



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

Erwin de Kreuk

SQL Day 2024



# Erwin de Kreuk

Principal Consultant – Lead Data & AI  
InSpark



**Erwin de Kreuk**

Principal Consultant - Lead Data & AI  
InSpark

 @erwinedekreuk

 [linkedin.com/in/erwinedekreuk](https://www.linkedin.com/in/erwinedekreuk)

 [erwinedekreuk.com](http://erwinedekreuk.com)

 [github.com/edkreuk](https://github.com/edkreuk)

 <https://sessionize.com/erwin-de-kreuk/>



Let's  
connect



# We Are InSpark

We help organizations  
accelerating their digital  
transformation with impactful  
Microsoft solutions & expertise



# 16 edycja konferencji SQLDay

13-15 maja 2024, WROCŁAW + ONLINE



partner platynowy



partner złoty



partner srebrny



# Objectives of today

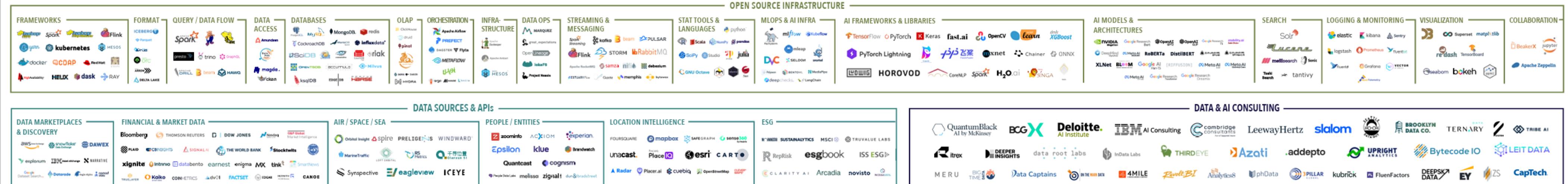
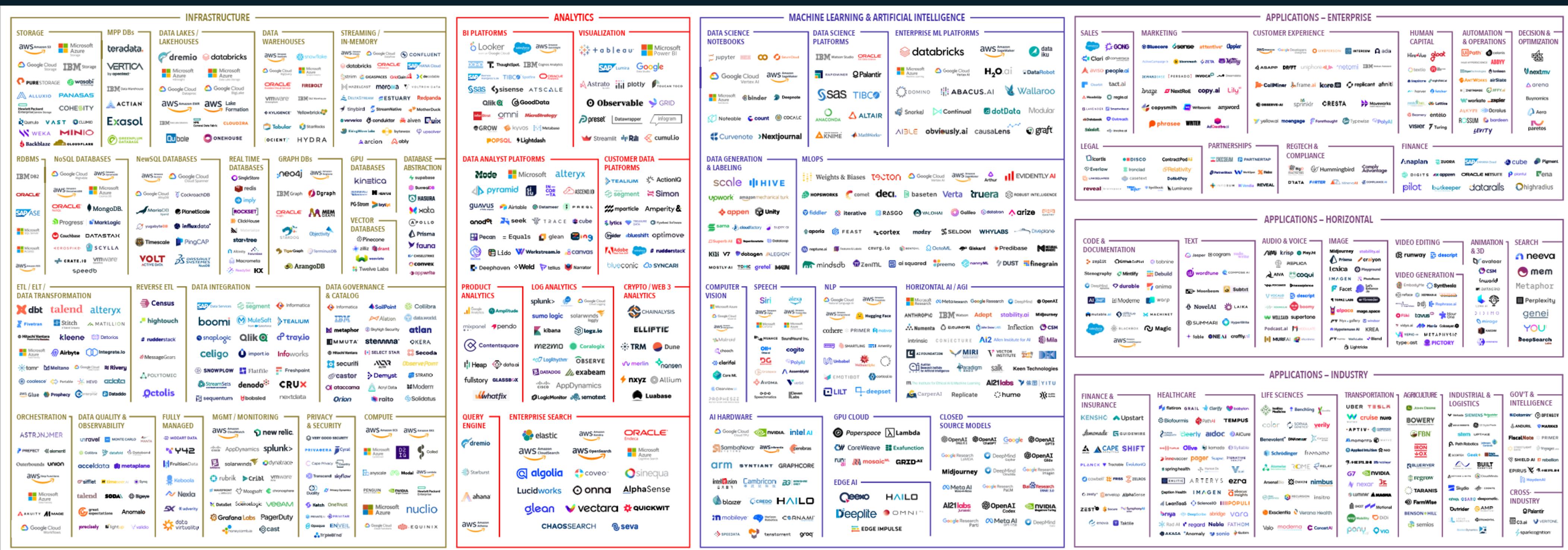
- Onelake
- Shortcuts
- External Data Sharing
- Recap



A world awash with data...

How do you translate data into  
competitive advantage?

# The 2023 ML, AI, and Data 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



How many are using Microsoft Fabric?

