



Microsoft Purview

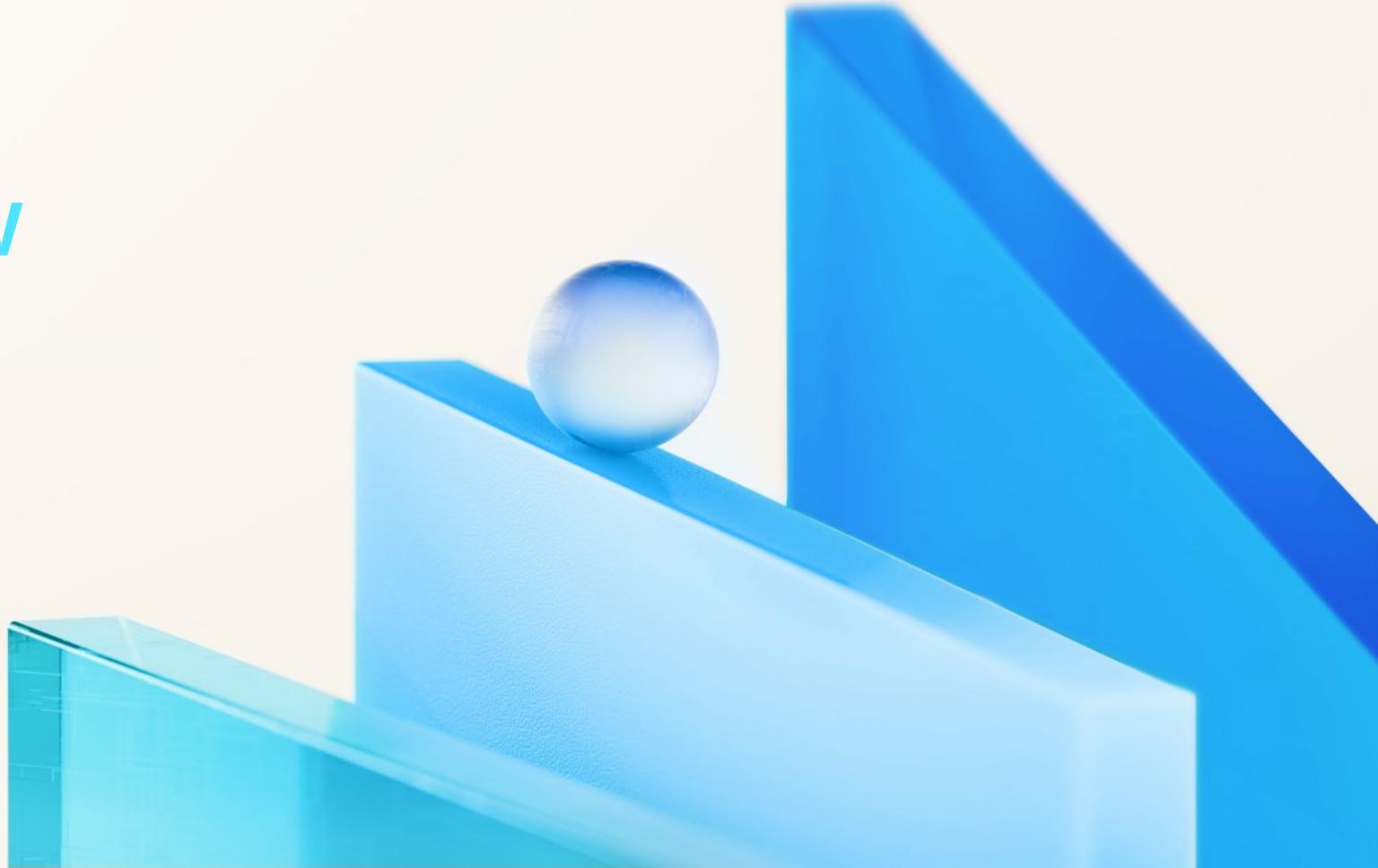
Andrea Benedetti

Sr Cloud Architect, Microsoft

/in/abenedetti

@anBenedetti

<https://github.com/anbened>



Today's data realities

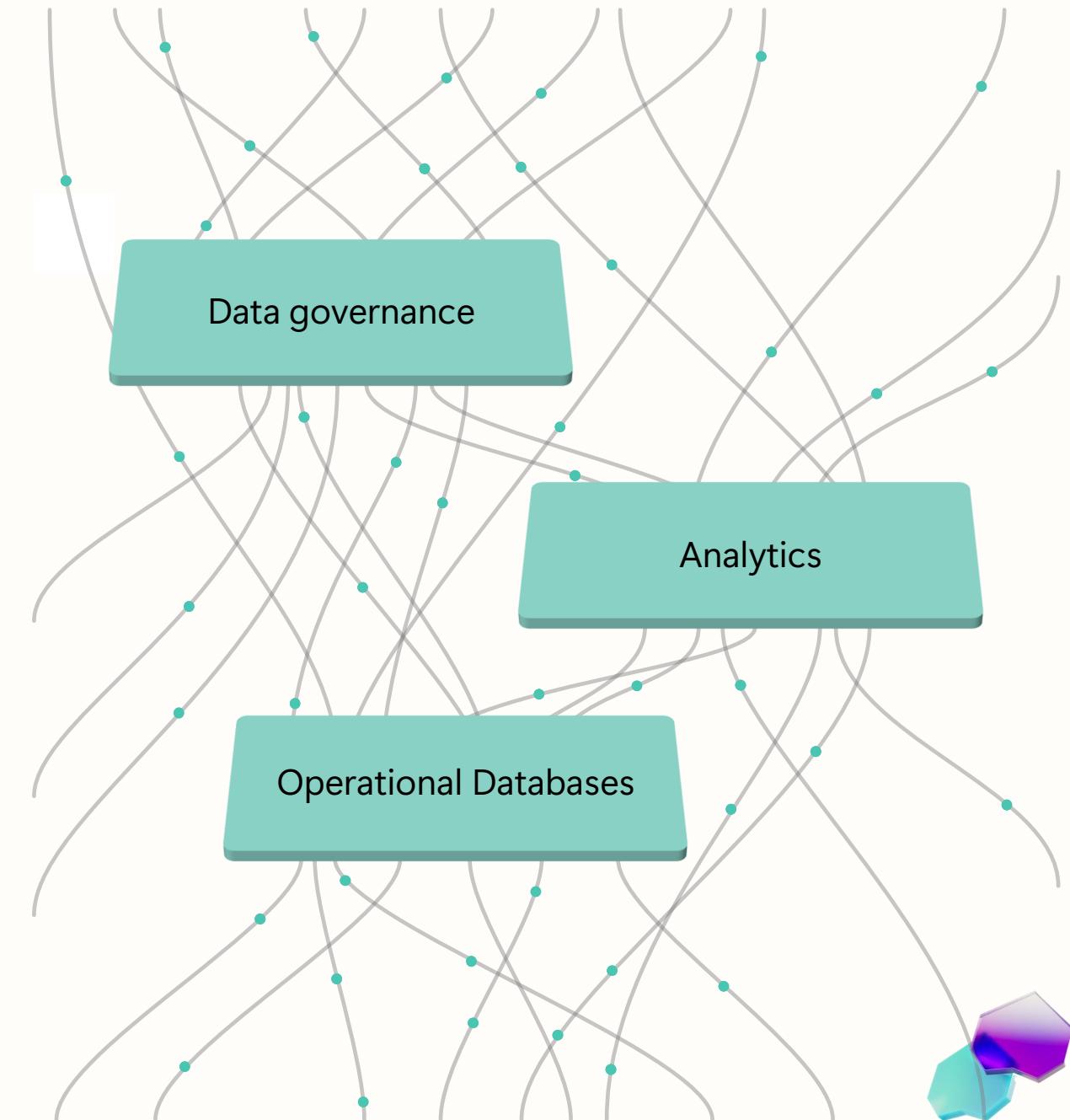
What **data** do I have?

Is it **trustworthy**?

Can people **access the data needed** to make the right decisions?

How can I enable faster **business insights**?

What's my **compliance** exposure?





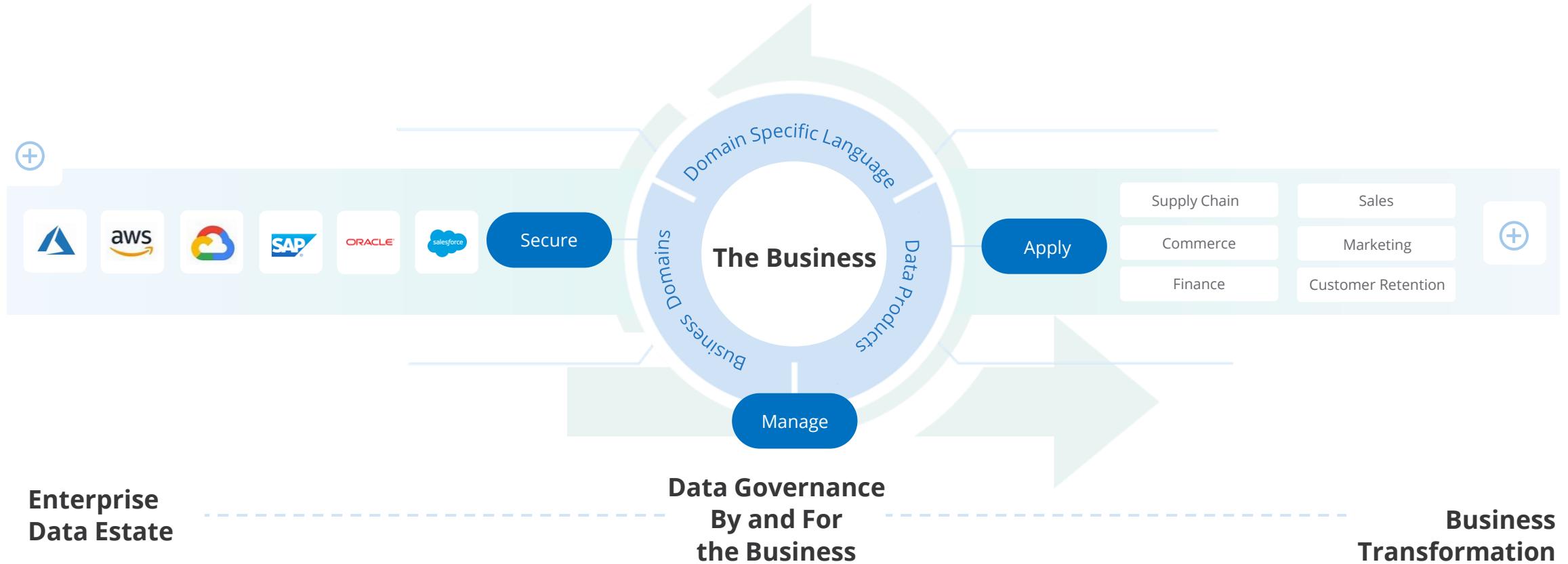
It is evident that Data Governance continues to be a principal concern

Modern architectures are increasingly federated

- a dispersed integration of data
- data products across various teams
- the critical nature of data quality
- the growing spread of data and AI

