

Let's connect



Erwin de Kreuk  
Principal Consultant  
Lead Data & AI InSpark

- @erwindekreuk.bsky.social
- linkedin.com/in/erwindekreuk
- erwindekreuk.com
- github.com/edkreuk
- <https://sessionize.com/erwin-de-kreuk/>
- Dutchfabricusergroup.com





Onelake with Fabric, The Data Lake-as-a-Service Platform

# THANK YOU, SPONSORS!!!!

Platinum



Gold



Silver



Bronze

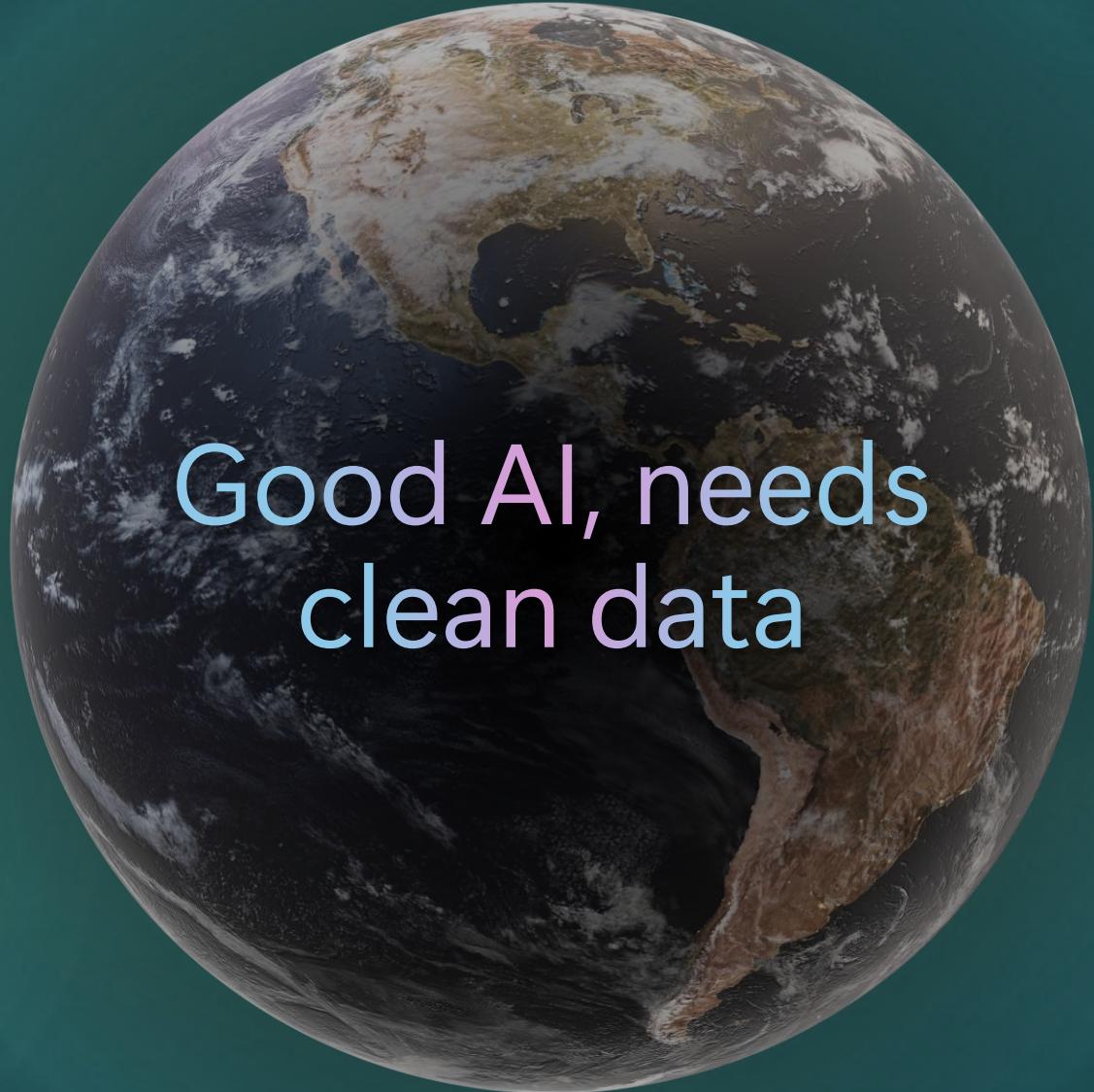


The background image shows a serene landscape with a calm lake in the foreground, a range of green and brown mountains in the middle ground, and a small cluster of houses nestled among trees. In the top right corner, there is a solid red vertical rectangle.

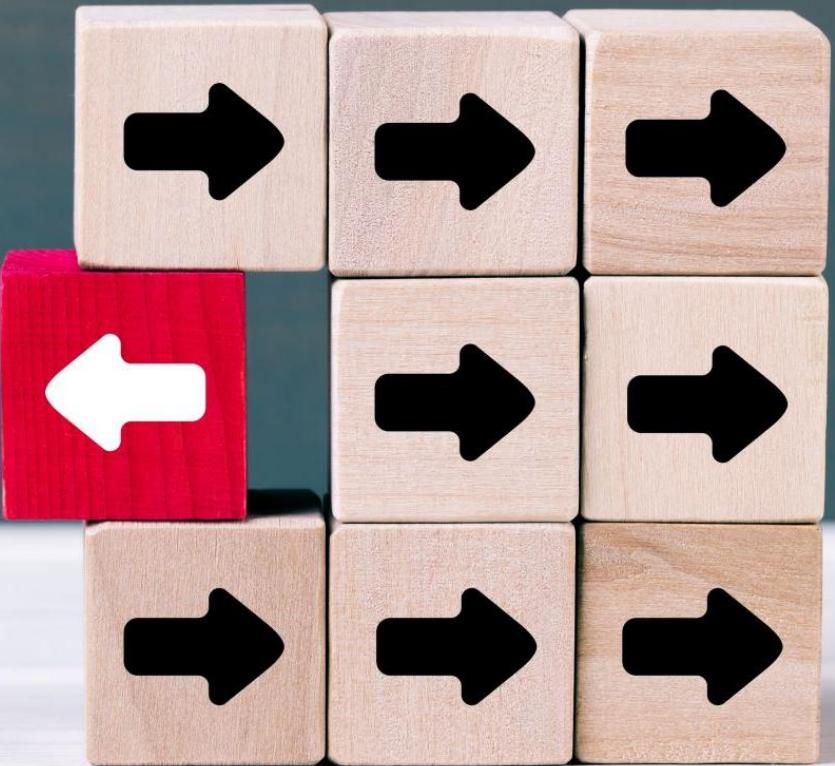
A world awash with **data...**



AI is transforming the world

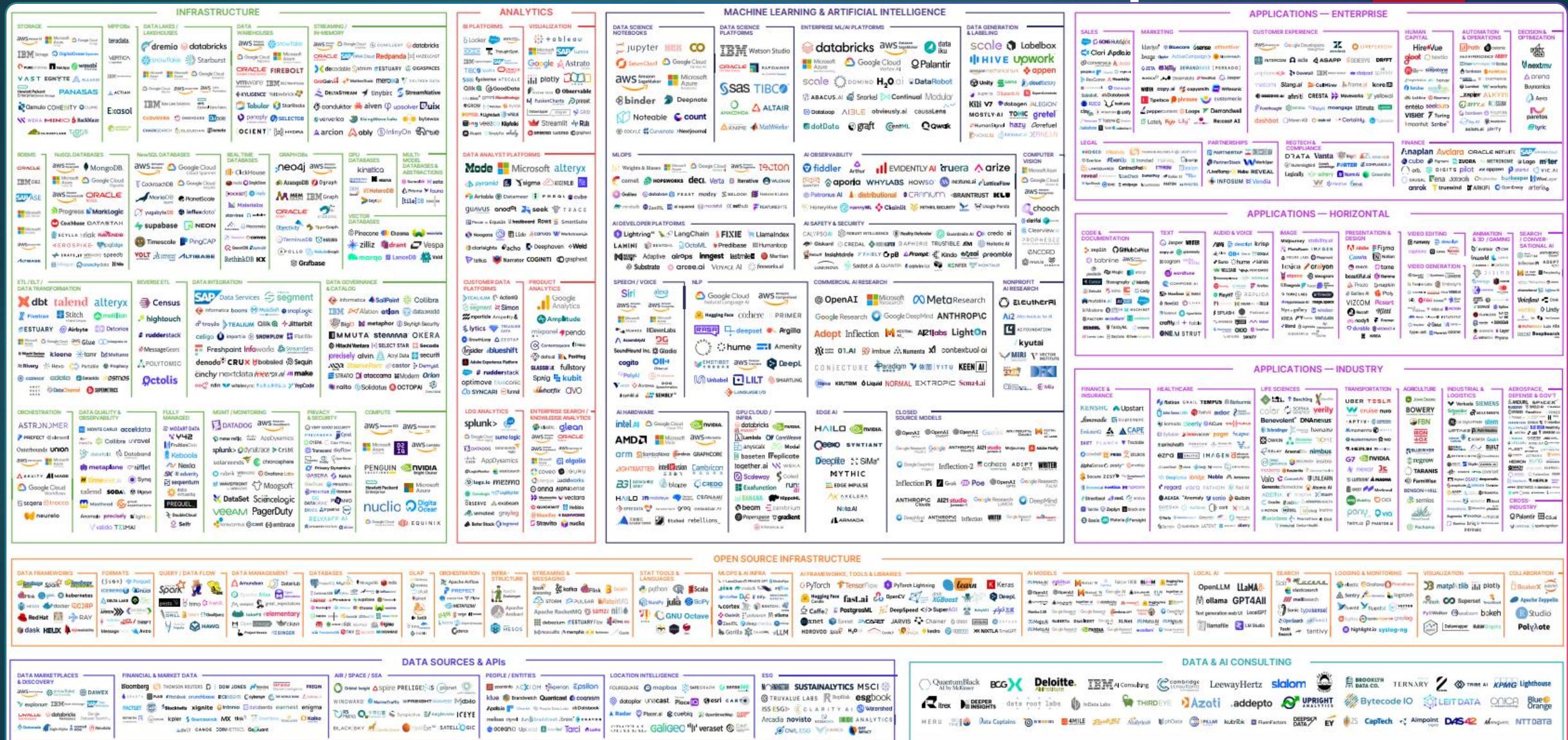


Good AI, needs  
clean data



How do you  
translate  
data into  
competitive  
advantage?

# The 2024 data and AI landscape



“

Unify,

I am the Chief Information Officer and don't  
want to be the Chief Integration Officer.

Help me unleash AI in my data estate.”

