

**HITACHI**



# Fabric in a Day

## Microsoft Fabric Fundamentals



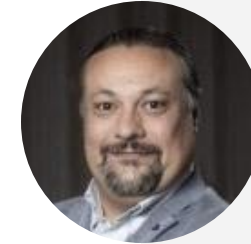
# Speakers



**Graeme Ord**  
Senior Architect  
Hitachi Solutions



**Anthony Sheldrake**  
Architect  
Hitachi Solutions



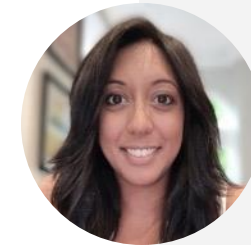
**Rumen Krastev**  
Global Senior Partner  
Solution Architect  
Microsoft



**Matt Holmes**  
Business Development  
Director, Commercial  
Hitachi Solutions



**Sarah Coward**  
Fabric Product Owner  
DEFRA



**Ayesha Jacob**  
Cloud and AI Platform Data  
and Analytics GTM Manager  
Microsoft

# Fabric in a Day: Agenda

HITACHI

**09:00 – 09:30**

## **Arrival and Welcome (breakfast provided)**

Settle in, meet the team, and get an overview of the day ahead.

---

**09:30 – 9:45**

## **Introduction to Hitachi Solutions**

Learn more about what we do and who we are.

---

**09:45 – 10:00**

## **Lessons from Microsoft:**

Unify and activate your data for AI Innovation

---

**10:00 – 10:45**

## **Choosing Your Storage and Compute Platform**

Explore the architecture of Fabric, including Lakehouse, Data Warehouse, and Real-Time Analytics, and learn how to select the right components for your needs.

---

**10:45 – 11:00**

## **Break**

---

**11:00 – 11:45**

## **Lab 1: Ingest and Load your Data**

Learn how to connect and ingest data from multiple sources using Dataflows, Data Factory, and Pipelines

---

**11:45 – 12:30**

## **Fabric Governance in Practice:**

Q&A with Sarah Coward, DEFRA Product Owner

---

**12:30 – 13:30**

## **Lunch Break**

---

**13:30 – 14:15**

## **Intro to Apache Spark**

---

**14:15 – 15:00**

## **Lab 2: Data Transformation Spark & Notebooks**

Hands-on session focused on preparing, cleansing, and transforming data using Apache Spark notebooks and business logic

---

**15:00 – 15:15**

## **Break**

---

**15:15 – 16:00**

## **Lab 3: Data Visualisation**

Build data models and bring insights to life using Power BI dashboards and KPIs

---

**16:00 – 16:45**

## **Real-Time Intelligence, AI, ML**

Explore how Fabric integrates with AI and Copilot to enable predictive analytics, automation, and intelligent insights

---

**16:45 – 17:00**

## **Wrap-Up, Key Takeaways & Closing remarks**

Recap key learnings and next steps for applying Fabric in your organisation

# WI-FI Details here



## Connection Instructions:

Make sure your wireless adapter is set to dynamically obtain an IP address.

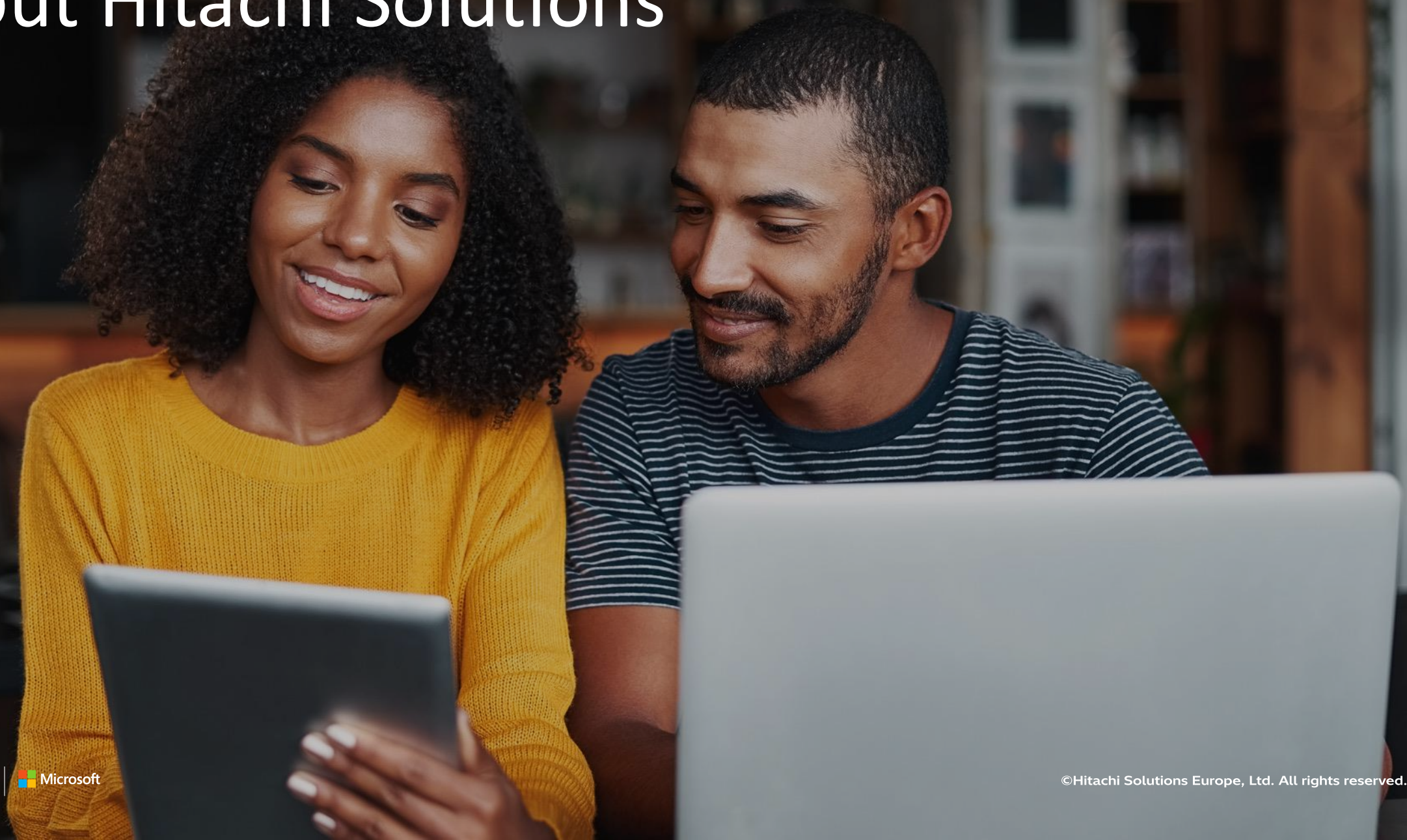
Connect to the wireless network **MSFTGUEST**.

Open a web browser and navigate to any website.  
Your browser should automatically redirect you to the  
Captive Portal.

Click on Event Attendee Code and enter the access code: **FAID\_MSFT\_HiSol**

# About Hitachi Solutions

HITACHI



# Hitachi has been a leader in innovation since 1910

All companies under the Hitachi name operate today based on the principle our founder Namihei Odaira recognized over 110 years ago: "Companies are their people."

Hitachi remains at the forefront of innovation today, delivering new innovations to benefit both society and organisations.

## Digital Innovation & Transformation

Accelerating digital and business transformation within organisations across the globe through AI, IoT, cloud computing, digital engineering and data analytics.

\$17B USD

107,000 employees

## Green Energy & Mobility

Sustainable innovation through renewable energy, smart grids, electric transportation, and advancing carbon neutrality and efficient infrastructure for a greener future.

\$20.2B USD

82,000 employees

## Connective Industries

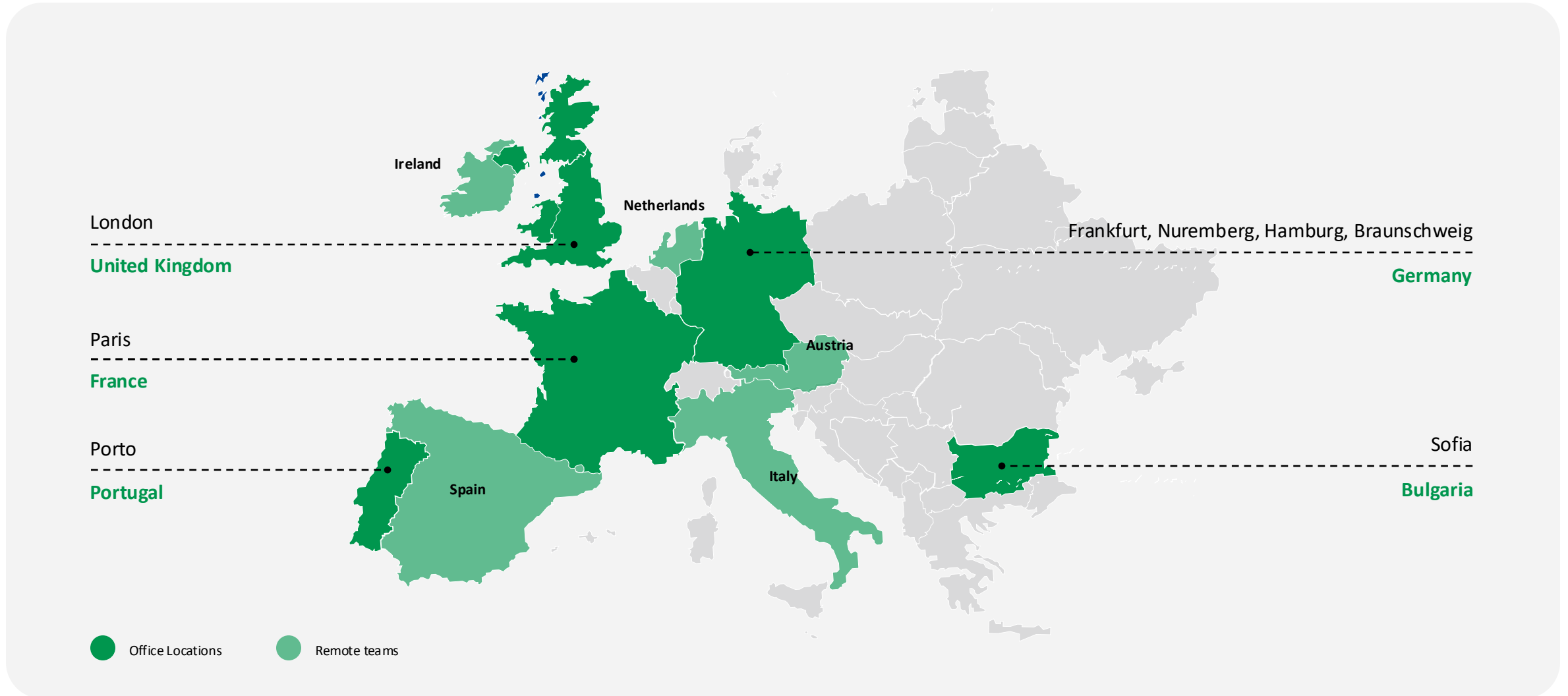
Integrates digital, IoT, and AI solutions to enhance efficiency, sustainability, and productivity across manufacturing, logistics, energy, and urban development.

\$21.2B USD

62,000 employees



# Europe in detail



# 20 years of partnership with Microsoft

We work exclusively alongside Microsoft, leveraging 20 years of collaboration to drive end-to-end digital transformation to organisations across the globe.

Our mission to contribute to society through the development of superior, original technology and products aligns with Microsoft's goal of empowering people and organizations worldwide



Our expanded partnership with Hitachi will bring together the power of the Microsoft Cloud – including Microsoft Copilot – with Hitachi’s industry expertise to **improve the productivity** of 270,000 Hitachi employees and help **address customers’ biggest challenges**, including sustainability.”

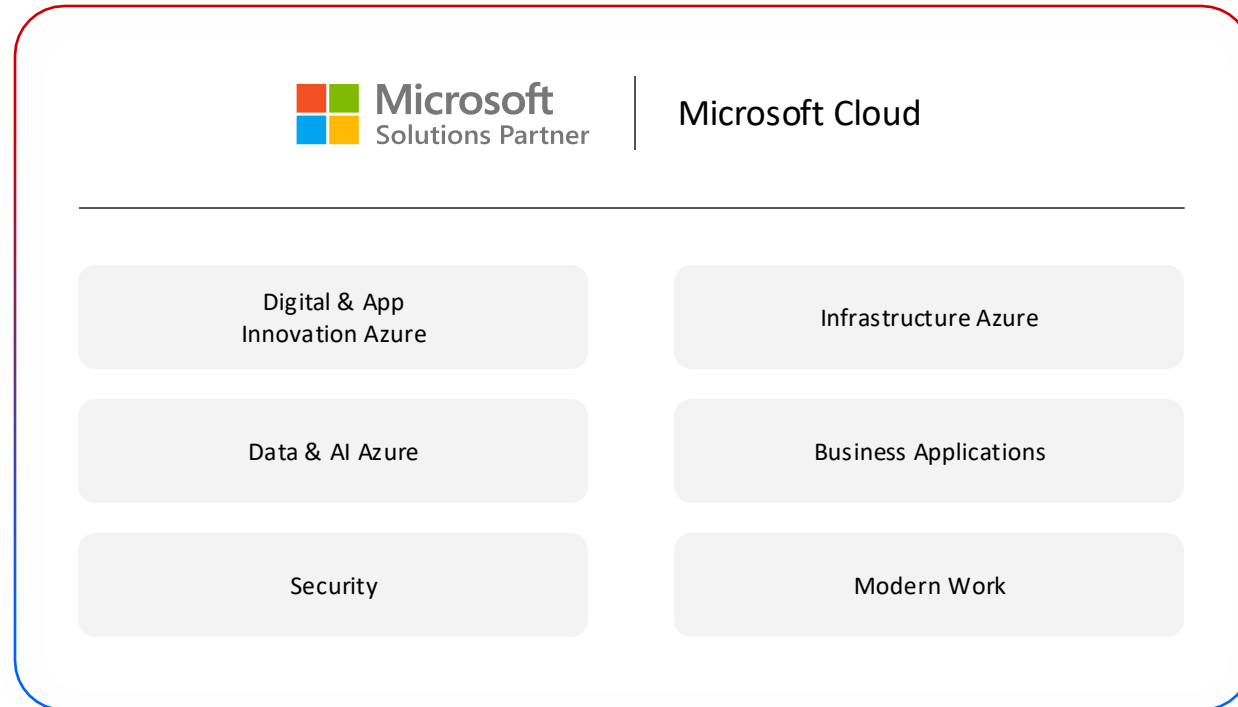


**Satya Nadella**  
Chairman and CEO, Microsoft

A handwritten signature in black ink that reads "Satya N.".

# Unrivalled Microsoft expertise

Our Microsoft Cloud Solutions Partner status is only available to a handful of organisations worldwide who meet **performance, skilling, and customer success criteria**. This demonstrates our ability to provide **comprehensive cloud services** to help businesses optimise operations, enhance security, and drive innovation.