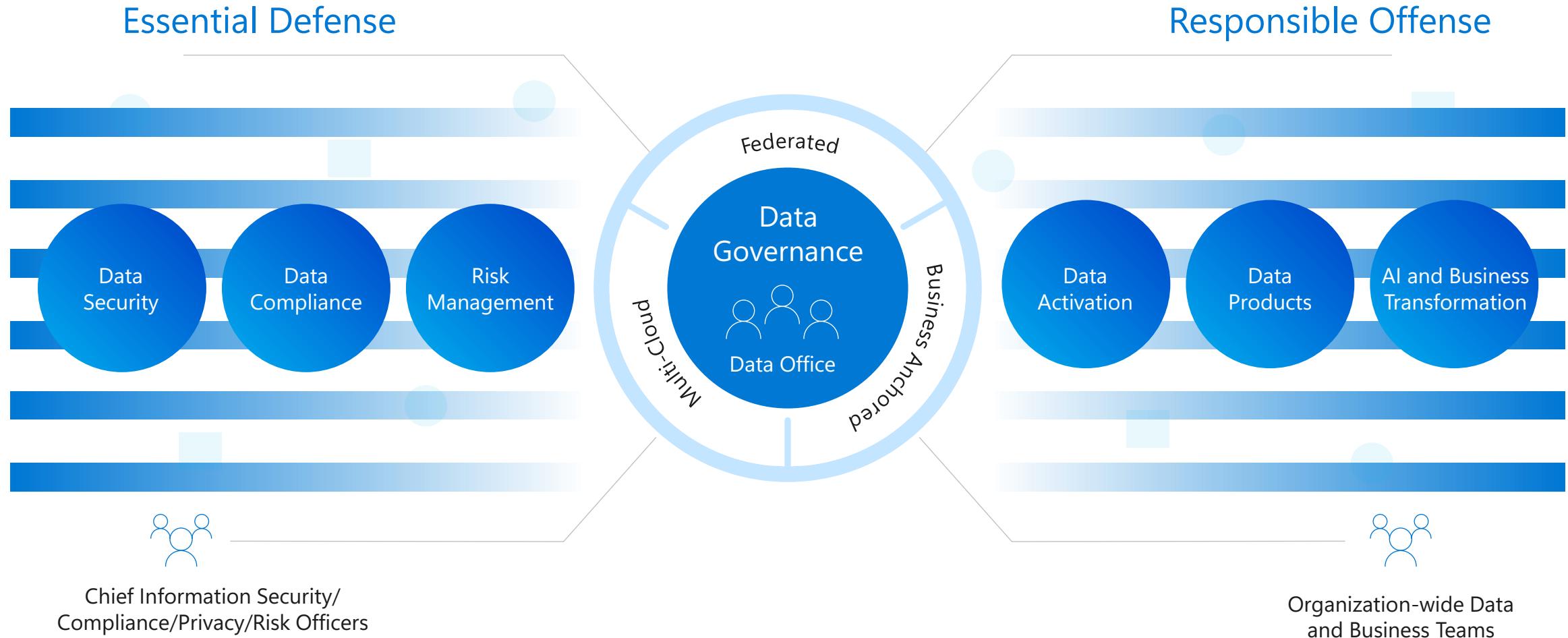


Data Governance by and for the business

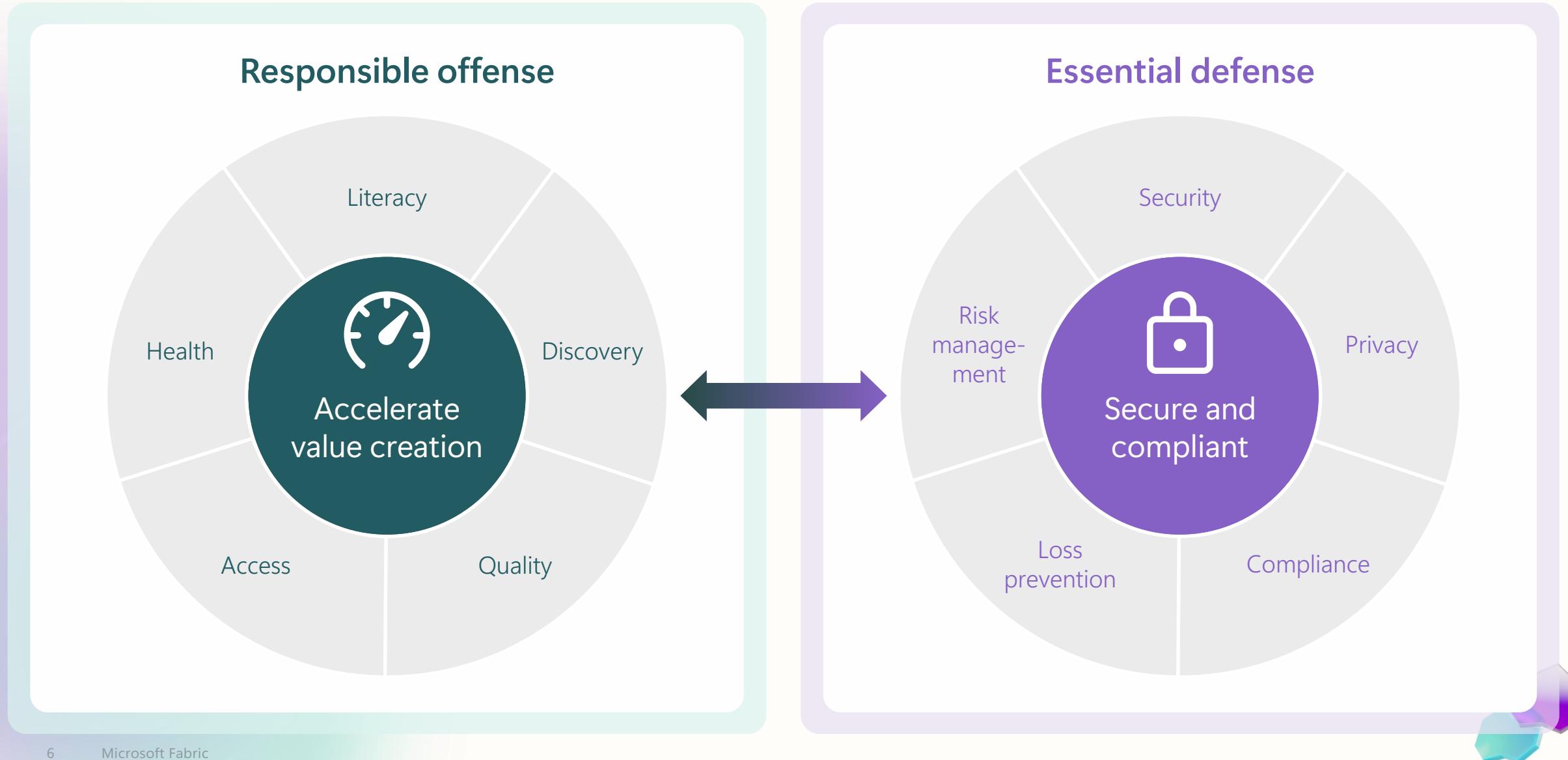


Data Governance for the Age of AI

The whole is greater than the sum of the parts



Data governance and security is evolving



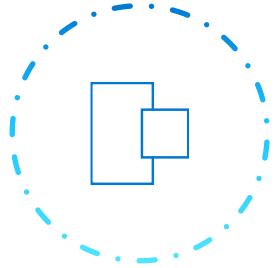


From a technical data catalog into a business-friendly data marketplace, teeming with a plethora of engaging features:

- Objectives and Key Results (OKRs),
- Business Domains,
- Data Products,
- Data Access,
- Data Quality,
- Data Estate Health



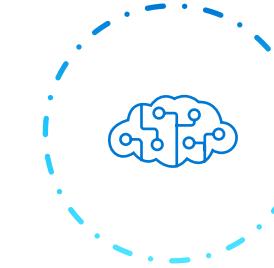
Our Approach to Data Governance



Integrated solution for
Federated Data
Governance



Practice anchored to
the business



Scaled with 'Human
in the loop'
Generative AI

What Data Governance done well can do for your organization



For Organization-wide Data Consumers

Data Discovery

Help you easily find and safely access the data you need for your purpose



For Organization-wide Data Owners and Stewards

Data Curation and Management

Help you deliver high quality data that is easy to discover, understand, and safely access for impactful organization-wide applications



For Data Office(r)s and CxO stakeholders

Data Value Creation

Help you help your organization maximize value creation from your data, while reducing operations spend

Data Understanding

Help you understand what you need to know about the data before and as you use it

Responsible Data Use

Help you ensure that your data is used by the intended users and for its intended purposes

Impact Analysis

Help you understand and action anomaly states impacting your data

Data Estate Standardization

Help you standardize common controls across your organization's data estate, with federated accountability for the health of your data estate