Every CIO, Every Enterprise

# Objectives of today

- Onelake
- Shortcuts
- Mirroring
- Catalog
- Security
- Recap



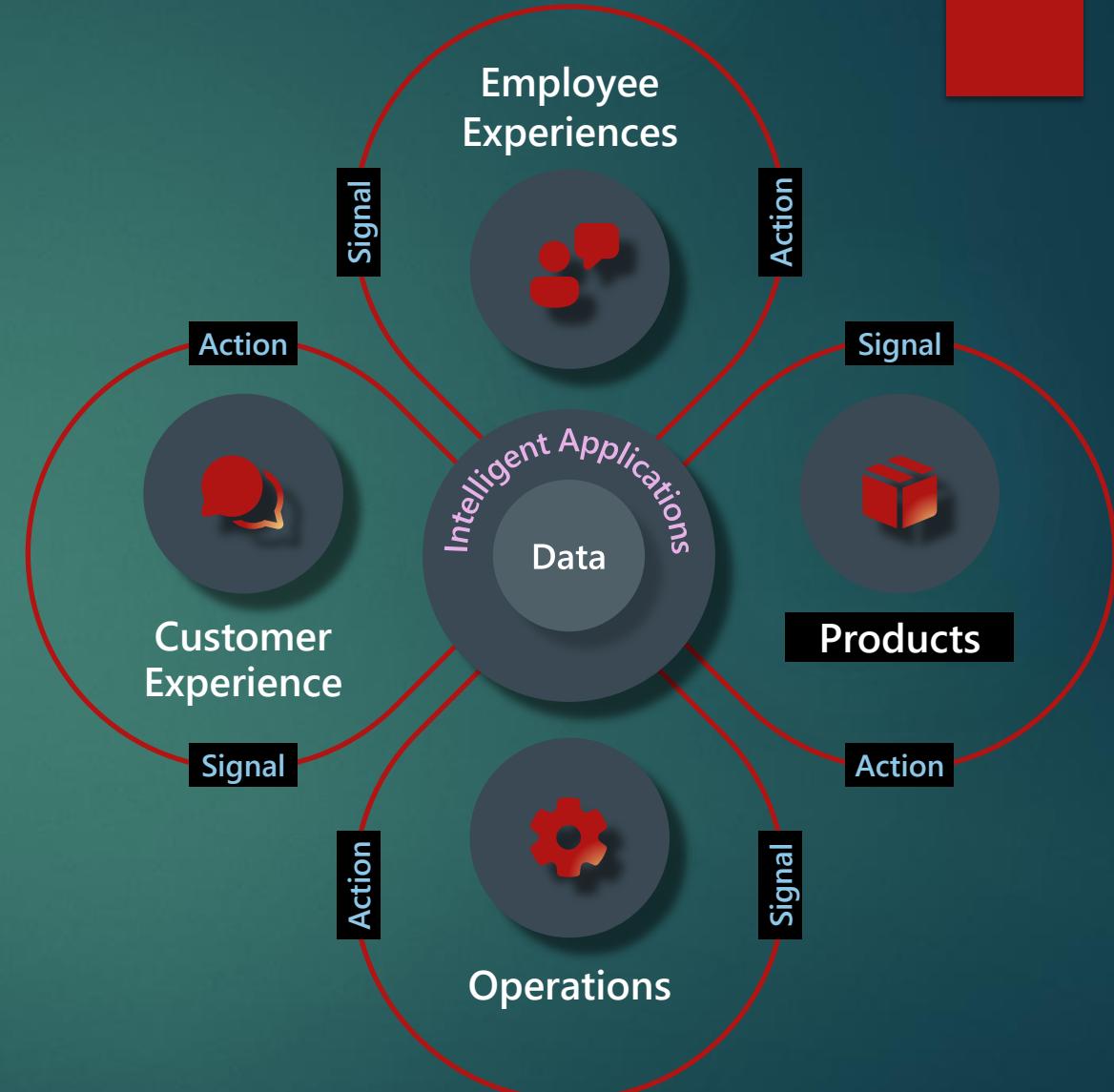


# Growing importance of Data in the era of AI

# Data is the oxygen of digital transformation

"A new kind of company — we call them insights driven businesses — has formed. They are growing at an average of more than 30% annually"

Forrester Analytics Business Technographics Global Data & Analytics Survey



# Today's data value creation challenges



Siloed systems and data



Fragmented toolset, Platforms  
requiring advanced skills



Costly integration and  
ongoing maintenance



Proprietary Platform, Vendor Lock in



Infrastructure not scalable  
as business grows

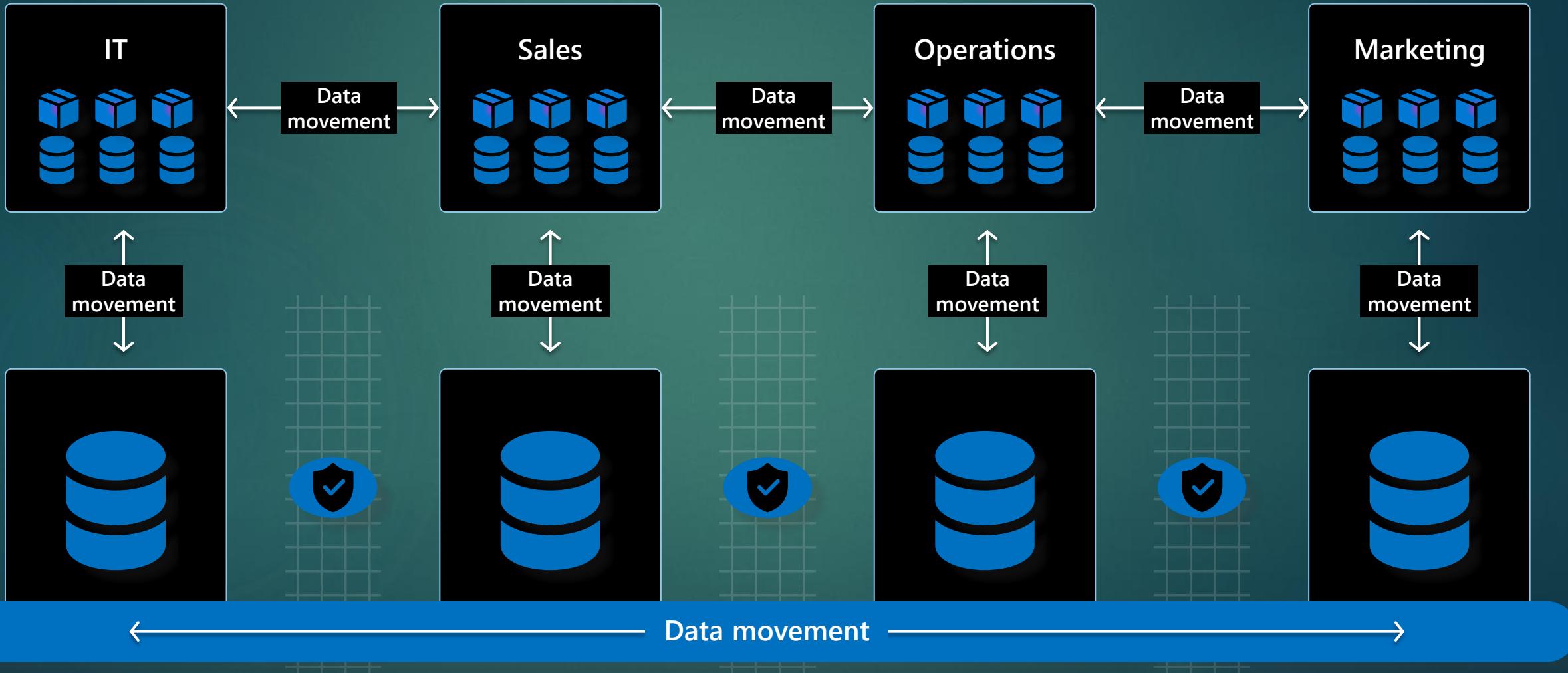


Gaining BI adoption to  
streamline data sharing, collaboration



Data and security risks,  
Governance & Compliance

Starting line: Multiple siloed lakes with lots of  
duplication





How many are using Microsoft Fabric?

# Microsoft Fabric

The unified data platform for AI transformation



Data  
Factory



Analytics



Databases



Real-Time  
Intelligence



Power BI



AI



OneLake



Governance

Fabric Platform

# Microsoft Fabric

The unified data platform for AI transformation



AI-powered  
data platform

Complete data platform

Unified, SaaS-ified

Secured and governed

Open and AI-ready  
data lake

OneLake

Multi-cloud

Open at every tier

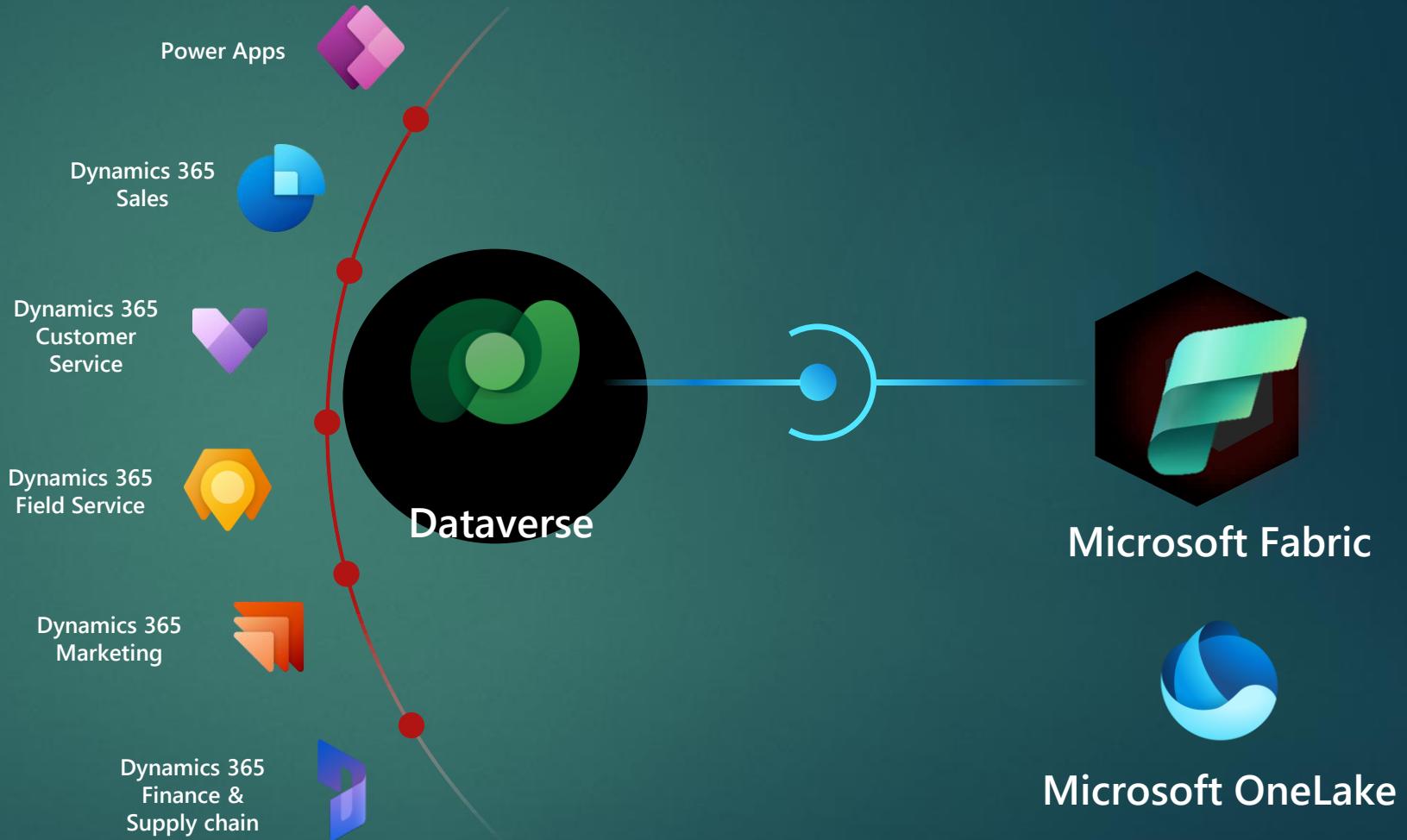
AI-enabled  
business users

For every business user

Built into Microsoft 365

M365 Copilot Integrated

# Bringing your data to Fabric



The value in  
your ERP data



# Why OneLake?

I already have <insert storage service here> !

With OneLake...

No more migration or adoption of new, centralized processes for storing and consuming data.

Build on the data you already access need and are empowered to use your preferred technologies.

Use one unified, logical lake that meets you where you are.

# OneLake data is available everywhere



# OneLake: A fully managed SaaS data lake

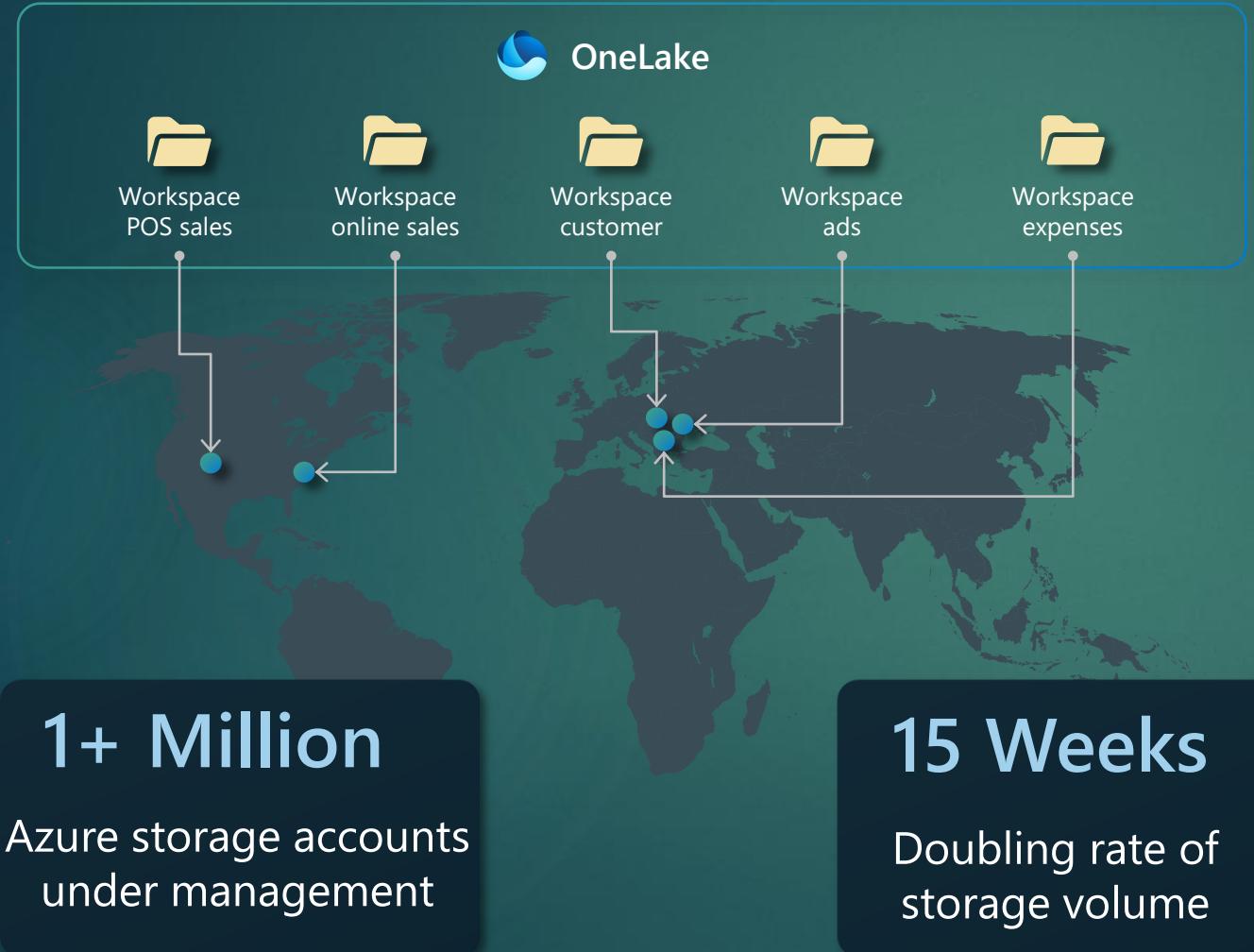
## Distributed ownership and isolation boundary with workspaces