# Microsoft Fabric

**Complete Analytics Platform**

Unified product, experience, and architecture

---

Delivered as SaaS

# Microsoft Fabric

End-to-end analytics data fabric  
From the data lake to the business user

## Complete Analytics Platform

Everything, Best of Breed  
Unified  
Azure OpenAI Copilots

## Lake centric and open

OneLake  
One Copy  
Always Synced

## Empower Every Office User

Familiar and intuitive  
Built into Office  
Insight to action

## Pervasive security and governance

End to end visibility  
Always governed  
Secure by default

# Microsoft Fabric

End-to-end analytics data fabric  
From the data lake to the business user

Complete  
Analytics Platform

Everything, Best of Breed  
Unified  
Azure OpenAI Copilots

Lake centric  
and open

OneLake  
One Copy  
Always Synced

Empower Every  
Office User

Familiar and intuitive  
Built into Office  
Insight to action

security  
and governance

End to end visibility  
Always governed  
Secure by default



# Microsoft Fabric

## The unified data platform for the era of AI



Data  
Factory



Synapse Data  
Engineering



Synapse Data  
Science



Synapse Data  
Warehousing



Synapse Real  
Time Analytics



Power BI



Data  
Activator



AI



OneLake



Purview

Unified  
architecture

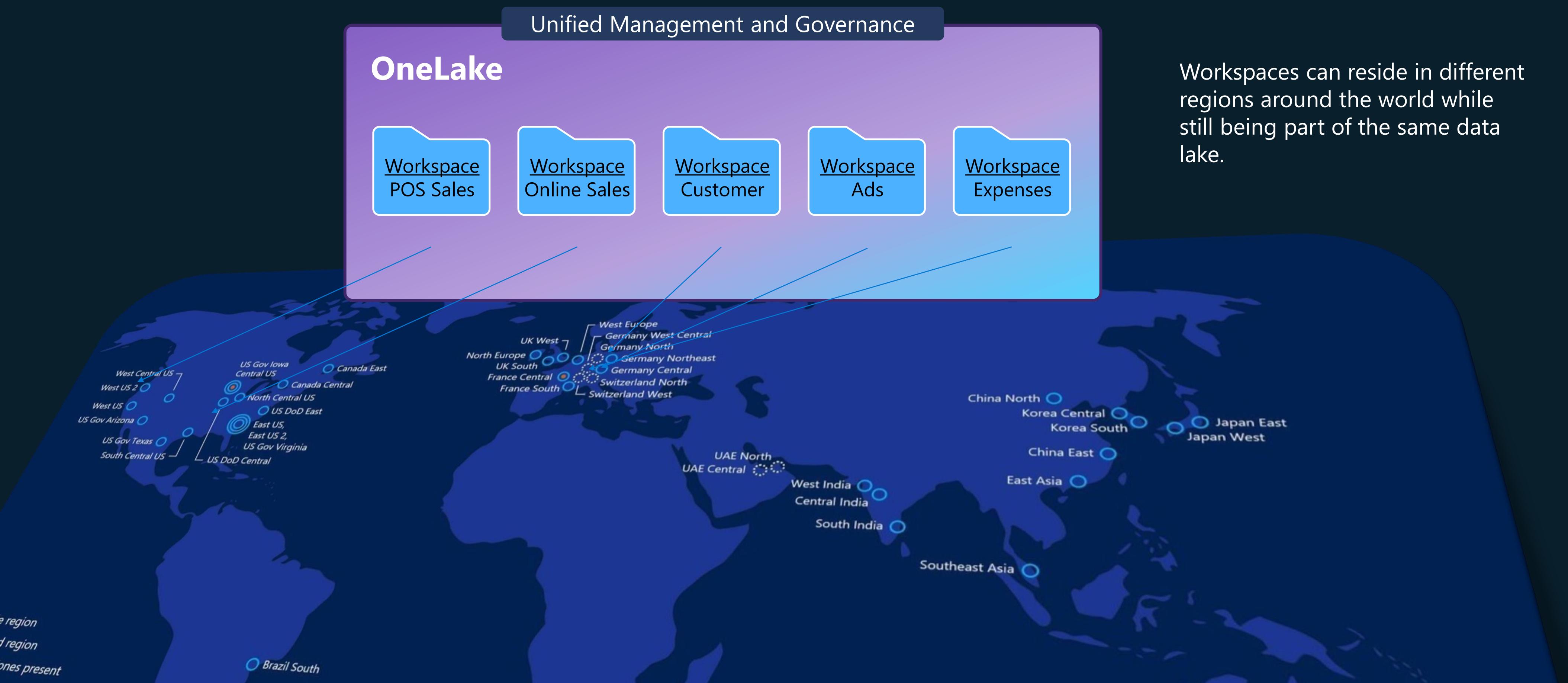
Unified  
experience

Unified  
governance

Unified  
business model

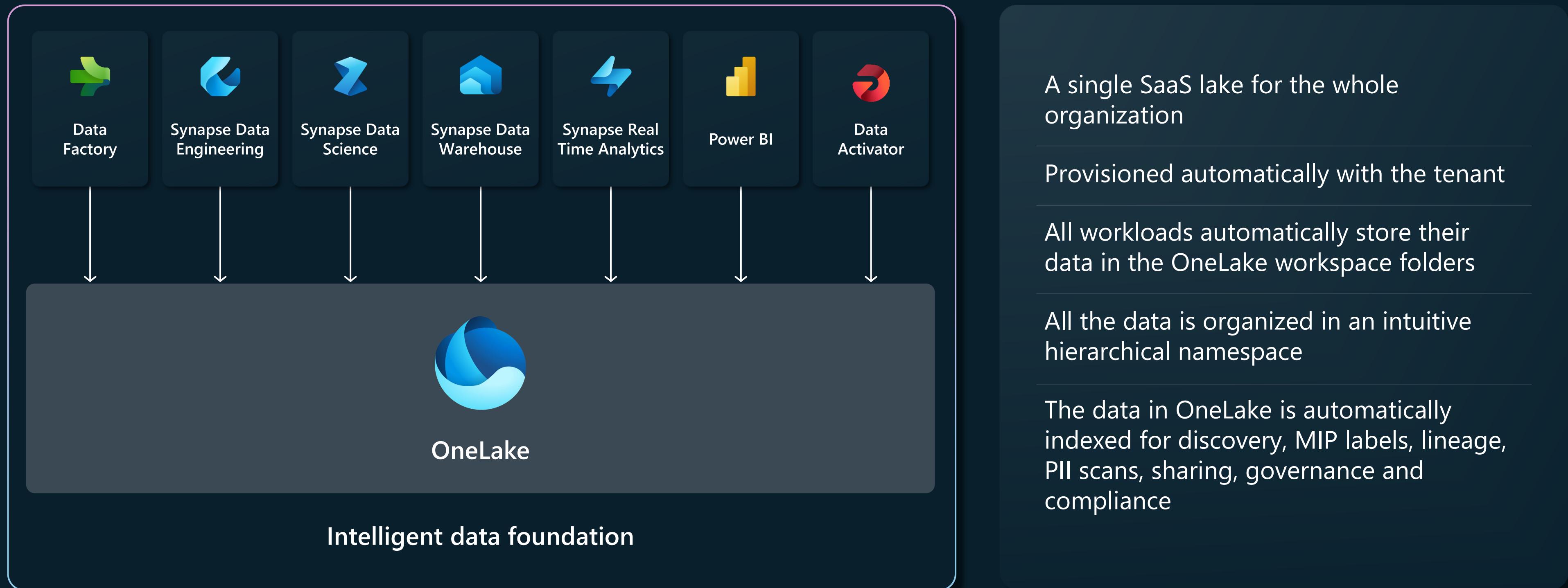
# OneLake which logically spans the world

“The OneDrive for Data”



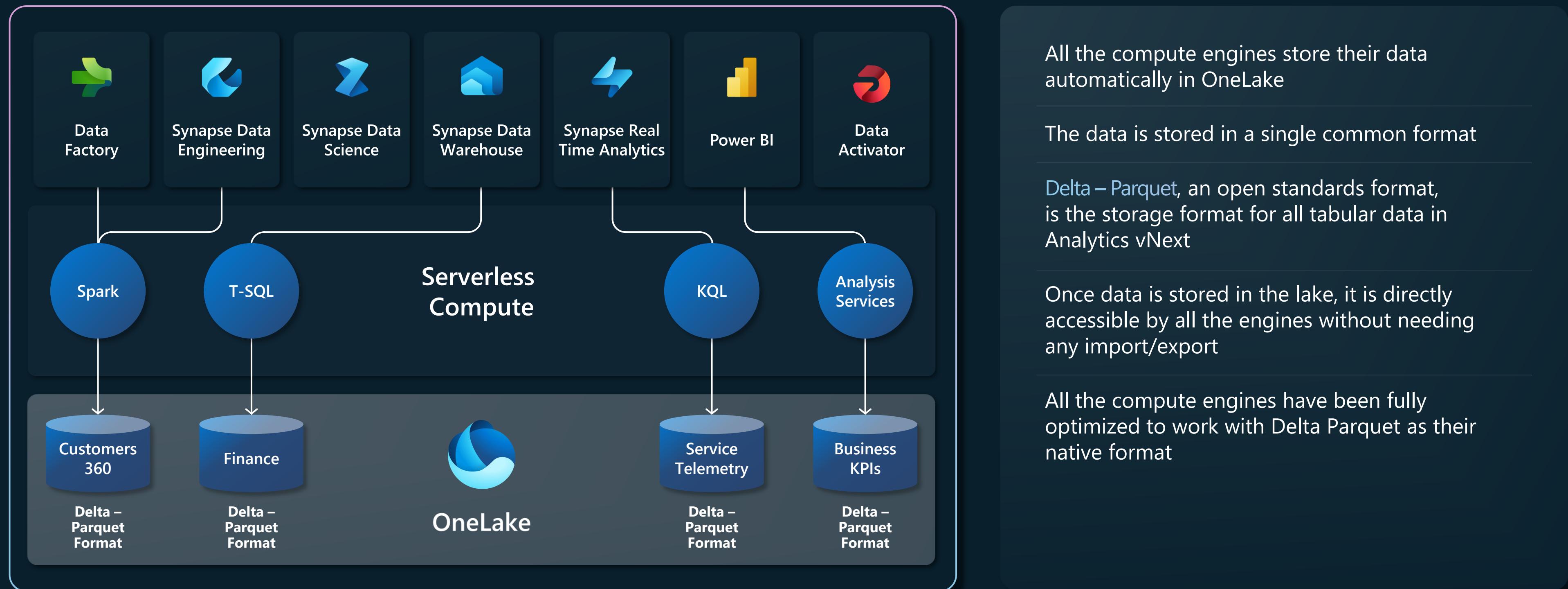
# OneLake for all Data

“The OneDrive for Data”



