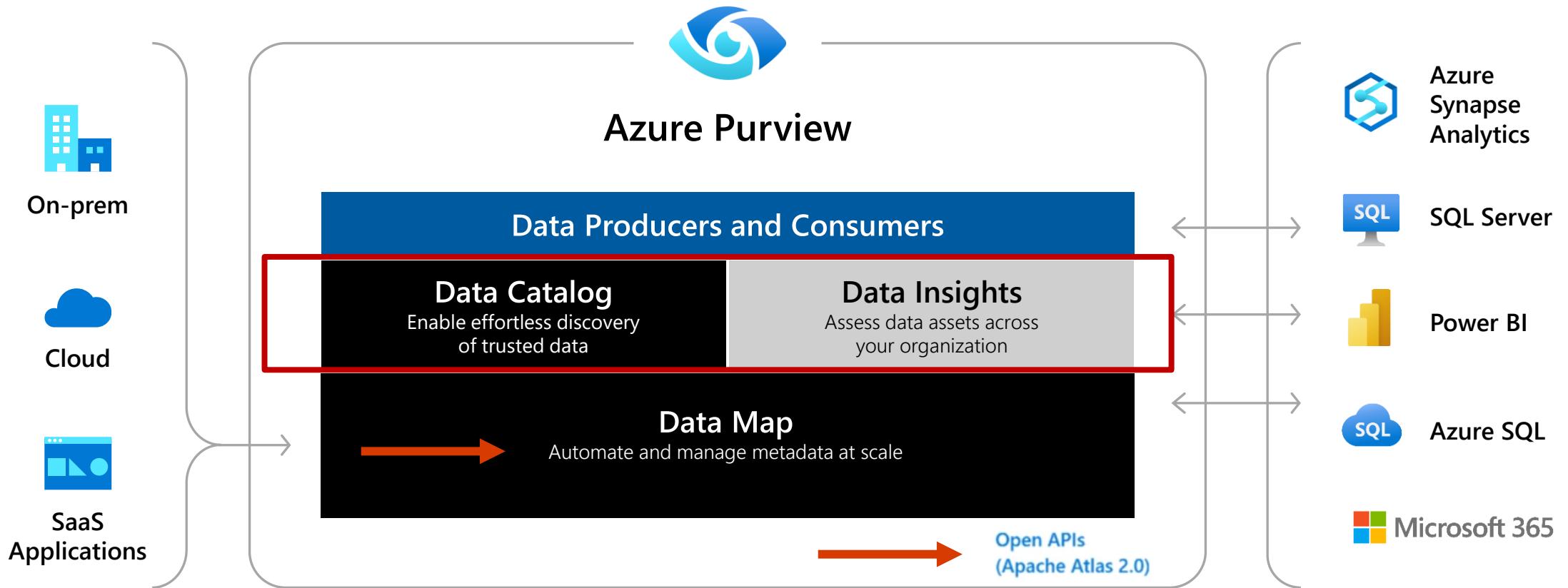


Azure Purview In Detail



Unified Data Governance with Azure Purview

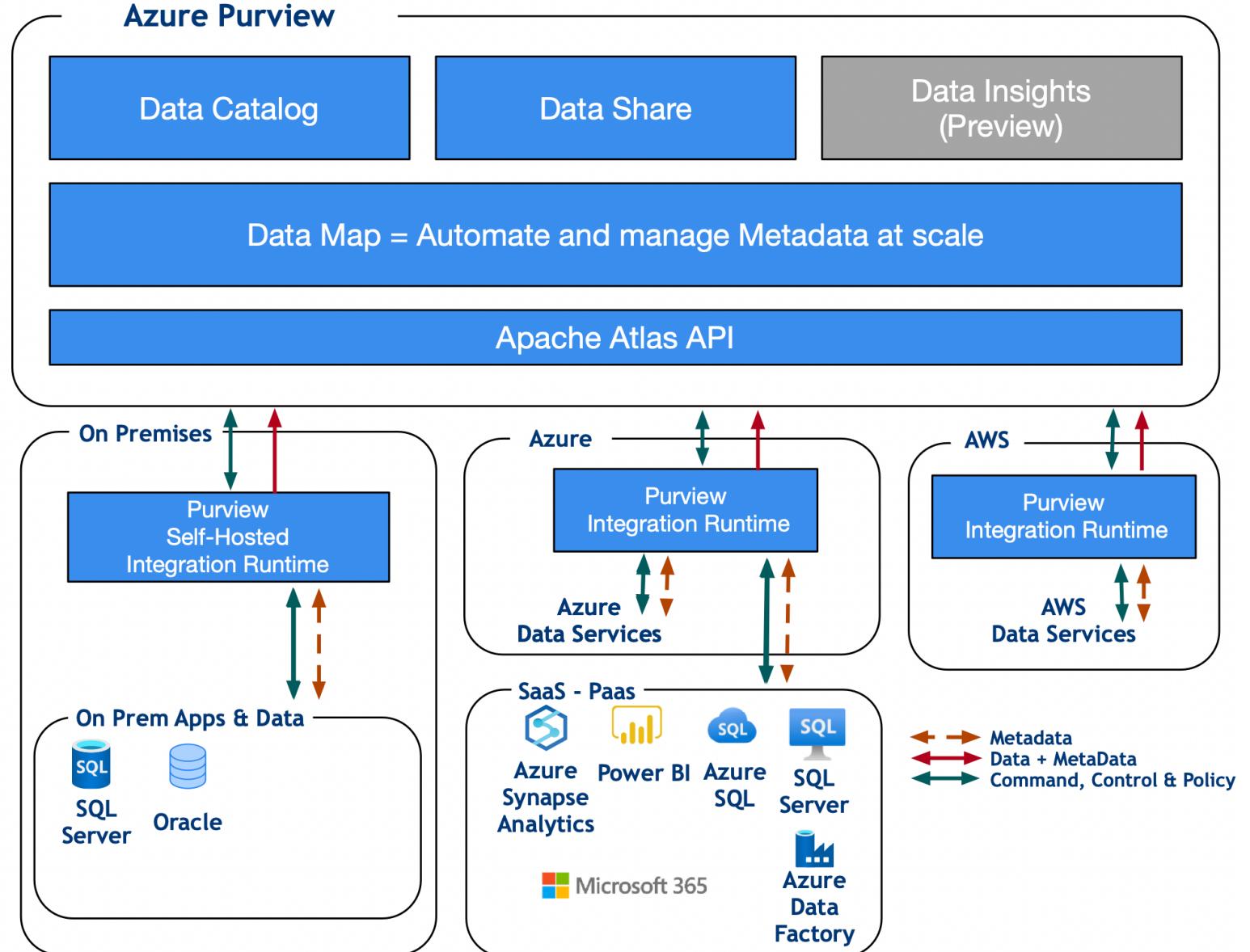
Generally Available
Preview



Azure Purview Data Overview

IMPORTANT!!