- Provisioned automatically with the tenant.
- All data is ultimately under the control of a tenant admin, making it easy to adhere to standardized policies and governance.
- OneLake can scale to the needs of any organization by enabling distributed ownership via workspaces.

# OneLake which logically spans the world

No infrastructure to manage



When you store data in OneLake, the underlying physical storage is virtualized away.

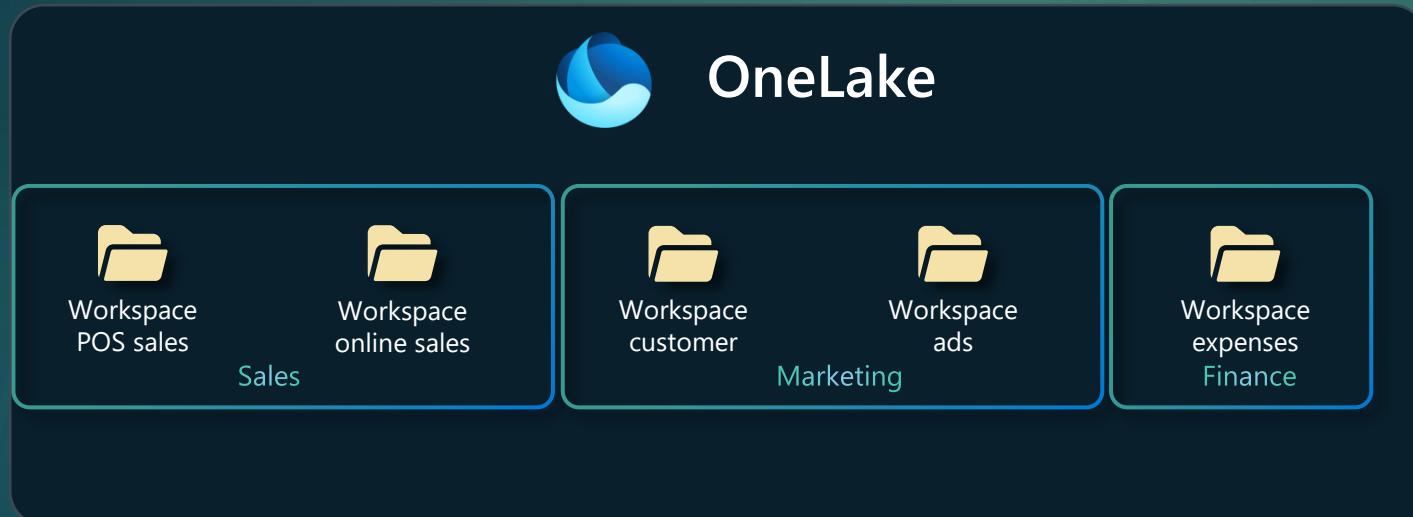
Data can reside in different regions without managing different storage resources.

Meet demand for any scale with dedicated storage resources for each workspace (capacity, throughput and IOPS).

All storage is zone redundant by default with an option for Geo redundancy.

# Everyone can work together in OneLake

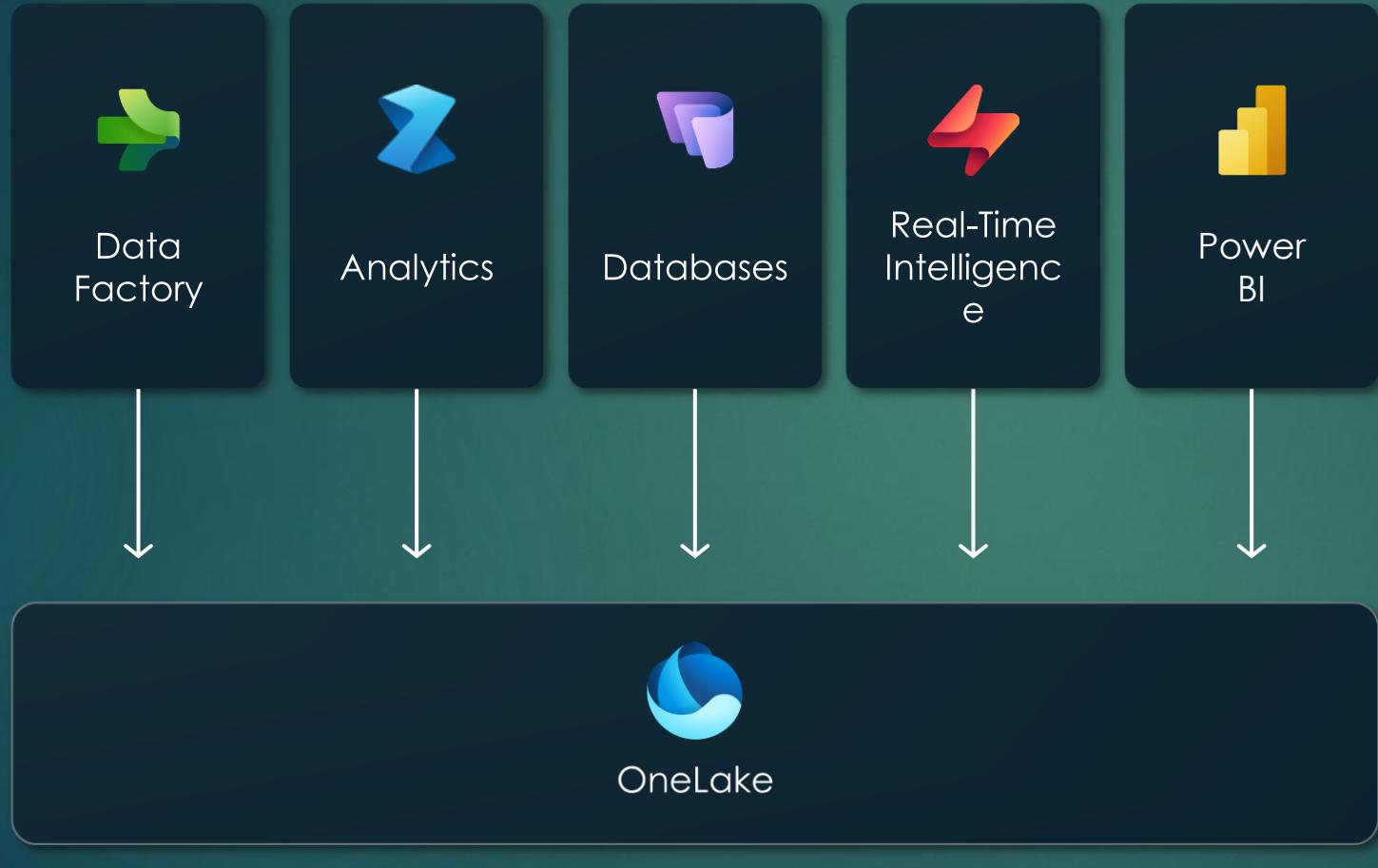
Distributed ownership and isolation boundary with workspaces



Within an organization:

- Some teams will be product developers that might use multiple workspaces and consume a lot of compute.
- Other teams might primarily blend data in a workspace for downstream consumption.

# OneLake is the OneDrive for data

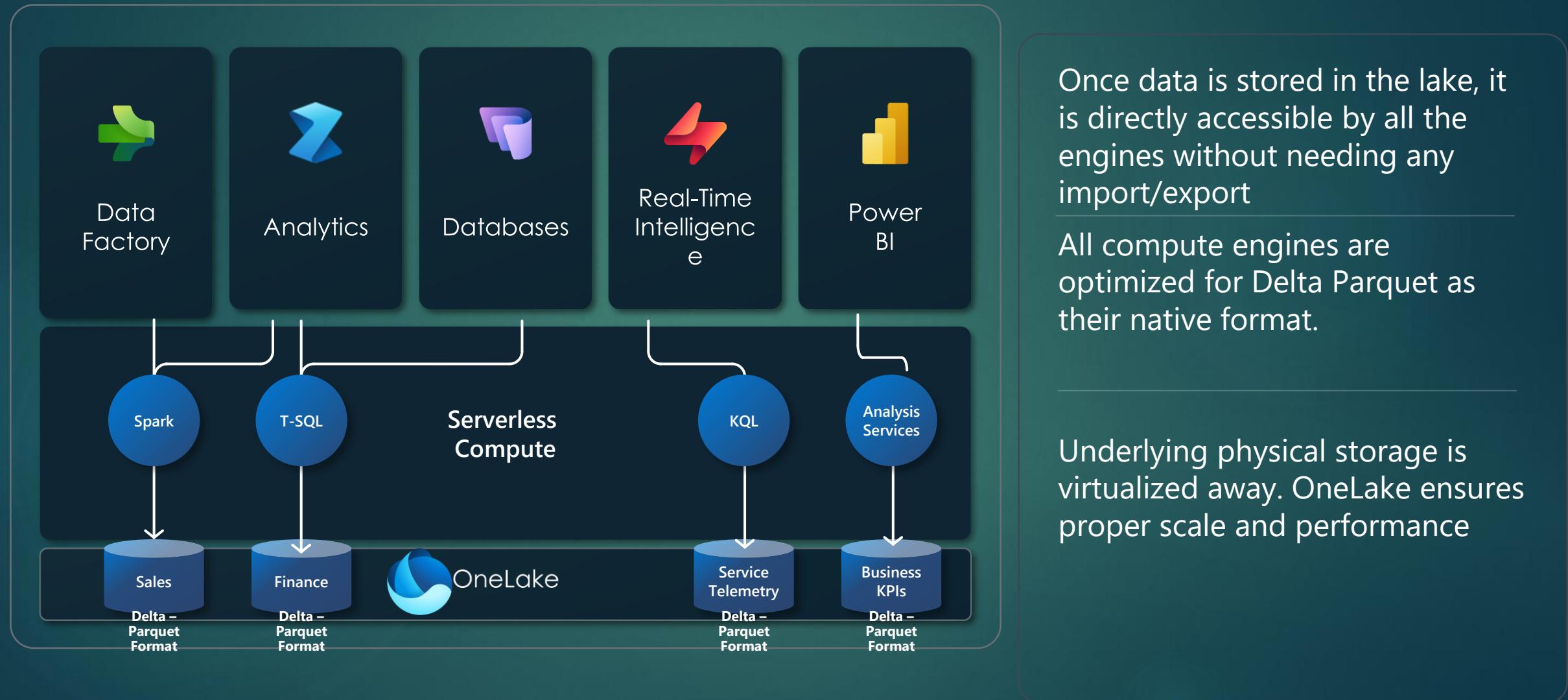


Single SaaS lake for  
the whole organization

Single open format used by  
all workloads to store and  
query data

Automatically indexed for  
discovery, MIP labels, lineage,  
PII scans, sharing, governance  
and compliance

# OneLake is the OneDrive for data



# Bringing data together with Shortcuts

No data movement or duplication



A shortcut is a symbolic link which points from one data location to another.