Bringing data security, governance, and compliance together



On-prem and multi-cloud

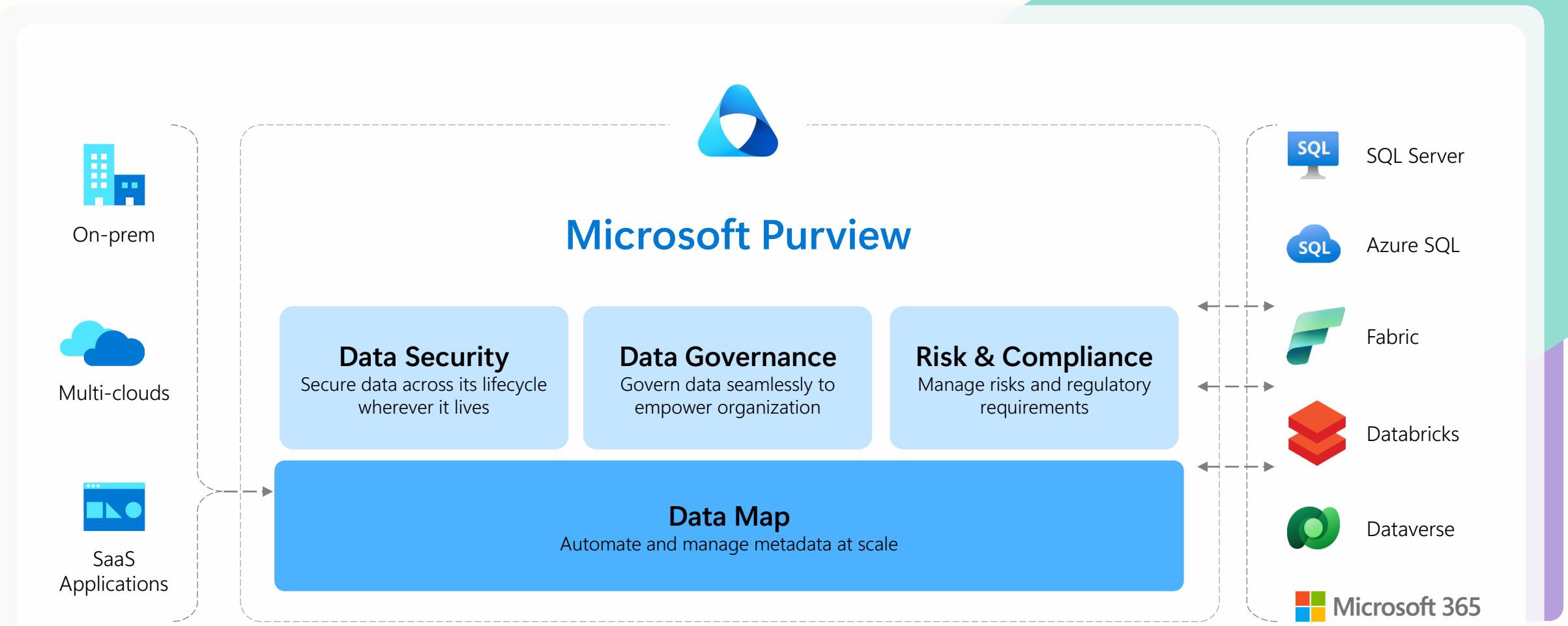


Unstructured & structured data



Across IaaS, and SaaS

Govern and secure your data estate





Build Business
Domain and
Curate Data
Products

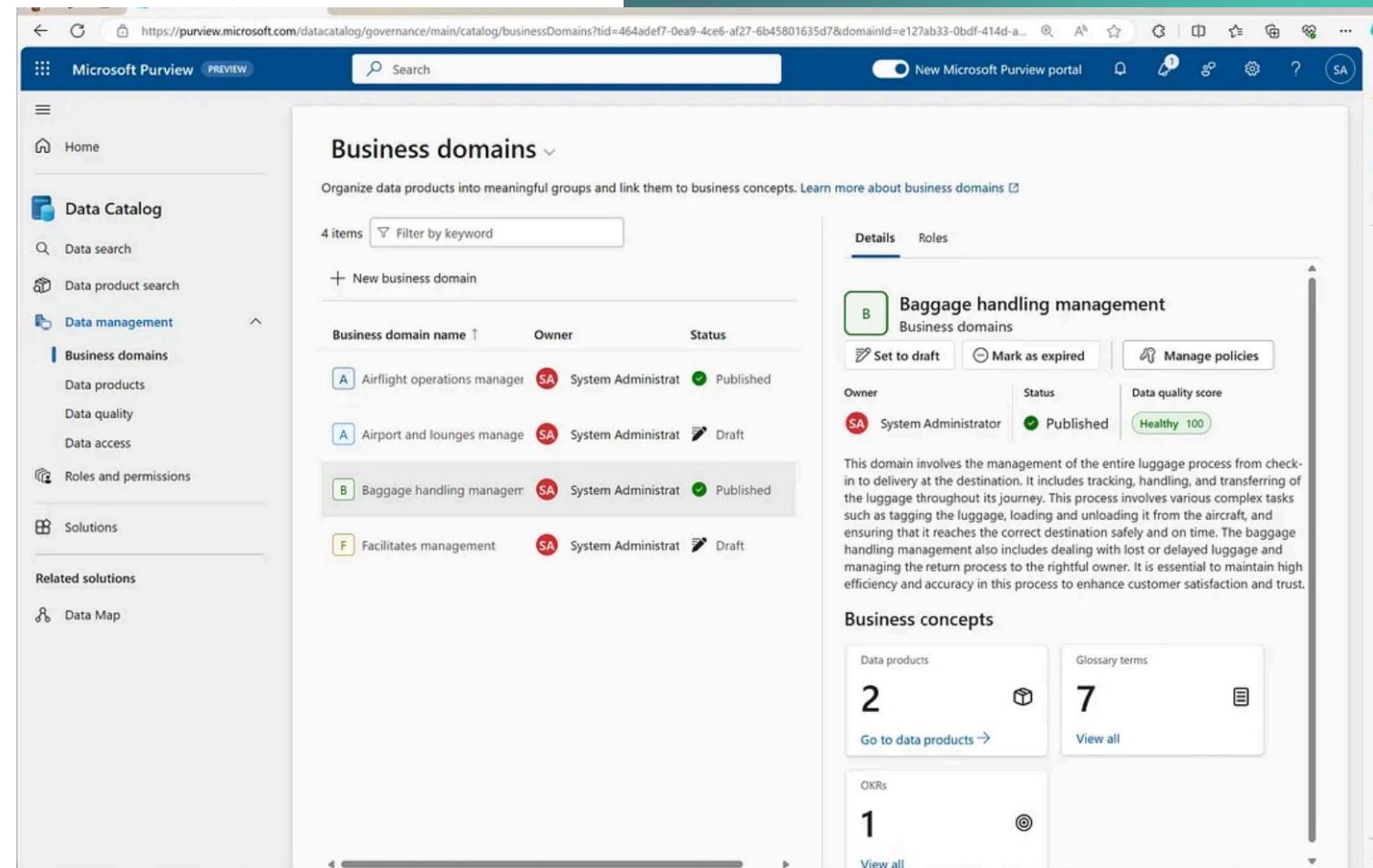
Business domains

Business domains allow enterprises to structure their catalog uniquely for individual teams.

Imagine a business domain as a framework within which you manage your business-related information in the catalog.

So, a smaller catalog nested within your larger one, encapsulating business details, data products, data quality, and so forth.

Each business domain introduces a new set of roles and delineates your business terms, OKRs, and how these align with your data products.



The screenshot shows the Microsoft Purview Data Catalog interface. On the left, a sidebar navigation includes Home, Data Catalog (with Data search, Data product search), Data management (with Business domains, Data products, Data quality, Data access), Roles and permissions, Solutions, and Related solutions (Data Map). The main content area is titled "Business domains" and displays a list of four items:

Business domain name	Owner	Status
Airflight operations manager	System Administrator	Published
Airport and lounges manager	System Administrator	Draft
Baggage handling manager	System Administrator	Published
Facilities management	System Administrator	Draft

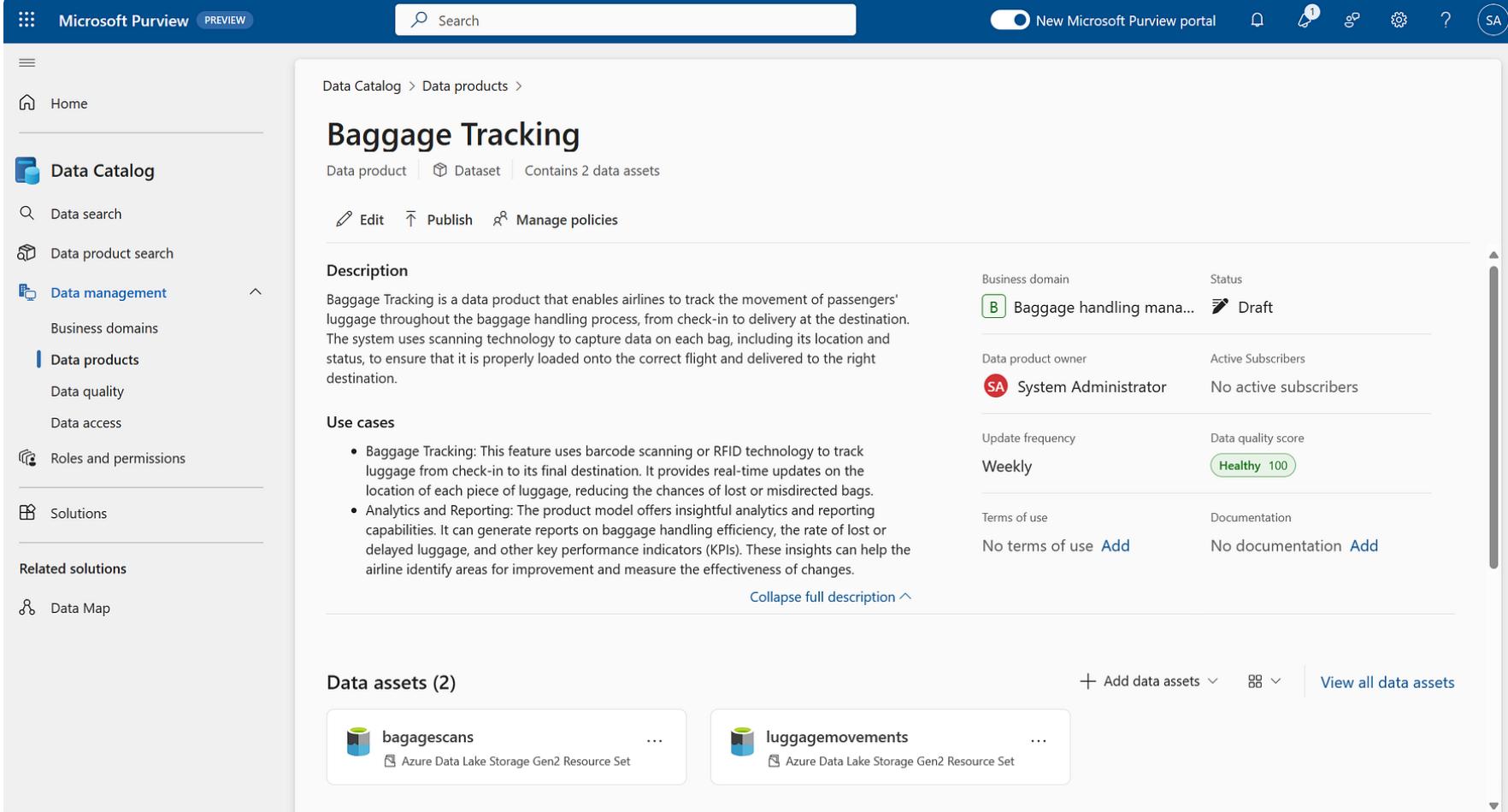
For the "Baggage handling manager" domain (highlighted with a green border), there is a detailed view on the right. It shows the domain's name, owner (System Administrator), status (Published), and a data quality score of "Healthy 100". A descriptive text block explains the domain's purpose: "This domain involves the management of the entire luggage process from check-in to delivery at the destination. It includes tracking, handling, and transferring of the luggage throughout its journey. This process involves various complex tasks such as tagging the luggage, loading and unloading it from the aircraft, and ensuring that it reaches the correct destination safely and on time. The baggage handling management also includes dealing with lost or delayed luggage and managing the return process to the rightful owner. It is essential to maintain high efficiency and accuracy in this process to enhance customer satisfaction and trust." Below this, sections for "Business concepts" show 2 data products and 7 glossary terms, and an OKRs section shows 1 item.

Data Products

Data Products is another new experience. It's a step forward from the existing catalog, enabling the bundling of data assets at the business level.

Suppose you wish to offer a mix of tables, files, and reports to your users — Data Products allow for this bundling.