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



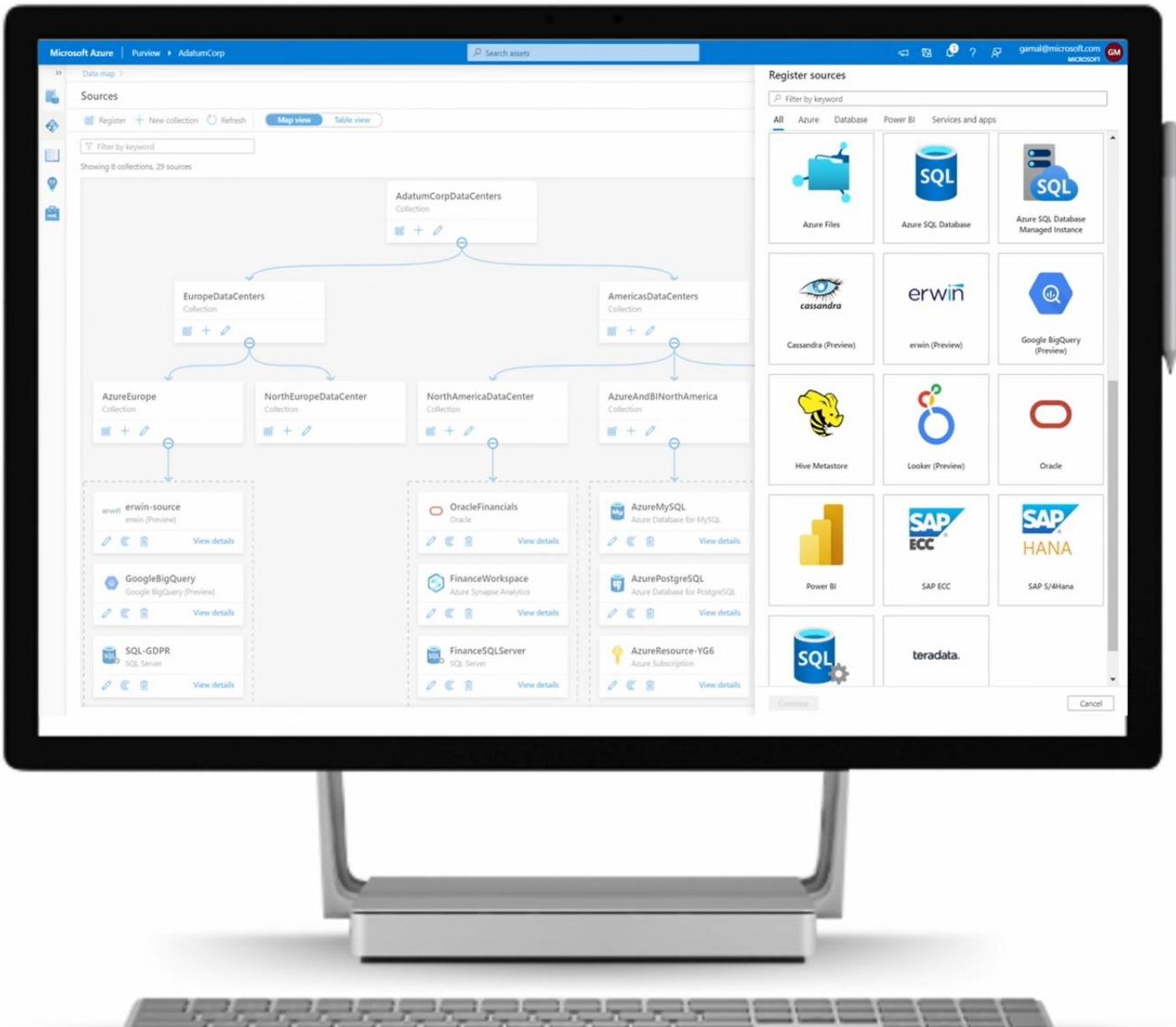
Automated Data Discovery using Purview Data Scan

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



Automated Data Discovery & Classification

Data Discovery

- Hybrid data scanning & classification service with central management
- Integration runtime: serverless in cloud, installable agent on prem
- Supports range of network topologies (on prem, VNET, etc.)
- 22 points of presence covering 12 geographies
- 35 supported data sources / file formats across cloud, on premises and SaaS. ~4 new sources added monthly, Preview= Google BigQuery, Erwin, Looker, Cassandra, MySQL

Data Classification & Labelling

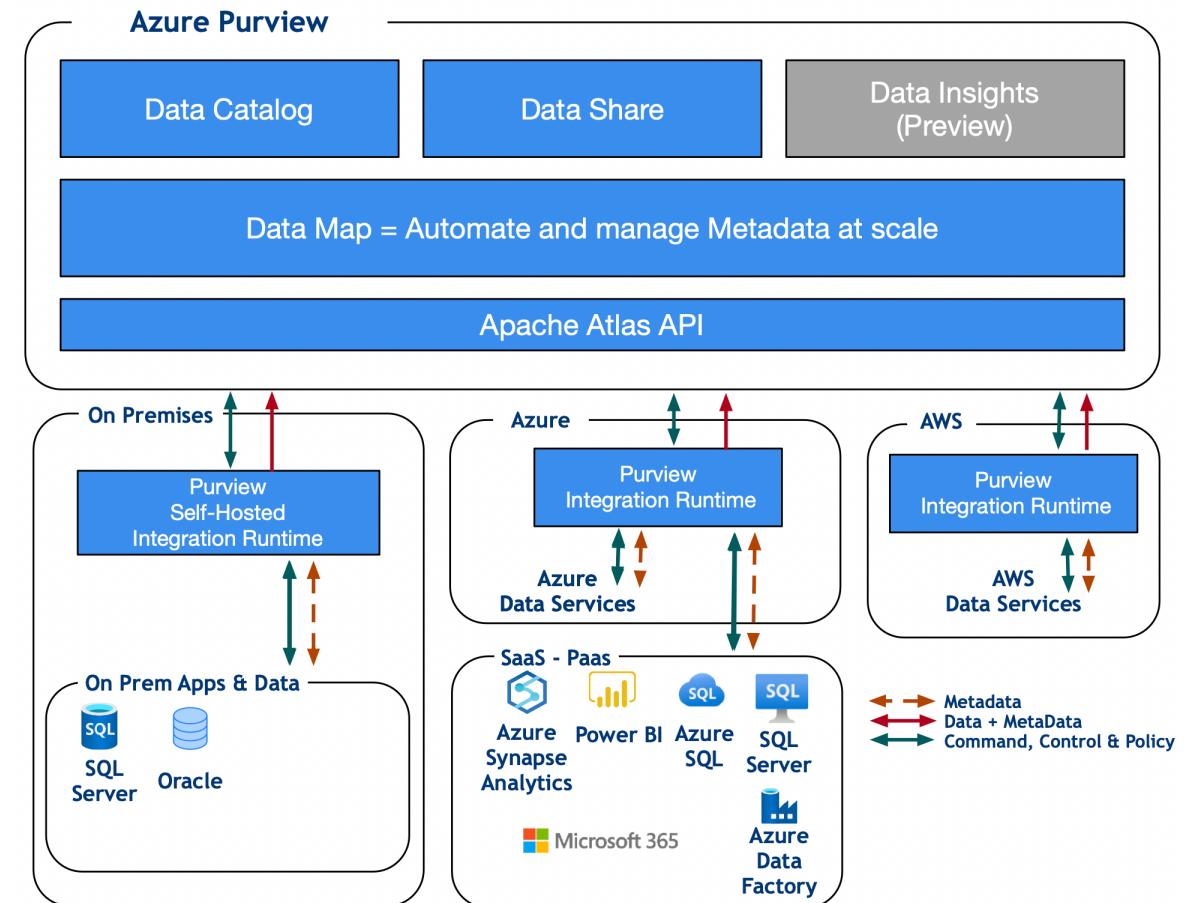
- 206 data classifiers across personal, government, security and finance domains.
- Custom data classifier support (data dictionary or expression) with tunable confidence levels & version support. What-if support
- Label sensitive data consistently across Office365, SQL Server, Azure and Power BI

Lineage Extraction

- Automated lineage extraction from Synapse pipelines, Azure Data Factory, Data Share, Power BI. Extending to Synapse Spark, SQL (tables, columns, scripts, views, stored procs), Power BI sub artifacts (datasets, tables, columns, measures)

Configuration & Management

- Visual, web-based configuration & management
- APIs & SDKs (Python & .NET)