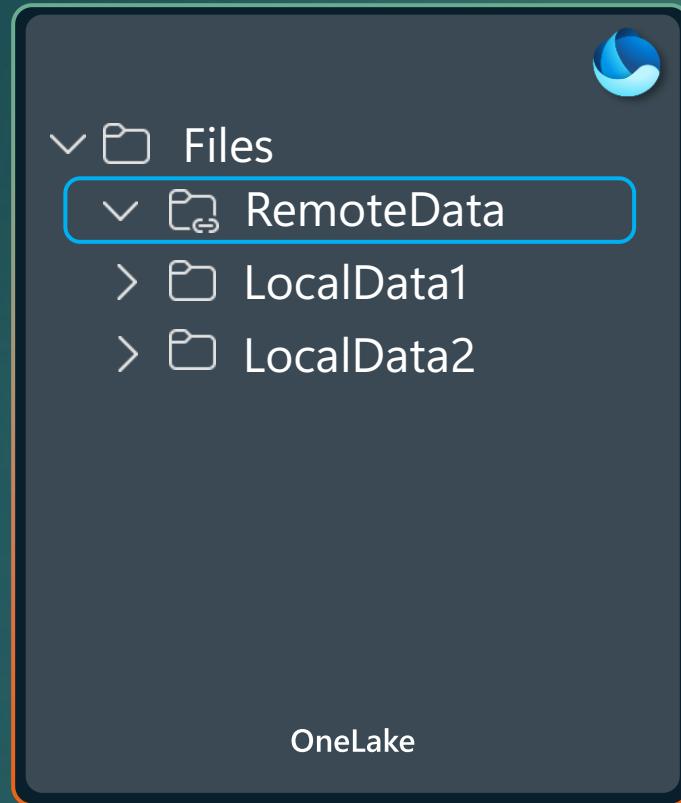
Create a shortcut to make data from a warehouse part of your lakehouse.

Create a shortcut within Fabric to consolidate data across items or workspaces without changing the ownership of the data. Data can be reused multiple times without data duplication.

Now different business team in the company can leverage data produced by domain owners or perhaps, centralized IT teams.

# Bringing data together with Shortcuts

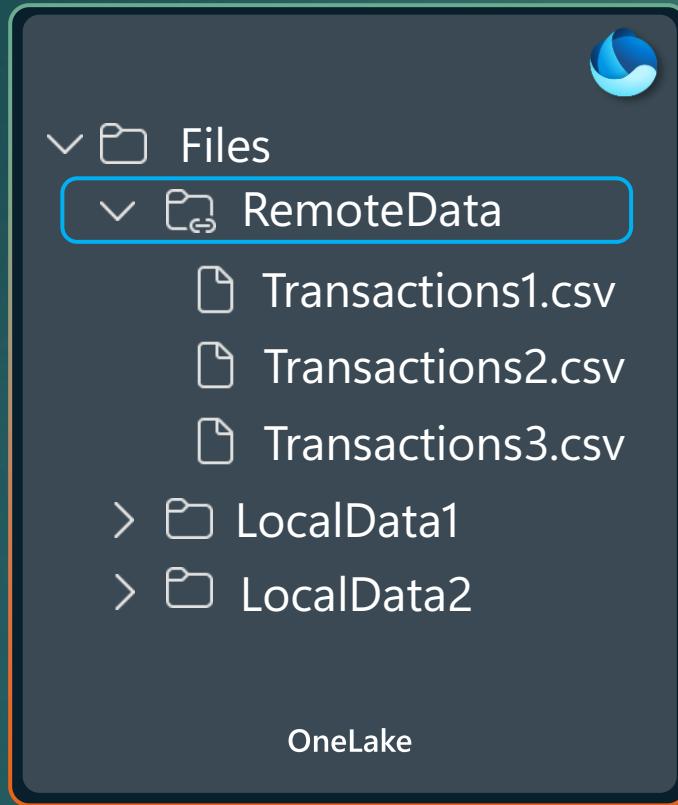
No data movement or duplication  
Shortcut Location