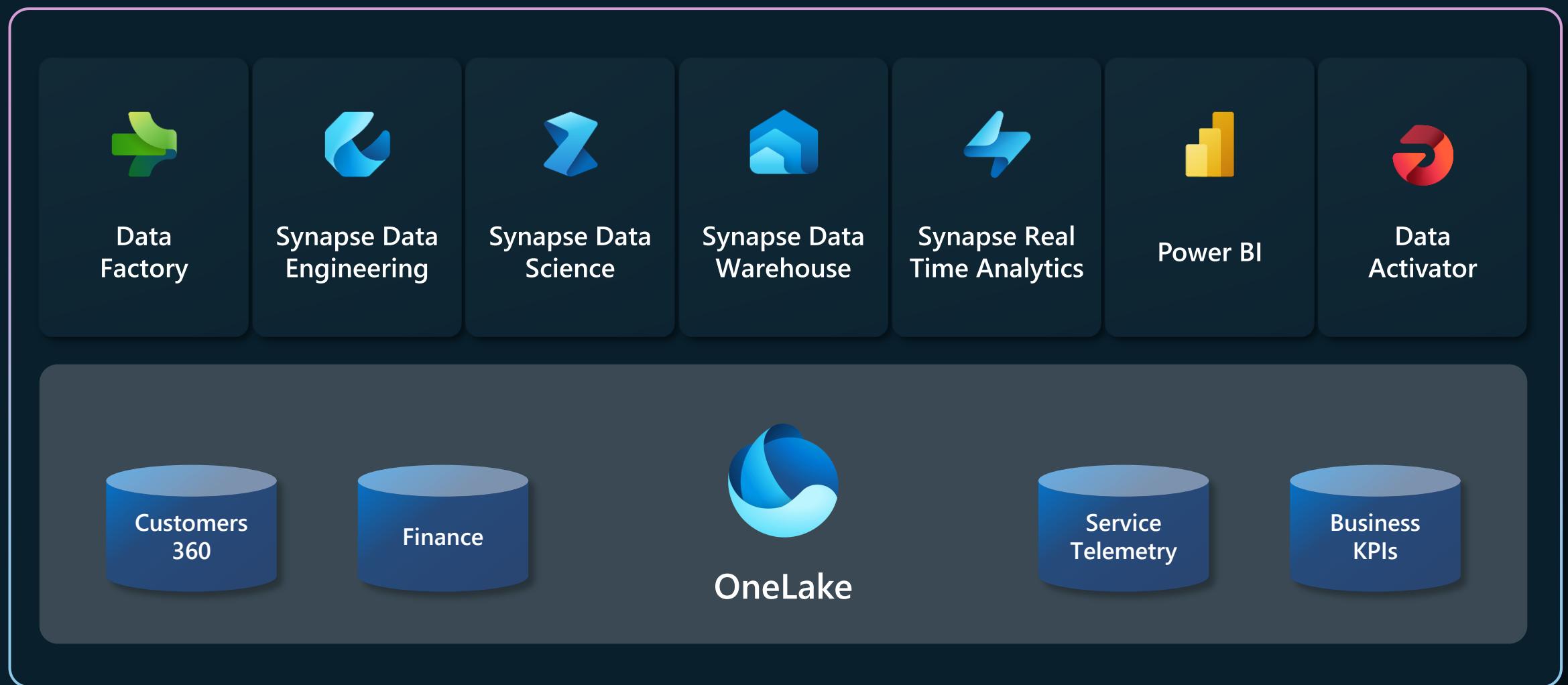
# One Copy for all computers

## Real separation of compute and storage



# Unifying data in OneLake

## Shortcuts



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

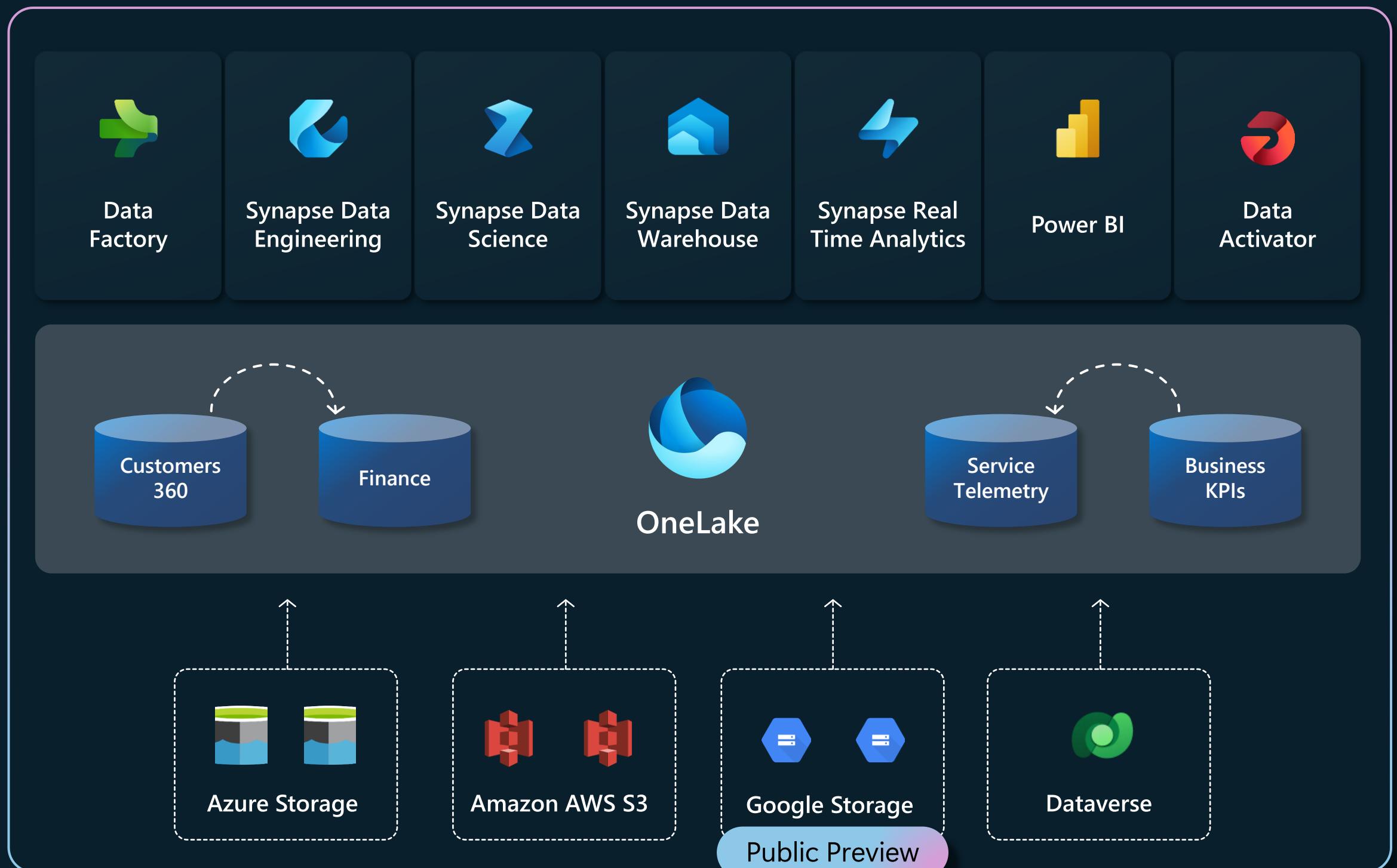
With shortcuts, data throughout OneLake can be composed together without any data movement

Shortcuts also allow instant linking of data already existing in Azure and in other clouds, without any data duplication and movement, making OneLake the first multi-cloud data lake

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

# Unifying data in OneLake

## Shortcuts



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

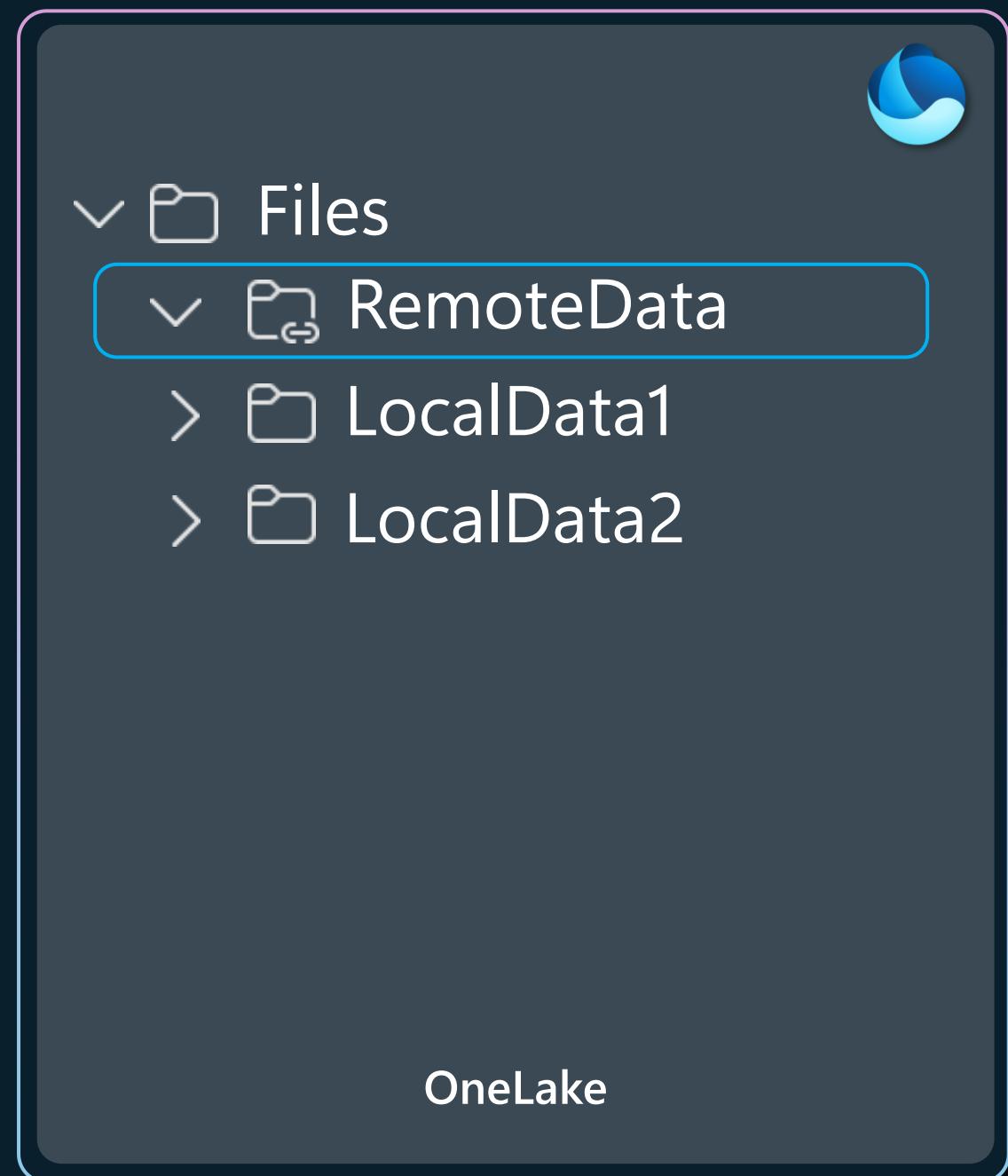
With shortcuts, data throughout OneLake can be composed together without any data movement