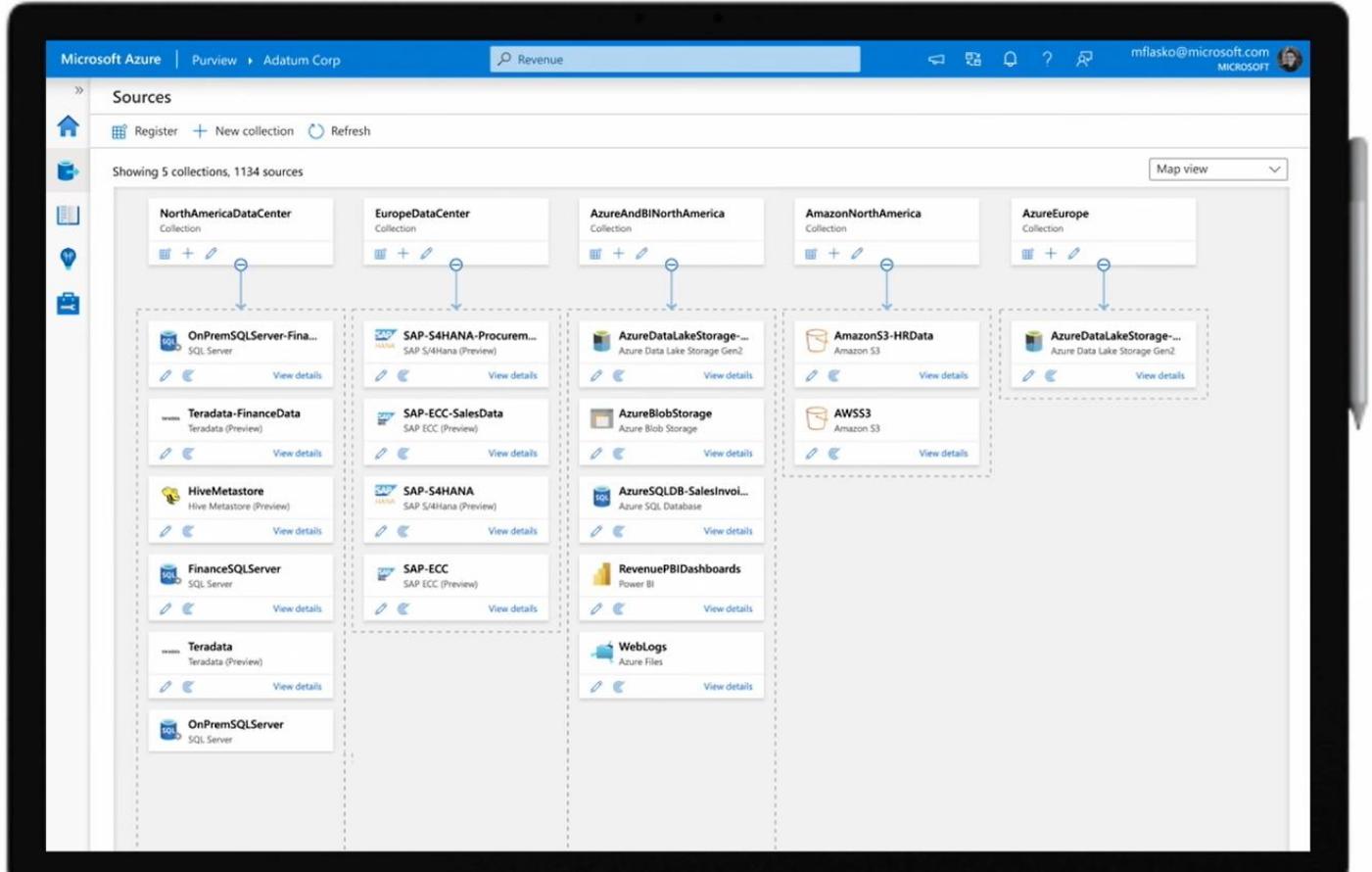


Make data meaningful using Purview Data Map

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

Fine-grained access controls

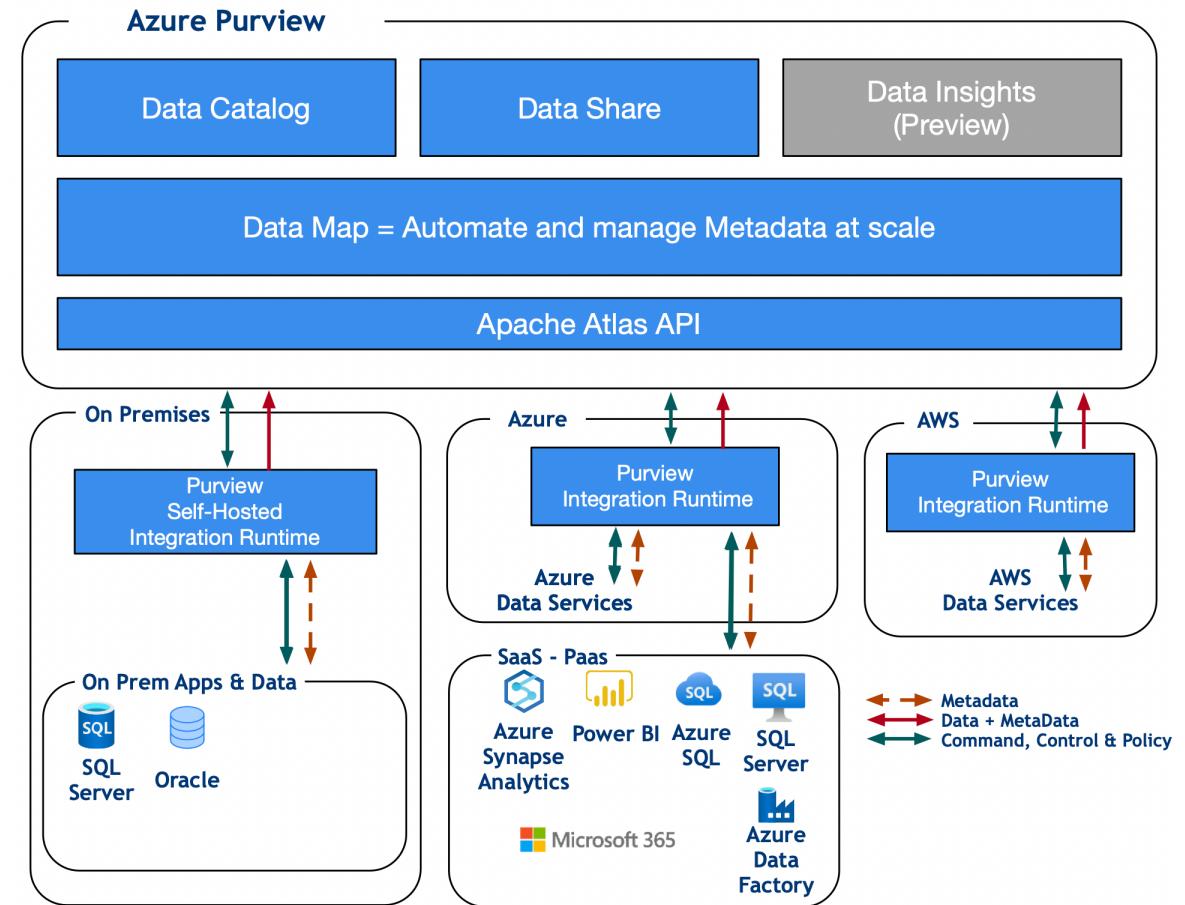
Open platform



Data Map

Data Map

- Inventory of data sources, data assets and their lineage
- Includes technical, semantic, business and crowd metadata
- Apache Atlas metamodel and APIs (REST and Messaging)
- Platform for all Purview user experiences -- data catalog, business glossary, data quality, etc.
- Hierarchical data ownership model and access control
- Private Endpoint support to securely access Purview.