# Bringing data together with Shortcuts

No data movement or duplication

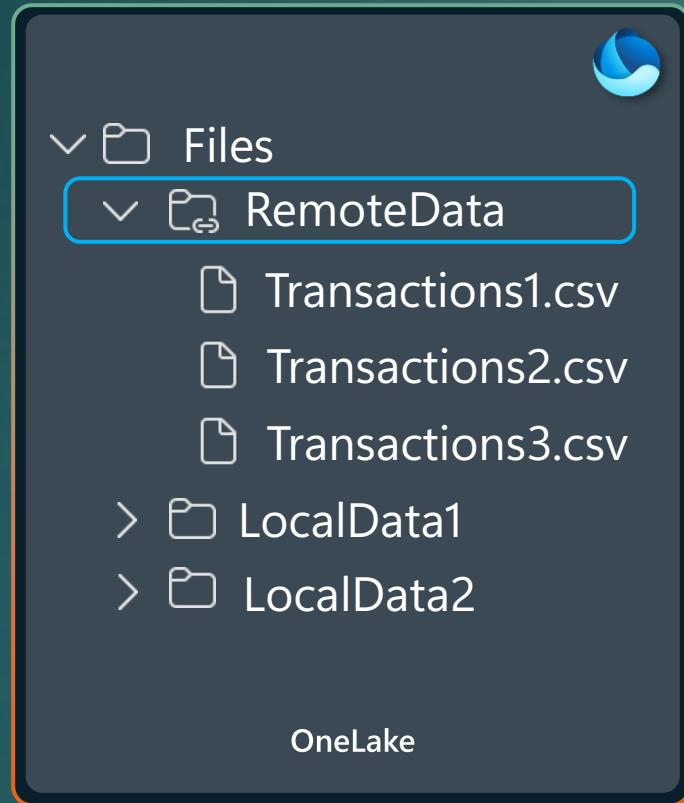
Shortcut Location



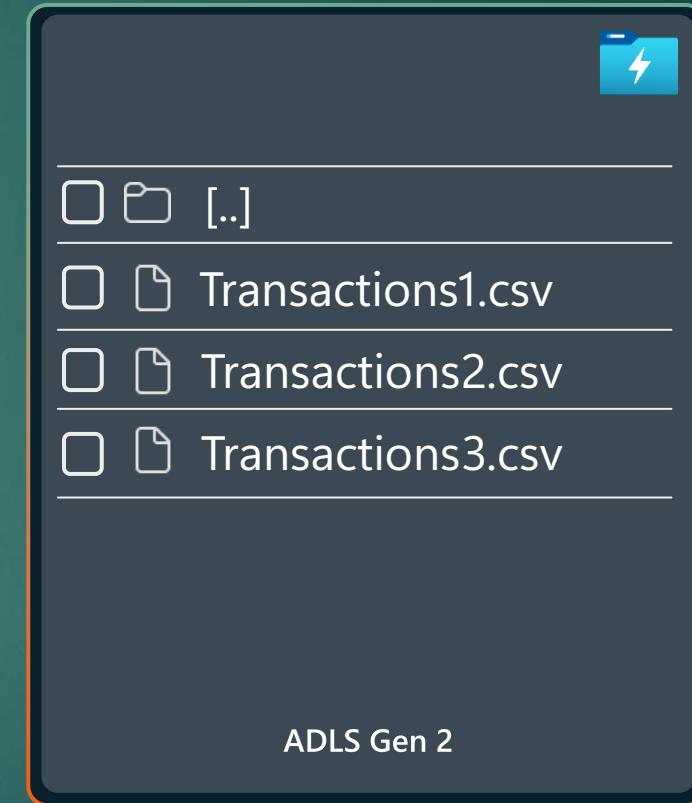
# Bringing data together with Shortcuts

No data movement or duplication

Shortcut Location

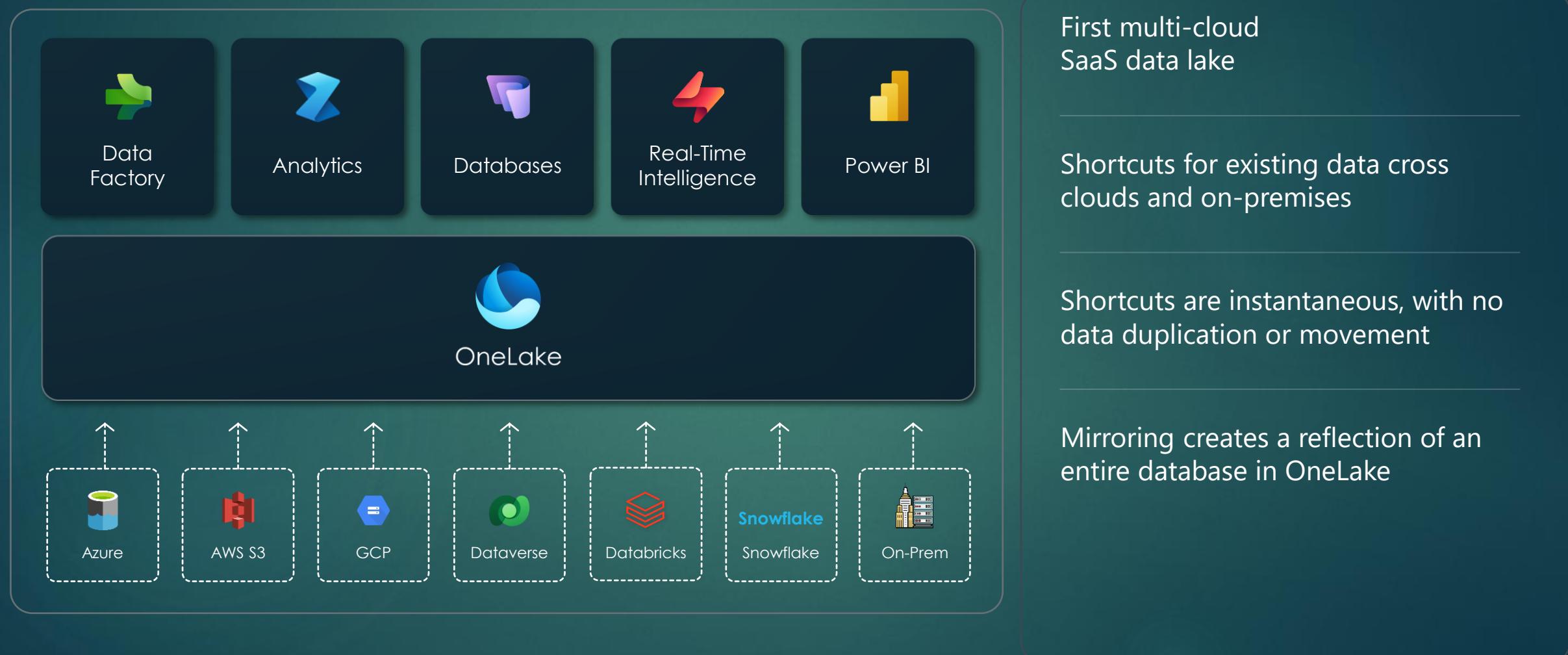


Shortcut Location

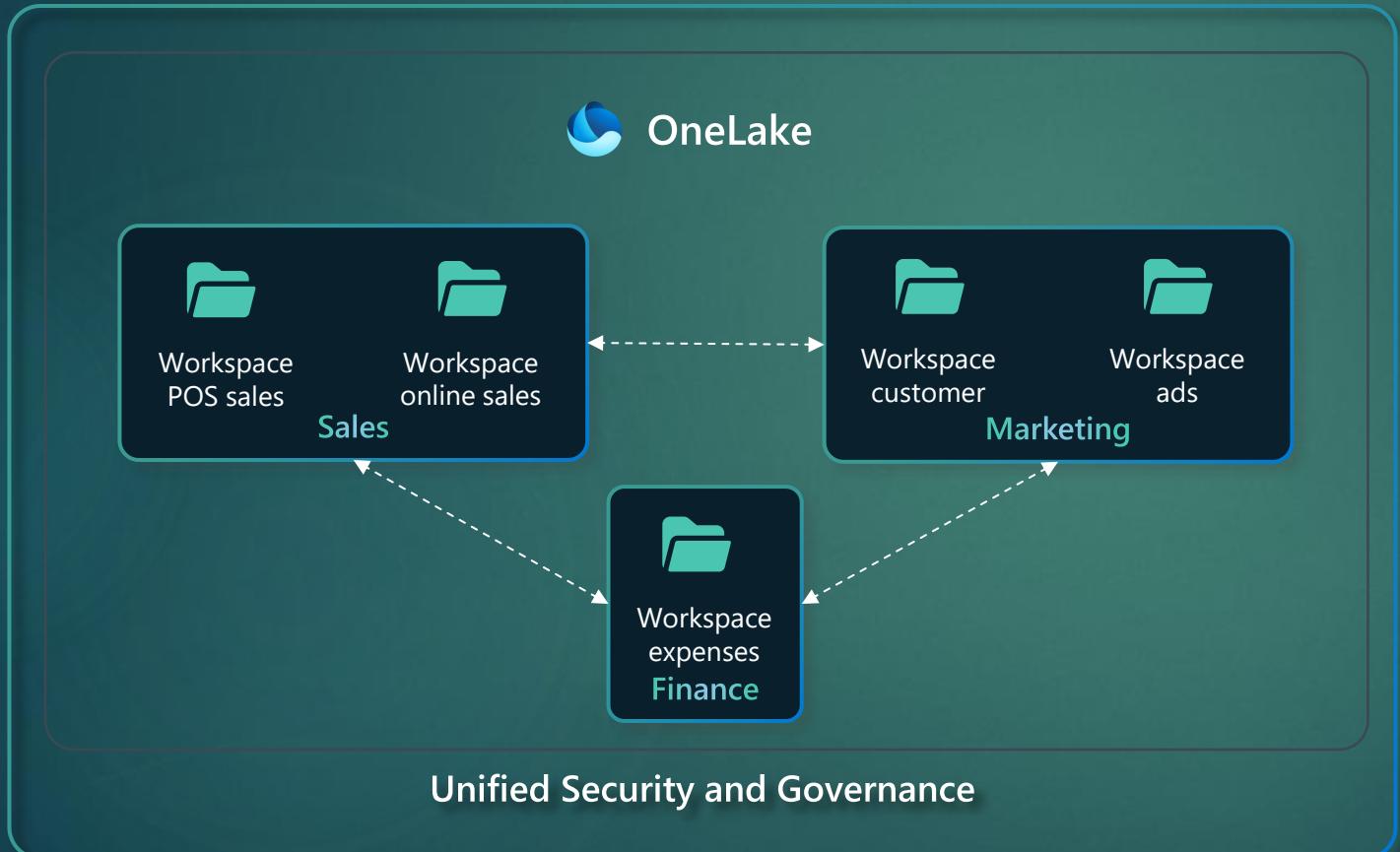


# Unifying data in OneLake

## Cross-cloud shortcuts & mirroring



# OneLake gives a true data mesh as a service



Group multiple workspaces into a single business domain and provide a governance boundary between the tenant admin and the workspace.

Avoid data swamps by endorsing certain data as certified or promoted, thus encouraging reuse.

OneLake Shortcuts provide the links between domains without data movement or duplication and without changing the ownership of the data.

# Building out a data mesh with capacities

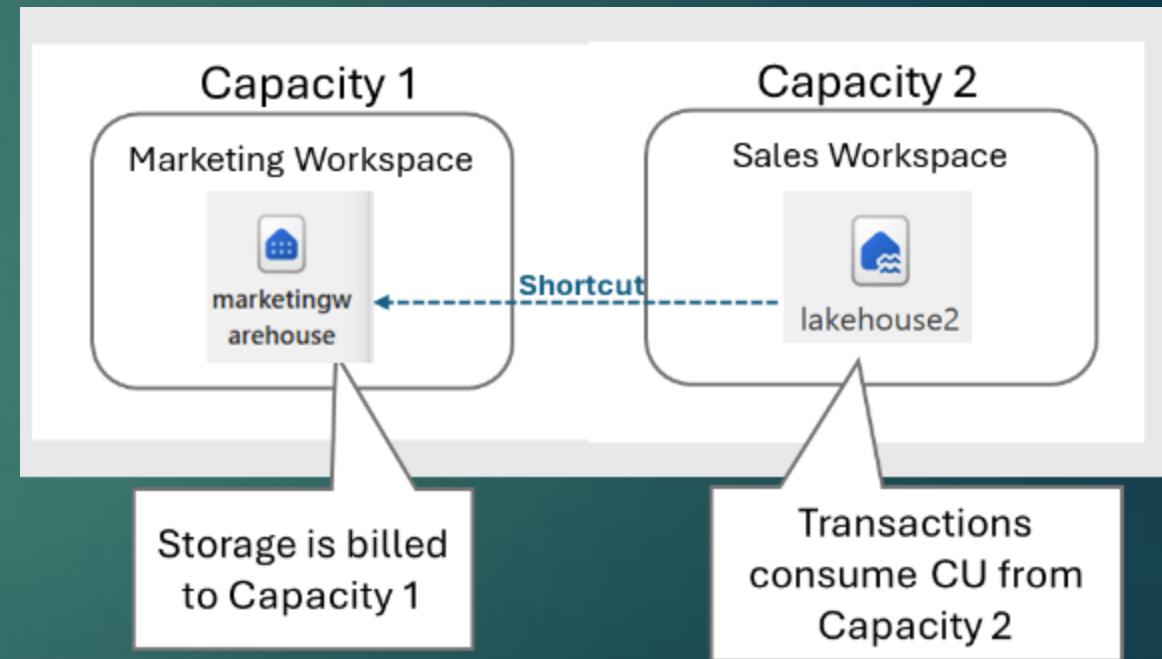
OneLake transactions (e.g., reads and writes) consume Fabric capacity.

Consumers pay with cross-capacity shortcuts

For example:

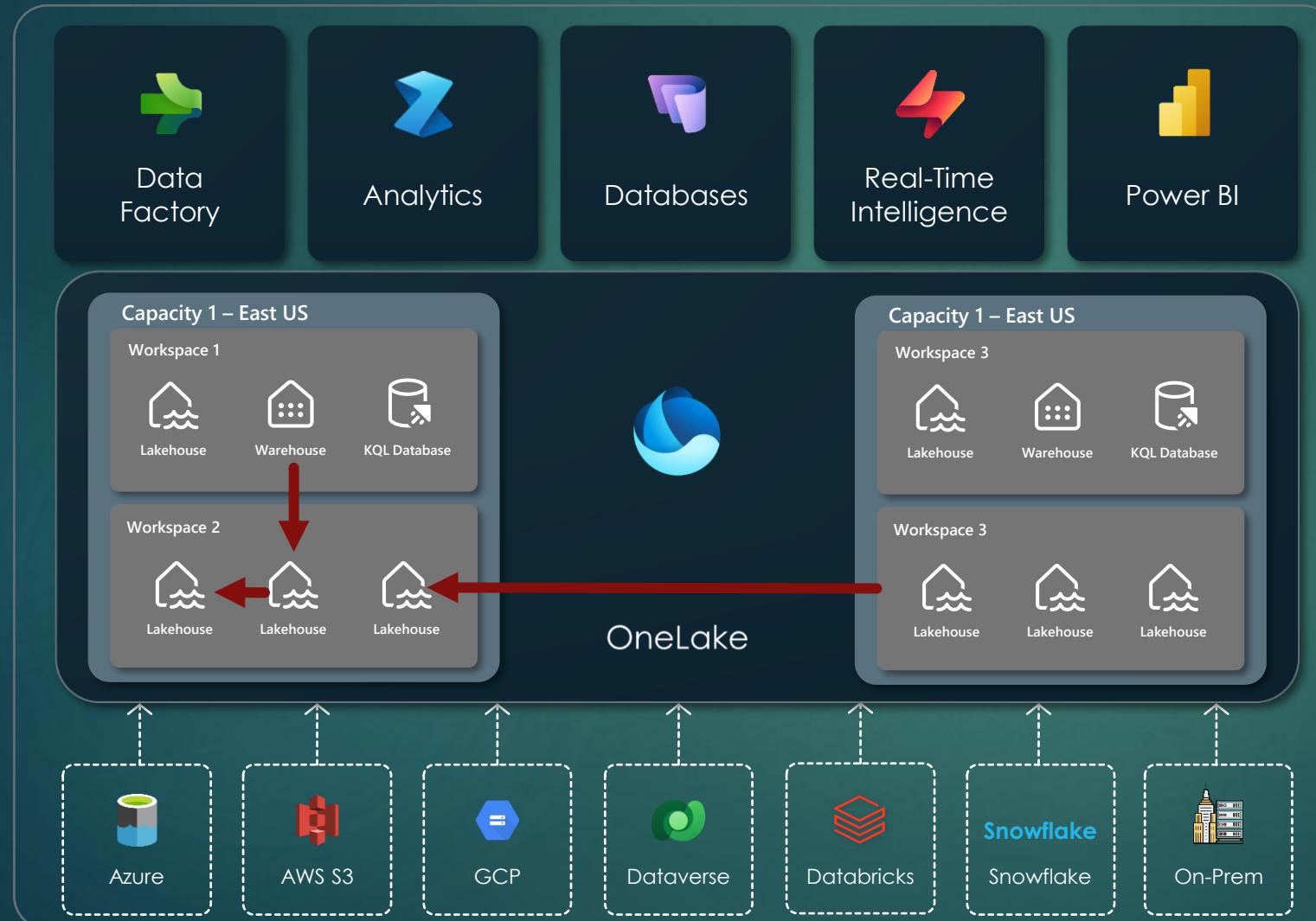
The marketing team can a capacity that powers data transformations and meets production SLAs.

Sales can shortcut to this data from a different capacity and not affect the marketing capacity, it can even be paused!



# Unifying data in OneLake

## Cross-cloud shortcuts & mirroring



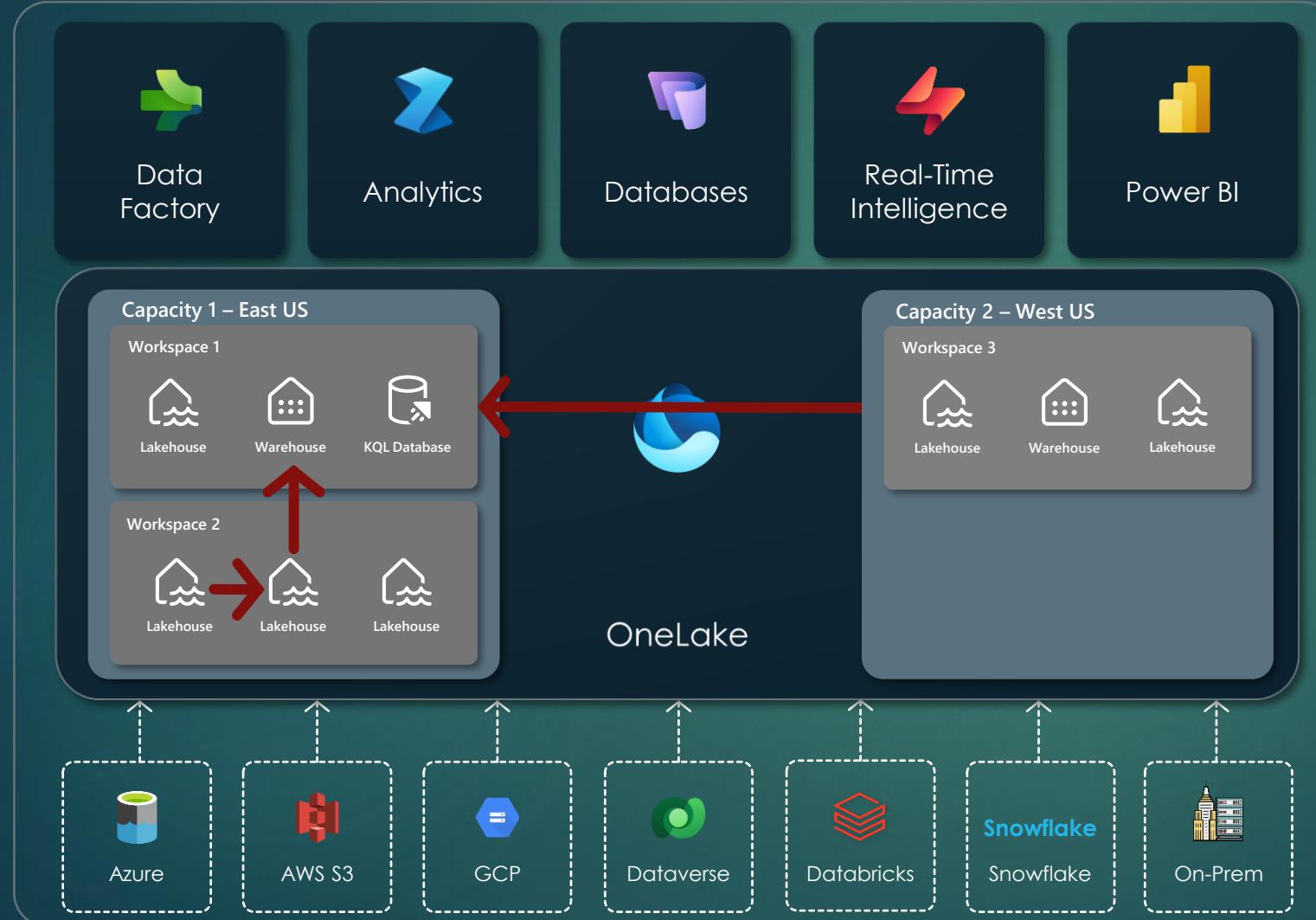
Sharing data in OneLake is as easy as sharing files in OneDrive, removing the needs for data duplication