Shortcuts also allow instant linking of data already existing in Azure and in other clouds, without any data duplication and movement, making OneLake the first multi-cloud data lake

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

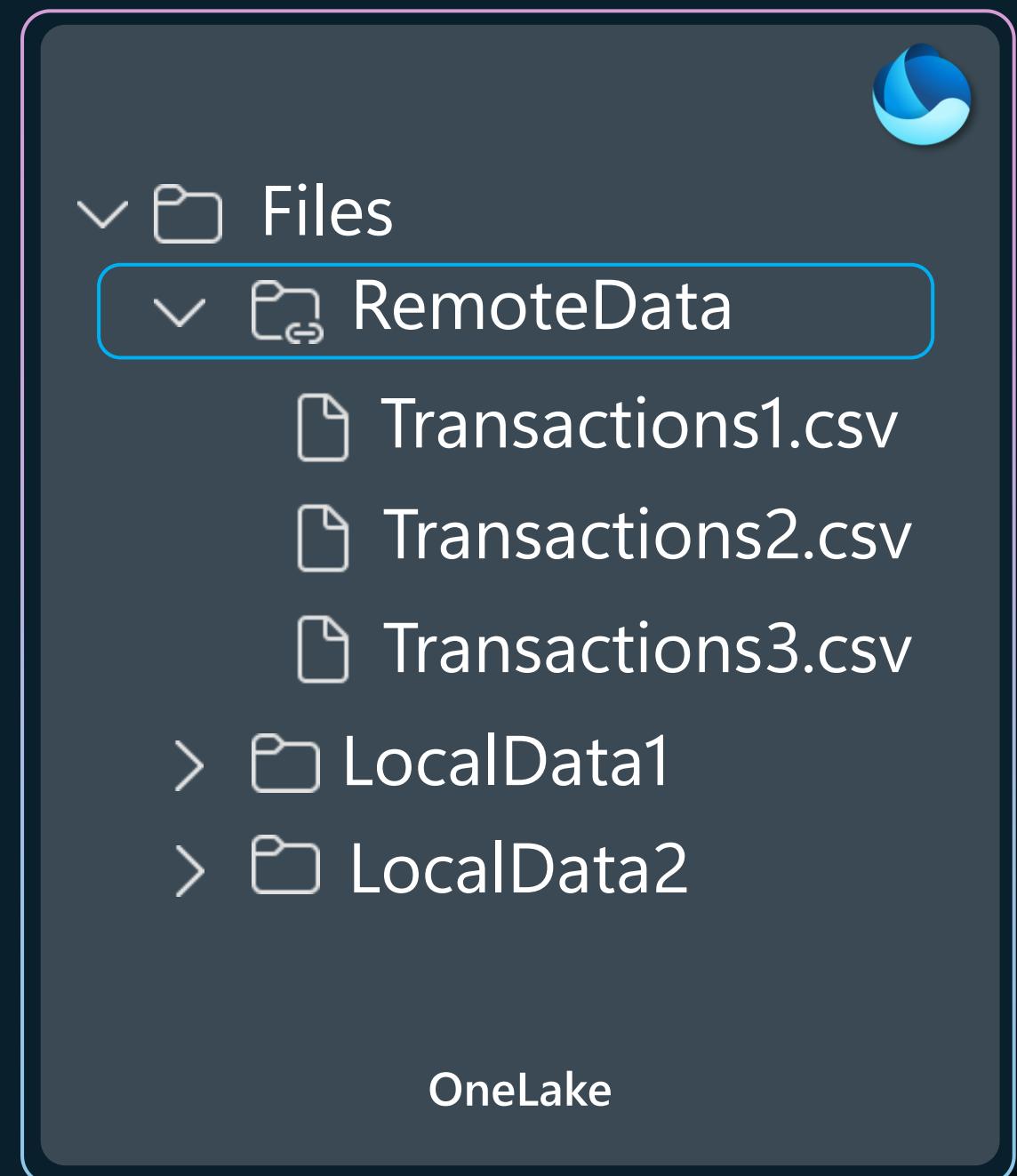
# What is a shortcut?