[Network isolation and Security: Use private endpoints for your Azure Purview account - Microsoft Tech Community](#)

>

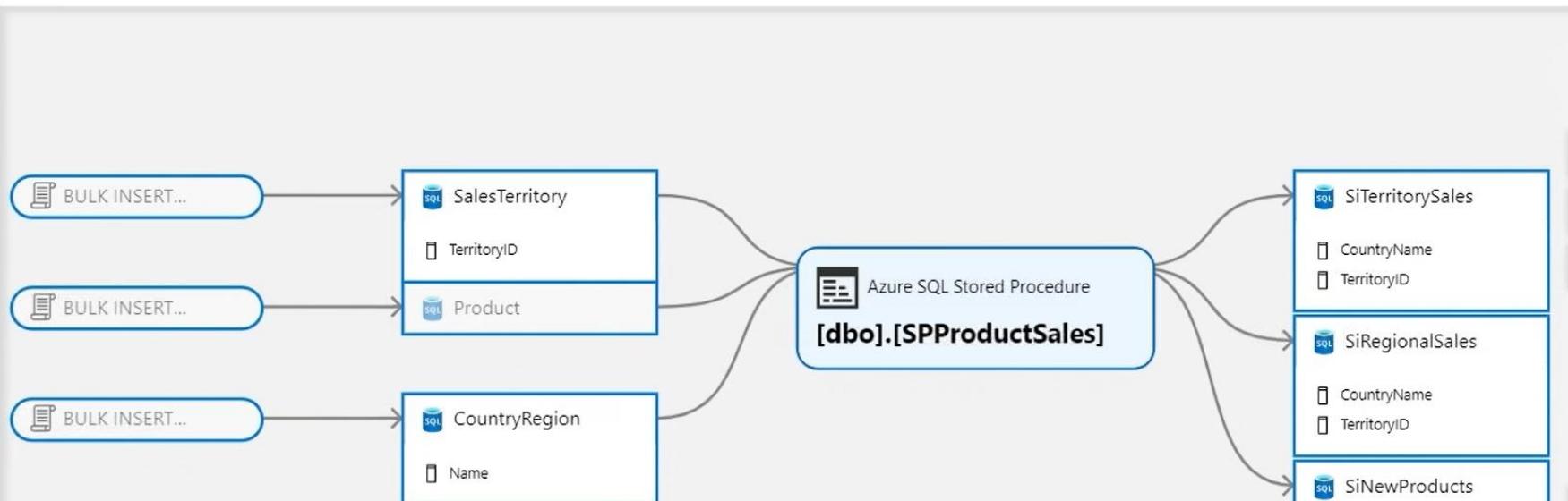
Data catalog >

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

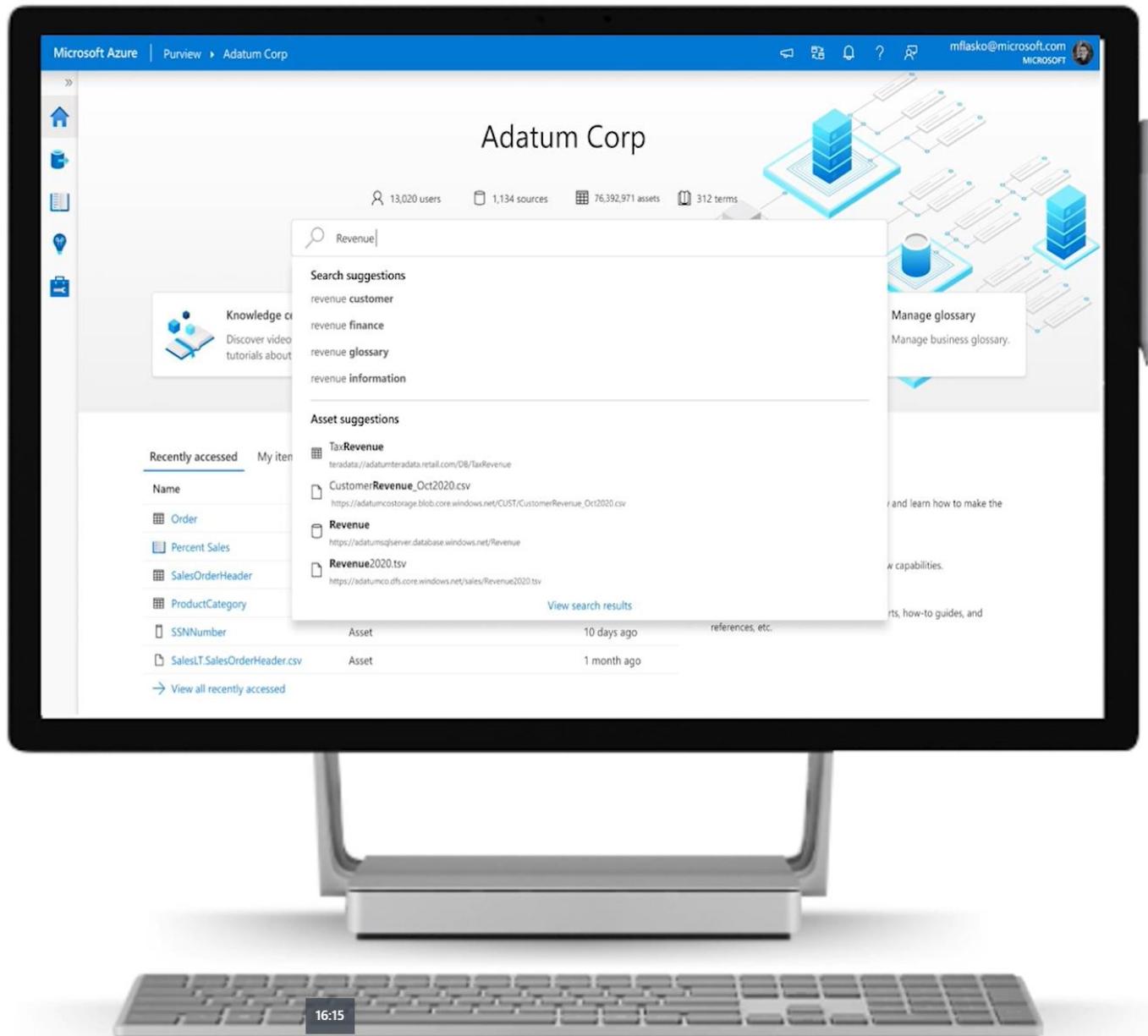
25.04

Find Trusted Data with Purview Data Catalog

Search, Browse & Curate

Intelligent Recommendations

Lineage Visualization



Data Catalog

Search & Browse

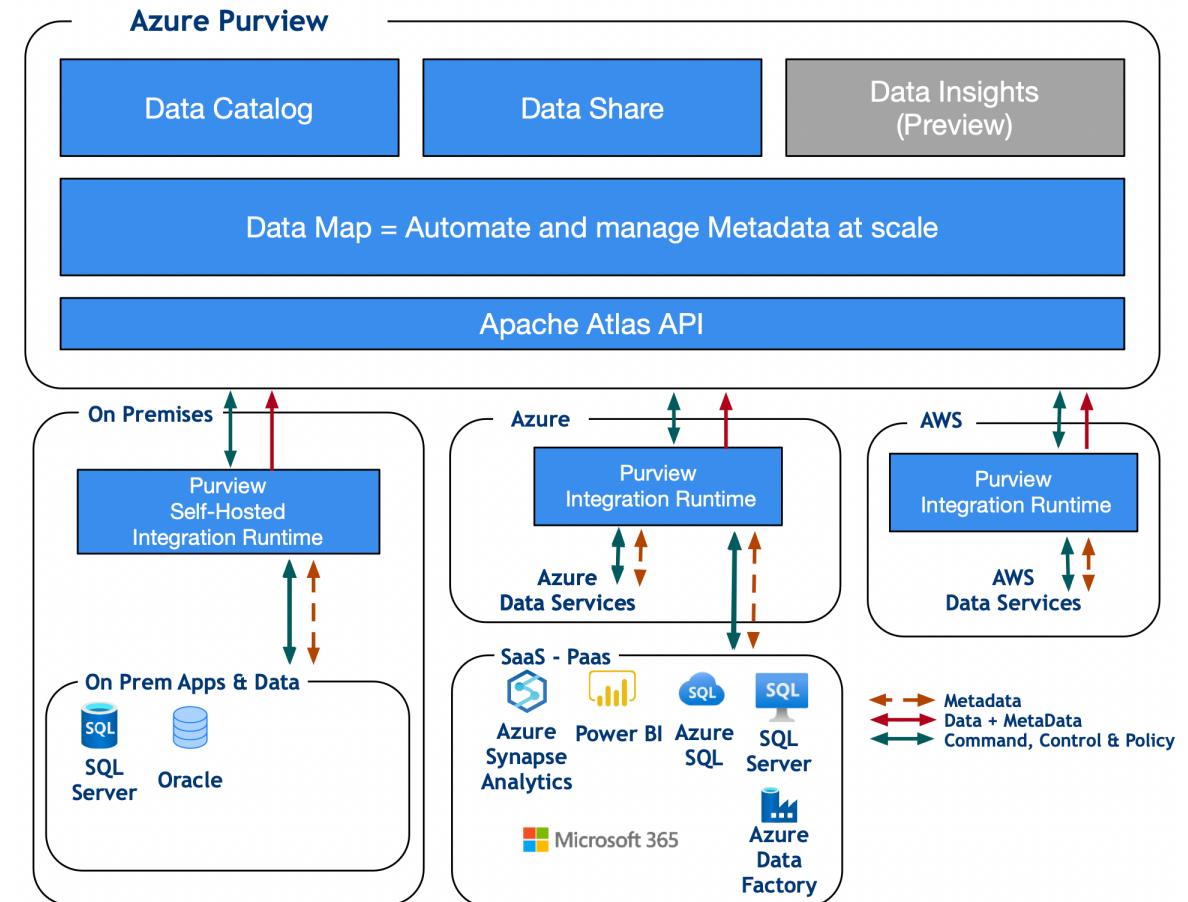
- Search for data using business terms, technical terms or data classification. Faceted filters also supported.
- Browse assets by hierarchy
- Embedded search in Synapse (preview) & Power BI (in progress)
- Open-in experience for catalog assets. Power BI supported. Power BI Desktop