## Specialisations

- ✔ Analytics on Microsoft Azure
- ✔ Azure Data Virtual Desktop
- ✔ Build and Modernise AI Apps with Microsoft Azure
- ✔ Cloud Security
- ✔ Data Warehouse Migration
- ✔ DevOps with GitHub
- ✔ Finance
- ✔ Identity and Access Management
- ✔ Information Protection and Governance
- ✔ Infra and Database Migration to Microsoft Azure
- ✔ Low Code Application Development
- ✔ Microsoft Azure Virtual Desktop
- ✔ Migrate Enterprise Apps to Microsoft Azure
- ✔ Sales
- ✔ Service
- ✔ Modernise Endpoints
- ✔ Supply Chain
- ✔ Threat Protection

# Our Microsoft awards

55x

Microsoft Partner of  
the Year Award Winner

100

Partner Contribution  
Index Score  
*out of 100*

19

Years on  
Microsoft's  
Partner Advisory Council

22

Years in  
Microsoft's  
Inner Circle



2025 Microsoft Partner of the Year

**Winner for Finance**

2025 Microsoft Partner of the Year

**Winner for Government**

2025 Microsoft Partner of the Year

**Finalist for Build and Modernize AI Apps Award**

2025 Microsoft Partner of the Year

**Finalist for Dynamics 365 Sales & Customer Insights Award**

2025 Microsoft Partner of the Year

**Finalist for Dynamics 365 Supply Chain Award**

2025 Microsoft Partner of the Year

**Finalist for Low Code Application Development Award**

2024 Microsoft Partner of the Year

**Winner for Low Code Application Development**

2024 Microsoft Partner of the Year

**Finalist for Intelligent Automation**

2023 Microsoft Partner of the Year

**Winner for Dynamics 365 Services**

2023 Microsoft Partner of the Year

**Winner for Dynamics 365 Supply Chain Management**

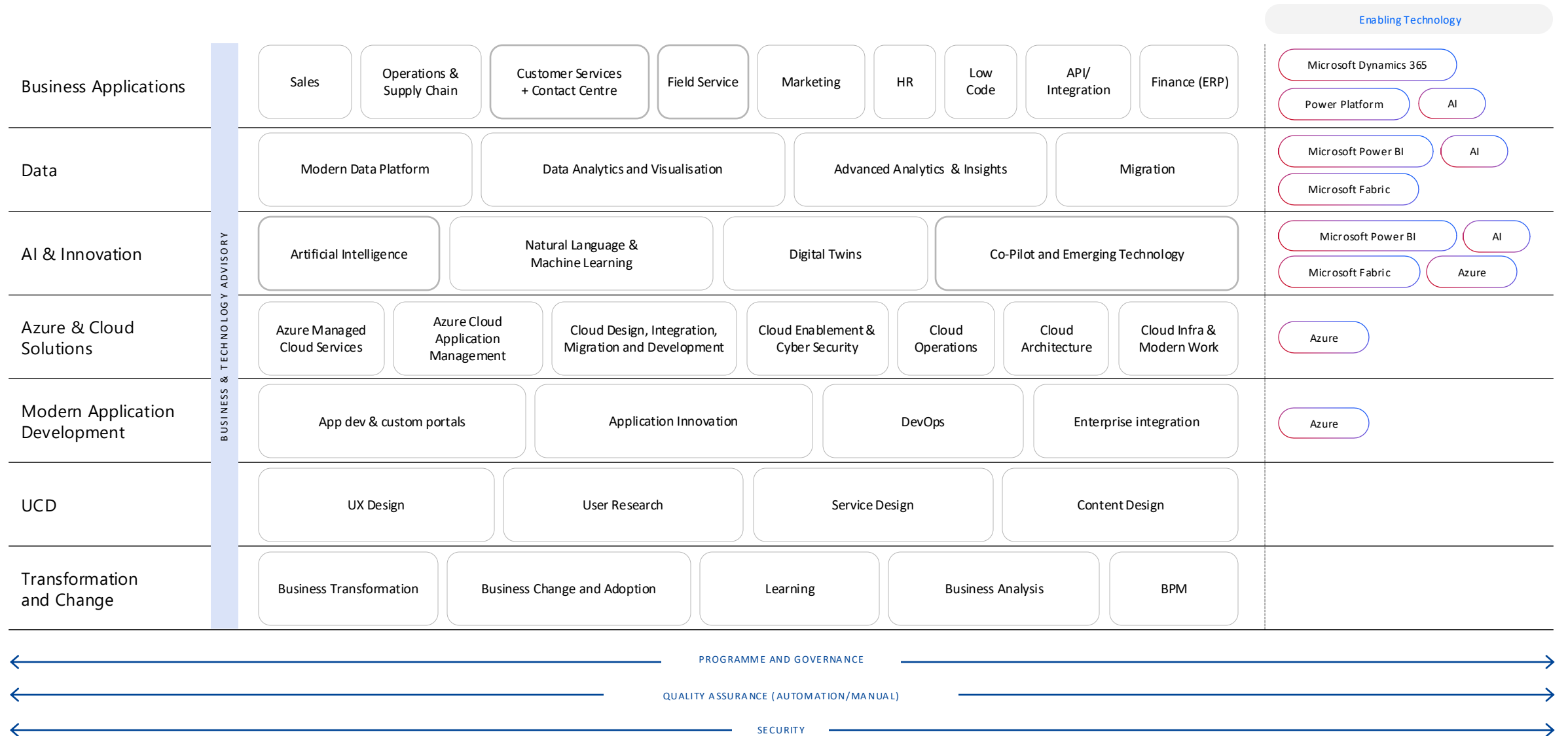
2022 Microsoft Partner of the Year

**Winner for Dynamics 365 Supply Chain Management**

2021 Microsoft Partner of the Year

**Winner for Dynamics 365 Field Service**

# Our overall services





# Unify and activate your data for AI innovation with Microsoft Fabric

Ayeshna Jacob  
Data & Analytics Lead, Microsoft UK&I

# Microsoft Fabric

The unified data platform for AI transformation



Data Factory



Analytics



Databases



Real-Time  
Intelligence



IQ



Power BI

Fabric Platform



AI



OneLake



Security & Governance

Accelerate  
innovation  
with Microsoft  
data platforms



# Activate your entire data estate for AI innovation

1

Unify your  
data estate

2

Transform data for  
powerful AI solutions

3

Empower everyone  
with analytics and  
insights

4

Seamlessly govern and protect your data

# OneLake data is available everywhere

## Productivity

M365 Copilot



Microsoft Excel



Microsoft Teams



Power Platform



## Data Platforms

Azure Databricks



Microsoft Fabric



Snowflake



## AI and ML

Azure ML



Copilot Studio



Azure AI Foundry



VS Code



OneLake

Multi-cloud shortcuts



Azure



AWS S3



GCP



Dataverse



Snowflake



On-Prem



Cross Tenant



Databricks



SQLDB



COSMOS DB



SQL MI

Database Mirroring

# Microsoft Fabric + Microsoft Foundry



Microsoft  
Fabric

Enhance data insights

Ground your AI on unstructured data

Analyze customer interactions

Customize machine learning models



Microsoft  
Foundry



OneLake serves as the connective tissue to build custom AI apps powered by your data

# Microsoft Intelligent Data Platform

Databases

Analytics



SQL Server on Azure Virtual Machines

IaaS



Azure SQL Database



Azure Cosmos DB



Azure Database for PostgreSQL



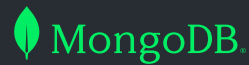
Azure Database for MySQL



Azure Databricks



Microsoft Fabric



Azure Managed Redis



SaaS

PaaS



Azure AI Foundry



Microsoft Defender



Microsoft Purview

AI

Security

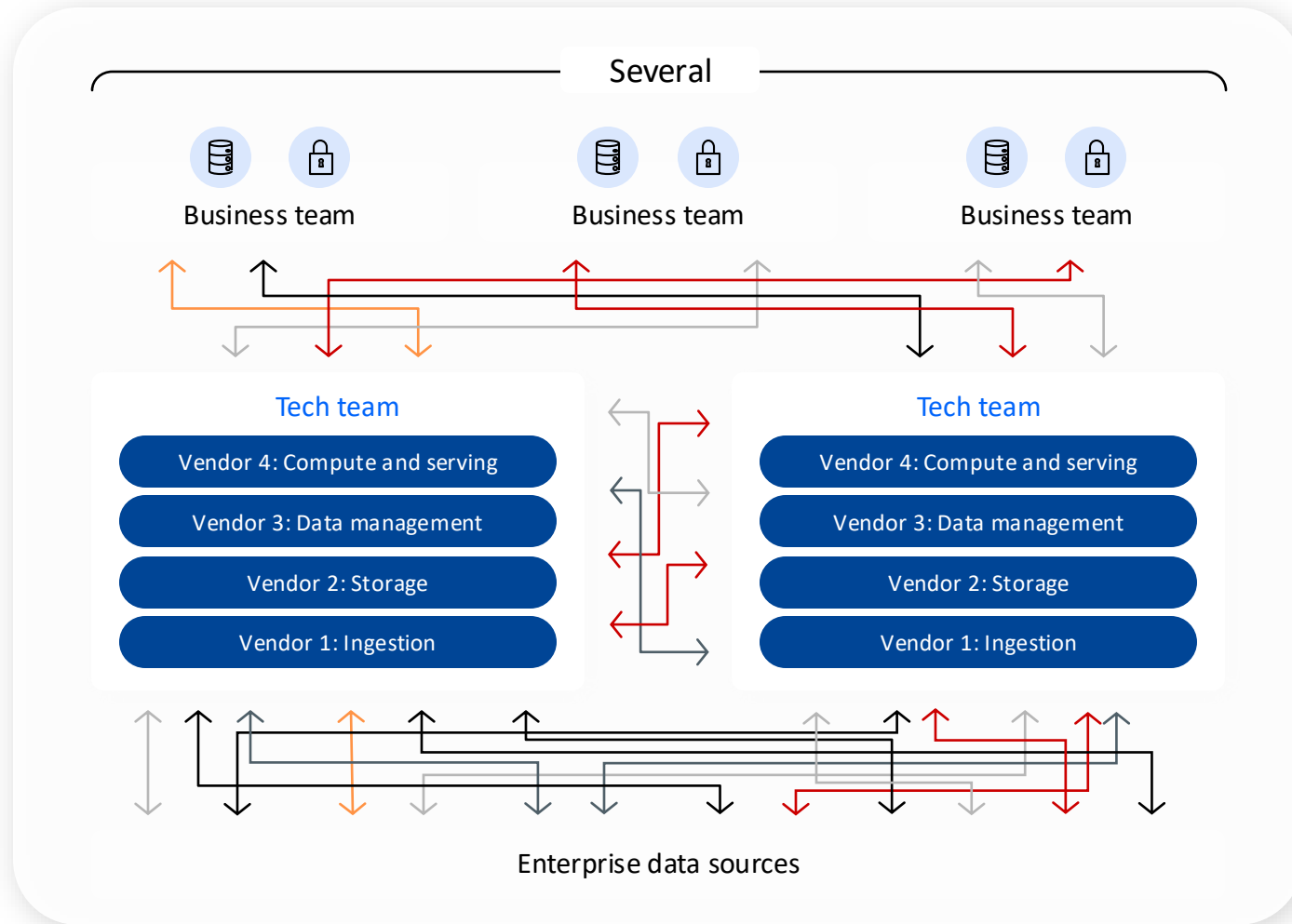
Governance

# Explore End-to-End Analytics and Analytics in the Era of AI



# Analytics is complex

A complex, organically evolved data estate



1

Data copies and infrastructure inefficiencies

2

Limited interoperability between vendor services

3

Data exposure risks

# There is Massive Fragmentation of the Modern Data Stack

The image displays a vast collection of logos for companies across various sectors of the modern data stack, organized into several main categories:

- INFRASTRUCTURE:** Includes STORAGE (AWS, Google Cloud, IBM, Microsoft Azure), MPP DBs (teradata, Vertica, Exasol), DATA LAKES / LAKEHOUSES (dremio, databricks), DATA WAREHOUSES (Snowflake, Oracle, Amazon Redshift), STREAMING / IN-MEMORY (Kafka, Flink, Spark), and GPU DATABASES (Neo4j, Dgraph).
- ANALYTICS:** Includes BI PLATFORMS (Looker, Tableau, Power BI), VISUALIZATION (Tableau, Power BI, Qlik), DATA SCIENCE NOTEBOOKS (Databricks, Jupyter), DATA SCIENCE PLATFORMS (Snowflake, Databricks), and ENTERPRISE ML PLATFORMS (Databricks, AWS SageMaker).
- MACHINE LEARNING & ARTIFICIAL INTELLIGENCE:** Includes DATA GENERATION & LABELING (Scale AI, Labelbox), MLOPS (Weights & Biases, MLflow), SPEECH (Siri, Alexa), NLP (OpenAI, Google), HORIZONTAL AI / AGI (Anthropic, OpenAI), AI HARDWARE (NVIDIA, Intel), GPU CLOUD (AWS, Azure), and CLOSED SOURCE MODELS (OpenAI, Google).
- APPLICATIONS - ENTERPRISE:** Includes SALES (Salesforce), MARKETING (HubSpot), CUSTOMER EXPERIENCE (Intercom), HUMAN CAPITAL (Workday), and AUTOMATION & OPERATIONS (UiPath).
- APPLICATIONS - HORIZONTAL:** Includes CODE & DOCUMENTATION (GitHub), TEXT (OpenAI), AUDIO & VOICE (OpenAI), IMAGE (OpenAI), VIDEO EDITING (Runway), and SEARCH (Elasticsearch).
- APPLICATIONS - INDUSTRY:** Includes FINANCE & INSURANCE (Kenshuc), HEALTHCARE (Tempus), LIFE SCIENCES (Moderna), TRANSPORTATION (Uber), AGRICULTURE (Bowers), INDUSTRIAL & LOGISTICS (Siemens), and GOVT & INTELLIGENCE (Palantir).
- OPEN SOURCE INFRASTRUCTURE:** Includes FRAMEWORKS (Kubernetes), QUERY / DATA FLOW (Spark), DATA ACCESS (Fivetran), DATABASES (PostgreSQL), OLAP (Snowflake), ORCHESTRATION (Airflow), INFRASTRUCTURE (Terraform), DATA OPS (Dagster), STREAMING & MESSAGING (Kafka), STAT TOOLS & LANGUAGES (Python), MLOPS & AI INFRA (MLflow), AI FRAMEWORKS & LIBRARIES (PyTorch), AI MODELS & ARCHITECTURES (OpenAI), SEARCH (Elasticsearch), LOGGING & MONITORING (Datadog), VISUALIZATION (Tableau), and COLLABORATION (Zoom).
- DATA SOURCES & APIS:** Includes DATA MARKETPLACES & DISCOVERY (Bloomberg), FINANCIAL & MARKET DATA (Bloomberg), AIR / SPACE / SEA (SpaceX), PEOPLE / ENTITIES (LinkedIn), LOCATION INTELLIGENCE (Esri), and ESG (Sustainalytics).
- DATA & AI CONSULTING:** Includes QuantumBlack, ECG, Deloitte, IBM AI Consulting, Cambridge Quantum, LeewayHertz, Slalom, Brookfield Data Co., Ternary, Truist AI, Meru, Seers Insights, Data Root Labs, Indura Labs, Thredye, Azati, Addepto, UpRight Health, Bytecode IO, L&T Data, and others.