Shortcut Location

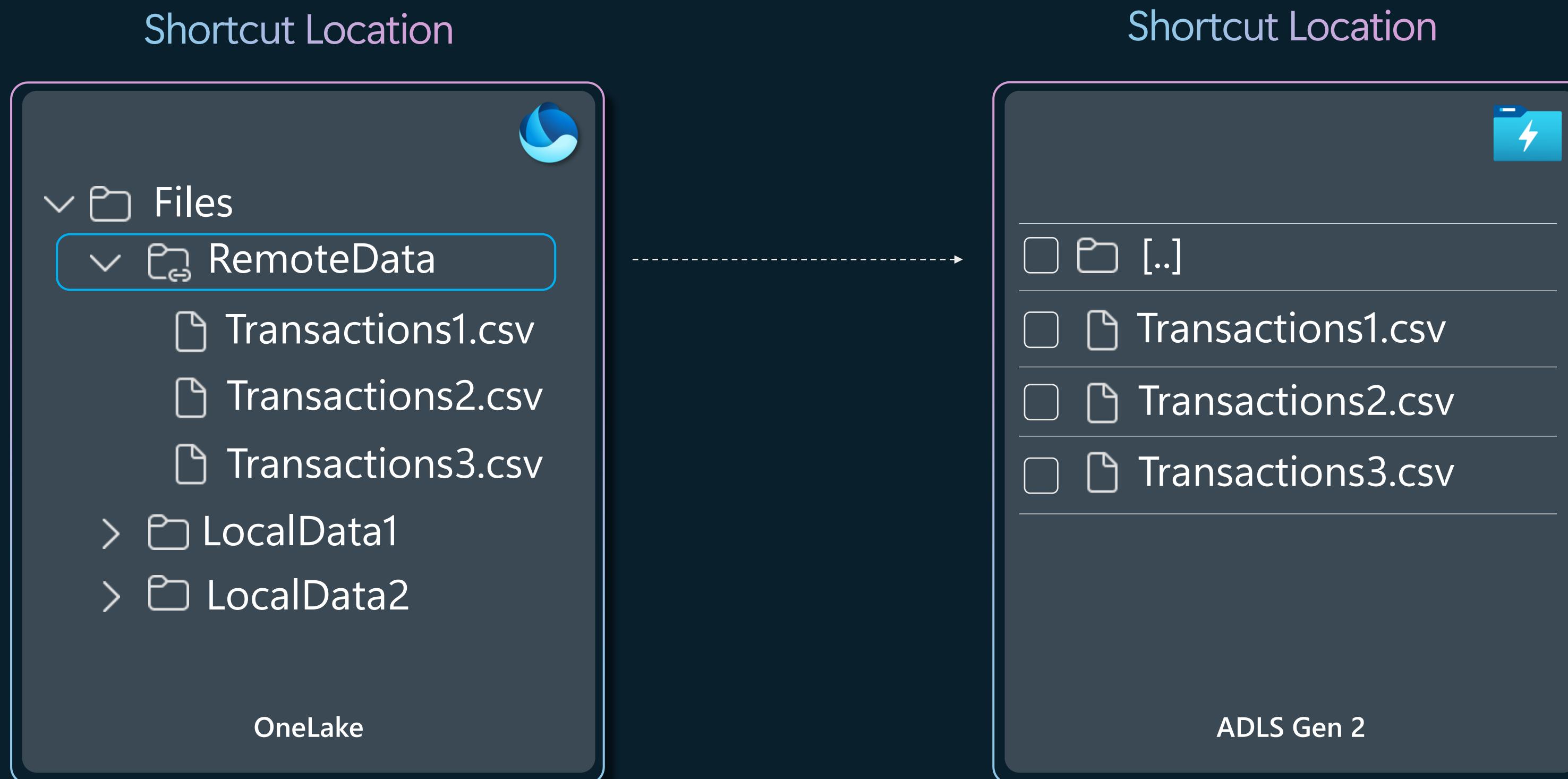


# What is a shortcut?

Shortcut Location



# What is a shortcut?

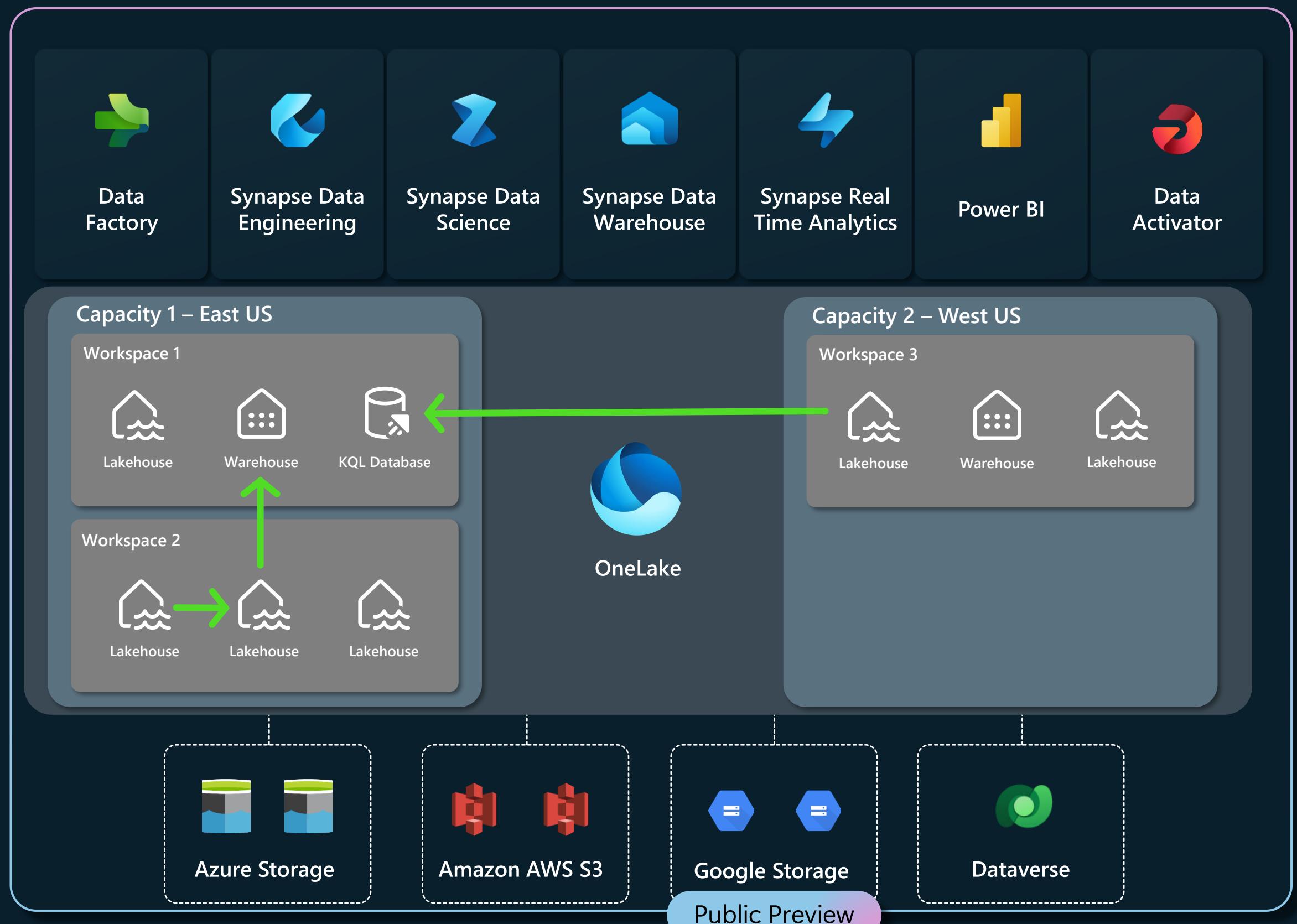




Any questions so far?

# Unifying data in OneLake

## Shortcuts



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

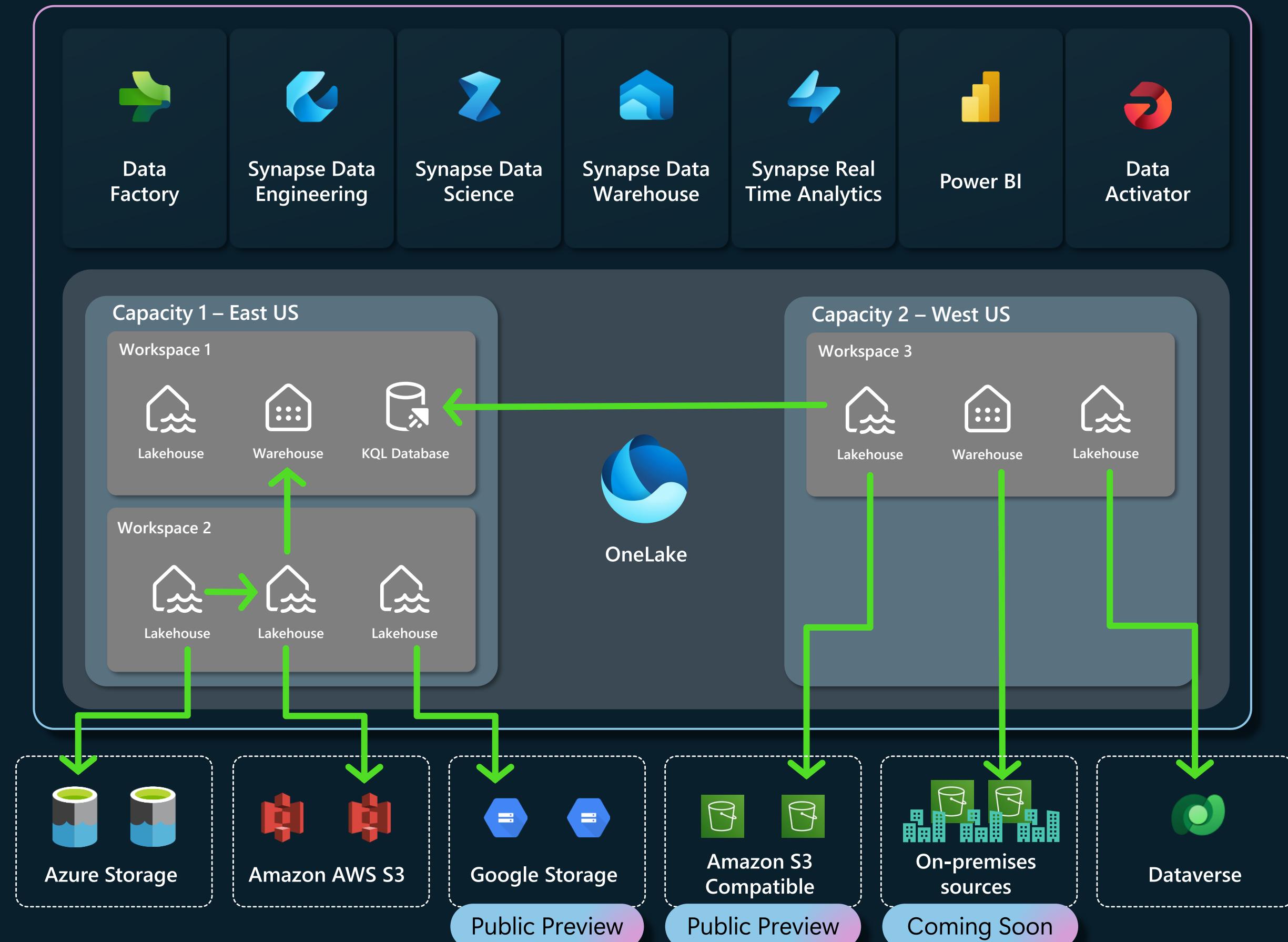
With shortcuts, data throughout OneLake can be composed together without any data movement