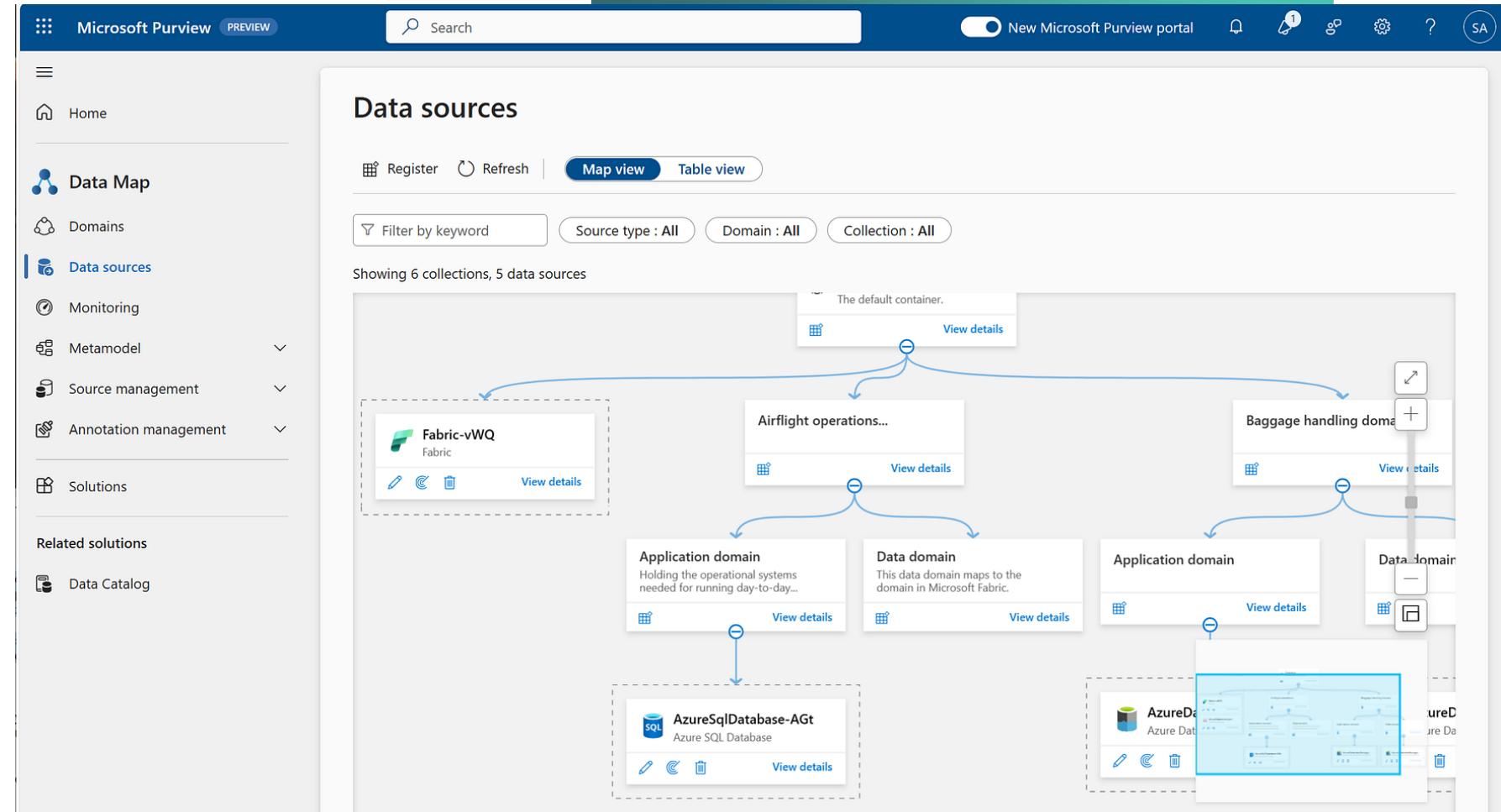
It's a one-stop-shop where users can discover information and request access to the data or artifacts.



The screenshot shows the Microsoft Purview Data Catalog interface. On the left, a sidebar navigation includes Home, Data Catalog (selected), Data search, Data product search, Data management (expanded to show Business domains, Data products, Data quality, Data access, Roles and permissions), Solutions, Related solutions, and Data Map. The main content area displays the 'Baggage Tracking' data product page. The title 'Baggage Tracking' is shown, along with a description stating it's a data product containing 2 data assets. Below the description are sections for 'Use cases' (listing Baggage Tracking and Analytics and Reporting) and 'Data assets (2)' (listing 'bagagescans' and 'luggagemovements' as Azure Data Lake Storage Gen2 Resource Sets). On the right, there are details for the data product, including Business domain (Baggage handling mana...), Status (Draft), Data product owner (System Administrator), Active Subscribers (No active subscribers), Update frequency (Weekly), Data quality score (Healthy 100), Terms of use (No terms of use Add), Documentation (No documentation Add), and a 'Collapsible full description' button.

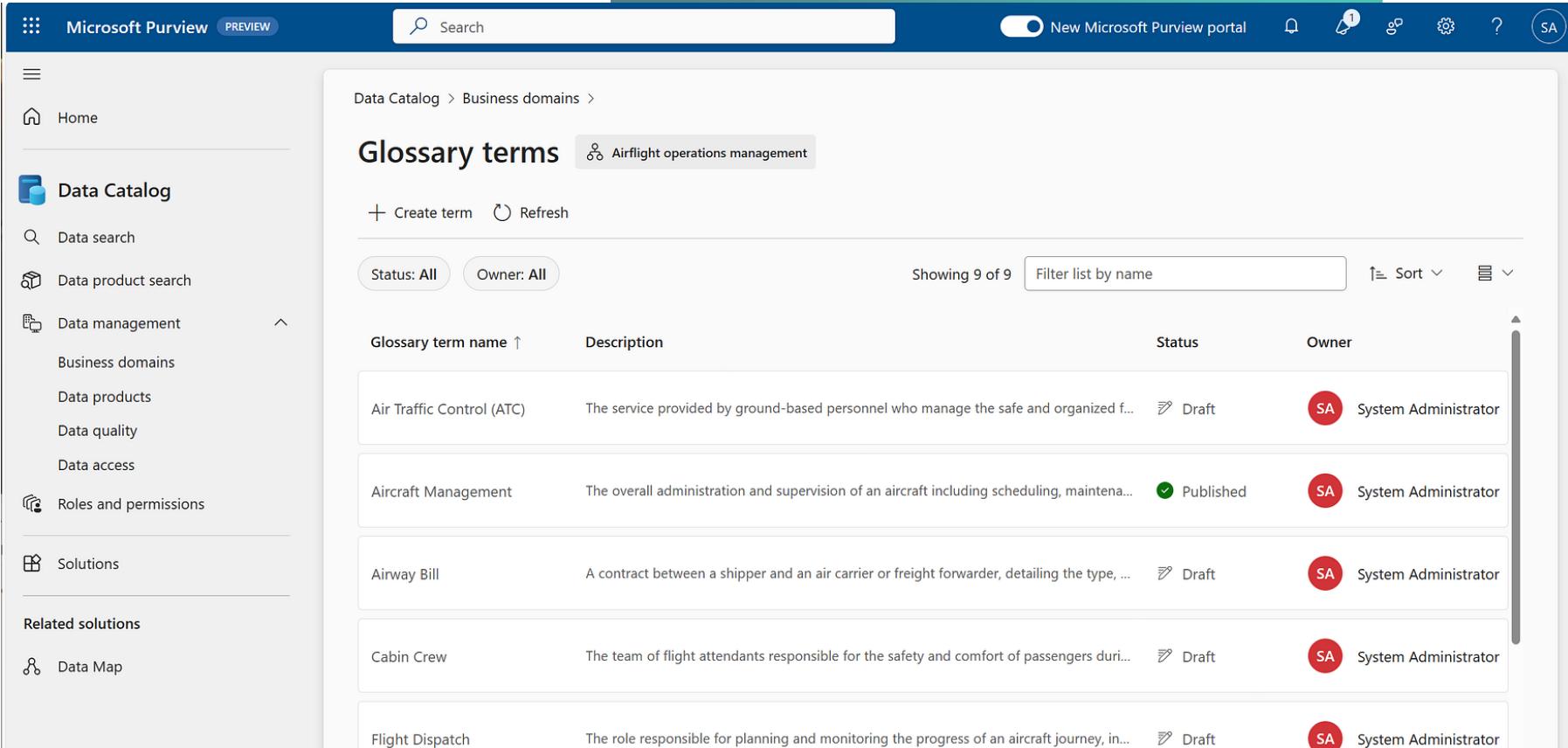
Collection Structure

Microsoft Purview uses (technical) domains for tasks like technical data discovery, data scanning, and classifications. It utilizes a Collection structure for management and logical groupings. Consider this as the solution space where applications and data platforms collaborate to achieve business objectives.



Glossaries for capturing business context

The business glossary as part of your business domain, with business terms linked to data attributes, can effectively highlight the overlap and differences in how the application domain and data domain, i.e. data products, relate in relation to the larger business domain



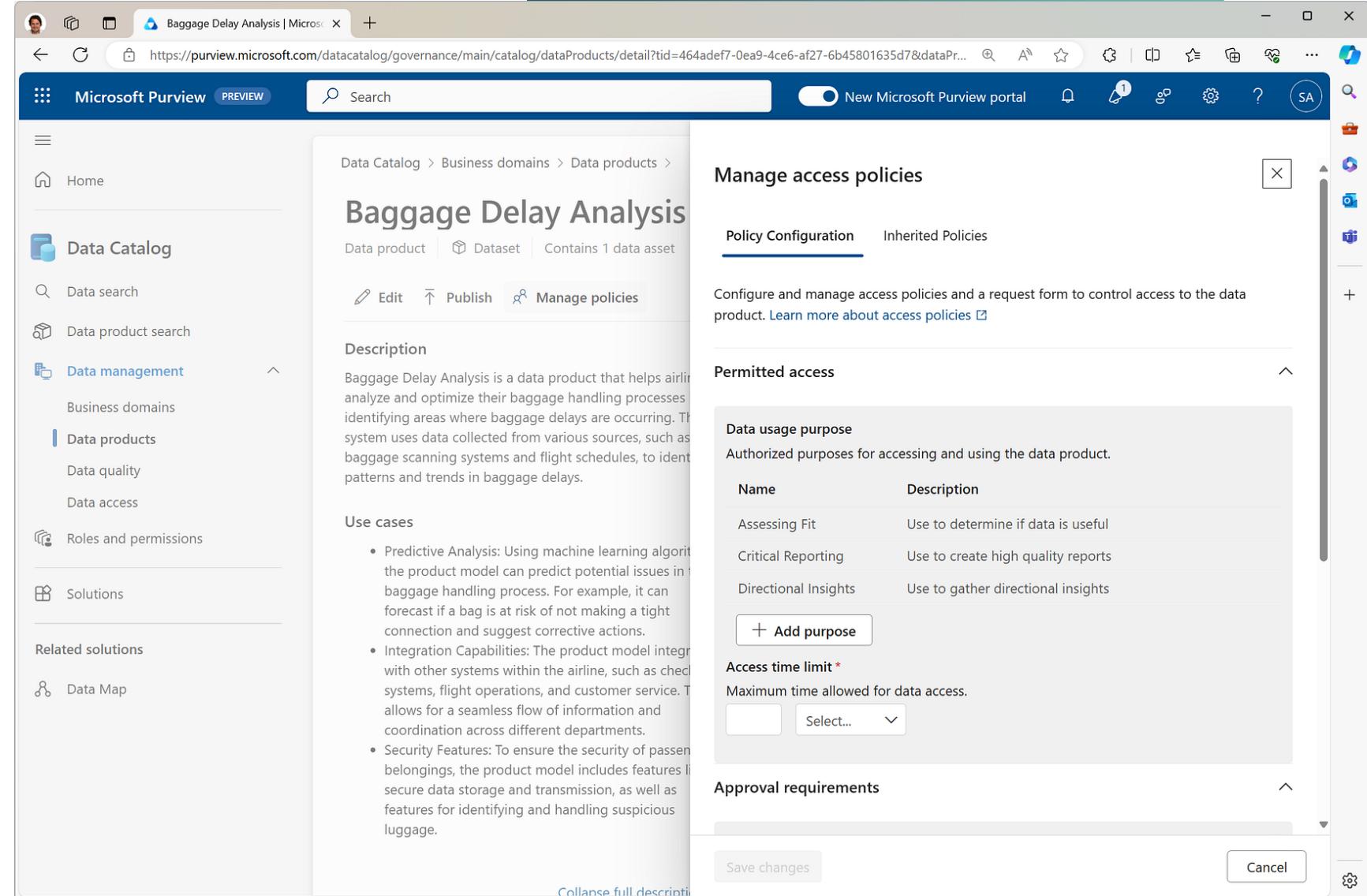
The screenshot shows the Microsoft Purview Data Catalog interface. The left sidebar navigation includes Home, Data Catalog (selected), Data search, Data product search, Data management, Business domains, Data products, Data quality, Data access, Roles and permissions, Solutions, and Related solutions (Data Map). The main content area displays a list of 'Glossary terms' under the heading 'Business domains'. A search bar at the top right shows 'Search' and a dropdown for 'Airflight operations management'. Below the search bar are buttons for '+ Create term' and 'Refresh'. Filter options 'Status: All' and 'Owner: All' are available. A table lists nine glossary terms:

Glossary term name ↑	Description	Status	Owner
Air Traffic Control (ATC)	The service provided by ground-based personnel who manage the safe and organized f...	Draft	SA System Administrator
Aircraft Management	The overall administration and supervision of an aircraft including scheduling, maintena...	Published	SA System Administrator
Airway Bill	A contract between a shipper and an air carrier or freight forwarder, detailing the type, ...	Draft	SA System Administrator
Cabin Crew	The team of flight attendants responsible for the safety and comfort of passengers duri...	Draft	SA System Administrator
Flight Dispatch	The role responsible for planning and monitoring the progress of an aircraft journey, in...	Draft	SA System Administrator

Data Access

Data products are accompanied by access policies.

These policies regulate user access requests, stipulates the data's terms of use, and outlines the approval process for access. These policies can be tailored to accommodate a variety of uses and are expected to support an even wider range of scenarios in the future.



The screenshot shows the Microsoft Purview portal interface. The left sidebar includes links for Home, Data Catalog (with sub-links for Data search, Data product search), Data management (Business domains, Data products, Data quality, Data access), Roles and permissions, Solutions, and Related solutions (Data Map). The main content area displays the 'Baggage Delay Analysis' data product details, including its description, use cases, and a 'Manage access policies' dialog box. The dialog box is titled 'Manage access policies' and contains tabs for 'Policy Configuration' (selected) and 'Inherited Policies'. It provides instructions for configuring access policies and managing a request form. The 'Permitted access' section includes a 'Data usage purpose' table:

Name	Description
Assessing Fit	Use to determine if data is useful
Critical Reporting	Use to create high quality reports
Directional Insights	Use to gather directional insights

Buttons for '+ Add purpose' and 'Select...' are also present. The 'Approval requirements' section is partially visible at the bottom. A 'Save changes' button and a 'Cancel' button are located at the bottom right of the dialog.



Associate Semantic Knowledge to Data

OKRs

The addition of Objectives and Key Results (OKRs) offers another fresh experience.

OKRs can serve as a linking pin between your business domains and data products.

They directly correspond to your objectives and are useful for identifying and monitoring progress on high-priority items.

For example, if you wish to establish a data strategy that includes significant objectives like enhancing customer satisfaction by over 5% using data, OKRs can facilitate measuring such goals.

What is an OKR?

An Objective and Key Result (OKR) is the goal or desired outcome of a business domain. For example, a 10% raise in sales, or a 3% reduction in support cases. Objectives should relate to everything an organization does, and should define how they're achieving their outcomes.

Parts of an OKR

Microsoft Purview's OKRs have several parts:

- **Owners** - the users responsible for maintenance of the objective
- **Definition** - An overarching description of your objective, which can be as broad as "empower every person and every organization on the planet to achieve more", or as specific as "use marketing campaigns to increase sales in the first quarter."
- **Target date** - when you expect to achieve your goal.
- **Key results** - these are the measurable, time-bound goals associated with your objective, and should be how you track progress toward your objective.

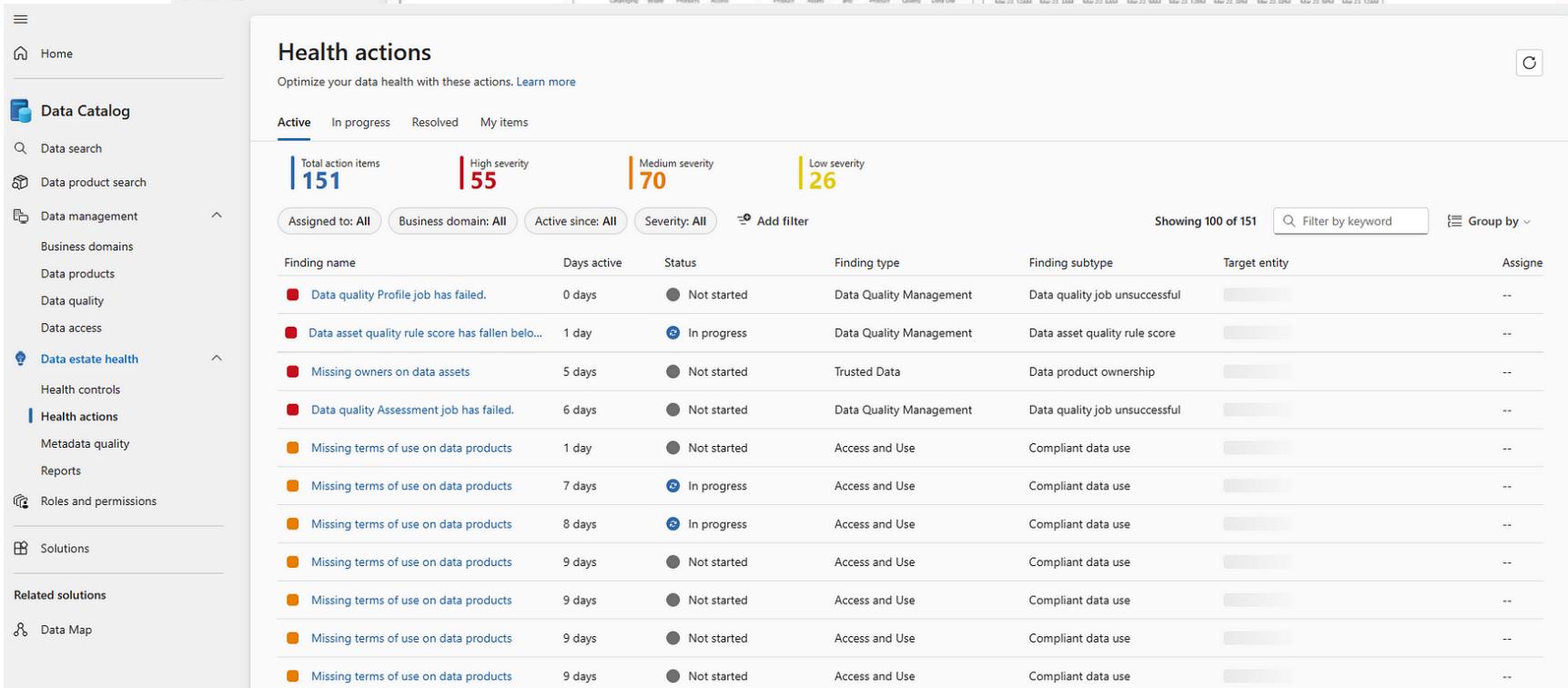
Your marketing team has developed an email campaign results data product that their team can use to track the impact of its email marketing campaigns. Your marketing business domain has a "Customer Response" OKR, with the goal of at least 20% customer response to email, telephone, and mail campaigns sent out.

Data Estate Health

Data Estate Health provides an elevated level of control and insight, enabling users to have a deeper understanding of data management activities across the entire organization.

This feature can be considered essential, particularly for Chief Data Officers tasked with overseeing the overall quality.

It provides them with the necessary tools to identify areas of concern and determine which teams need to be involved in improving overall data quality.





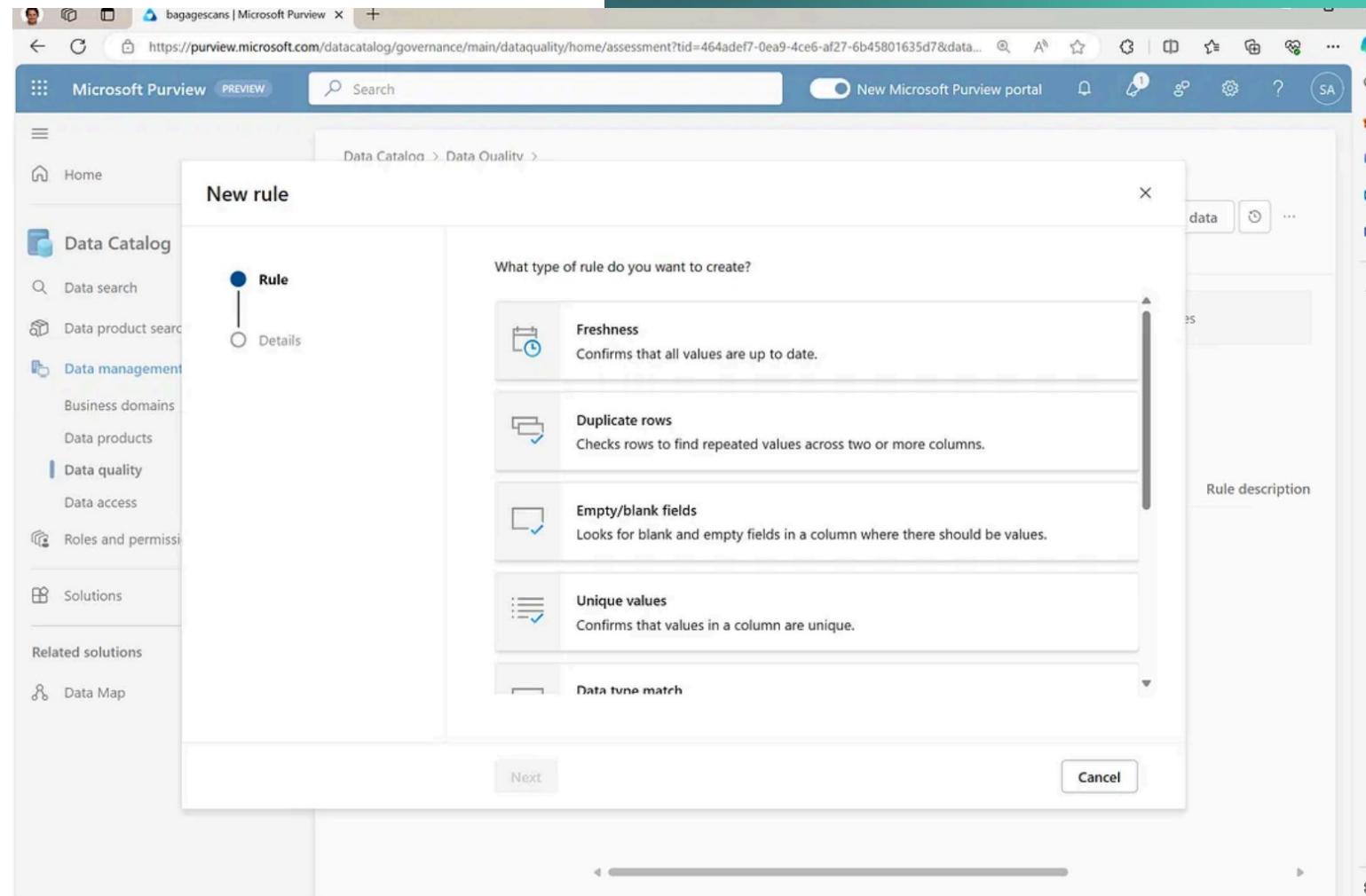
Improve the Data Quality and Remove Data Issues