Version 1.0 - Feb 2023 | © Matt Turck (@mattturck), Kevin Zhang (@ykevinzhang) & FirstMark (@firstmarkcap) | Blog post: mattturck.com/MAD2023 | Interactive version: MAD.firstmarkcap.com | Comments? Email MAD2023@firstmarkcap.com | FIRSTMARK EARLY STAGE VENTURE CAPITAL

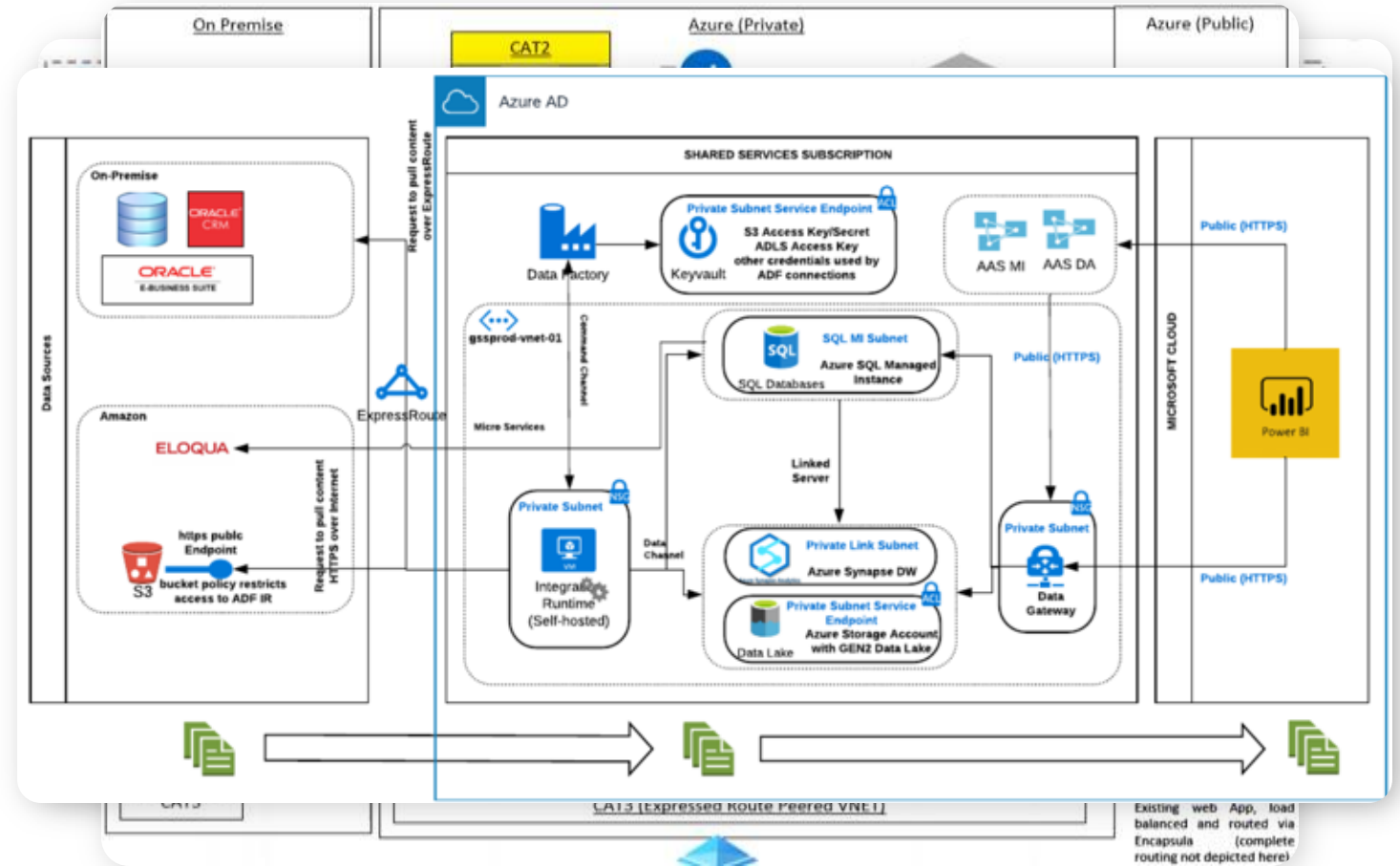
# Analytics is complex and fragmented

Every project has **many subsystems**

Every subsystem need a **different class of product**

Products often comes from **multiple vendors**

Integration is **complex, fragile and expensive**



Microsoft Cloud  
Dynamics 365 | Microsoft 365 | Power Platform

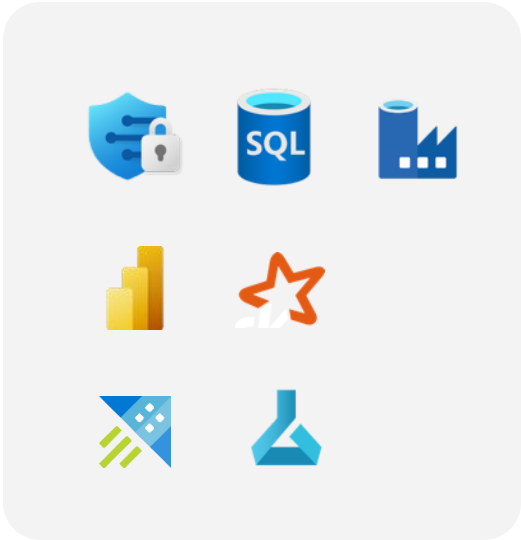
Data governance

Analytics

Operational databases

Microsoft Intelligent Data Platform  
Microsoft Fabric | SQL Server • Azure SQL | Azure Cosmos DB |  
Azure Databricks | Microsoft Purview | Azure AI

Leverage everything Microsoft has to offer



## Microsoft Fabric



Data Factory



Real-Time  
Intelligence



Databases



Analytics



Industry  
Solutions



Power BI



Partner  
Solutions



Copilot in Fabric



OneLake



Microsoft Purview

# Microsoft Fabric Capabilities

HITACHI



## Unify your analytics on a complete, governed platform

Reduce the cost and effort of integration with a unified, secure, and governed platform



## Empower every business user

Empower everyone to uncover insights accessible data, easy-to-use tools, and visuals embedded apps they use everyday



## Establish a trusted data foundation

Connect to any data from a single, multi-cloud data lake and use a single copy of data across engines

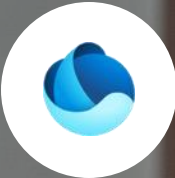


## Fuel your AI innovation

Adopt a data platform infused with AI at every layer to help you get more done, faster

# OneLake overview

HITACHI



# OneLake for all data

The OneDrive for data



**OneDrive**  
for documents

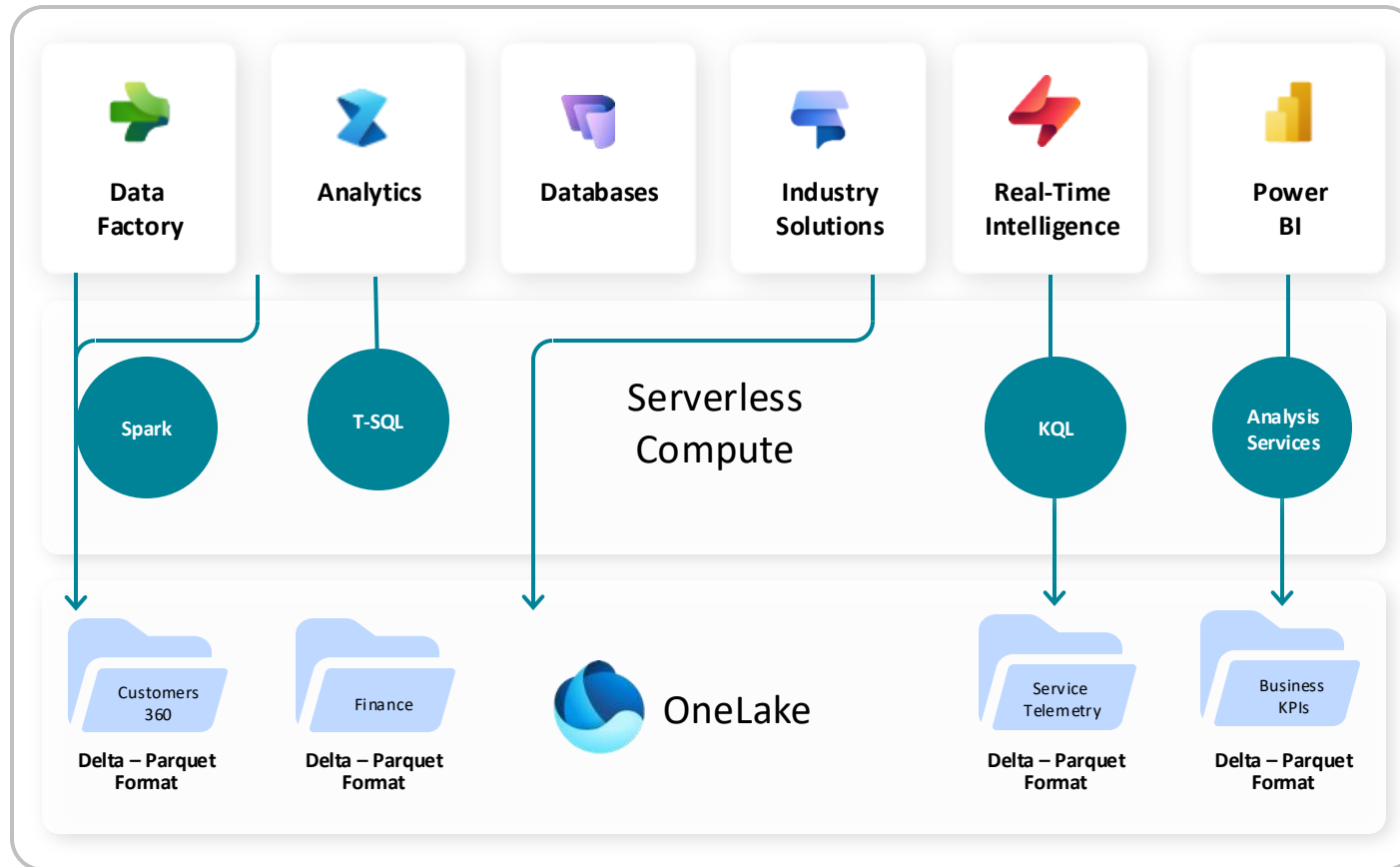


**OneLake**  
(One Drive)  
for data

**OneLake** provides a data lake  
as a service **without**  
**you needing to build it**



# At the center OneLake... “The OneDrive for Data”



All the compute engines store their data automatically in OneLake

The data is stored in a single common format

**Delta – Parquet**, an open standards format, is the storage format for all tabular data

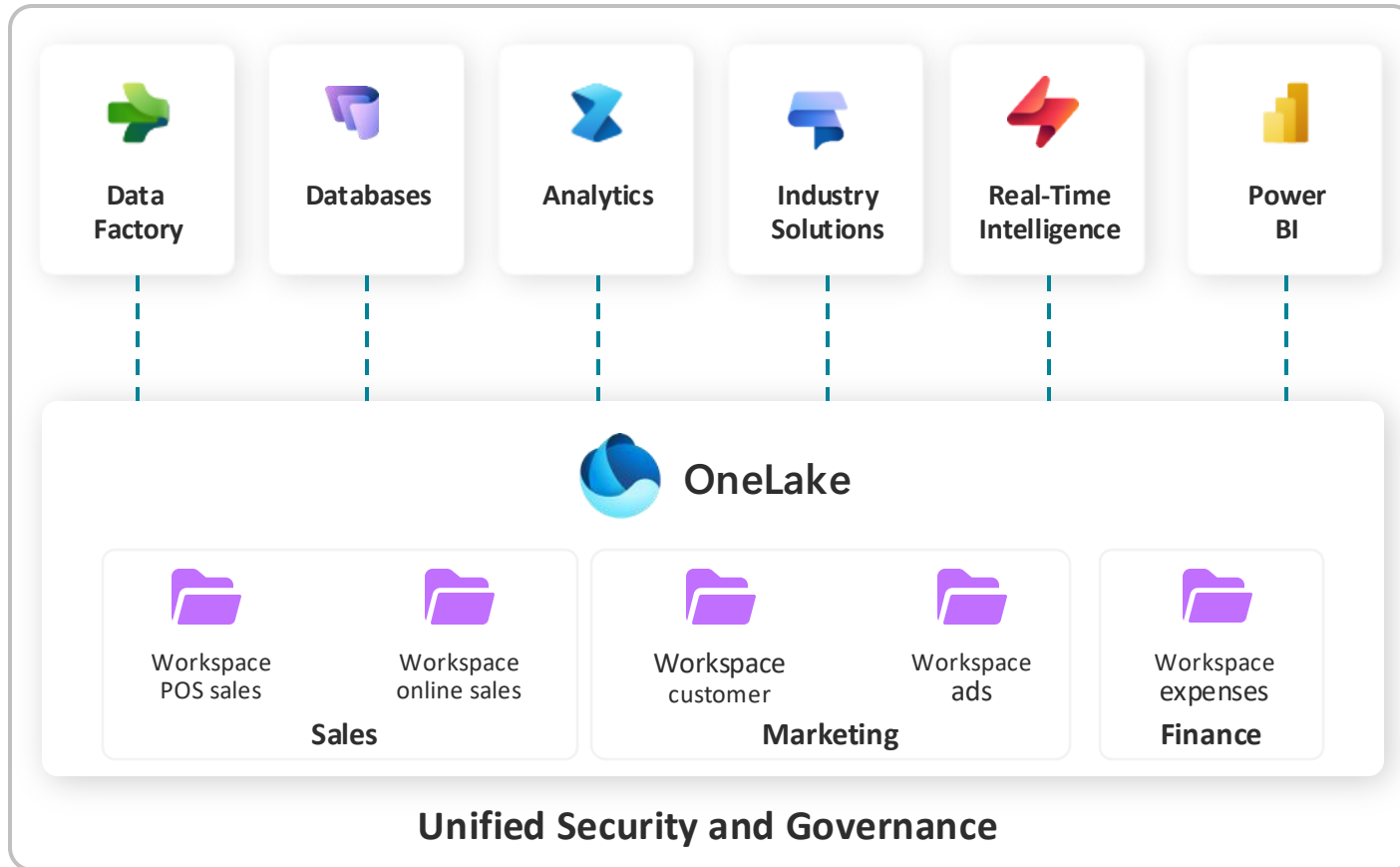
Once data is stored in the lake, it is directly accessible by all the engines without needing any import/export

All the compute engines have been fully optimized to work with Delta Parquet as their native format

Shared universal security model is enforced across all the engines

# OneLake for all domains

OneLake gives a true data mesh as a service



Introducing domains as an integral part of Fabric:  
A **domain** is a way to logically group together all the data in an organization relevant to an area or field, according to business needs.

Domains are defined with domain admins and contributors who can associate workspaces and group them together under a relevant domain.

Federated governance can be achieved by delegating settings to domain admins, thus allowing them to achieve more granular control over their business area.

Domains simplify discovery and consumption of data across the organization, thus allowing business optimized consumption.