Shortcuts also allow instant linking of data already existing in Azure and in other clouds, without any data duplication and movement, making OneLake the first multi-cloud data lake

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

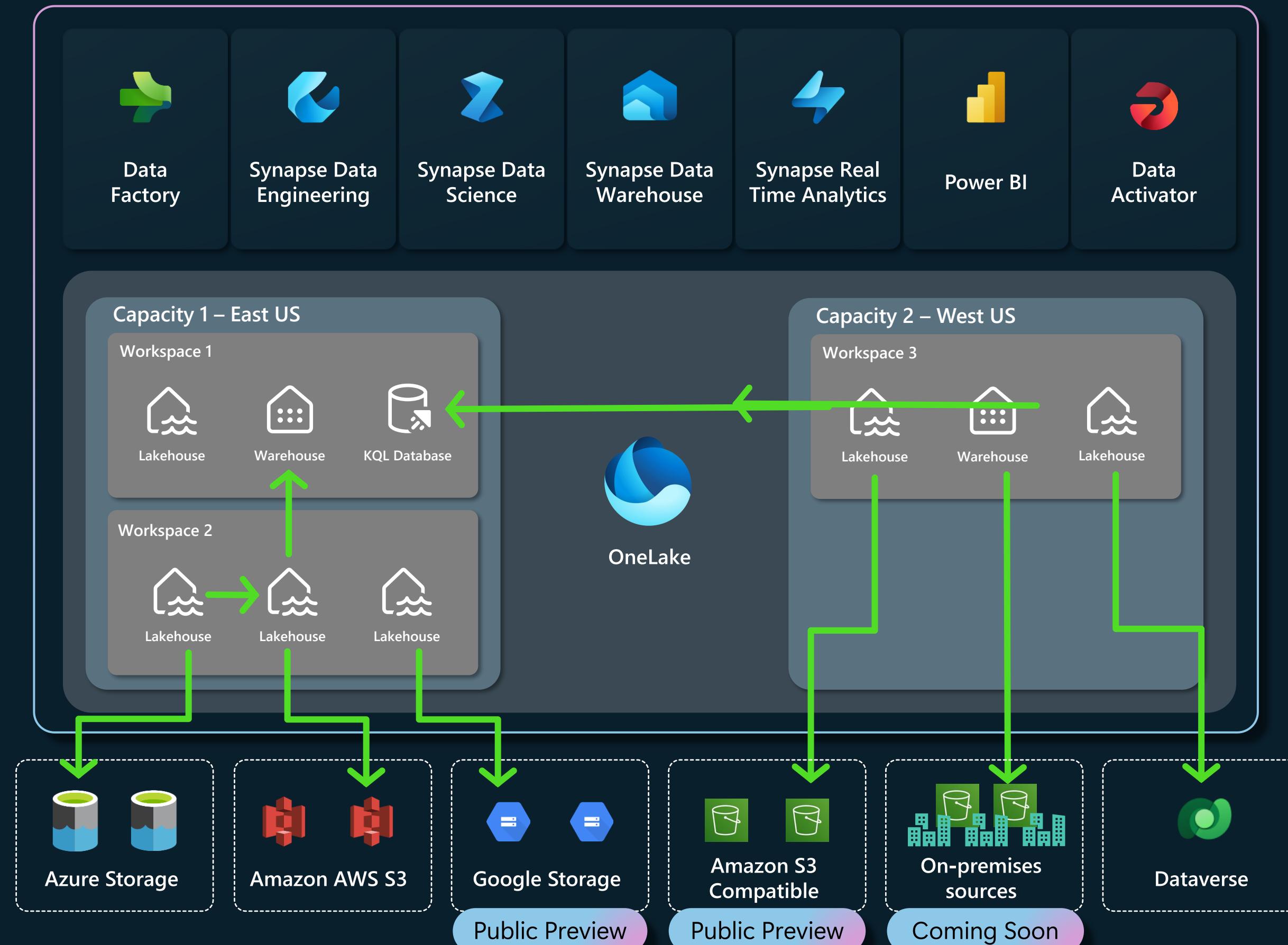
# Unifying data in OneLake

## Shortcuts



# Unifying data in OneLake

## External Data Sharing (Preview)



# Unifying data in OneLake

## External Data Sharing (Preview)

### 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](#)



Enabled

 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](#)



Enabled

 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.

# Mirroring in Fabric

Connect your databases and data warehouses to OneLake seamlessly with Zero ETL (Preview)



Azure  
Cosmos DB



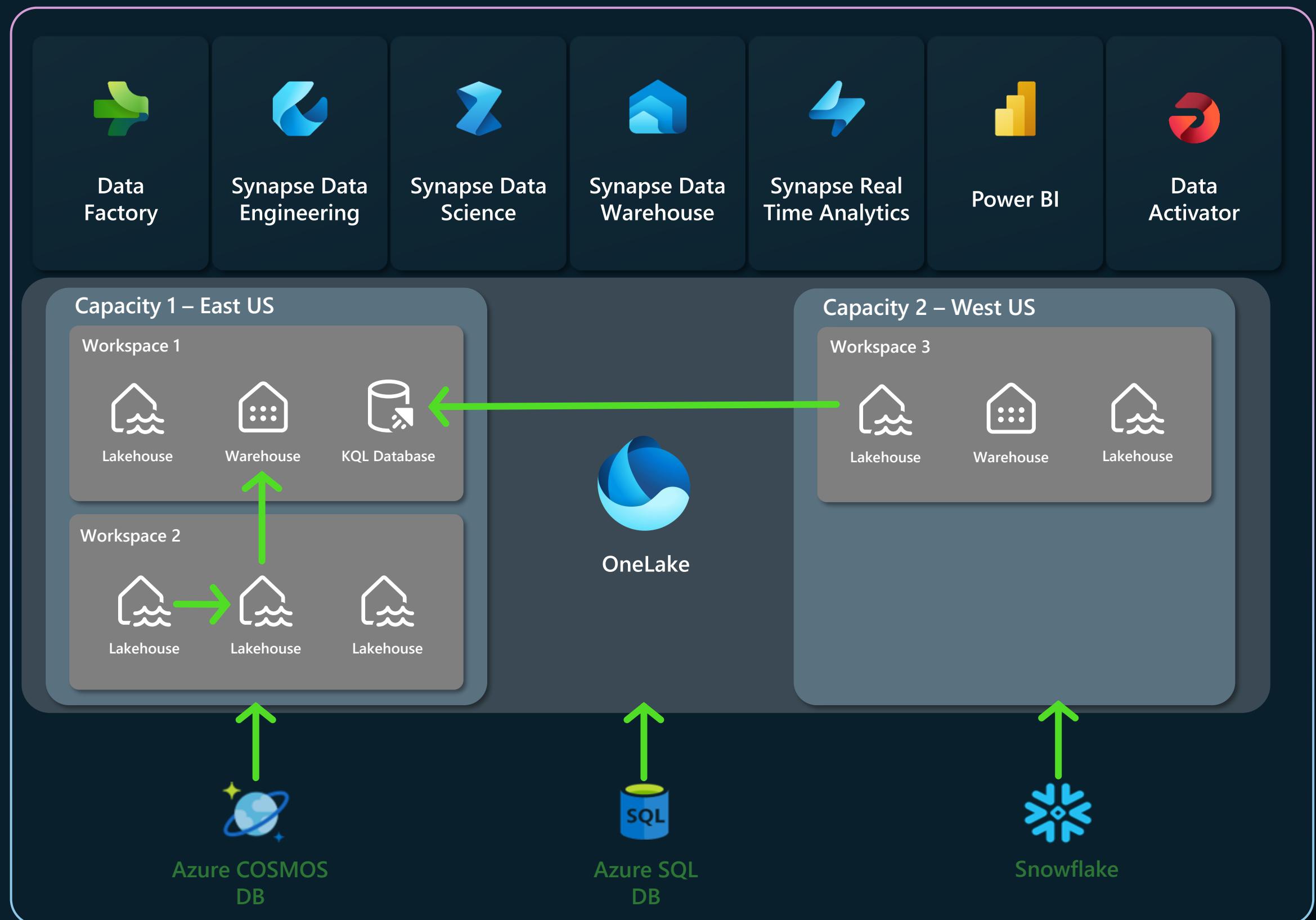
Azure  
SQL DB



Snowflake

# Mirroring in Fabric

Connect your databases and data warehouses to OneLake seamlessly with Zero ETL (Preview)