Data Quality

As previously mentioned, the importance of data quality is paramount.

It's vital for large-scale data usage. The newly introduced data quality service is equipped with a rich set of features, including numerous roles, dashboards, scheduling options, scoring options, rule and scanning types, and more.

There's even a Copilot to assist in rule creation.





Demo

The screenshot shows the Microsoft Data Catalog interface. At the top, there's a search bar with placeholder text "Search data products or ask a question". Below the search bar, there are several navigation links: "Data Catalog", "Data product", "Data management", "Roles and permissions", and "Solutions". On the left, there's a sidebar with "Recent" and "My daily access" sections, and a "Quick access" section listing various data products. The main content area is titled "Data product search" and "Explore data products by business domain". It lists several data products with their types and descriptions:

Data product name	Type	Business domain	Description
Employee data	Operational	Human Resources	Our employee profile includes key elements of an employee's details at the company along with their current position. This data does not include benefits or compensation information as it can be used to determine if an employee is in a position for assignment.
Claim Center	Business System	Claims	Claim Center to track care and data source for all relevant claim data. Policy ID and Identifier data used in the claims process.
Hospital Capacity	Analytical	Care Providers	Determining the optimal number of hospital beds is a complex and challenging endeavor and requires models and techniques which are sensitive to the model itself, constraints, and dynamic variables involved. This study identifies and characterizes such models.
Know your customer	Operational	Customer	The data product is intended to provide an aligned and centralized set of customer data to ensure an organization is positioned for potentially using or facilitating customer.
Opportunities conversion to sales	Operational	Sales	If a user needs a customer and contact, an individual organization that operates an internal or external product or service, could be specifically interested through the release of an offering conversion or through a direct response to advertising or political.
Customer Master List	Master data sets	Sales	One master master customer list for all Sales audits and marketing use cases. In addition, sales will be using the list to calculate incentives for the sales team and partners. The data for this report –
Product Master List	Master data sets	Commerce Functions	This data product contains all of products and offerings. To do so to understand the update of our offerings and releases or anything with a product dimension point will meet this data. We can specifically request this data for any decision to be able to reference it.
Organization Data	Master data sets	Commerce Functions	The data is complete from external sources that have been stored and disseminated to combine and validate our internally collected data. The data contains company specific data from multiple company sites, financial accounting, and strategic and analytical.
Plan Outcomes Program	Operational	Finance	In summary, a plan oriented program involves strategic planning across multiple dimensions, ensuring a holistic approach to organizational success. Sustaining shareholders, customers, and regulators is crucial. Steps may include streamlining new products, markets,

Get started with Fabric and Purview for free



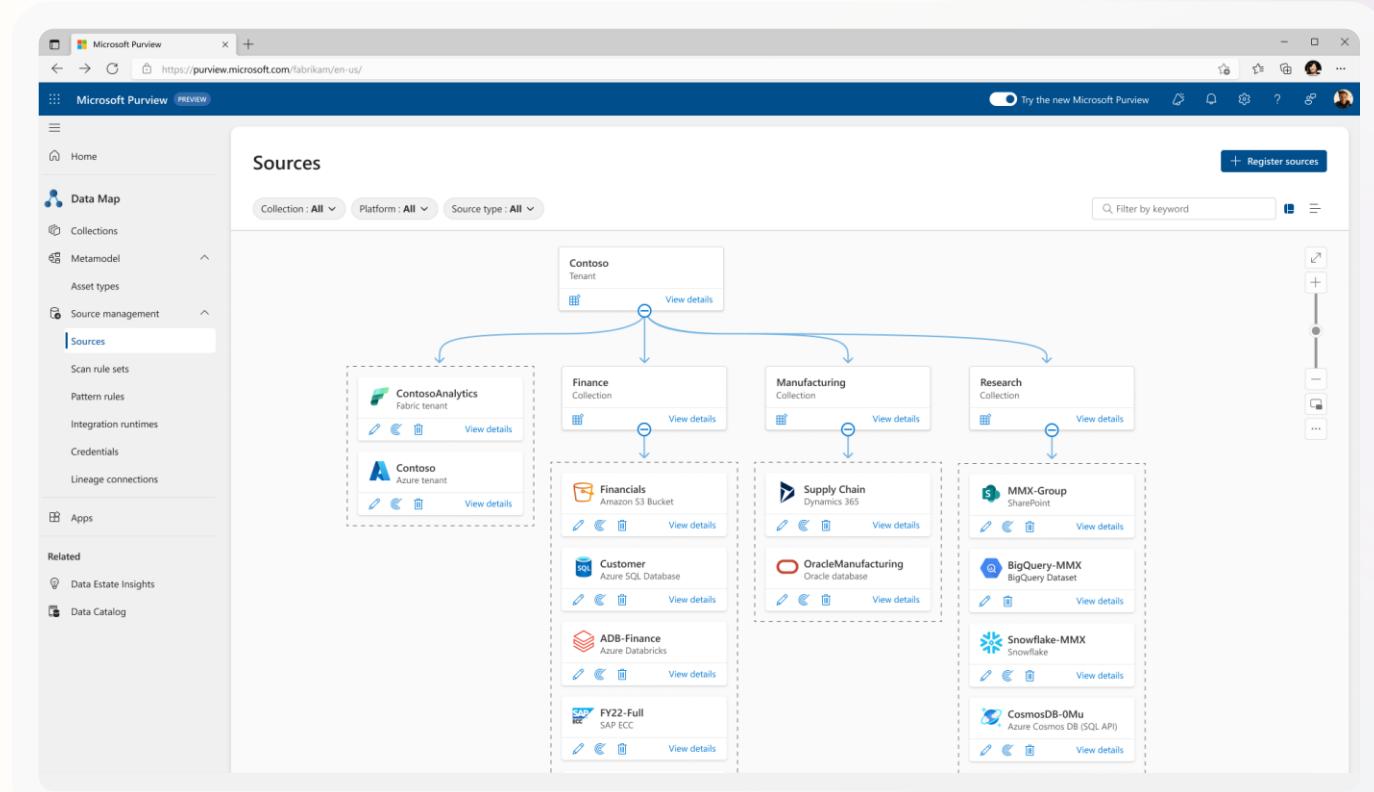
Try out Purview easily with **no set-up required** with Purview Data Map automatically provisioned



Scan and discover data assets across your data estate with data discovery, sensitive data classification, and end-to-end data lineage



Search and explore data assets in the Purview Data Catalog including the technical properties, sensitive data classification, and data lineage



Govern entire data estate with Microsoft Purview Data Governance



Data Discovery

Discover, understand, and get access to data products containing Fabric data in the Purview data catalog

Data Curation

Curate data across entire data estate (including Fabric data) into **business domains** and **data products** for business-friendly data discovery and understanding

Data Management

Manage access and data quality on data products; **Master data management** through partners

Data Estate Health

Monitor and improve data estate health for data products across your organization including Fabric data

Partner Integrations



Microsoft Purview Data Catalog

Create business domains, glossary terms and OKRs to manage data estate (multi-cloud, on-prem, SaaS)

Create data products with data from hybrid sources including auto-attached Fabric to drive consumption and business value generation

Apply data governance policies to govern data access and data use

*Additional Microsoft Purview purchase required
Currently onboarding preview by invitation*



The screenshot shows the Microsoft Purview Data Catalog interface. On the left, there's a sidebar with navigation links: Home, Data Catalog, Data search, Data product search, Data management, Business domains, Data products (which is selected), Data quality, Data access, Data sharing, Data estate health, Summary, Reports, Health controls, Actions, Solutions, Related, Data Lifecycle Management, and Data Map. The main content area displays the 'Marketing Leads' data product page. At the top, it says 'Marketing Leads' and 'Analytics data product'. Below that are buttons for 'Edit', 'Publish', and 'Manage access'. A 'Description' section explains that Marketing Leads is a data-driven lead generation solution for financial marketers. It lists three 'Business use' steps: 1. Data Exploration, 2. Lead Segmentation, and 3. Lead Scoring. Under 'Data assets (3)', there are three entries: 'Customers' (Power BI Dataset), 'Marketing leads' (Power BI Dashboard), and 'Conversion rates' (Power BI Report). A 'Glossary terms (3)' section is also present. On the right side, there are sections for 'Business domain' (Marketing), 'Status' (Draft), 'Owner' (Ralph Ballinger), 'Subscribers' (No active subscribers), 'Health' (No active health controls), 'Action items' (No action items), 'Terms of use' (No terms of use Add), 'Documentation' (No documentation Add), and a 'View all data' button. A large green decorative shape is visible in the background.

Microsoft Purview Data Discovery

Browse, search, and get access to data products across data estate (multi-cloud, on-prem, SaaS) including Fabric data

Get intelligent recommendations on data products best fit business needs

Understand data product's use case, terms of use, and data quality; explore included data asset metadata and lineage

Self-service request access to data product

*Additional Microsoft Purview purchase required
Currently onboarding preview by invitation*



The screenshot shows the Microsoft Purview Data Discovery interface. On the left, a sidebar menu includes Home, Data Catalog, Data search, and Data product search (which is selected). Other options in the sidebar are Data management, Business domains, Data products, Data quality, Data access, Data sharing, Master data, Data estate health, Summary, Reports, Health controls, Solutions, Related, Data Lifecycle Management, and Data Map. The main content area has a header "Data product search" with a sub-header "Discover, understand, and access data products from across your organization." It features a search bar and a "Browse data products by business domain" section with cards for Finance, Human Resources, Marketing, Sales, and Purchasing. Below this is a "Quick access" section with a "Recent" tab and a "My access requests" tab (which is selected). The "My access requests" table lists various data products with columns for Name, Type, Business domain, Owner, Access expires, and Status. Examples include "Customer Data Optimization Initiative" (Analytics data product, Operations, Allan Munger, 9/2/2023, 3:30 PM, Approved), "Data-Backed Customer Journey Mapping" (Reference data product, Human Resources, Celeste Burton, 9/4/2023, 5:00 PM, Approved), and "Sales Commissions Q2 2021" (Analytics data product, Sales, Mona Kane, 10/28/2023, 10:00 AM, Approved).