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

# Fabric Capacities

HITACHI



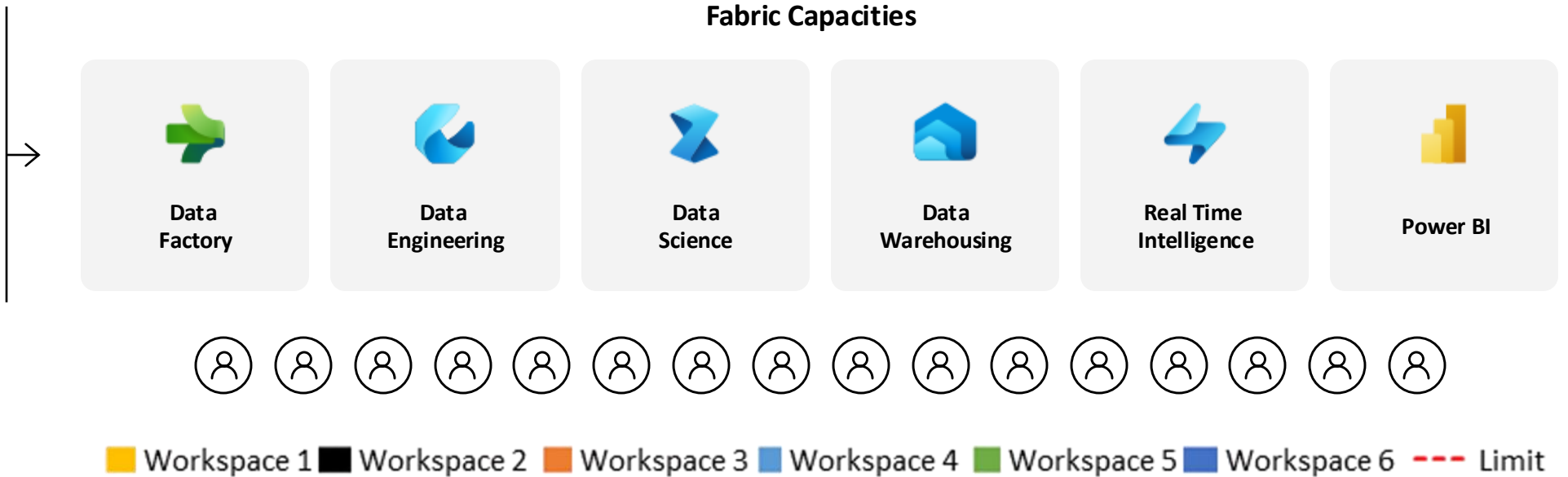
# Understanding Fabric Capacities



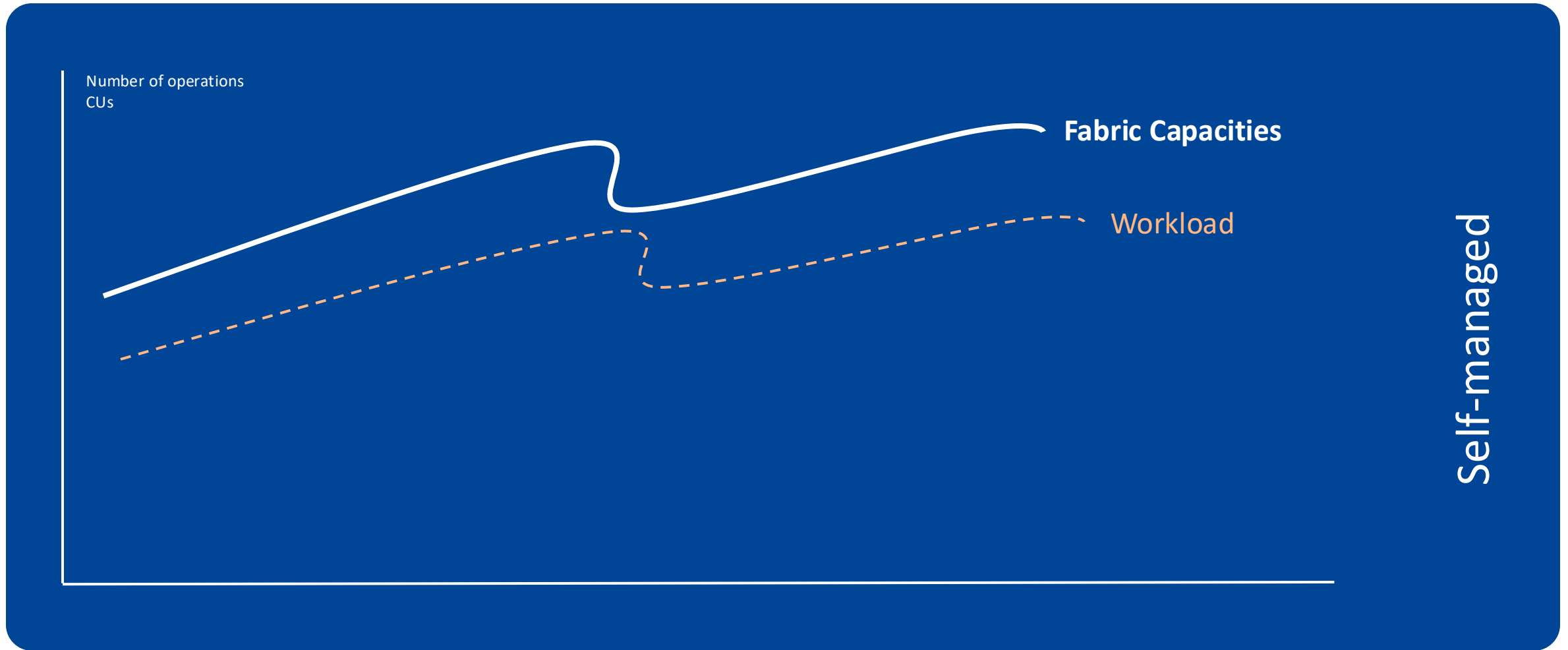
How fast can you go?



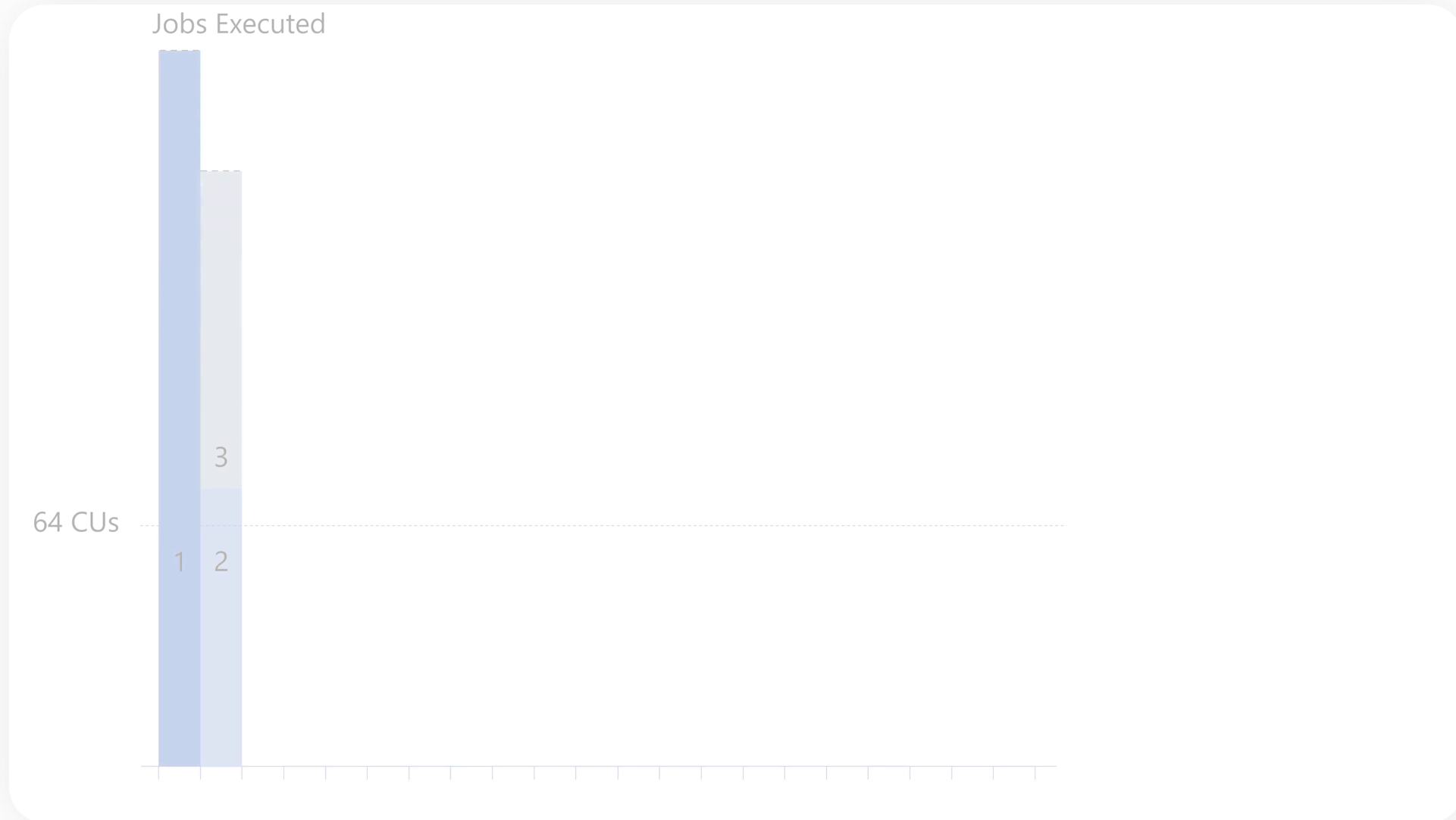
How far can you go?



# Understanding Capacities as Shared resources



# Bursting and Smoothing

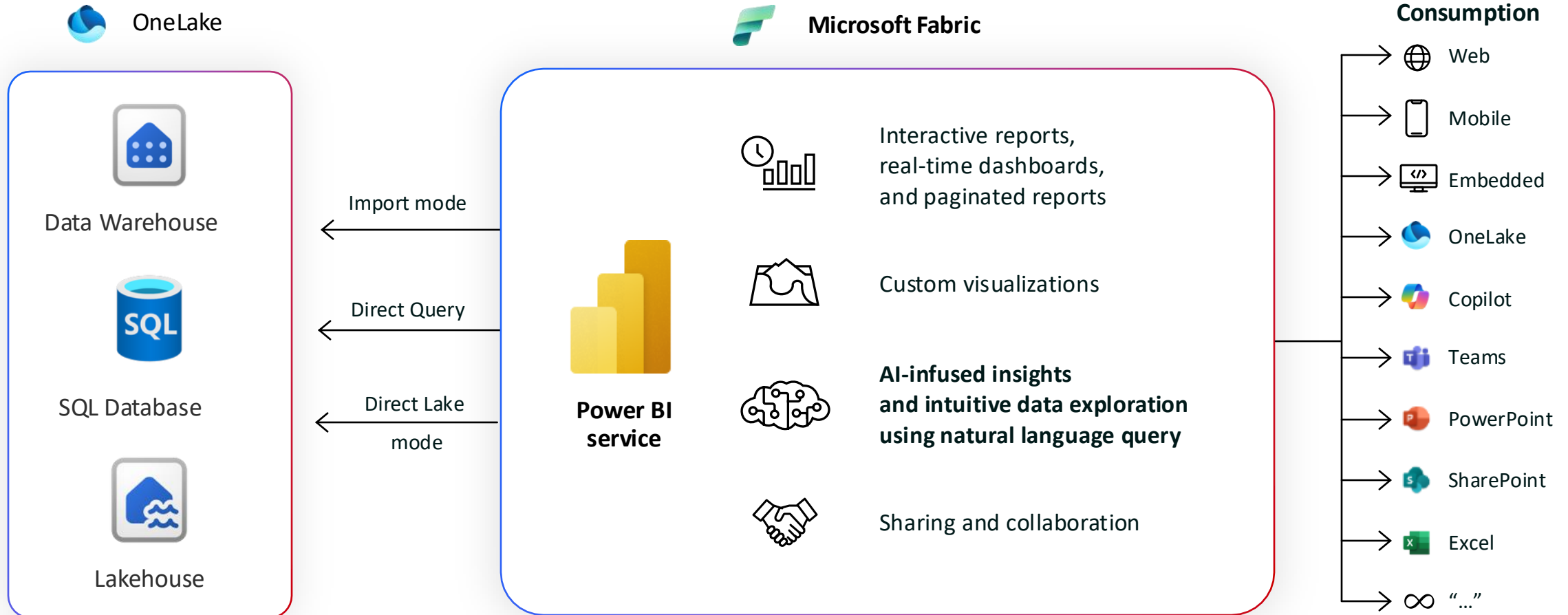


# Power BI



HITACHI

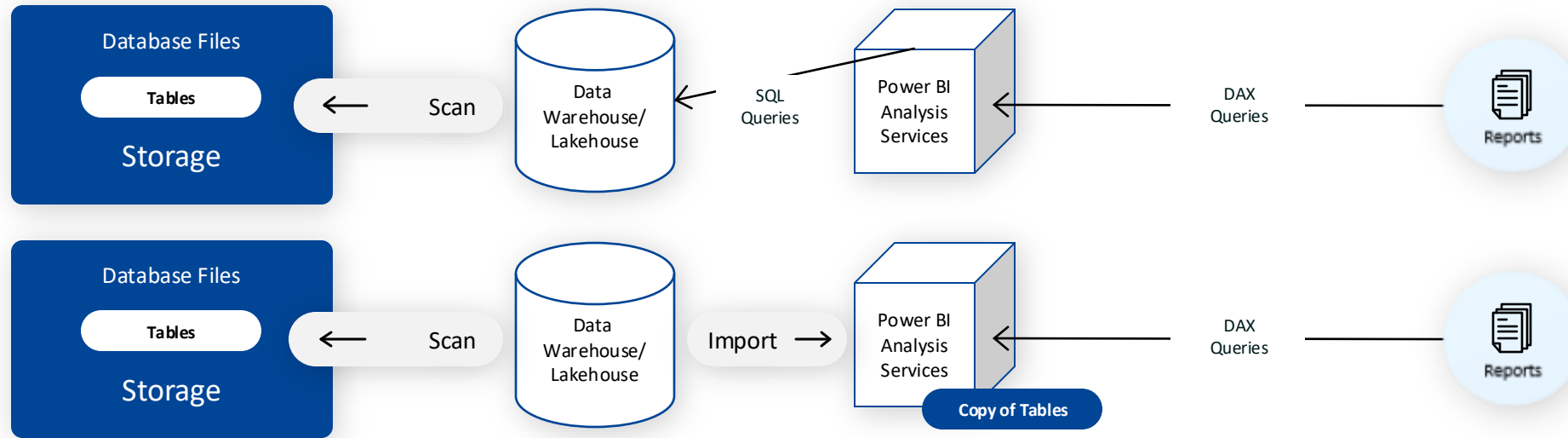
# Power BI: The bridge between data and decisions