Business Glossary

- Define business terms, their definitions, related terms, owners & stewards
- Bulk import existing data dictionaries
- Bulk assign glossary terms to assets
- Hierarchical business glossaries

Lineage Visualization

- Interactive lineage visualization from raw data through ETL processing to BI
- Object and column level lineage to do root cause analysis, impact analysis



Business Glossary

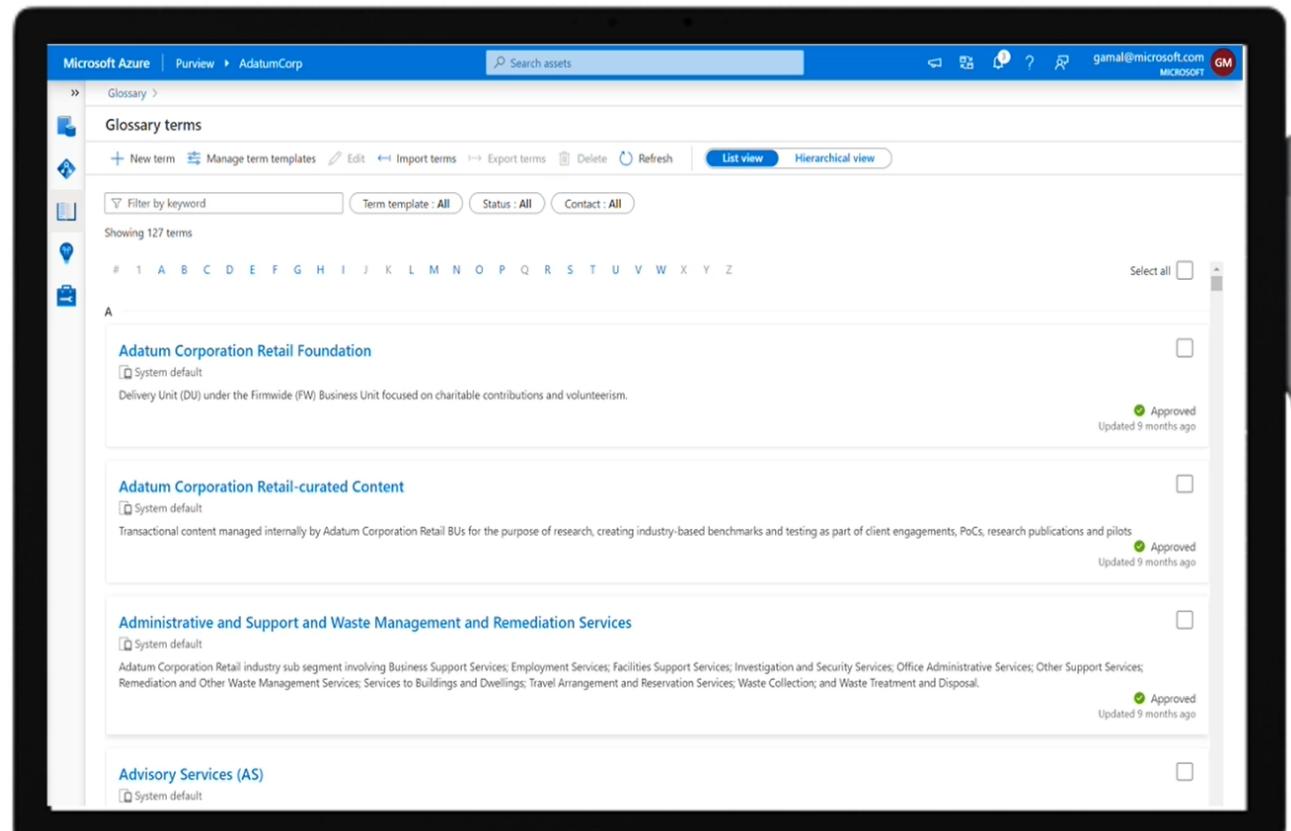
Make data relevant with business context