With shortcuts, data throughout OneLake can be composed together without any data movement

With support for industry standard APIs, OneLake data can be directly accessed by any application or service

# Unifying data in OneLake

## Cross-cloud shortcuts & mirroring



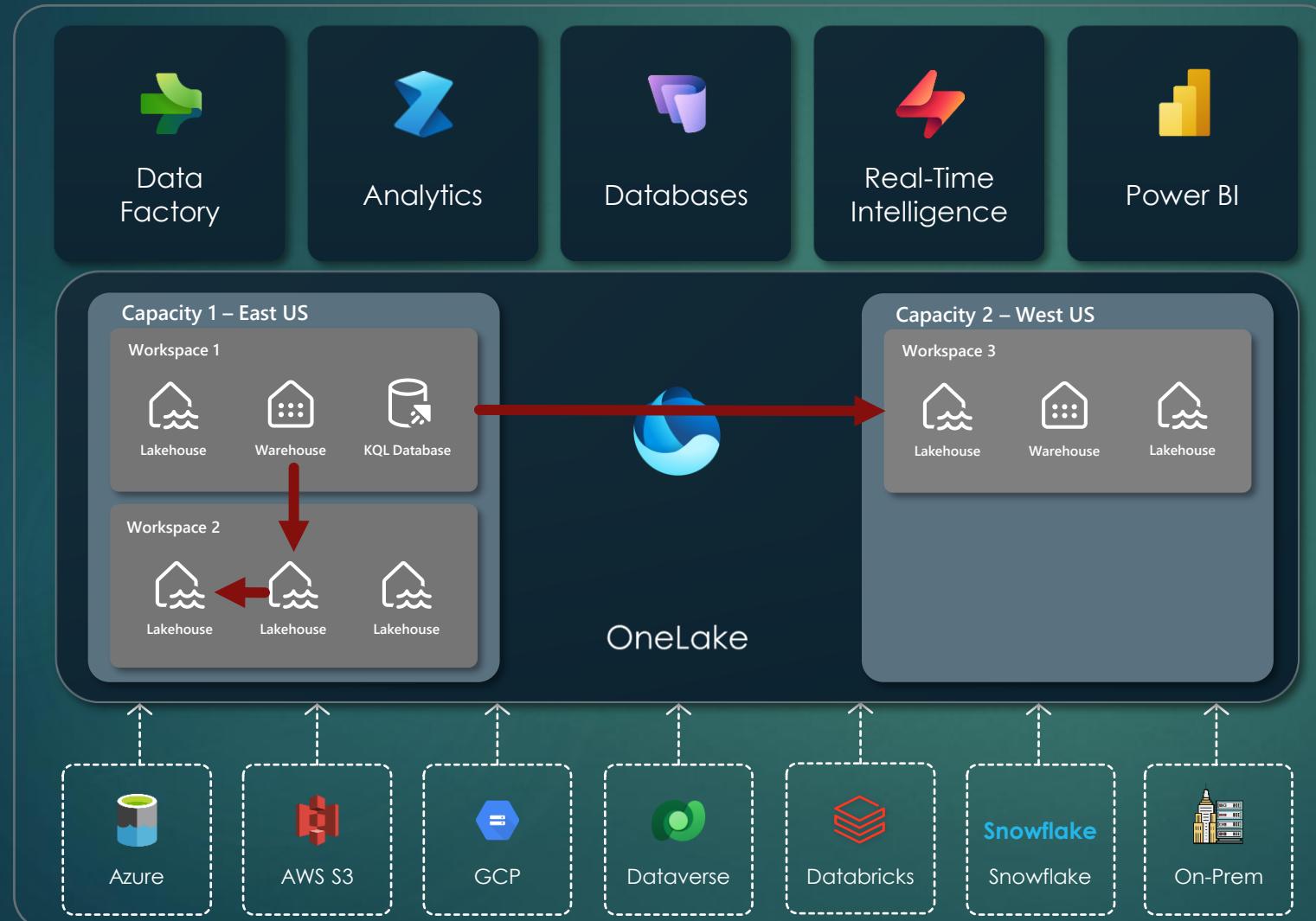
Sharing data in OneLake is as easy as sharing files in OneDrive, removing the needs for data duplication

With shortcuts, data throughout OneLake can be composed together without any data movement

With support for industry standard APIs, OneLake data can be directly accessed by any application or service

# Unifying data in OneLake

## Cross-cloud shortcuts & mirroring



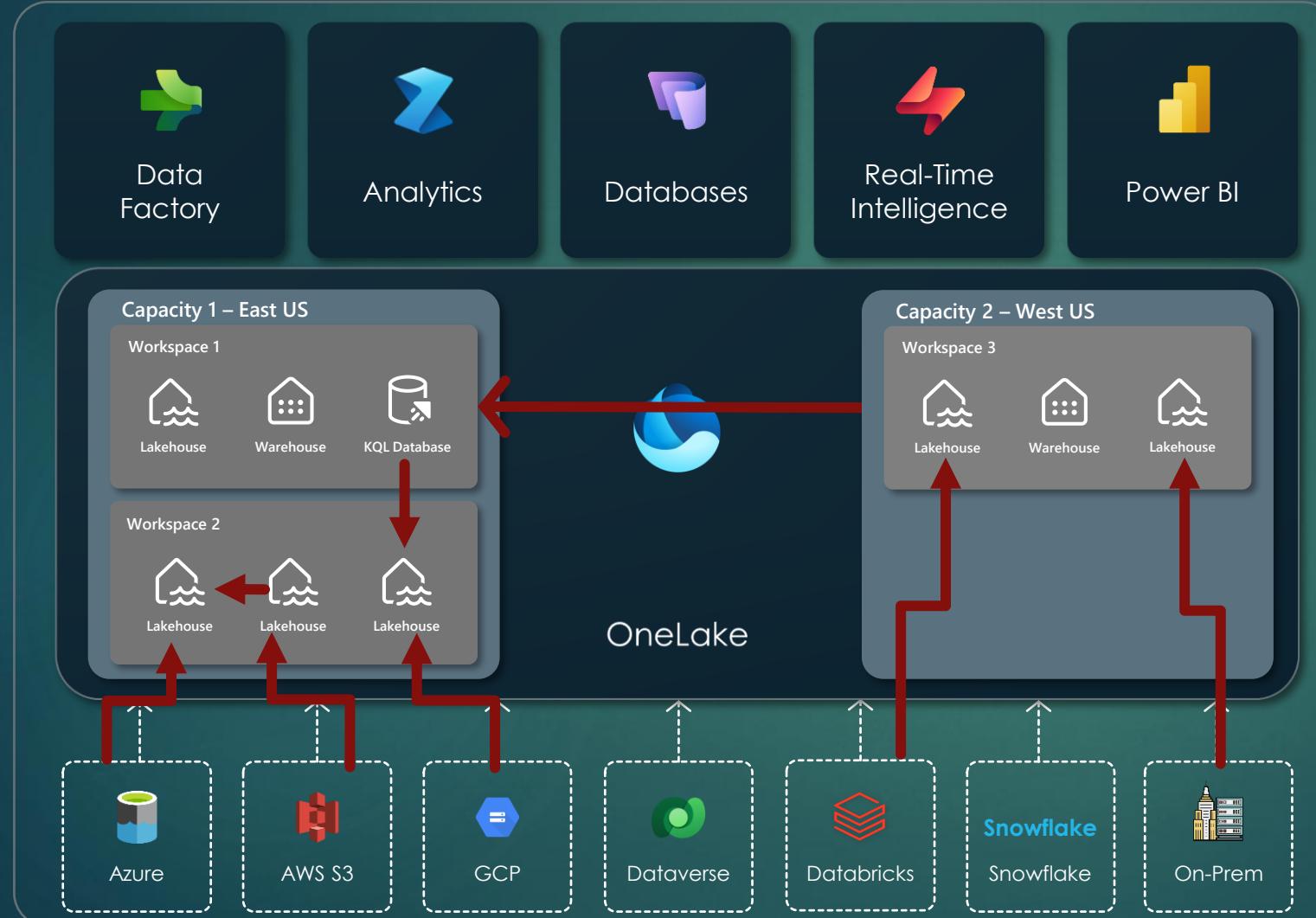
Sharing data in OneLake is as easy as sharing files in OneDrive, removing the needs for data duplication

With shortcuts, data throughout OneLake can be composed together without any data movement

With support for industry standard APIs, OneLake data can be directly accessed by any application or service

# Unifying data in OneLake

## Cross-cloud shortcuts & mirroring



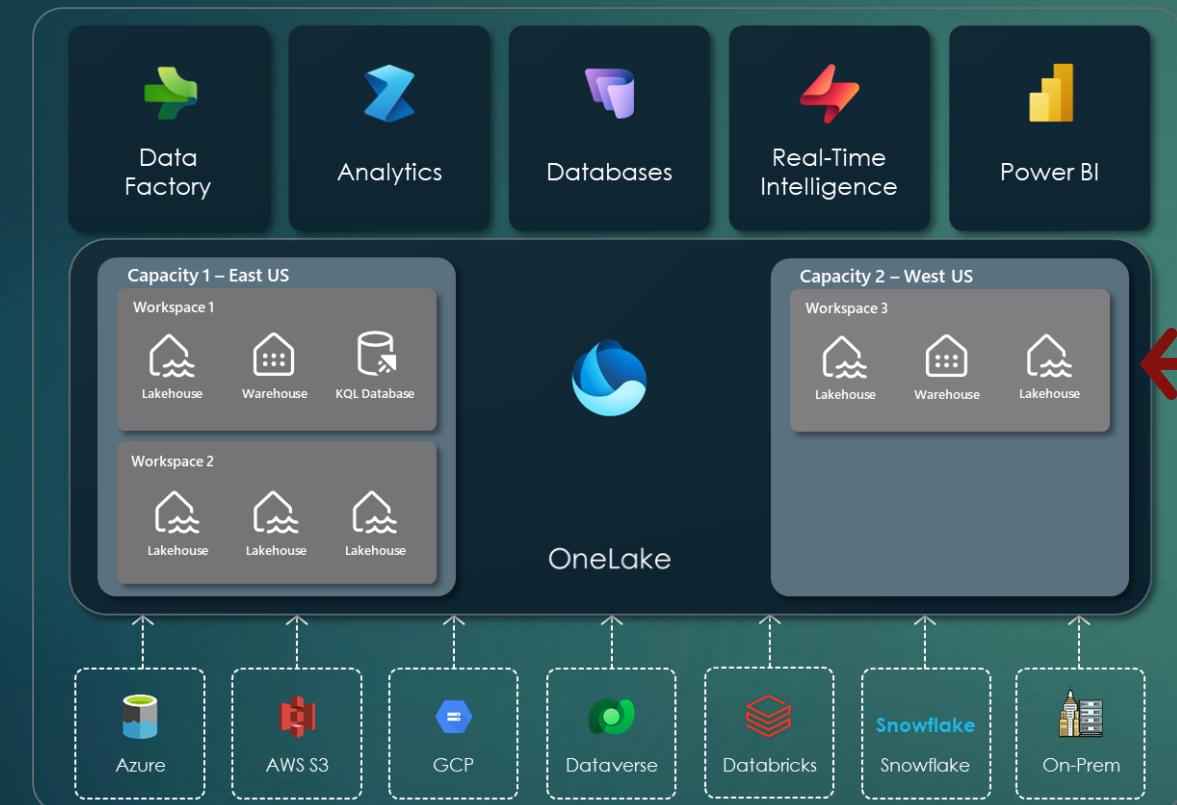
Sharing data in OneLake is as easy as sharing files in OneDrive, removing the needs for data duplication

With shortcuts, data throughout OneLake can be composed together without any data movement

With support for industry standard APIs, OneLake data can be directly accessed by any application or service

# Unifying data in OneLake

## External Data Sharing



Tenant Customer A

Tenant Customer B

# Unifying data in OneLake

## External Data Sharing

### ▫ External data sharing (preview)

*Enabled for the entire organization*

Users can share a read-only link to data stored in OneLake with collaborators outside your organization. When you grant them permission to do so, users can share a link to data in lakehouses and additional Fabric items. Collaborators who receive the link can view, build on, and share the data both within and beyond their own Fabric tenants, using their organization's licenses and capacities. [Learn More](#)



⚠️ Usage of shared data within a recipient tenant is not restricted by policies within your tenant, such as Conditional Access, Information Protection, Microsoft Cloud App Security, etc. Shared data may also be processed outside the region where it's stored.