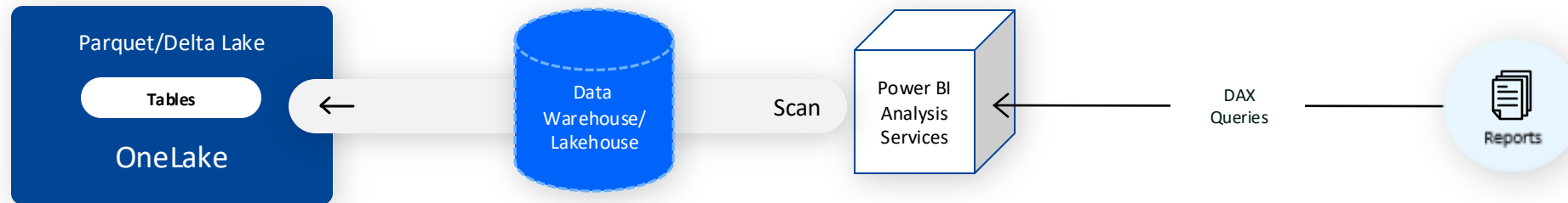
Powered by AI with Copilot in Microsoft Fabric

# Power BI : Direct Lake Mode

## Direct Query Mode. Slow, but real time



## Direct Lake Mode. Fast and real time



# Capacities compared to Power BI Premium



## Power BI Premium

- License is provisioned through Microsoft Office
- The smallest provisioned capacity provides 64 capacity units (CU)
- Power BI Premium SKUs cannot be paused but can be scaled through Office
- It can be used for all Fabric workloads

VS



## Microsoft Fabric

- Fabric capacity is provisioned through the Azure Portal
- The smallest provisioned capacity provides 2 capacity units (CU)
- Fabric capacity can be paused, resumed, and scaled via the Azure Portal or API

# Capacities compared to Power BI Premium

Microsoft Fabric		
Power BI Premium per capacity SKUs	Fabric Capacity Reservation SKUs	Fabric Capacity pay-as-you-go SKUs
<p>From</p> <p><b>\$4,995</b></p> <p>Per month for P1 SKU</p> <p>License your organization for access to Microsoft Fabric in a unified product experience that uses the same compute capacity and storage.<sup>3</sup></p> <ul style="list-style-type: none"> <li>Includes all the features available in Power BI Premium per user.</li> <li>Gain access to the rest of Microsoft Fabric workloads through a unified product experience and capacity.</li> </ul>	<p>From</p> <p><b>\$5,000</b></p> <p>Per month for F64 (P1 equivalent)</p> <p>License your organization for access to Microsoft Fabric in a unified product experience that uses the same compute capacity and storage.<sup>3</sup></p> <ul style="list-style-type: none"> <li>Smaller entry-level compute starting at F2.</li> <li>Microsoft Azure Consumption Commitment (MACC) eligible.</li> <li>Gain access to full Microsoft Fabric workloads through a unified product experience and capacity.</li> </ul>	<p>From</p> <p><b>\$8,409.60</b></p> <p>Per month for F64 SKU (P1 equivalent)</p> <p>License your organization for access to Microsoft Fabric in a unified product experience that uses the same compute capacity and storage.<sup>3</sup></p> <ul style="list-style-type: none"> <li>Small entry-level compute starting at F2.</li> <li>Microsoft Azure Consumption Commitment (MACC) eligible.</li> <li>Gain access to full Microsoft Fabric workloads through a unified product experience and capacity.</li> <li>Enjoy pay-as-you-go pricing, which you can scale up or down dynamically and even pause with no usage commitment.</li> </ul>
Power BI in Microsoft Fabric free account	Power BI Pro	Power BI Premium
<p>Free</p> <p>Create rich, interactive reports that put visual analytics at your fingertips using Power BI, included in your free account for Microsoft Fabric Preview.</p> <ul style="list-style-type: none"> <li>No credit card required.</li> <li>Upgrade to Power BI Pro or Power BI Premium to share reports.</li> </ul>	<p>Per user</p> <p><b>\$10</b></p> <p>Per user/month</p> <p>License individual users with modern, self-service analytics for publishing reports and dashboards and viewing content across your organization.</p> <ul style="list-style-type: none"> <li>Publish and share Power BI reports.</li> <li>Power BI Pro is included in <a href="#">Microsoft 365 E5</a>.</li> <li>Available to buy now with a credit card.<sup>1</sup></li> </ul>	<p>Per user</p> <p><b>\$20</b></p> <p>Per user/month<sup>2</sup></p> <p>License individual users with access to larger model sizes, more frequent refreshes, XMLA read/write, deployment pipelines, and other enterprise-scale features.</p> <ul style="list-style-type: none"> <li>Includes all the <a href="#">features</a> available with Power BI Pro.</li> <li>See additional Power BI Premium <a href="#">features</a>.</li> <li>Available to buy now with a credit card.<sup>1</sup></li> </ul>



**1<sup>st</sup> January 2025**

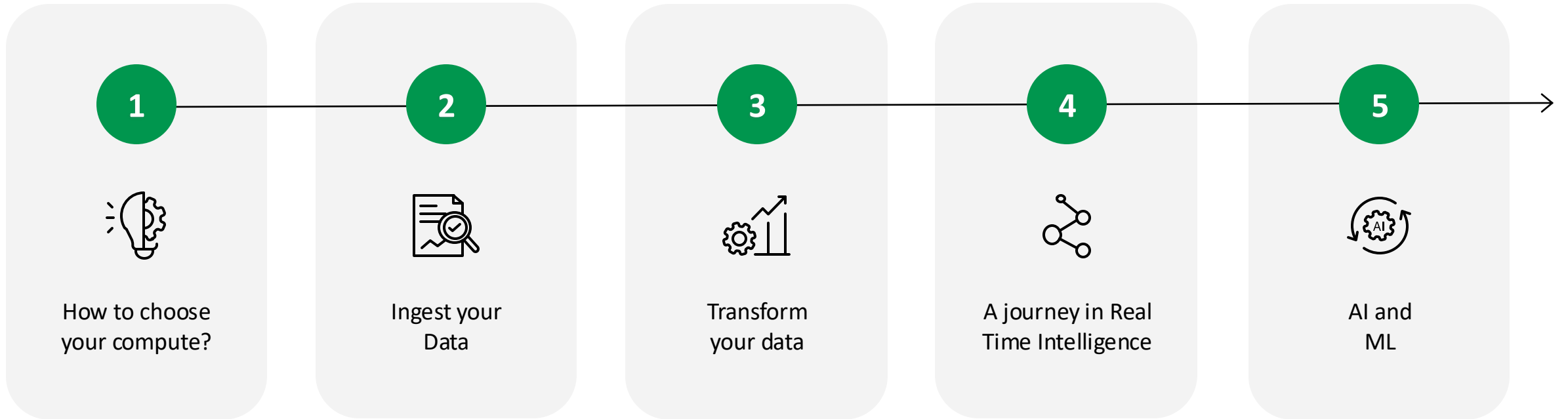
Power BI Premium capacities will have to be replaced by Fabric

- Capacity usage in Fabric is almost identical to capacity usage in Power BI Premium Per Capacity (PPC)
- If you currently have Power BI Premium Per Capacity, such as P1, you essentially have everything you need to begin using Fabric
- However, with Fabric comes the introduction of a new license type, the F SKU

# Capacities compared to Power BI Premium

HITACHI

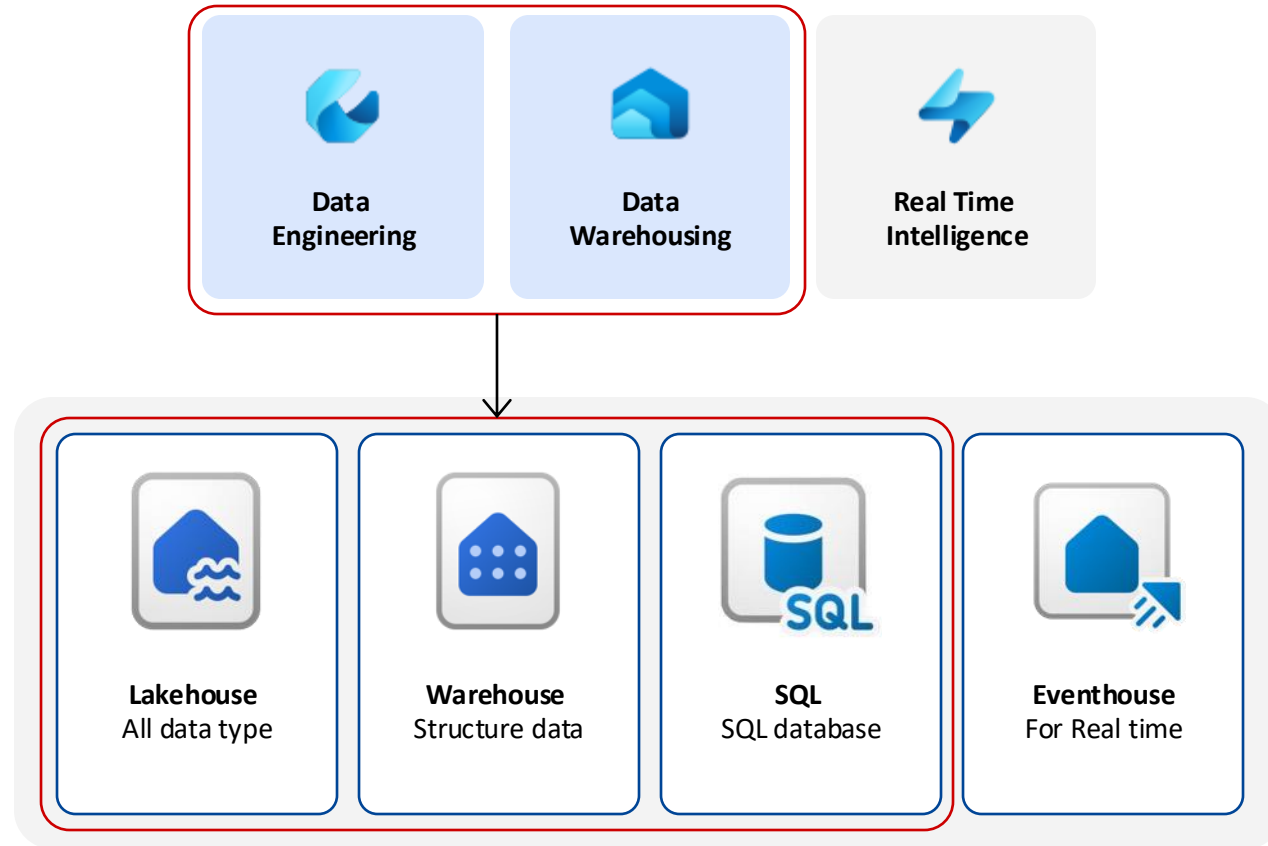
SKU <sup>+</sup>	Capacity Units (CU)	Power BI SKU	Power BI v-cores
F2	2	-	0.25
F4	4	-	0.5
F8	8	EM/A1	1
F16	16	EM2/A2	2
F32	32	EM3/A3	4
F64	64	P1/A4	8
Trial	64	-	8
F128	128	P2/A5	16
F256	256	P3/A6	32
F512	512	P4/A7	64
F1024	1024	P5/A8	128
F2048	2048	-	256



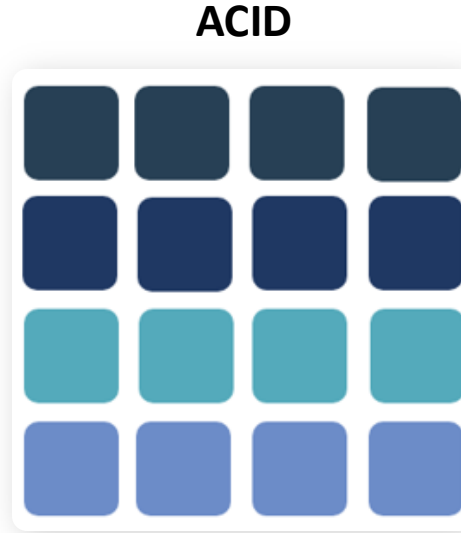
# Choose a compute

HITACHI

# Data Engineering and Data Warehouse



# What is a Lakehouse ?



Data lake



Warehouse



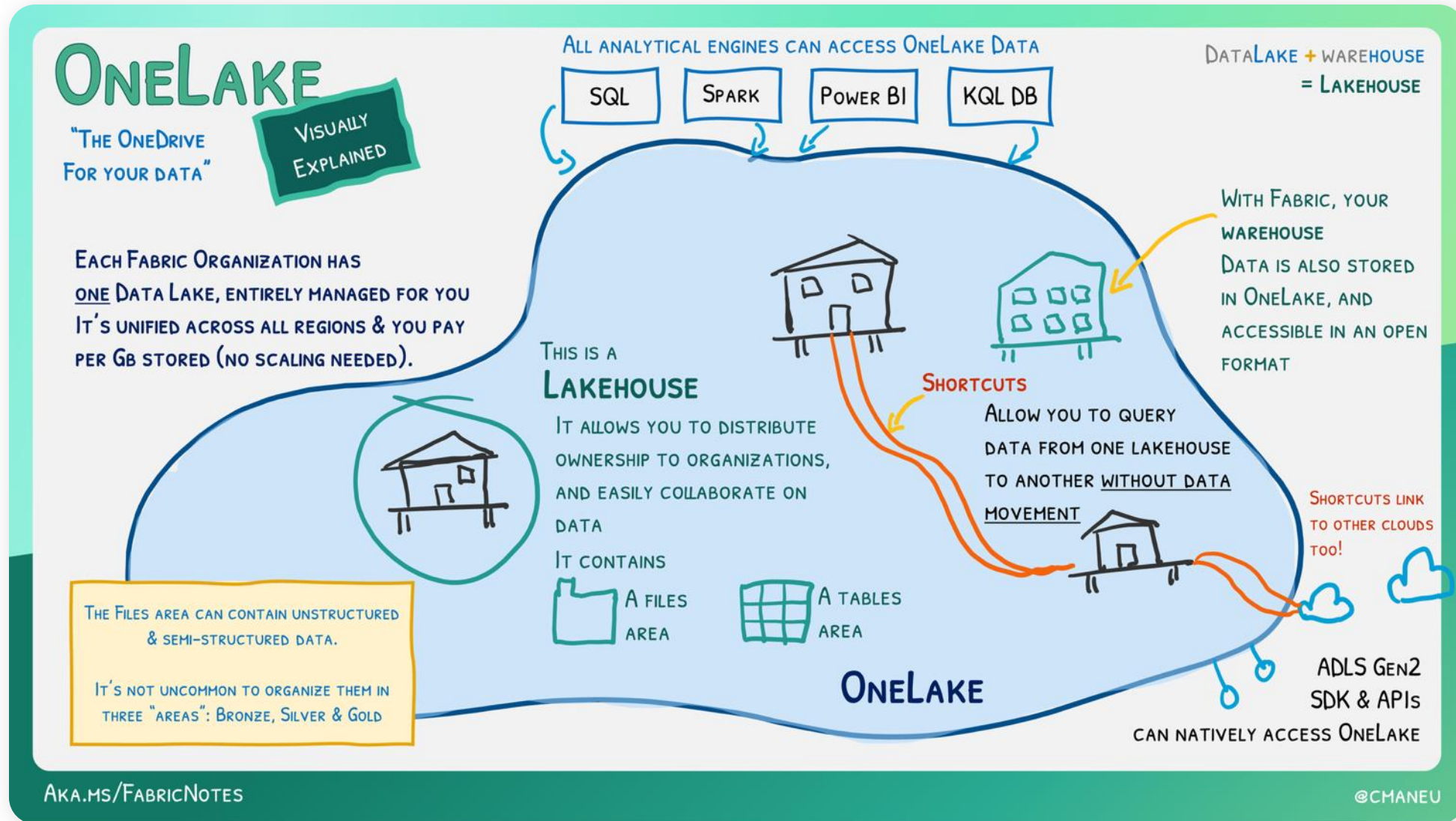
Consistency, Atomicity, isolation, and durability