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

A full live querying experience of the source database is available for the Mirrored database

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

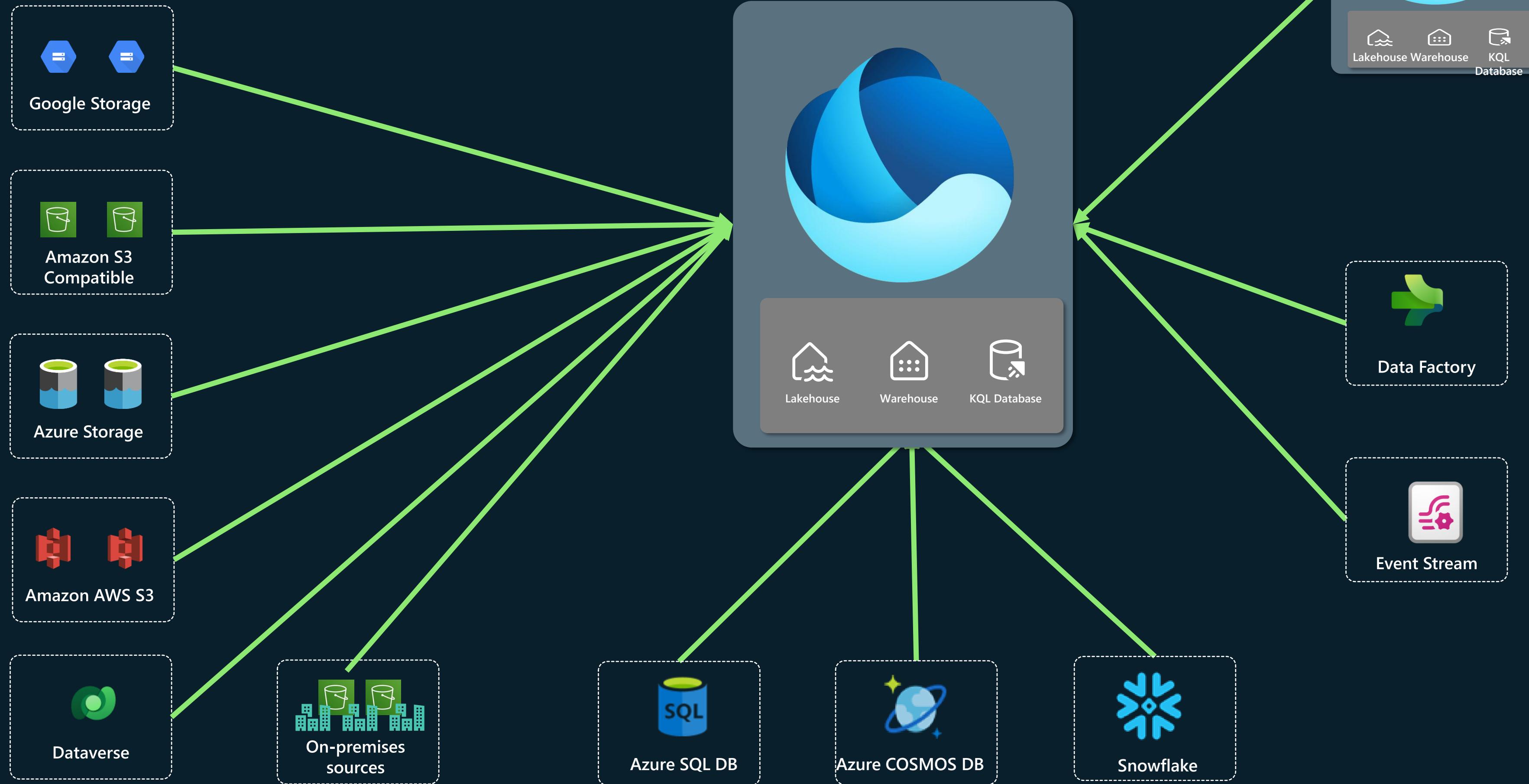
All the Fabric experiences instantly work with the OneLake replica, including data warehousing, Spark, BI experiences

Analysts and Data Scientists can work with real-time data

The replica protects operational databases from analytical queries

# Unifying data in OneLake

The heart of your organization



Innovate to accelerate

# DEMO

# Unify analytics and governance with Fabric and Purview



**Microsoft  
Fabric**



**Microsoft  
Purview**

Reshape how you access, manage, and act on data and insights by connecting every data source and analytics service together in Fabric

Keep your data safe and governed with unified data governance, information protection, and risk and compliance solutions

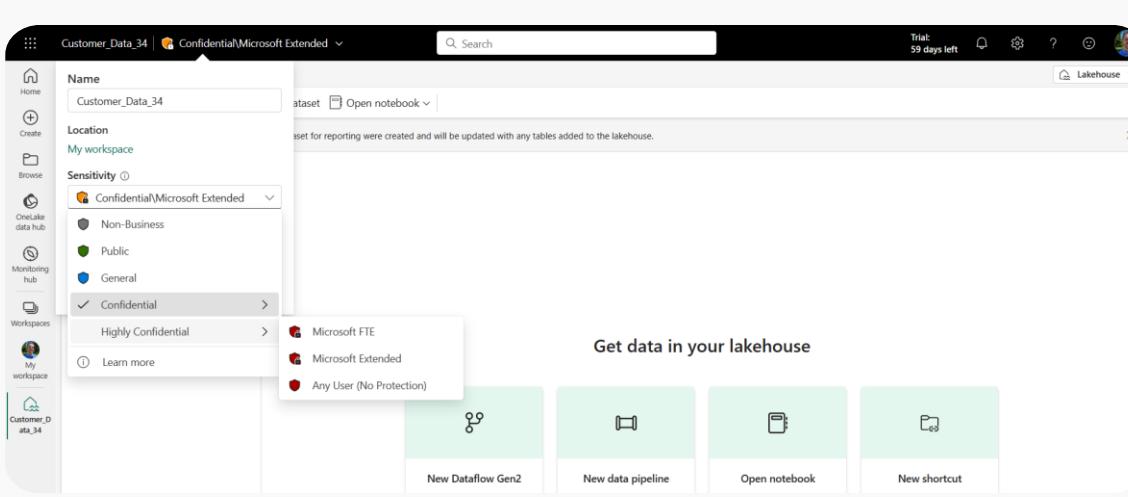
Data ownership

Unmatched security & compliance

Unified governance

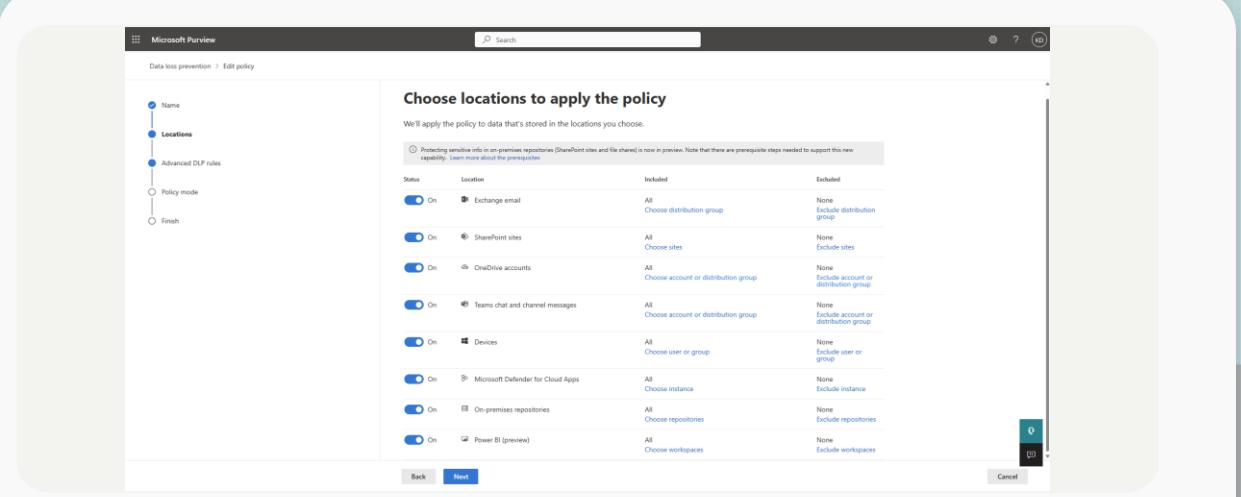
# Use Purview security and compliance features to protect your data in OneLake