Data product name	Type	Business domain	Owner	Access expires	Status
Customer Data Optimization Initiative	Analytics data product	Operations	Allan Munger	9/2/2023, 3:30 PM	Approved
Data-Backed Customer Journey Mapping	Reference data product	Human Resources	Celeste Burton	9/4/2023, 5:00 PM	Approved
Sales Commissions Q2 2021	Analytics data product	Sales	Mona Kane	10/28/2023, 10:00 AM	Approved
Marketing Plan Q2	Master data product	Marketing	Cecil Folk	11/3/2023, 9:30 AM	Approved
AI-Driven Fraud Detection System	AI/ML data product	Administration	Andre Lawson	—	Pending
Churn Prediction Model	Analytics data product	Research	Cecil Folk	—	Pending
Data Governance Enhancement Project	Analytics data product	Supply Chain Management	Charlotte Watson	—	Pending
Data Insights Accelerator	Master data product	Research	Cecil Folk	—	Pending
E-Commerce Personalization Project	Master data product	Marketing	Cecil Folk	—	Pending

Microsoft Purview Data Management

Define data access policy at business domain, glossary term, and data product level

Set data quality rules and monitor data quality for data products including Fabric data

Master data management via partners

Additional Microsoft Purview purchase required
Currently onboarding preview by invitation



The screenshot shows the Microsoft Purview Data Management interface. On the left, a sidebar menu includes options like Home, Data Catalog, Data management, Business domains, Data products, Data quality (which is selected), Data access, Data sharing, Master data, Data estate health, Summary, Reports, Health controls, Solutions, Related (Data Lifecycle Management, Data Map), and Profile. The main content area displays the "Pipeline Data Q1" assessment under "Opportunity Insights Audit". It shows a "Data quality score" of 64 (yellow), "Business domain" (Sales), "Data Product" (Opportunity Insights Audit Reference), and "Data asset" (Pipeline Data Q1 Fabric Delta Lake table). A "Quality score trend" chart shows a score of 77, a change of +102%, and previous scores of 38. The chart tracks data from July 15 to August 15. Below the chart, it says "Last 20 scores". The overall interface has a modern design with dark mode options and various status indicators.

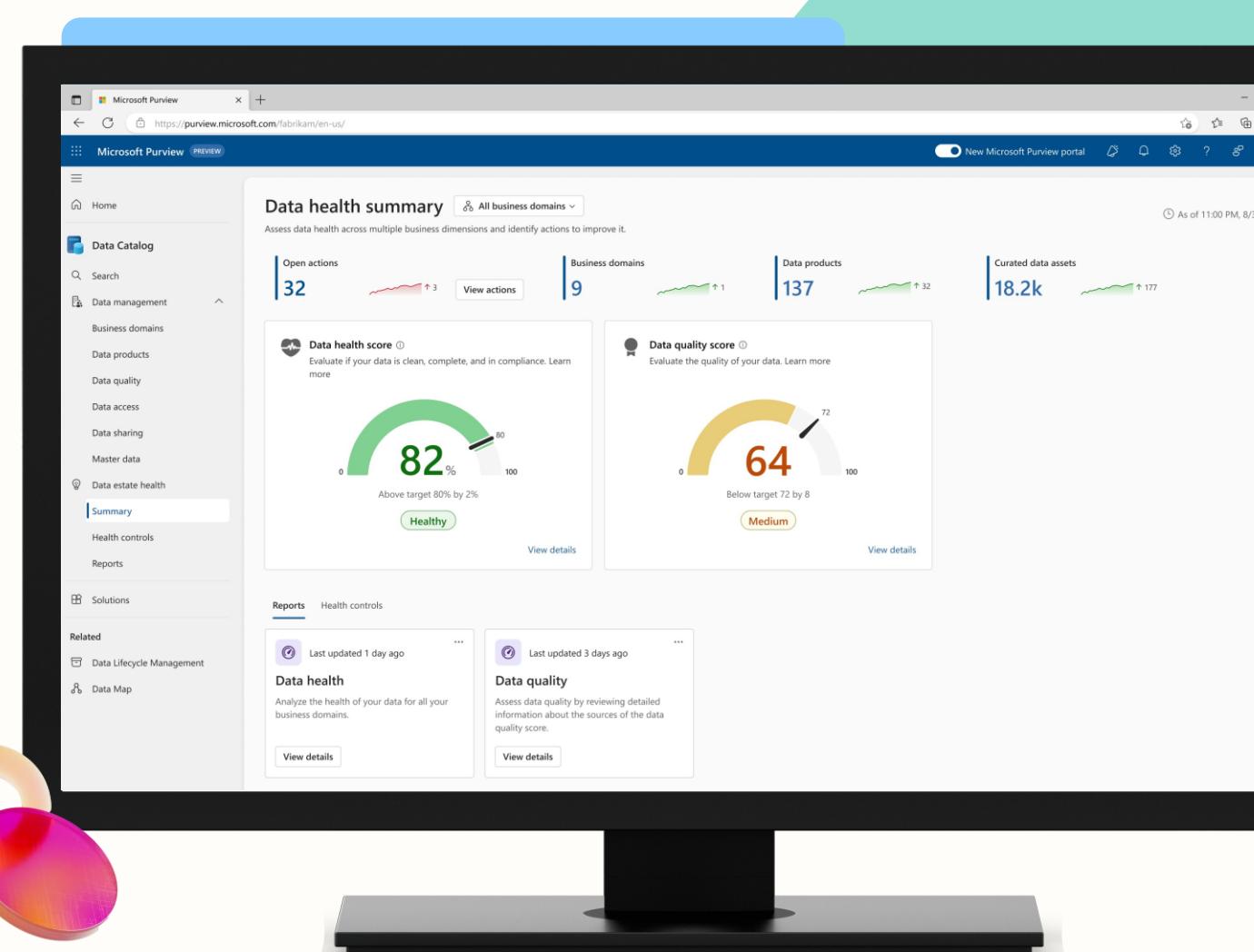
Microsoft Purview Data Estate Health

Gain rich, built-in, insights into organization's data estate health including Fabric data

Leverage actionable insights to improve organization's overall data estate health

Customize dashboards to cater to organization's specific needs

*Additional Microsoft Purview purchase required
Currently onboarding preview by invitation*



Product look ahead...



Microsoft Purview Data Governance Public Preview Rollout

Enterprise Data Governance in the Age of AI



Microsoft Purview Data Governance Public Preview Rollout¹

	Wave 1 By 4/15/24	Wave 2 By 4/29/24	Wave 3 By 5/13/24	Wave 4+ To be announced
North Europe	✓			
South Central US	✓			
France Central	✓			
UK South	✓			
East US2		✓		
Southeast Asia		✓		
Australia East		✓		
Canada Central		✓		
West Europe			✓	
East US			✓	
UAE North				✓
Japan East				✓
Brazil South				✓
Central India				✓
More regions to come!				

¹ Private Preview customer product availability: 4/9

© Copyright Microsoft Corporation. All rights reserved.



Microsoft Purview Data Governance Roadmap

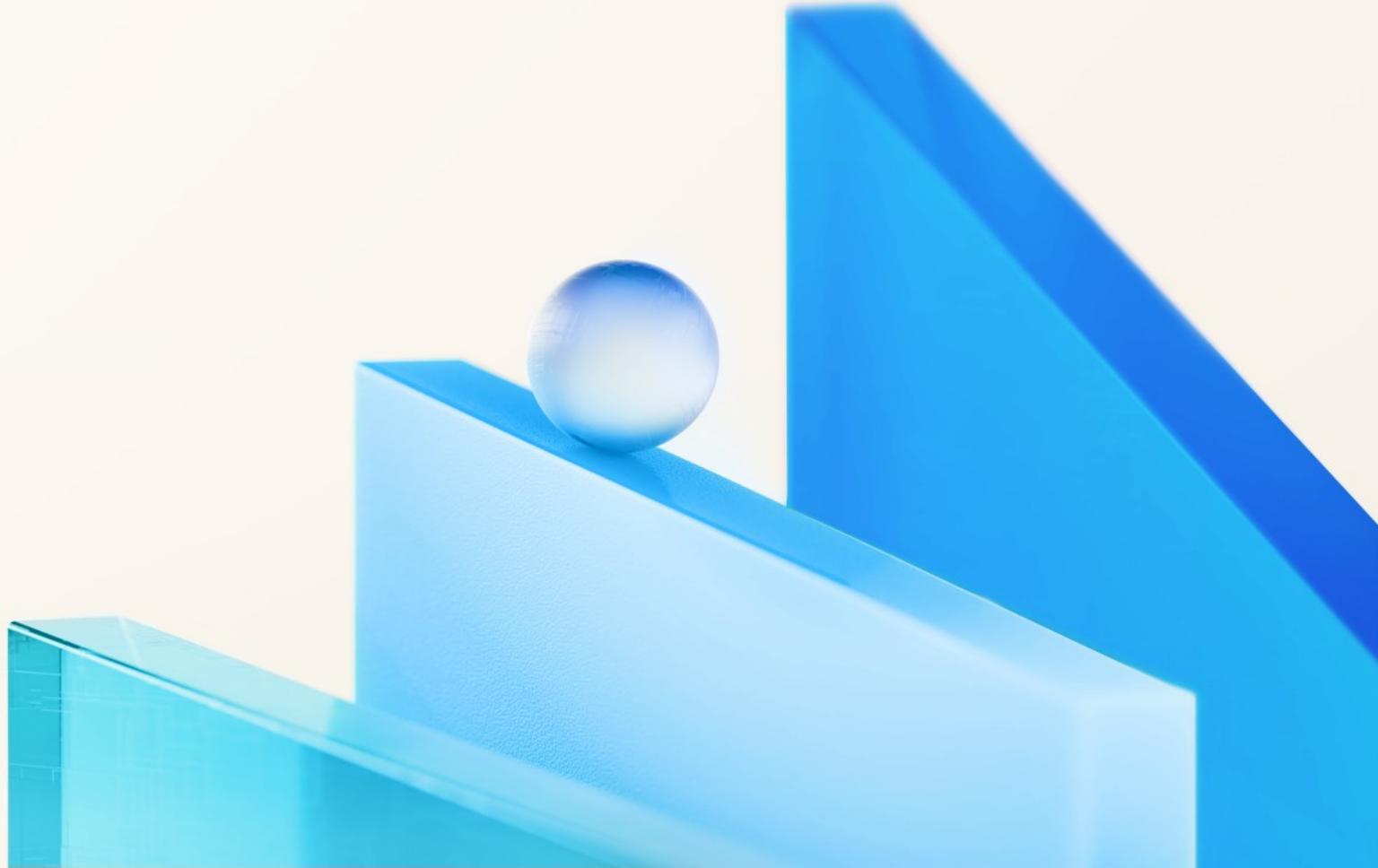
Enterprise Data Governance in the Age of AI