Very expensive for high volume  
Multiple format



# What is a Lakehouse ?



# Fabric Warehousing

Fabric Warehouse is not a traditional enterprise Data Warehouse, it's a lake warehouse that supports two distinct warehousing items: **the Fabric data warehouse and the SQL analytics endpoint.**

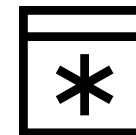
**Both utilize T-SQL.**

The Lakehouse SQL endpoint is strictly **read-only**, allowing for data querying without modification

Name	Type	Owner
Demo Workspace	Lakehouse	Chris Q. Public
Demo Workspace	Semantic model (default)	Contoso Admin
Demo Workspace	SQL analytics endpoint	Contoso Admin
Getdata	Data pipeline	Chris Q. Public
Notebook 1	Notebook	Chris Q. Public
SecurityTest	Warehouse	Chris Q. Public
SecurityTest	Semantic model (default)	Contoso Admin

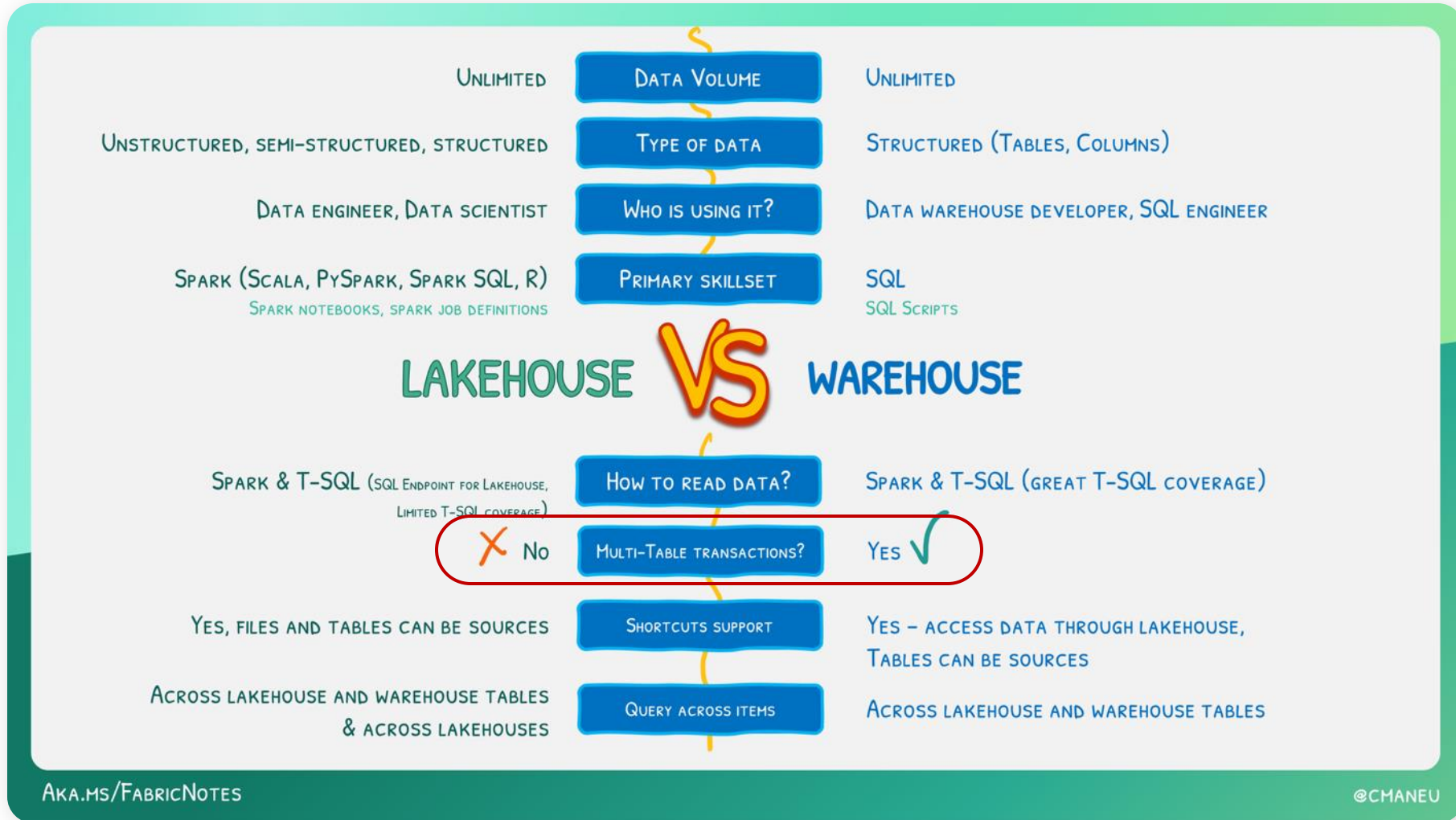


Data Warehouse



SQL analytics endpoint of the Lakehouse

# When do you choose Lakehouse?



# Break

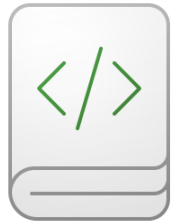
HITACHI

# Ingest and Transform your Data

HITACHI

# Data Extraction and transformations

The choice of tools



Notebooks



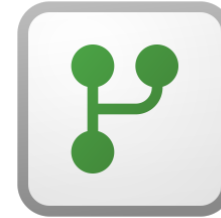
Spark Job  
Definitions



TSQL  
Stored  
Procedures



Pipelines



Dataflows  
Gen2



Copy  
Job

---

Code-based

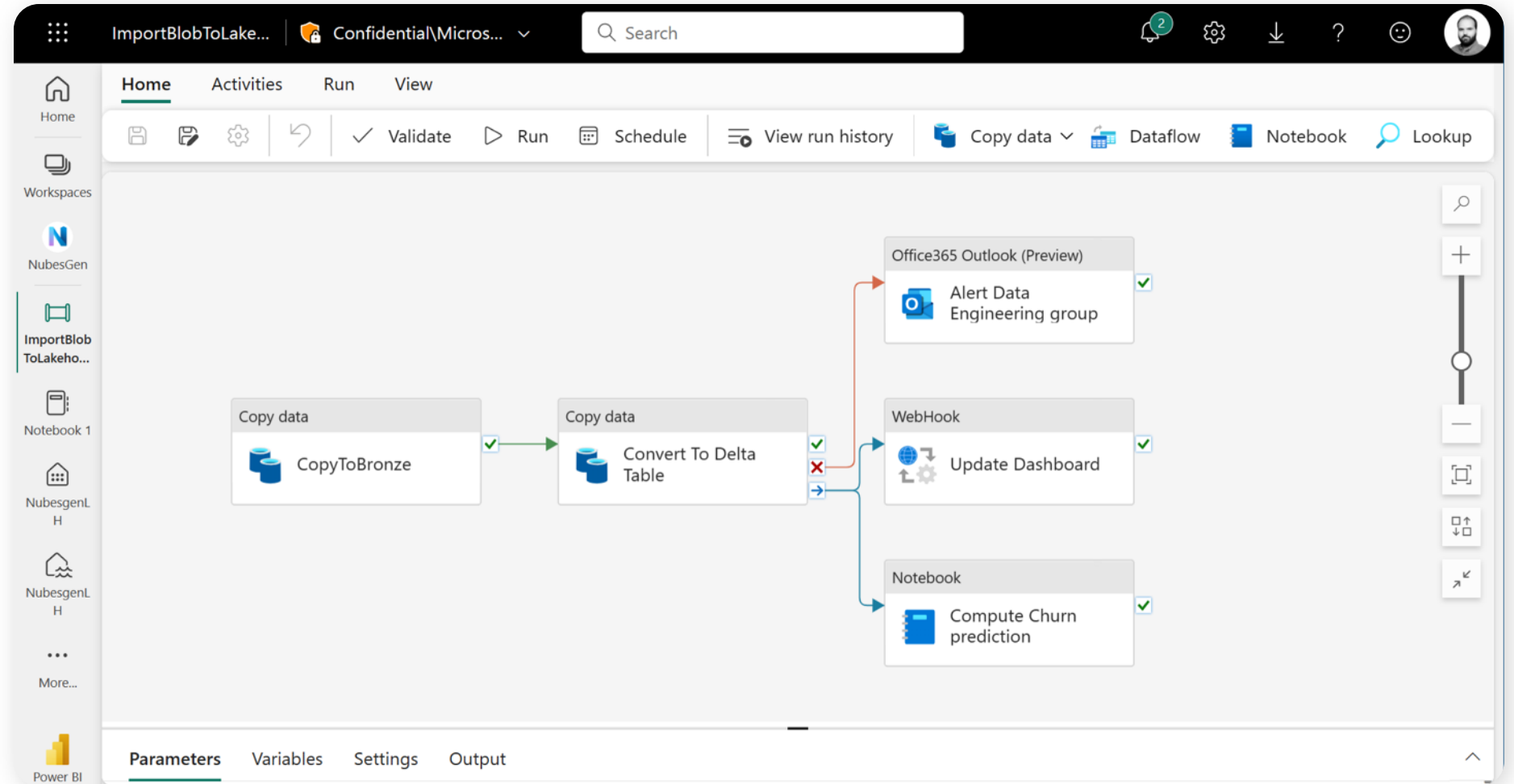
---

UI-based (Low code/no code)

# Pipelines



Pipelines



# Pipelines: Data Connectors



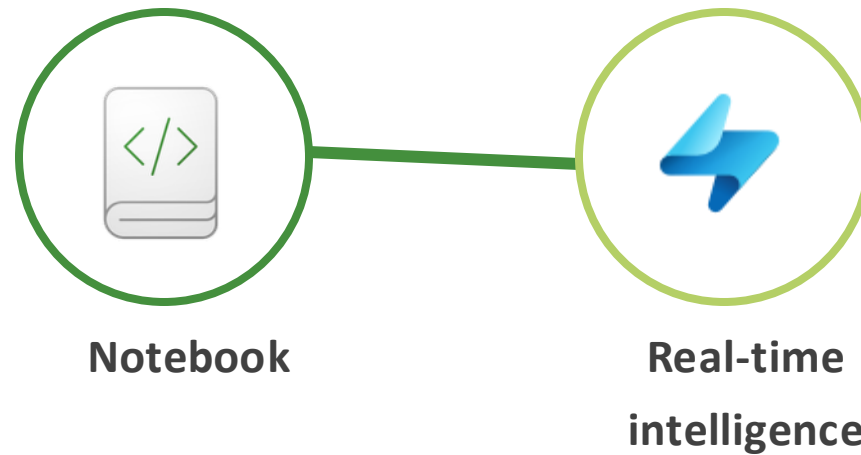
Pipelines

The screenshot displays the Microsoft OneLake data hub interface. At the top, there are navigation tabs: Home, OneLake data hub, Sample data, **New** (highlighted), and Azure. Below the navigation is a search bar. A filter bar contains tabs for All, File, Database, Power Platform, Azure, Online services, and Other. The main area is a grid of data connector cards, each with an icon, name, and category. The connectors are arranged in a grid that is 10 columns wide and 8 rows high, with the last cell in the bottom row empty.

Connector Name	Category
Folder	File
SQL Server database	Database
Oracle database	Database
IBM Db2 database	Database
MySQL database	Database
PostgreSQL database	Database
SAP HANA database	Database
Snowflake	Database
Google BigQuery	Database
Amazon Redshift	Database
Vertica	Database
Dataverse	Power Platform
Azure SQL database	Azure
Azure Synapse Analytics (S...	Azure
Azure Blobs	Azure
Azure Tables	Azure
Azure Data Explorer (Kusto)	Azure
Azure Data Lake Storage Ge...	Azure
SharePoint Online list	Online services
Salesforce objects	Online services
OData	Other
Odbc	Other
FTP	File
Oracle Cloud Storage	File
SFTP	File
Amazon RDS for SQL Server	Database
Azure Database for Postgre...	Database
Azure SQL Managed Instance	Database
MariaDB for Pipeline	Database
MongoDB Atlas for Pipelines	Database
MongoDB for Pipeline	Database
Azure Cosmos DB for Mong...	Azure
Azure Cosmos DB v2	Azure
Azure Database for MySQL	Azure
Azure Files	Azure
Amazon S3	Other
Amazon S3 Compatible	Other
Dynamics 365	Other
Dynamics AX	Other
Dynamics CRM	Other
Google Cloud Storage	Other
Http	Other
Microsoft365	Other
REST	Other

# Notebooks and Fabric experiences

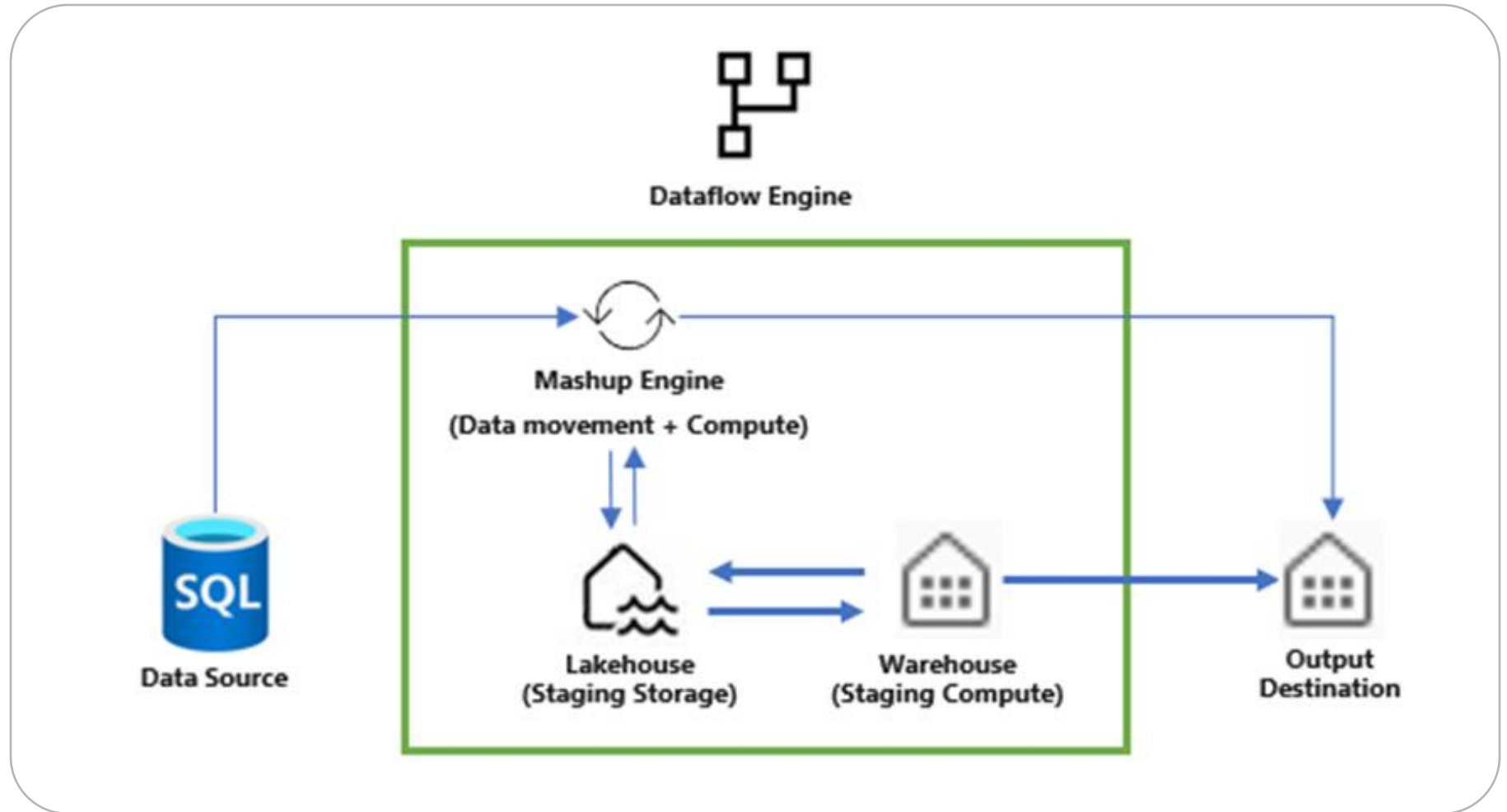
Fabric Notebook is a powerful tool that can be used across all the experiments available in fabric.  
The execution of notebooks can be orchestrated using pipelines