Secure and protect data across your organization



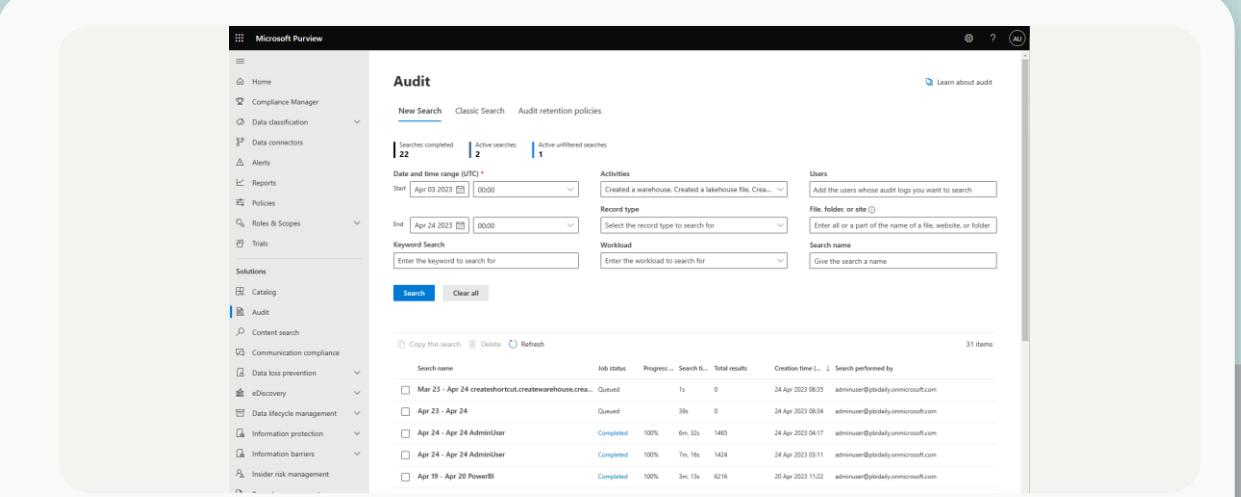
## Information Protection labels\*

Classify sensitive Fabric data using the same **sensitivity labels** that are used in Microsoft 365—enforced even when the data is exported



## Data Loss Prevention policies\*

Automatically detect the upload of sensitive data such as PII and trigger automatic risk remediation actions such as alerts



## Audit\*

Log user activities from Microsoft Fabric to support security, forensic, and internal investigations



\*Additional Microsoft Purview purchase required



# Sensitivity Labels

## Information protection

### Allow users to apply sensitivity labels for content

Enabled for the entire organization

With this setting enabled, Microsoft Purview Information Protection **sensitivity labels** published to users by your organization can be applied. All [prerequisite steps](#) must be completed before enabling this setting.

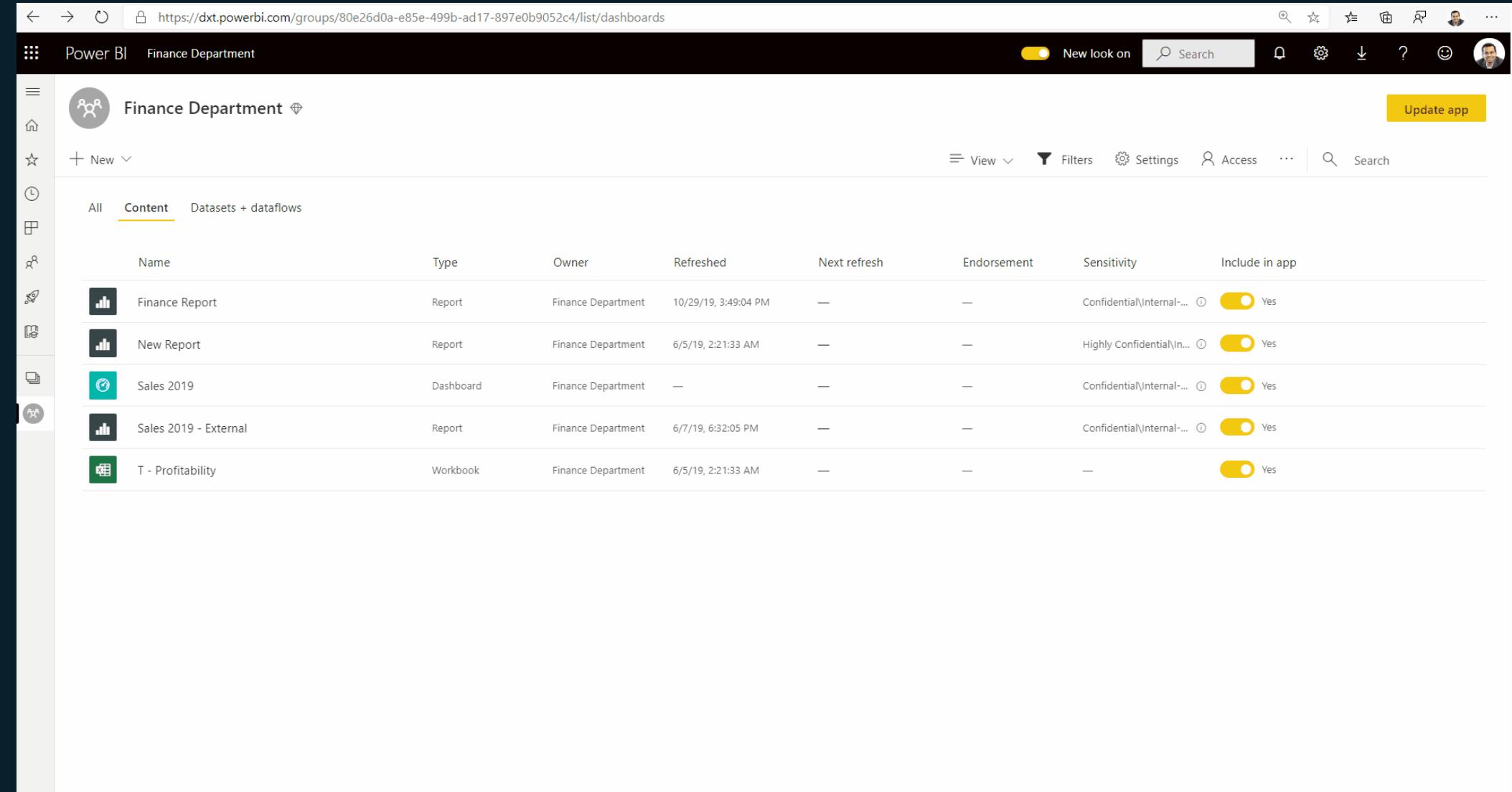
Note: **Sensitivity** label settings, such as encryption and content marking for files and emails, are not applied to content. [Learn More](#)

Visit the [Microsoft Purview compliance portal](#) to view **sensitivity** label settings for your organization.

Note: **Sensitivity** labels and protection are only applied to files exported to Excel, PowerPoint, or PDF files, that are controlled by "Export to Excel" and "Export reports as PowerPoint presentation or PDF documents" settings. All other export and sharing options do not support the application of **sensitivity** labels and protection.

 Enabled

 The setting below determines which users in the organization can apply and change sensitivity labels. All other users in the organization can only view the labels.



The screenshot shows a Microsoft Power BI dashboard titled 'Finance Department'. The 'Content' tab is selected, displaying a list of five items:

Name	Type	Owner	Refreshed	Next refresh	Endorsement	Sensitivity	Include in app
Finance Report	Report	Finance Department	10/29/19, 3:49:04 PM	—	—	Confidential\Internal...	
New Report	Report	Finance Department	6/5/19, 2:21:33 AM	—	—	Highly Confidential\In...	
Sales 2019	Dashboard	Finance Department	—	—	—	Confidential\Internal...	
Sales 2019 - External	Report	Finance Department	6/7/19, 6:32:05 PM	—	—	Confidential\Internal...	
T - Profitability	Workbook	Finance Department	6/5/19, 2:21:33 AM	—	—	—	

Innovate to accelerate

# DEMO



Any questions left?

# Thank you



**Erwin de Kreuk**

Principal Consultant - Lead Data & AI  
InSpark

 @erwinedekreuk

 [linkedin.com/in/erwinedekreuk](https://www.linkedin.com/in/erwinedekreuk)

 [erwinedekreuk.com](http://erwinedekreuk.com)

 [github.com/edkreuk](https://github.com/edkreuk)

 <https://sessionize.com/erwin-de-kreuk/>



Let's  
connect





# 16 edycja konferencji SQLDay

13-15 maja 2024, WROCŁAW + ONLINE



partner platynowy



partner złoty



partner srebrny