Azure Purview Business Glossary

Search by Business Term

Hierarchy Support

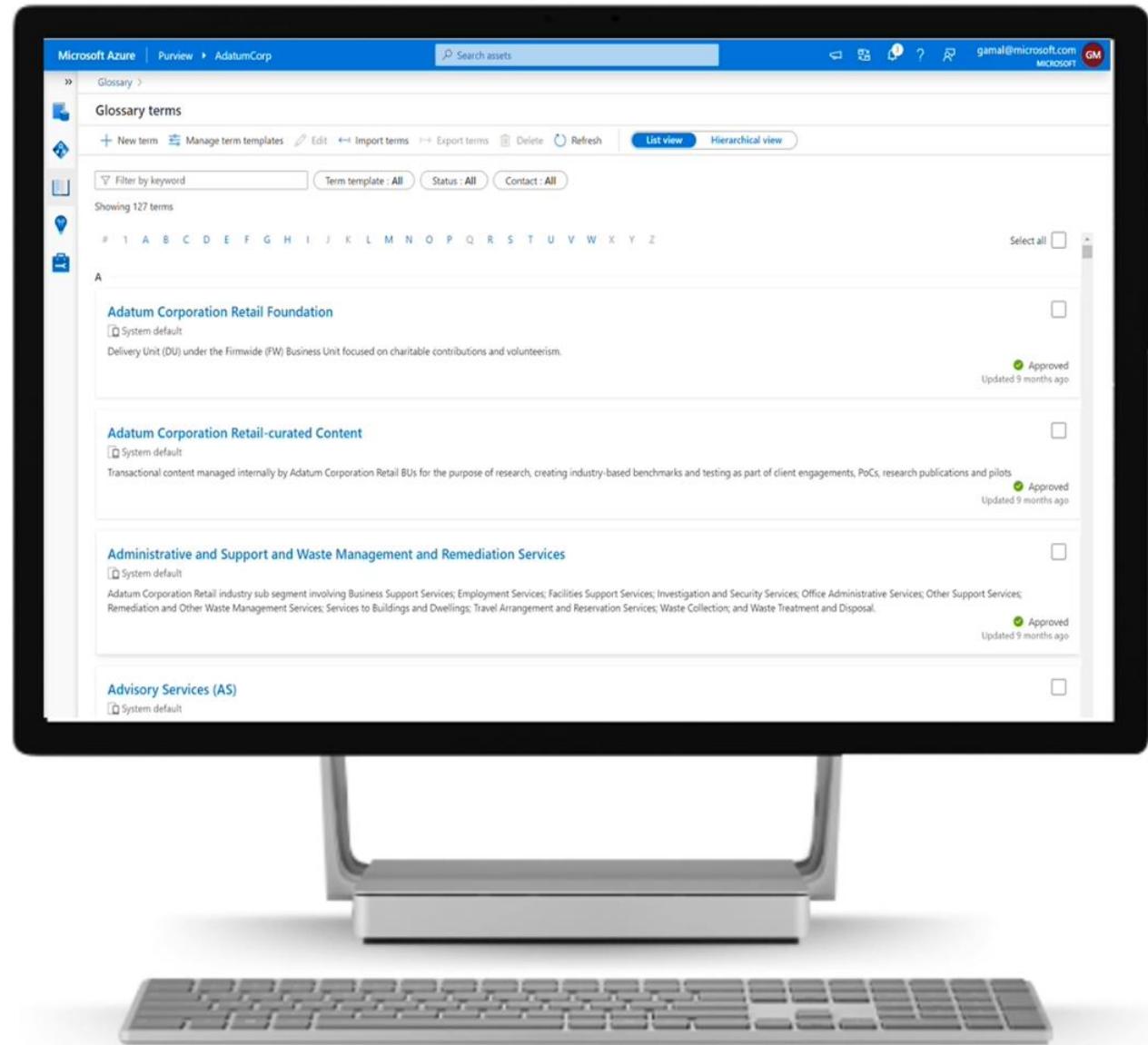
Integrated Approval Workflows



Business Glossary

Lineage Visualization

- Understand the business context associated with data in the organization
- Bulk import/export existing business glossaries
- Create glossary terms and attributes per business domains in organizations
- Define parent-child relationship to organize glossary terms in hierarchies
- Manage lifecycle of business terms
- Integrated Approval Workflows



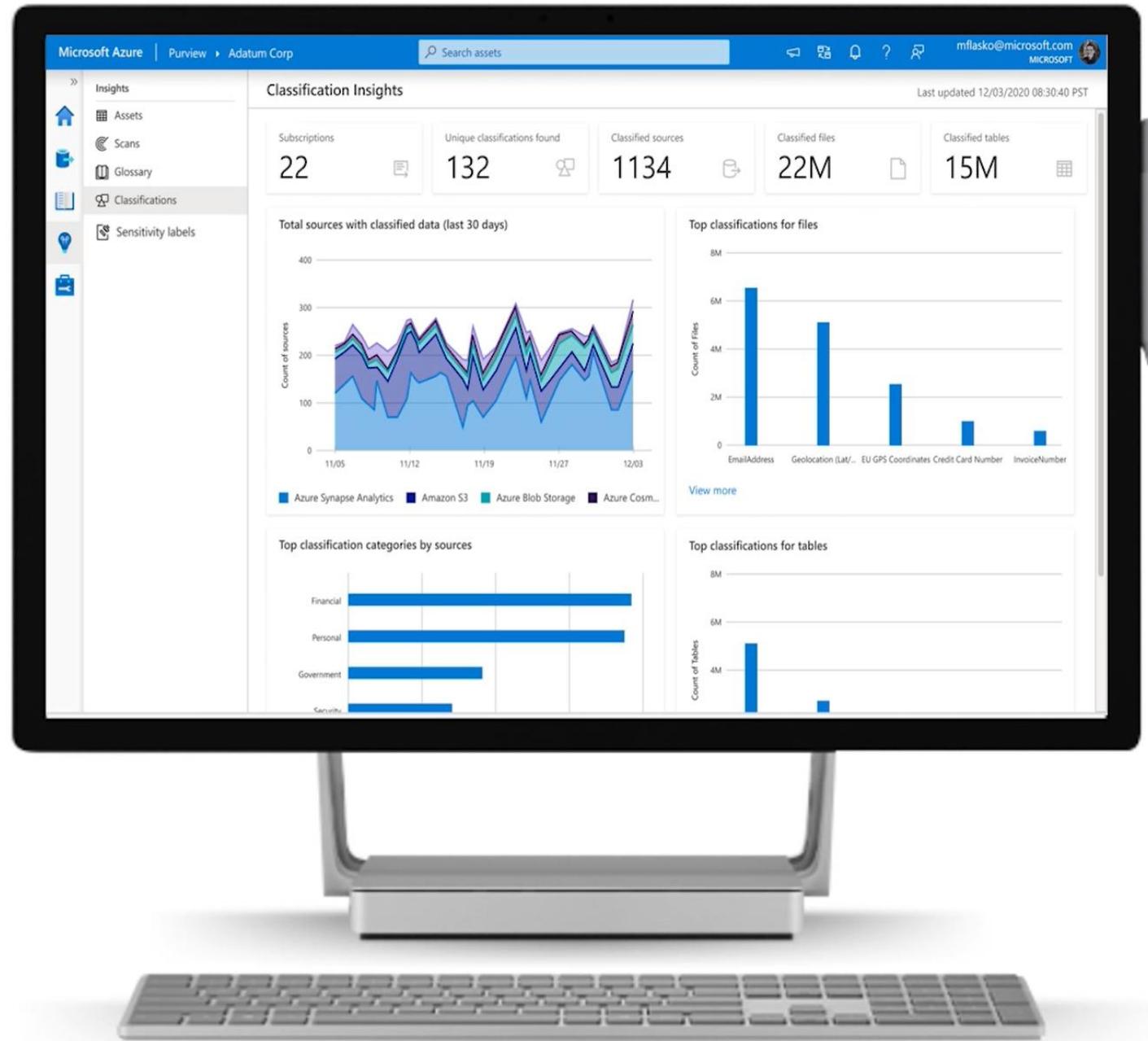
Get a bird's-eye view of your data landscape with Purview Data Insights

Data Asset distribution

Sensitive Data

Data Scanning Coverage

Business Glossary Utilization



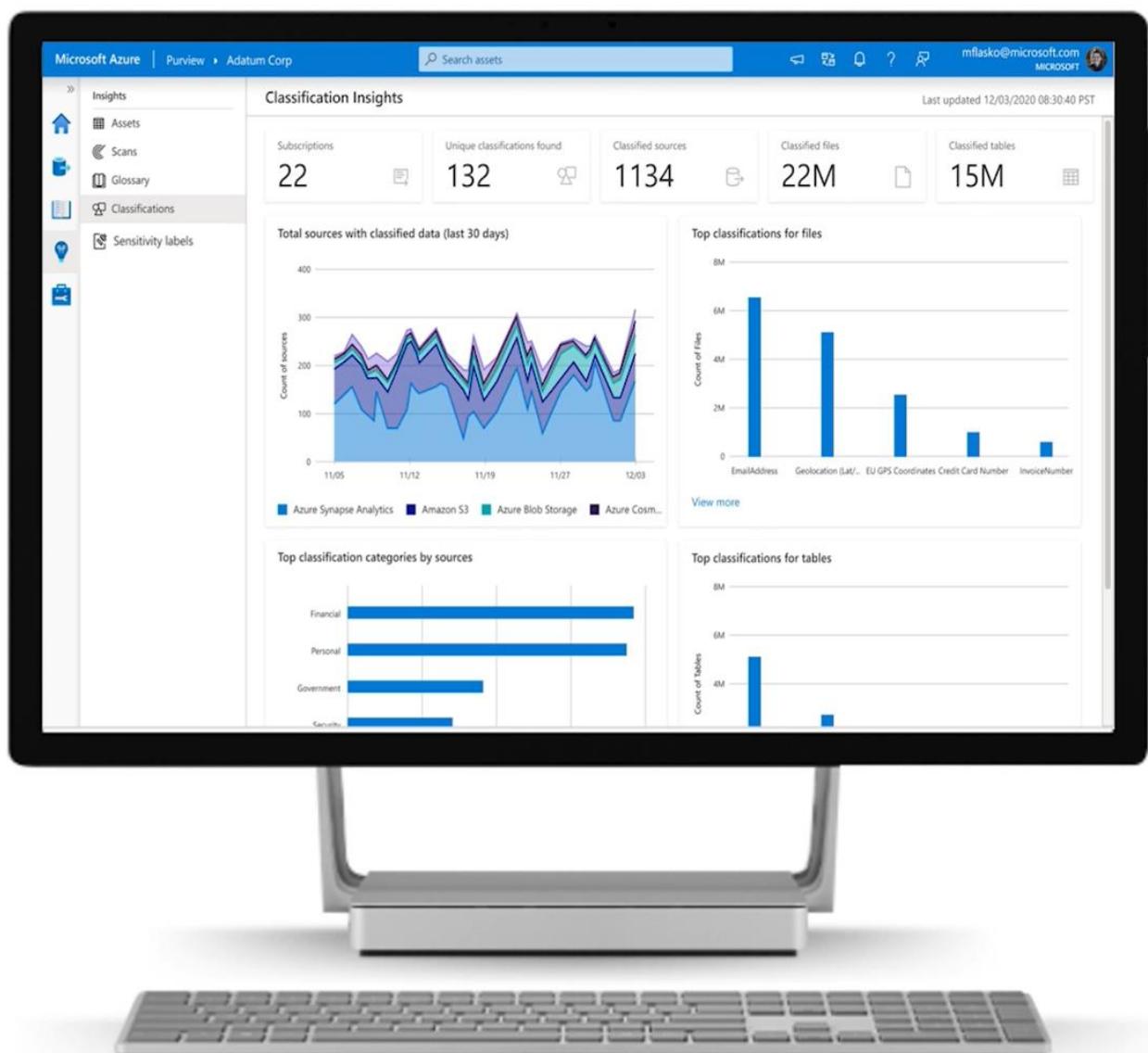
Data Insights (Preview)