Microsoft Purview Roadmap

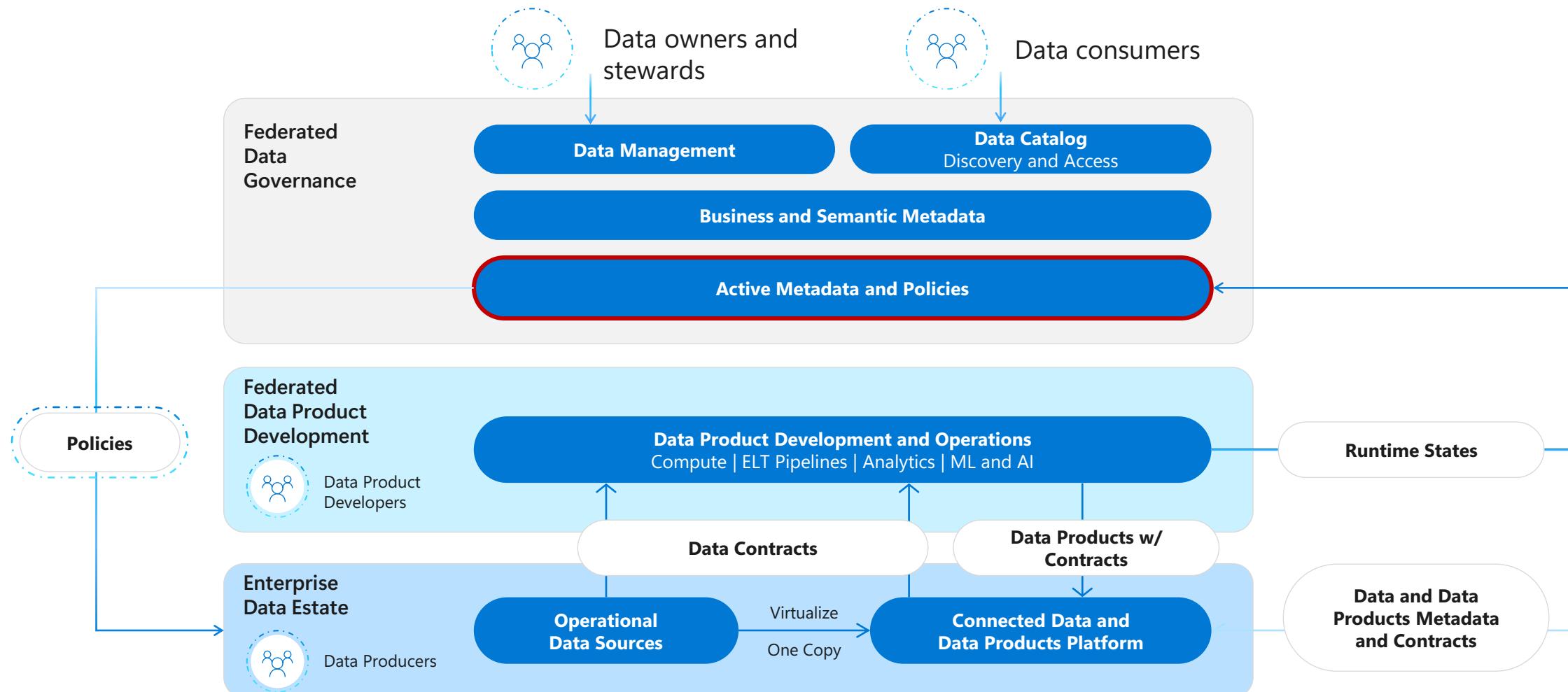


Thank you

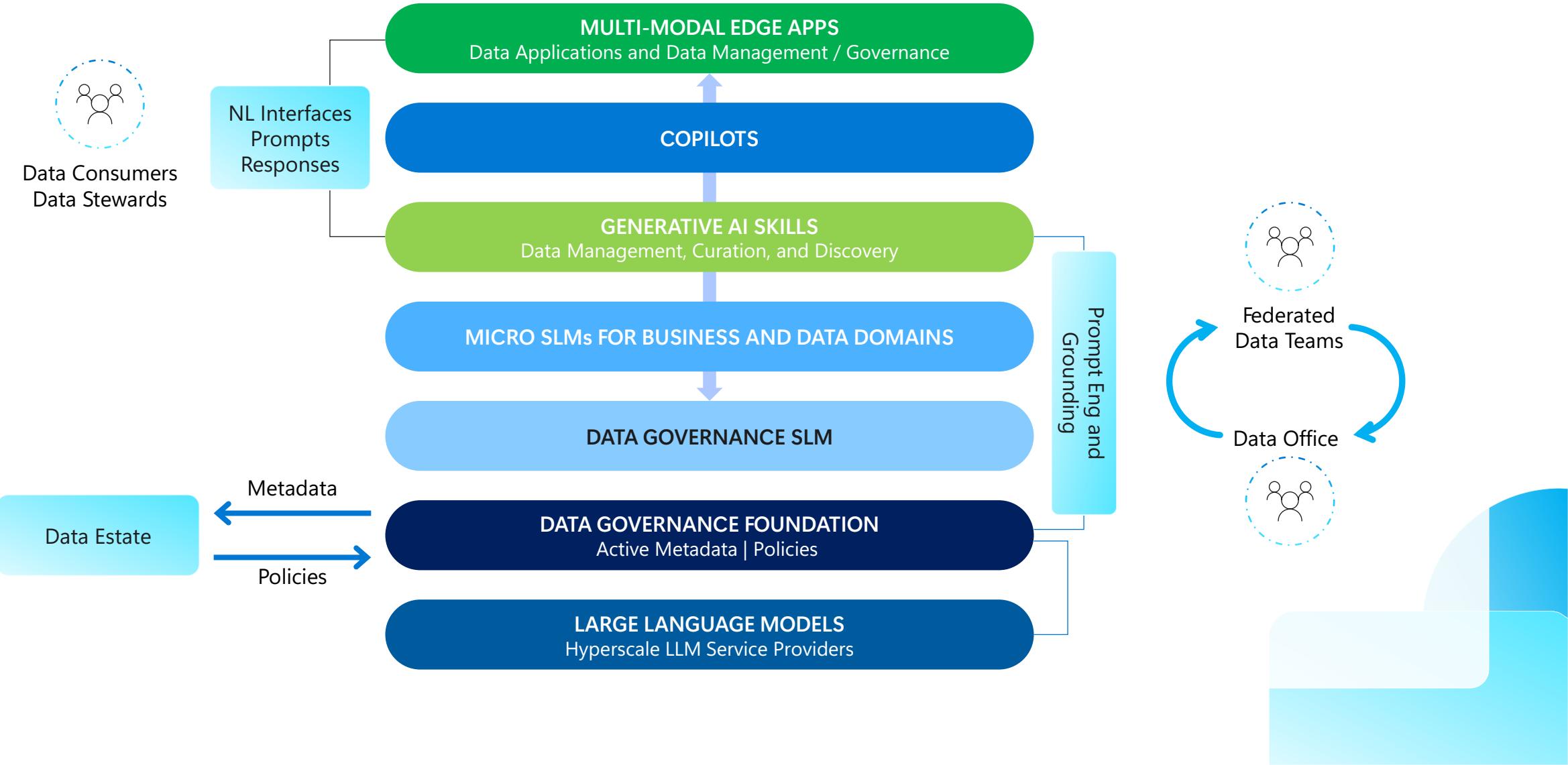


Add On

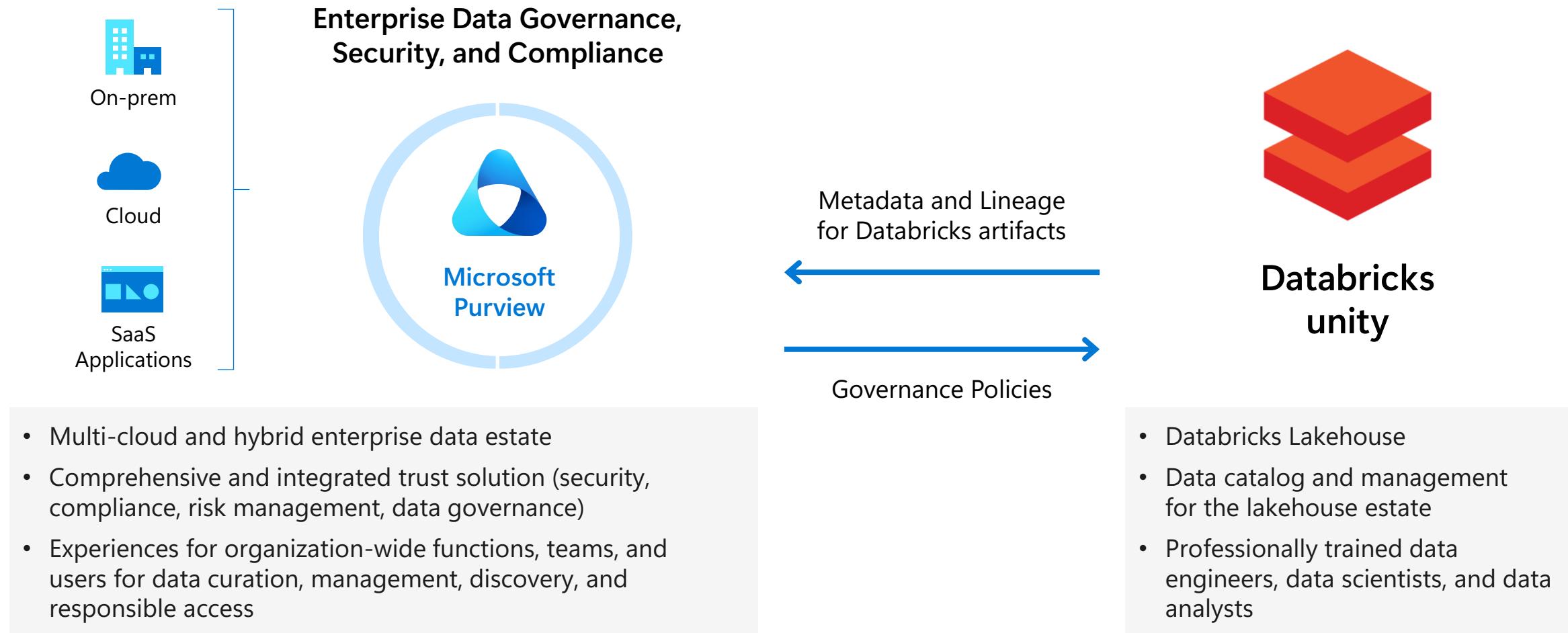
The Data Governance Stack Unpacked



Generative AI for data governance



Purview and Databricks Unity Co-opt



Microsoft Fabric and Microsoft Purview