# Dataflow Gen2



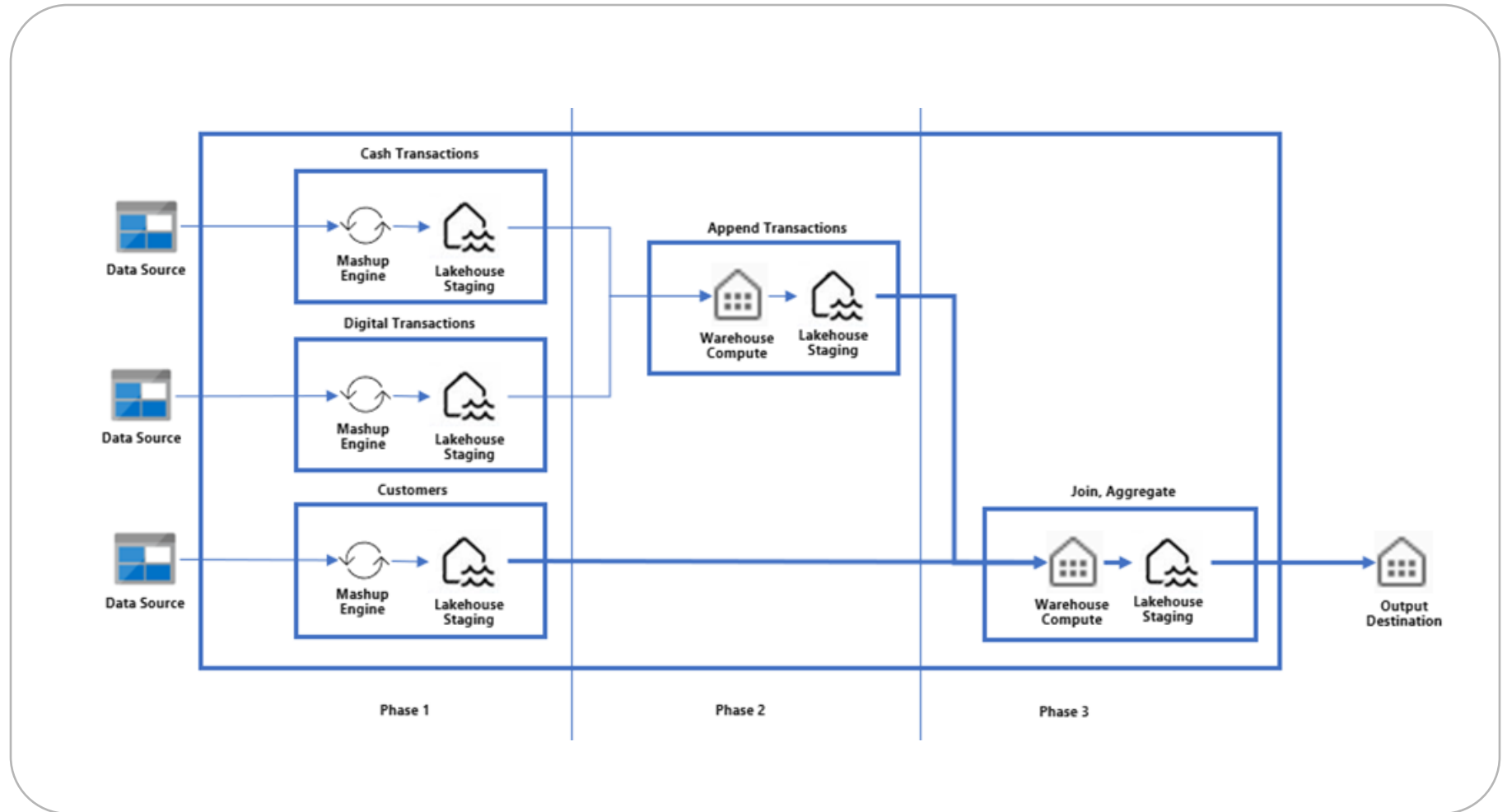
Dataflow Gen2



# Dataflow Gen2



Dataflow Gen2



# Dataflow Gen2: Data Connectors



Dataflow  
Gen2

**Choose data source**

Select a connector or directly drag a file from your computer.

All categories File Database Power Platform Azure Online services Other

Search

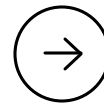
Excel workbook File	Text/CSV File	XML File	JSON File
Folder File	PDF File	Parquet File	SharePoint folder File
SQL Server database Database	Access Database	SQL Server Analysis Services Database	Oracle database Database
IBM Db2 database Database	MySQL database Database	PostgreSQL database Database	Teradata database Database
SAP HANA database Database	SAP BW Application Server Database	SAP BW Message Server Database	Snowflake Database
Google BigQuery Database	Amazon Redshift Database	Impala Database	Vertica Database
Dataflows Power Platform	Power BI dataflows (Legacy) Power Platform	Dataverse Power Platform	Common Data Service (Legacy) Power Platform
Azure SQL database Azure	Azure Synapse Analytics (SQL DW) Azure	Azure Analysis Services Azure	Azure Blobs Azure
Azure Tables Azure	Azure Data Explorer (Kusto) Azure	Azure Data Lake Storage Gen2 Azure	Azure HDInsight Spark Azure
SharePoint Online list	Microsoft Exchange Online	Salesforce objects	Salesforce reports

## Simple



### Dataflow Gen2

Use **Dataflow Gen2** if you're looking for a simpler, more accessible way to prepare data, particularly for Power BI or similar analytics tools. It's ideal for scenarios that don't require extensive custom coding or **automation**.



## Complex and automation



### Pipelines

Use **Pipelines** if your data integration requirements are complex, involve multiple sources and transformations, or require detailed control and automation. It's the go-to solution for large-scale, enterprise-level data workflows that demand robustness and flexibility

# What is Copy Job?

Simple steps to move your data ...



Copy Job

- 
- 1** Create a copy Job

---

  - 2** Choose source type and connect

---

  - 3** Select your destination

---

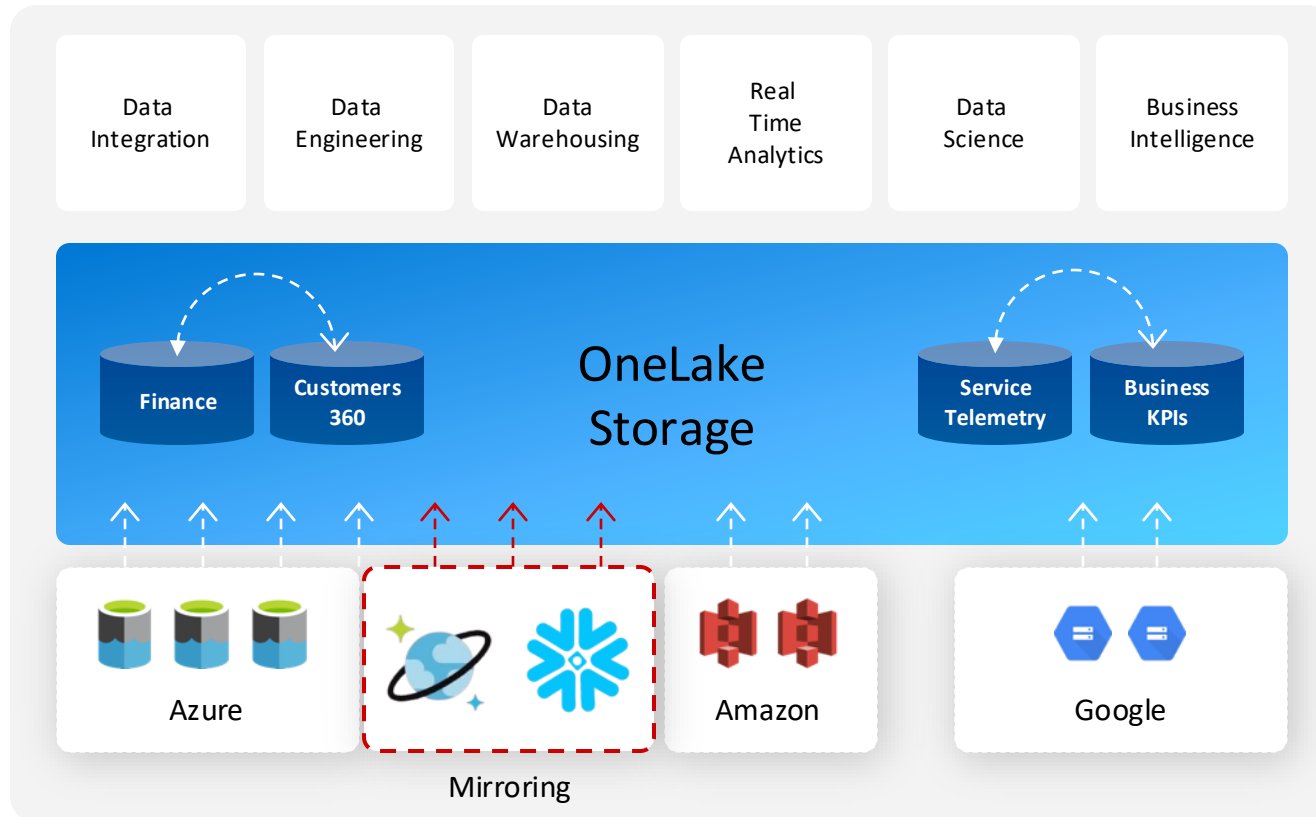
  - 4** Configure table or column mapping if you need

---

  - 5** Choose the copy mode, either a one-time full data copy, or continuous incremental copying.

---

# Fabric Shortcuts and Mirroring



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

With **shortcuts or Mirroring**, data throughout OneLake can be composed together without having to write complex pipelines.

Shortcuts also allow instant linking of data already existing in Azure and in other clouds, without any data duplication and movement, making **OneLake a 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 in Fabric?

## ONELAKE SHORTCUTS

SHORTCUTS ALLOW YOU TO CREATE A VIRTUALIZED DATA LAKE, ELIMINATING COPIES OF DATA BETWEEN ORGANIZATION DOMAINS, ANALYTICAL ENGINES, OR CLOUDS

### INTERNAL

FROM A LAKEHOUSE FOLDER TO A LAKEHOUSE FOLDER

FROM A LAKEHOUSE TABLE TO A LAKEHOUSE TABLE

FROM A WAREHOUSE TABLE TO A LAKEHOUSE TABLE

FROM A KQL DATABASE TABLE TO A LAKEHOUSE FOLDER

### EXTERNAL

FROM AN ADLS GEN2/S3 FOLDER TO A LAKEHOUSE FOLDER

FROM AN ADLS GEN2/S3 FOLDER TO A TABLE

S3 SHORTCUTS ARE READ-ONLY

AKA.MS/FABRICNOTES @CMANEU

The diagram illustrates various types of OneLake shortcuts. Internal shortcuts include: 1) From a Lakehouse Folder to a Lakehouse Folder (e.g., ContosoThreadsLH staging to FinanceLH staging). 2) From a Lakehouse Table to a Lakehouse Table (e.g., ContosoThreadsLH fact\_sale to FinanceLH fact\_sale). 3) From a Warehouse Table to a Lakehouse Table (e.g., ContosoThreadsDWH Agg\_NewCustomersPerMonth to FinanceLH Agg\_NewCustomersPerMonth). External shortcuts include: 4) From a KQL Database Table to a Lakehouse Folder (e.g., NubesgenRT events to ContosoThreadsLH events). 5) From an ADLS Gen2/S3 Folder to a Lakehouse Folder (e.g., ContosoThreadsLH agg\_customermmonth to ContosoThreadsLH agg\_customermmonth/). 6) From an ADLS Gen2/S3 Folder to a Table (e.g., ContosoThreadsLH s3lake to ContosoThreadsLH s3lake). A warning icon indicates that S3 shortcuts are read-only.

# Mirroring Types

No ETL

## Database mirroring

Allows replication of **entire databases and tables**, allowing you to bring data from various systems together into a single analytics platform

## Metadata mirroring

**Synchronizes metadata** (such as catalog names, schemas, and tables) **instead of physically moving the data**.

This approach leverages [shortcuts](#), ensuring the **data remains in its source** while still being easily accessible within Fabric.

## Open mirroring

is designed to **extend mirroring based on open Delta Lake table** format.

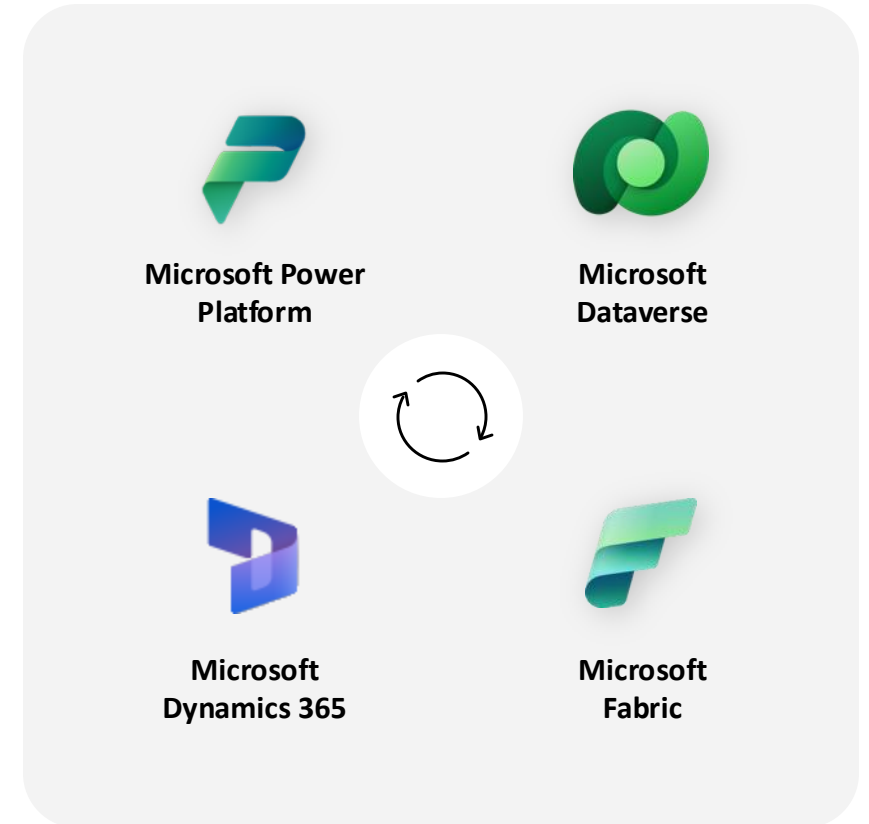
This capability enables any developer to write their application's change data directly into a mirrored database item in Microsoft Fabric, based on the open mirroring approach and **public APIs**.