Lineage Visualization

- Insights provides customers, a single pane of glass view into their catalog and further aims to provide specific insights to the data source administrators, business users, data stewards, data officer, and security administrators.

Azure Purview currently has the following reports available:

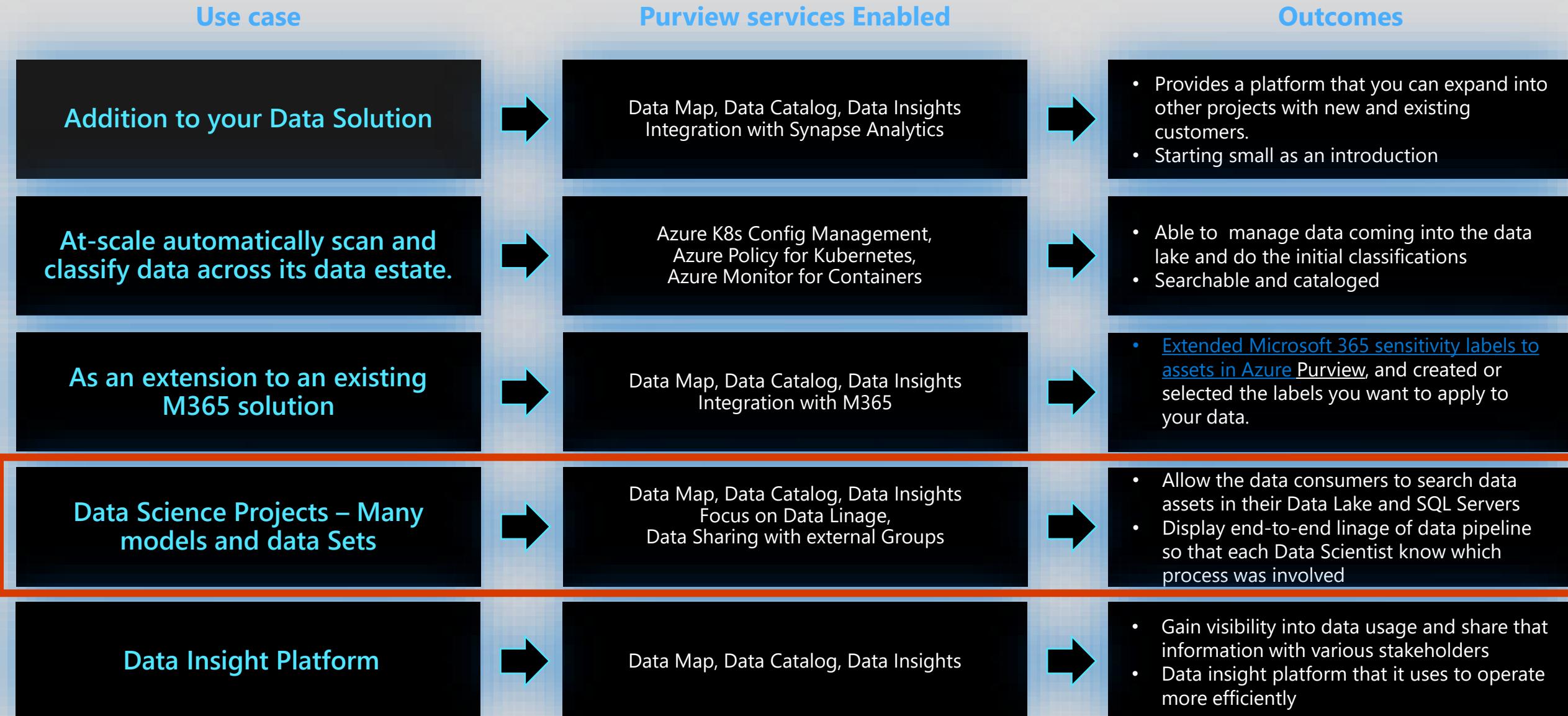
- Assets
- Scans
- Glossary
- Classification
- Sensitivity Labels