### ▫ Users can accept external data shares (preview)

*Enabled for the entire organization*

Users can accept a read-only link to data from another organization's Fabric tenant. Users who accept an external share link can view, build on, and share the data, both inside and outside of your organization's tenant. For more information about the security limitations of this preview feature, view the feature documentation. [Learn More](#)



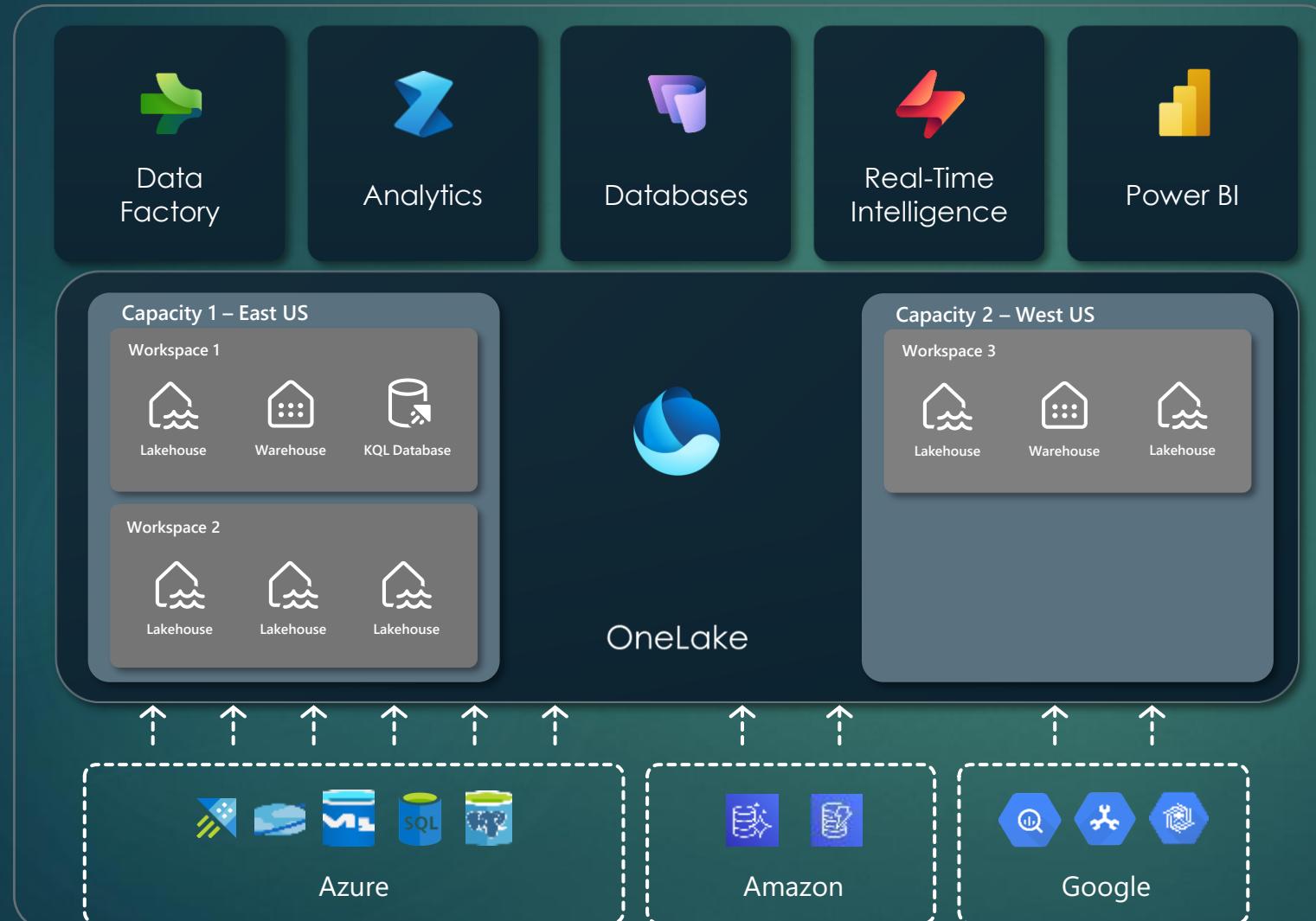
⚠️ Your organization is solely responsible for use of this feature. Potential risks include malicious intent on the part of the data provider, changes in the data at source, and legal liability arising from the accessing of regulated, proprietary, or sensitive data, responsibility for licensing enforcement, or the processing of data across regions.



Any questions so far?

# Unifying data in OneLake

## Cross-cloud shortcuts & mirroring



Fabric Mirroring allows adding existing databases and data warehouses to Fabric without any ETL with a new item type Mirrored database

Data is replicated into OneLake in Delta format and kept up-to-date in near-real-time via change data feed technology

Open mirroring framework allows mirroring of virtually every database type

The replica protects operational databases from analytical queries

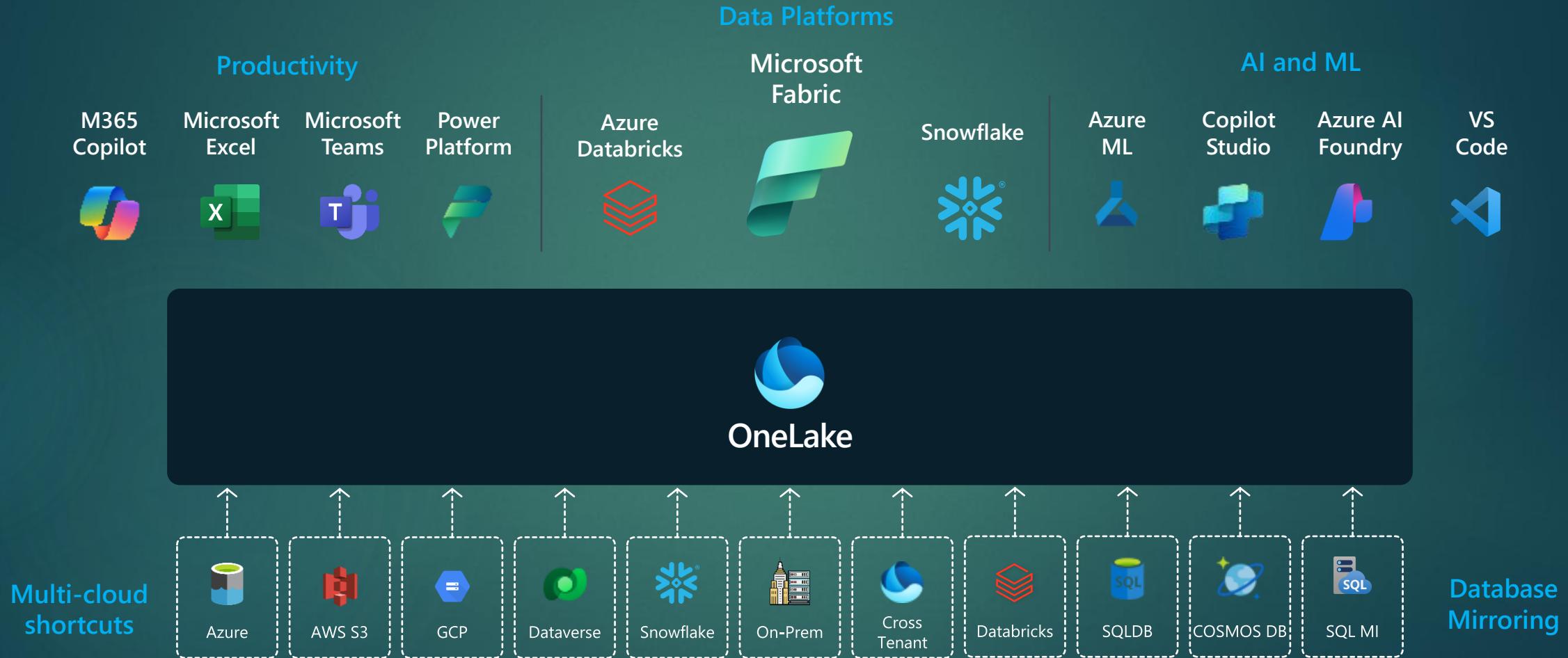
**Storage for mirroring is free!**

# Unify data in OneLake with zero ETL

## Shortcut & mirroring sources



# OneLake data is available everywhere



# **Announcing Shortcut Transformations: from files to Delta tables. Always in sync, no pipelines required**

## **Why use Shortcut transformations**

Shortcut transformations help users:

- Accelerate time-to-insight by skipping complex ETL pipelines.
- Reduce operational overhead and data movement.
- Enable self-service data transformation for business users.
- Maintain governance with built-in monitoring and lineage tracking.

customer\_support.csv

Environmental Movement Extended

Search

New item

Open file

Recent items

Manage OneDrive item access (preview)

View all items

File

Home

Get help

New item

Recent items

Open file

Manage OneDrive item access (preview)

View all items

File

Recent

customer\_support.csv

customer\_support.csv is located in the region East US 2 (USAP). Any data sources through this document will be processed in the same region.

Recent changes

Download to Office

Selected data

Auto-transform applied Transform from CSV to Delta

This was auto-transformed since it was in the wrong format. You can revert changes if you don't want to apply the transform.

Delimiter

CSV (comma-separated)

Custom

Use first row as headers

On

Next

Cancel

customer\_support.csv

customer\_support.csv



A photograph of a person's hands holding a smartphone, set against a background of a desk with a laptop, a tablet, and a coffee cup. A white callout box with a five-star rating icon is overlaid on the image, containing the word "DEMO".

DEMO



# OneLake Catalog

# Explore your data w/ OneLake catalog

## Complete catalog for data discovery, governance & insights in Fabric

The screenshot shows the Power BI OneLake catalog interface. On the left, there's a sidebar with various navigation options like Home, Copilot, OneLake, Apps, Metrics, Monitor, Real-Time, Workloads, Workspaces, My workspace, and Sales Performance. The main area is titled "OneLake catalog" and shows a "Domain: Sales". It has tabs for "Explore" (selected) and "Govern (preview)". There are filters for "All items by" (Data types: All), "Tags", and a search bar. A large preview window shows a colorful arrangement of circular objects. Below it, a modal window for the semantic model "SalesGrowthFY2024" is open, displaying its description: "Provides a structured view of sales performance for FY2024, integrating sales transactions, customer insights and more." It includes an "Open" button and tabs for Overview, Lineage, Monitor, and Permissions. The Overview tab shows details like Location (Sales > Sales Data), Refreshed (3/23/25, 1:58:42 PM), Owner (Mona Kane), and Tags (FY2024, LT KPIs, Sales Conf). At the bottom, there's a table titled "Tables" with columns Name, Type, and Description, showing entries for Customer\_Insights (Table), Age (Column), and Customer\_Buying\_Group (Column).

Unified view of org data estate to boost discovery & collaboration across item types

Easily explore & manage your data centrally incl' lineage, metadata, sensitivity, permissions

OLC everywhere: Integrated into MS365 – Teams & Excel (Insiders Fast), MS Copilot Studio, Azure Portal.

Coming next:

Curate your data with Copilot for auto-description & auto-summary.

Explore your data with Copilot to efficiently get insights from your data.

Roadmap: OneLake catalog Public APIs.

# Govern your data w/ OneLake catalog

## Complete catalog for data discovery, governance & insights in Fabric

The screenshot shows the Power BI OneLake catalog interface. At the top, there's a search bar and a trial status (19 days left). On the left, a sidebar lists various Power BI services: Home, Copilot, OneLake, Apps, Metrics, Monitor, Real-Time, Workloads, Workspaces, and My workspace. The main area displays 'Your governance status at a glance' with four cards: 'Summary' (6 Domains, 10 Workspaces, 146 Items), 'Items you own by type' (Semantic model: 28, Report: 15, Lakehouse: 10, SqlAnalyticsEndpoint: 10, Dashboard: 10, KustoDatabase: 5, KustoQueryWorkbench: 5, KustoDashboard: 5, Pipeline: 5, DataExploration: 5), 'Items you own by last refresh date' (0: 104, Last month: 42), and 'Items with description' (With: 39, Without: 68). Below this is a 'View more' button. A 'Recommended actions' section lists three items: 'Increase sensitivity label coverage', 'Increase the percentage of endorsed items', and 'Apply relevant tags to your items'. Under 'Top solutions', there are four cards: 'Monitor' (View and track the status of activities across all the Fabric workspaces you can access, with an 'Open' button), 'Data lineage and impact analysis' (View lineage relationships between content items and see how changes impact downstream, before you make them, with links to Lineage in Fabric—Microsoft Learn and Impact analysis—Microsoft Learn), 'Endorsement' (Encourage data discovery and reuse by endorsing quality content as Promoted, Certified, or Master data, with a link to Endorsement—Microsoft Learn), and 'Tags' (Apply relevant tags to your items so they're easier to discover and manage, with a link to Tags—Microsoft Learn).

Govern & gain **insights** on your data incl' your item inventory, refresh, item curation, MIP labels.

Get **Recommended actions**, with information about the problem and instructions how to fix it

Find all the **Tools , solutions and links** available for you in Fabric

Coming next :

Additional **Insights** [external shares, permissions,..]

From **Insights** into **immediate action**

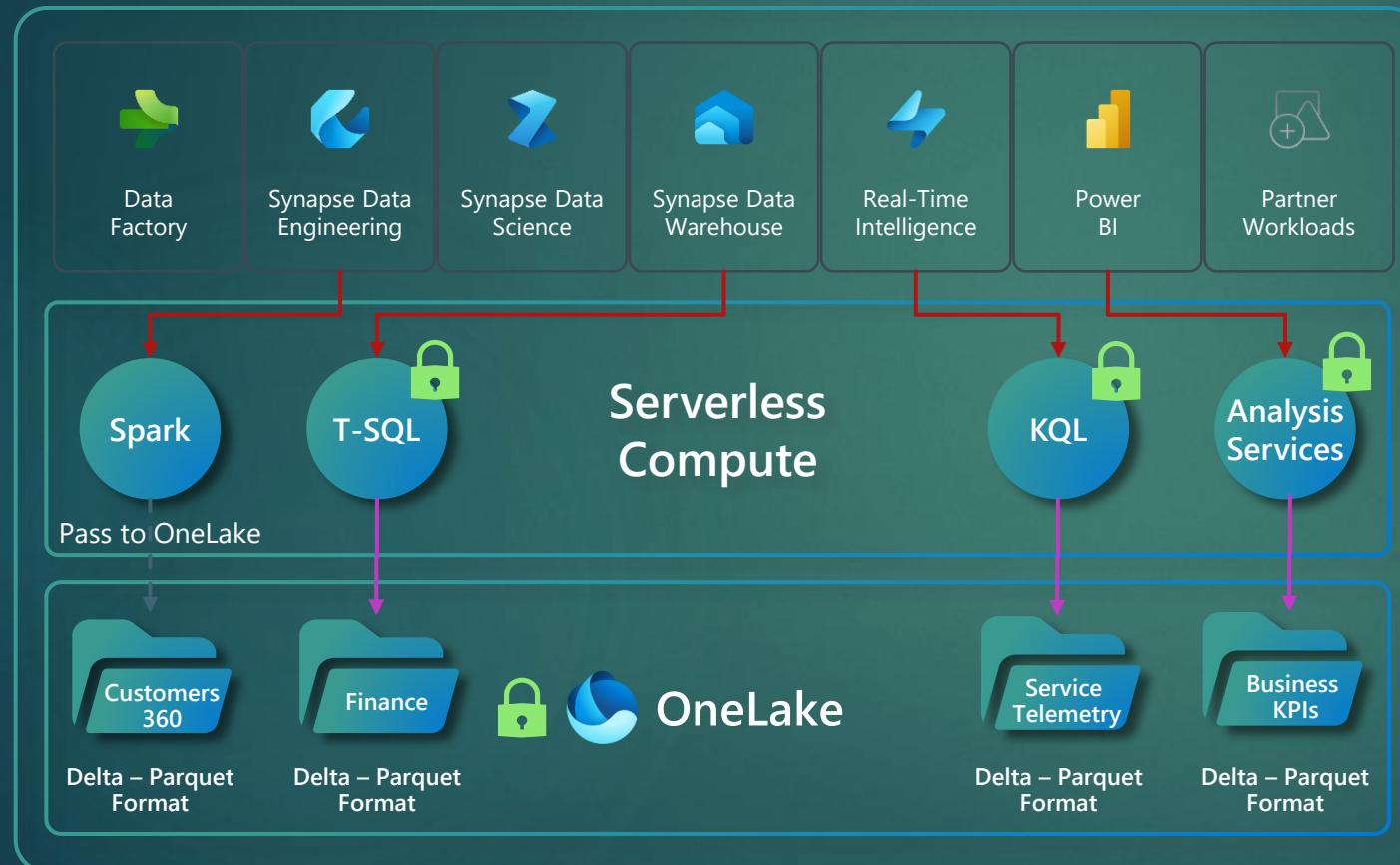
Roadmap: Extend Govern for **tenant Admins**



# OneLake security

# One Copy, many security definitions

Enabling fine-grained security at each engine



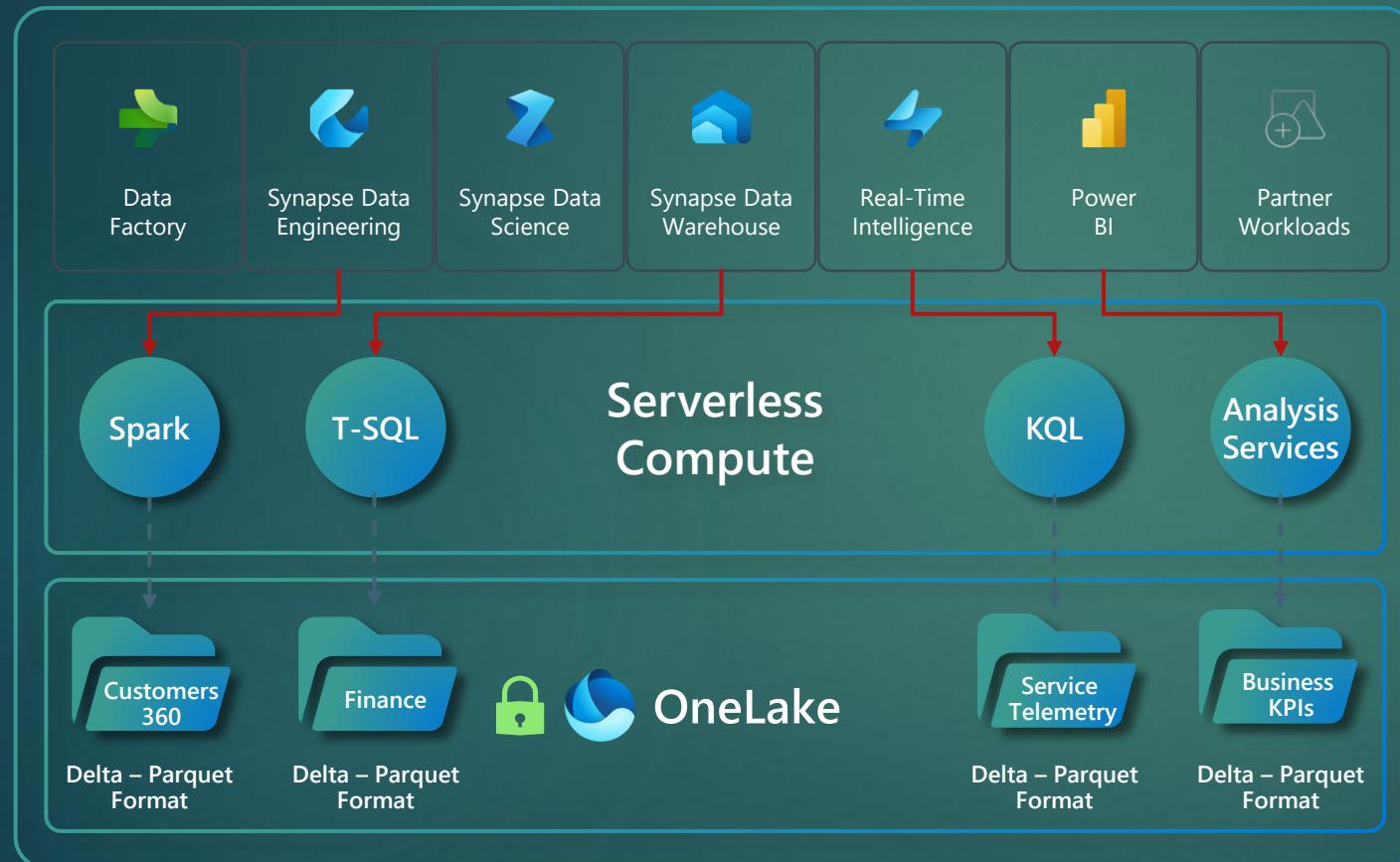
Each engine in Fabric provides its own robust set of security capabilities.

OneLake also brings data lake security, applicable to engines like Spark.

Each Fabric engine relies on one of two modes when reading data from OneLake.

# One Copy, One Security Definition

A single place to define and manage security



OneLake now has a complete set of fine-grained security features:

Table/folder security  
Column level security  
Row level security  
+ more coming soon...

Engines can then switch to using passthrough mode as the primary security model.

Security is now centrally defined and enforced consistently across all engines.

# Sign up for OneLake security preview

OneLake security - early access  
signup form



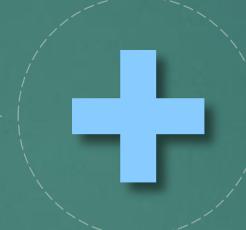
# Manage, govern, and secure data in era of AI

Go further by integrating Fabric with Microsoft Purview



**Microsoft  
Fabric**

Manage, govern, and secure your data locally in Microsoft Fabric with built-in capabilities



**Microsoft  
Purview**

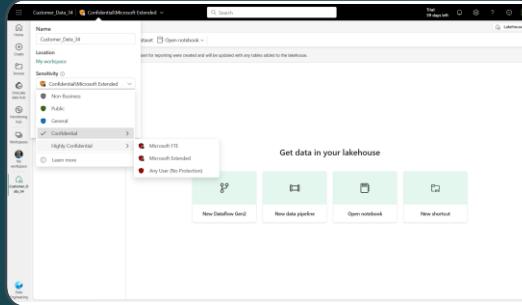
Employ advanced and global governance and security, including Microsoft Fabric

Local governance

Simplified administration

Global governance

# Purview Data Security & Compliance integration with Fabric

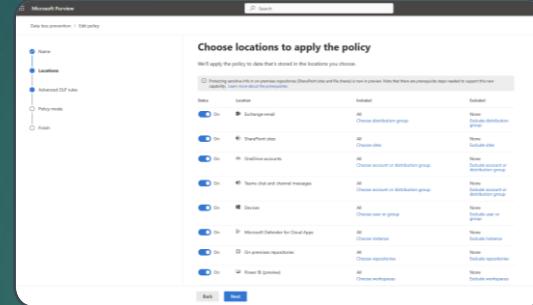


## Information protection sensitivity labels and protection policies

Classify sensitive Fabric data using the same **sensitivity labels** that are used in Microsoft 365—the label and protection travels with the data within Fabric, and enforced even when the data is exported to Office.

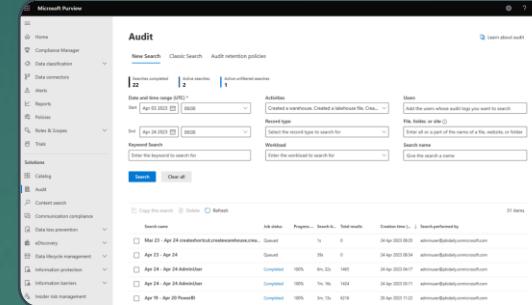


*Additional Microsoft Purview purchase required*



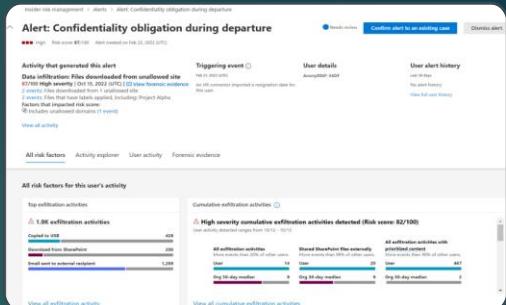
## Data loss prevention policies

Automatically detect sensitive data such as PII or SSN in Lakehouse or Semantic Models and trigger automatic risk remediation actions such as alerts or restrict access.



## Audit

Log user activities from Microsoft Fabric in MS Purview Audit to support security, forensic, and internal investigations.



## Insider Risk Management

Ingest audit logs from Fabric in addition to other millions of signals to identify potential malicious or inadvertent insider risk.





Feedback



Any questions left?

Let's connect



# Thank you

Erwin de Kreuk  
Principal Consultant  
Lead Data & AI InSpark

- @erwindekreuk.bsky.social
- linkedin.com/in/erwindekreuk
- erwindekreuk.com
- github.com/edkreuk
- <https://sessionize.com/erwin-de-kreuk/>
- Dutchfabricusergroup.com