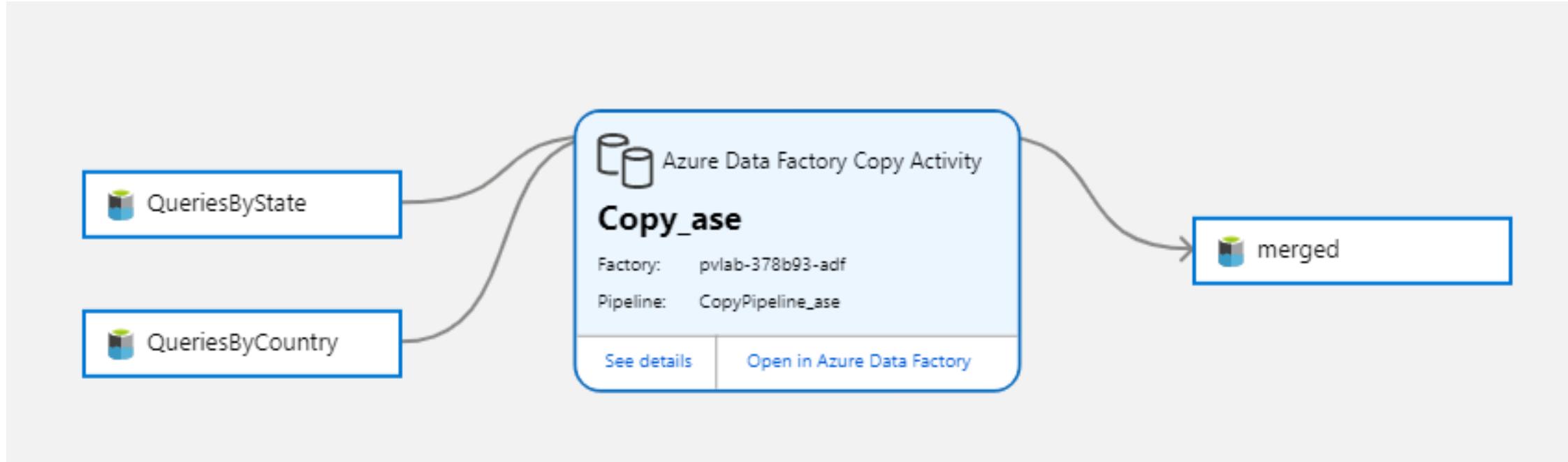
Partner Opportunities and Use Cases



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

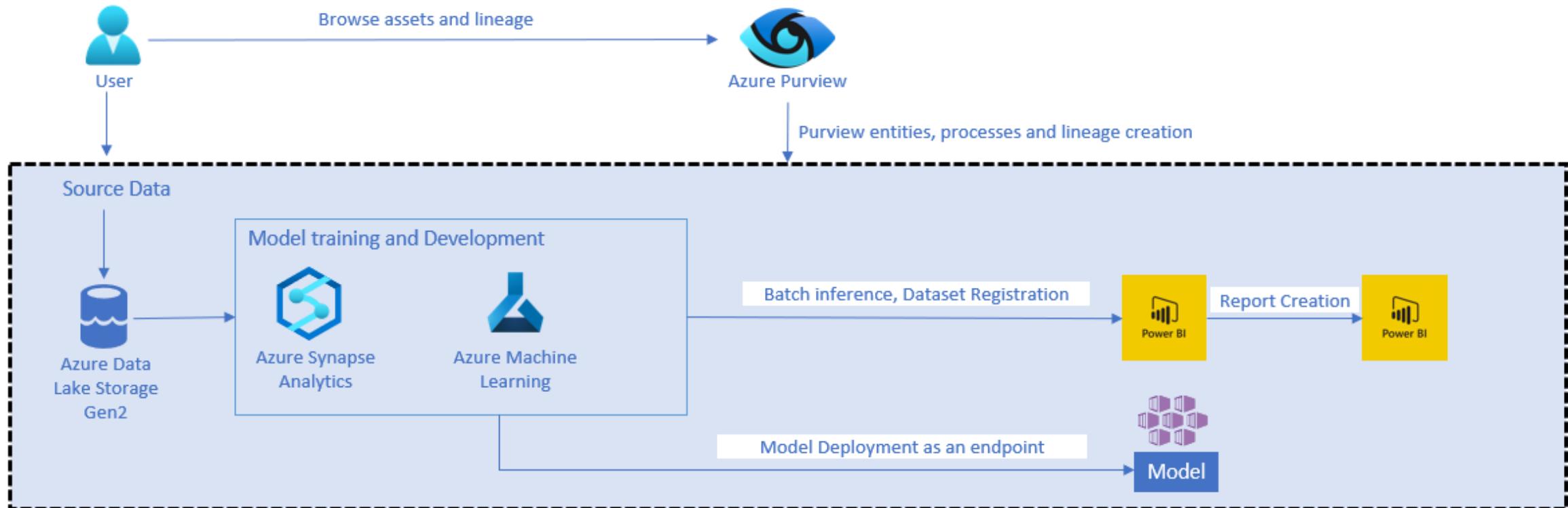
EARLY ACCESS ENGINEERING



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)

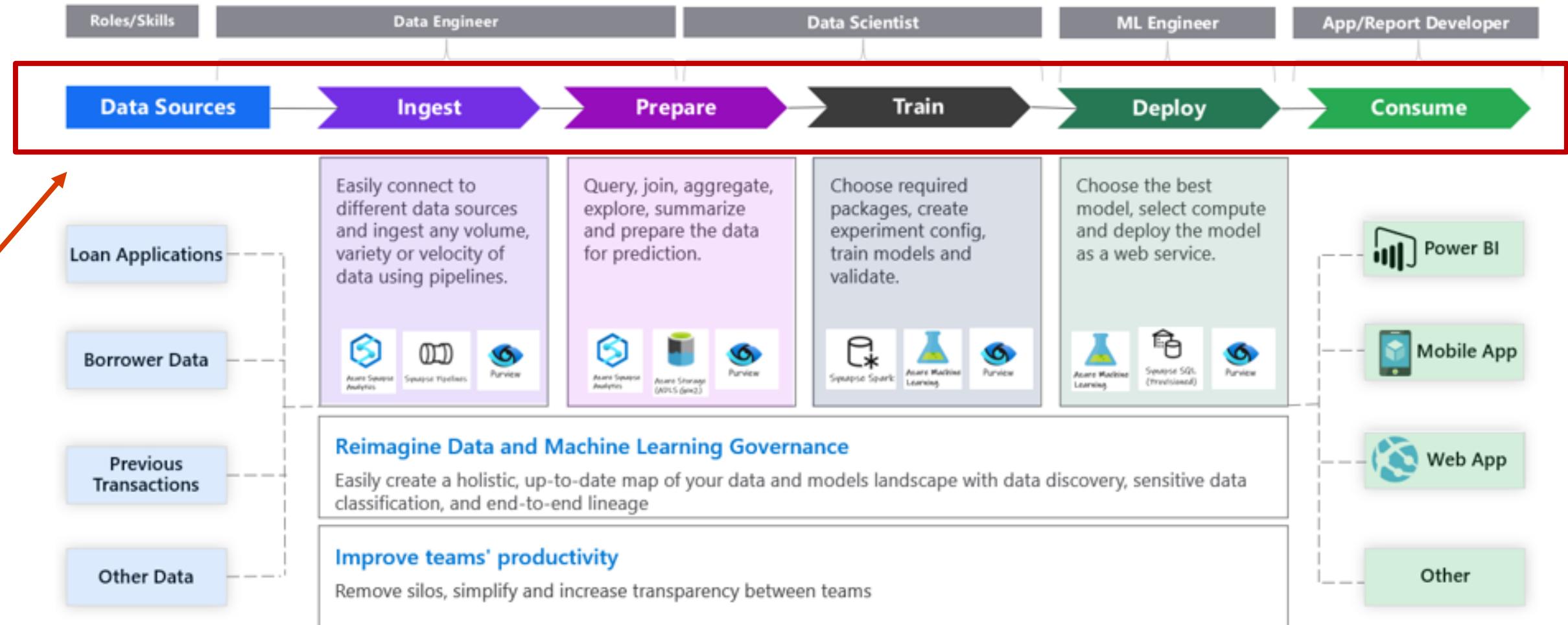
Use Case - Machine Learning Lineage



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

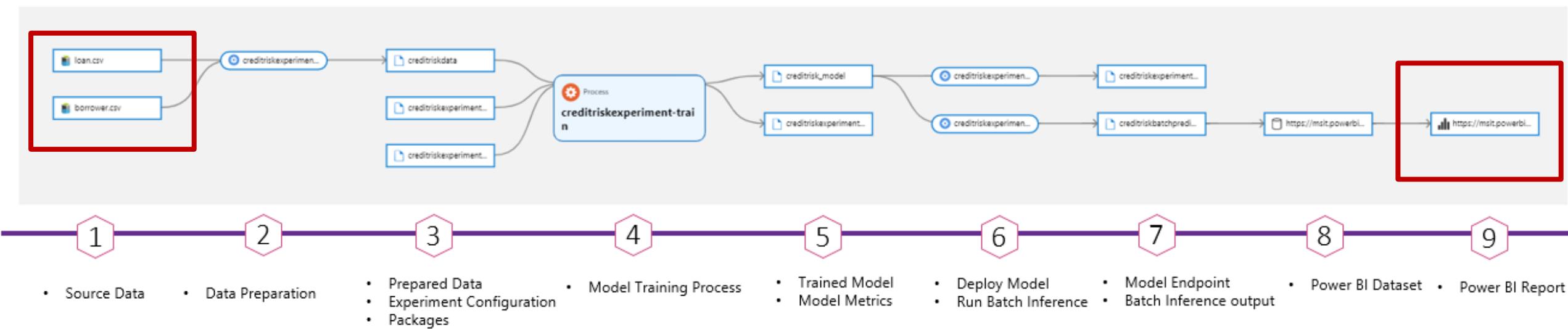


Source:

[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



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](#)
- [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](#)

Resources - References in Slides

Links to Use Case References

- [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](#)