# Dataverse makes it easy ...

To connect Power Platform and Dynamics 365 to Fabric

HITACHI

- No Copy. No ETL.
- Direct Connection via Dataverse
- Insights democratized to all low code apps and business using Fabric's 7 core workloads
- Makers informed by insights improves quality of applications
- Data is governed



# Speaker



**Sarah Coward**  
Fabric Product Owner

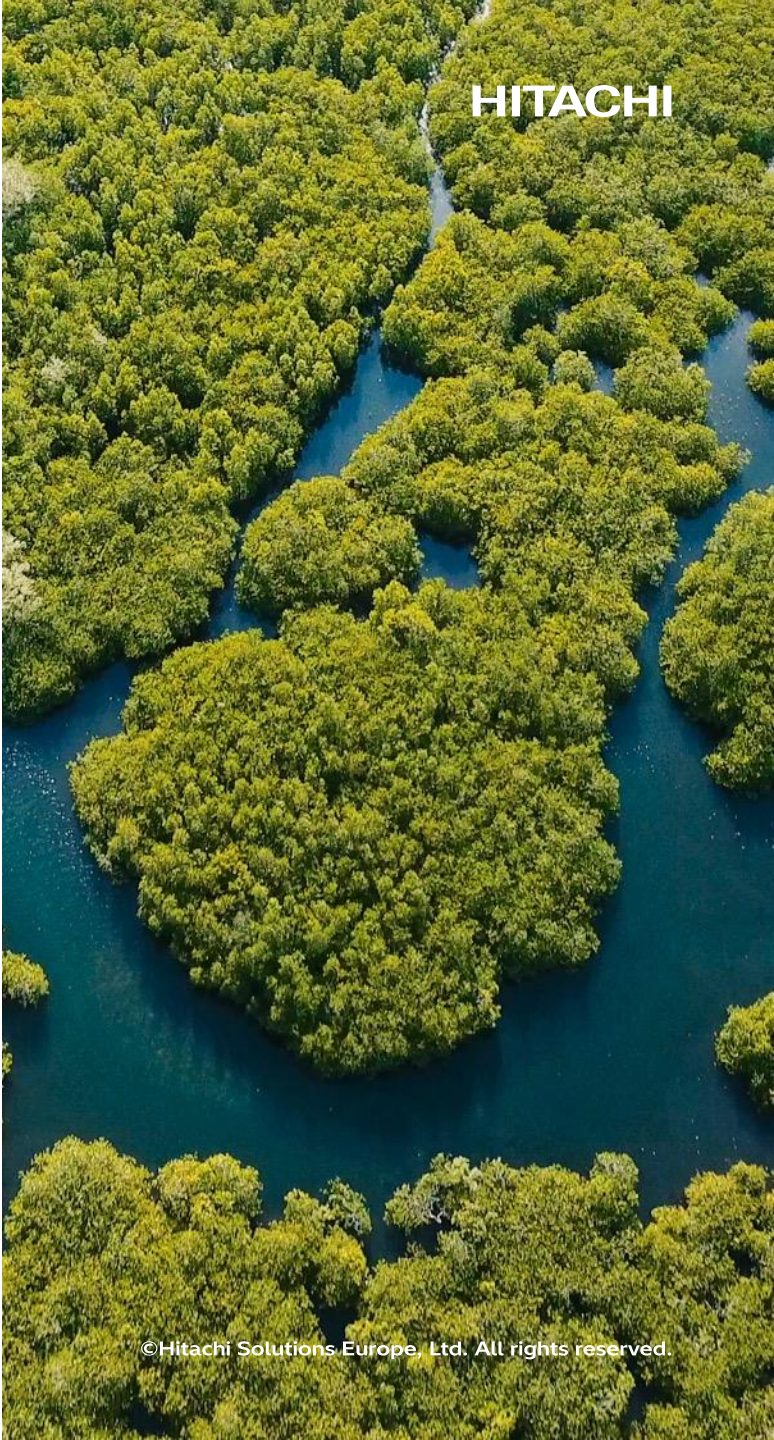


Department  
for Environment  
Food & Rural Affairs



Environment  
Agency

HITACHI



# Lab 1: Bring data into a Lakehouse in Fabric



# Lunch

HITACHI

# Apache Spark

HITACHI



# Introduction to Apache Spark



**Apache Spark** is an open-source framework for **large-scale and distributed data processing**.



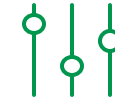
It's popular for "**big data**" processing and is available in multiple Microsoft platforms:

**Azure HDInsight**

**Azure Synapse Analytics**

**Azure Databricks**

**Microsoft Fabric**



Spark in Fabric is used for **data ingestion, processing, analysis and Machine learning** with lakehouse at the core.

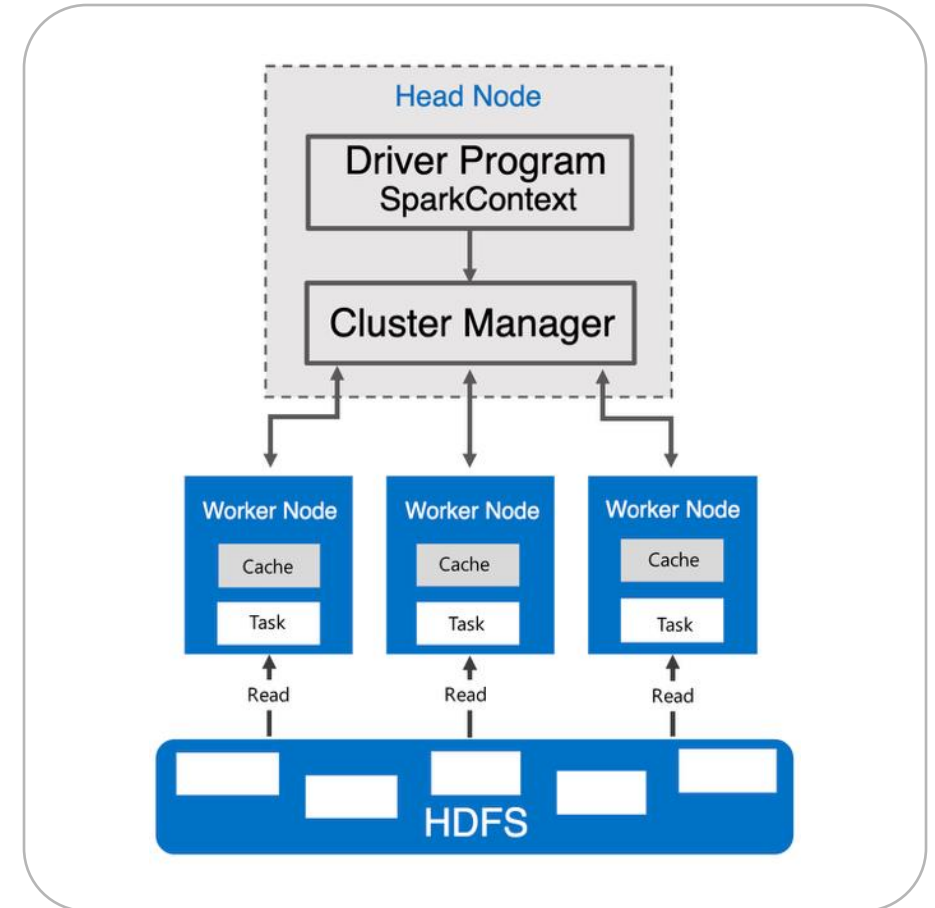
# Basics of Apache Spark

Uses a **"divide and conquer"** strategy by spreading the workload across multiple machines.

Distribution of tasks and results aggregation is managed by Spark.

## Microsoft Fabric's Spark pool approach:

- Starter pools available by default, spins up in 30 seconds
- Custom pools with number of nodes and cluster sizes can be created
- Easy creation and configuration of Spark environments



# Scaling Spark Clusters



**Workspace settings**

- General
- License info
- Azure connections
- System storage
- Git integration
- OneLake
- Workspace identity
- Network security
- Monitoring
- Power BI
- Data Engineering/Science
- Spark settings**

**Spark settings**

Configure and manage settings for Spark workloads and the default environment for the workspace.

Pool Environment Jobs High concurrency Automatic log

**Default pool for workspace**

Use the automatically created starter pool or create custom pools for workspaces and items in the capacity. If the setting Customize compute configurations for items is turned off, this pool will be used for all environments in this workspace.

StarterPool

Pool details		
Node family	Node size	Number of nodes
Memory optimized	Medium	1 - 10

**Customize compute configurations for items**  On

When turned on, users can adjust compute configuration for individual items such as notebooks and Spark job definitions. [Learn more about Customize compute configurations for items](#)

The screenshot displays a Databricks Notebook titled "Sales Analytics Notebook". The interface includes a top navigation bar with "Home", "Edit", "AI tools", "Run", and "View" tabs. Below this is a toolbar with "Run all", "Standard session", "PySpark (Python)", "Environment", "Workspace default", "Data Wrangler", and other icons. The left sidebar shows the "Explorer" view with a tree structure under "fabric\_lakehouse" containing "Tables" and "Files" folders, with "orders" under "Files".

The main content area shows a code cell with the following PySpark code:

```
1 from pyspark.sql.types import *
2
3 orderSchema = StructType([
4     StructField("SalesOrderNumber", StringType()),
5     StructField("SalesOrderLineNumber", IntegerType()),
6     StructField("OrderDate", DateType()),
7     StructField("CustomerName", StringType()),
8     StructField("Email", StringType()),
9     StructField("Item", StringType()),
10    StructField("Quantity", IntegerType()),
11    StructField("UnitPrice", FloatType()),
12    StructField("Tax", FloatType())
13 ])
14
15 df = spark.read.format("csv").schema(orderSchema).load("Files/orders/2019.csv")
16
17 display(df)
```

Below the code cell, a status bar indicates: "[2] ✓ 3 sec - Command executed in 3 sec 488 ms by System Administrator on 3:37:44 PM, 5/19/25".

The output of the code is a table view with 9 columns and 1000 rows. The visible data is as follows:

ABC SalesOrderNumber	123 SalesOrderLineNumber	ABC OrderDate	ABC CustomerName	ABC Email	ABC Item	123 Quanti
1 SO43701	1	2019-07-01	Christy Zhu	christy12@...	Mountain-...	1
2 SO43704	1	2019-07-01	Julio Ruiz	julio1@adv...	Mountain-...	1
3 SO43705	1	2019-07-01	Curtis Lu	curtis9@ad...	Mountain-...	1
4 SO43700	1	2019-07-01	Ruben Prasad	ruben10@...	Road-650 ...	1

- Ideal for **interactive exploration** and analysis using Spark
- Allows for a mix of text, images, and code in multiple languages
- Consists of cells that can contain **markdown content** or **executable code**
- Code can be **executed interactively** with immediate results
- Facilitates collaboration and sharing

# Notebooks

Notebooks can be used by multi skills teams each one using the language they know best.  
In Fabric you can create new notebooks or import existing ones.



Notebooks



Pyspark  
(Python)



Spark  
(scala)



Spark  
SQL



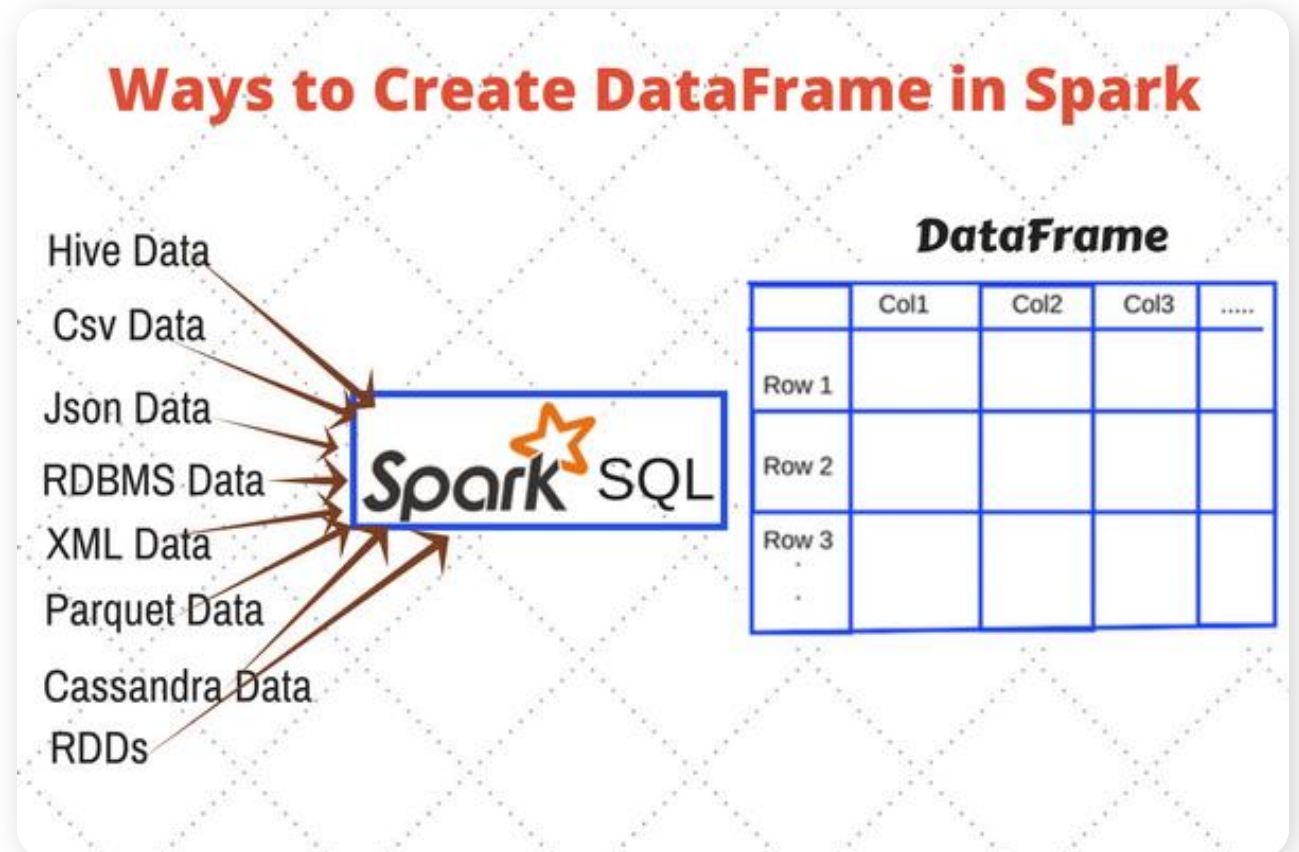
SparkR

# Working with Dataframes

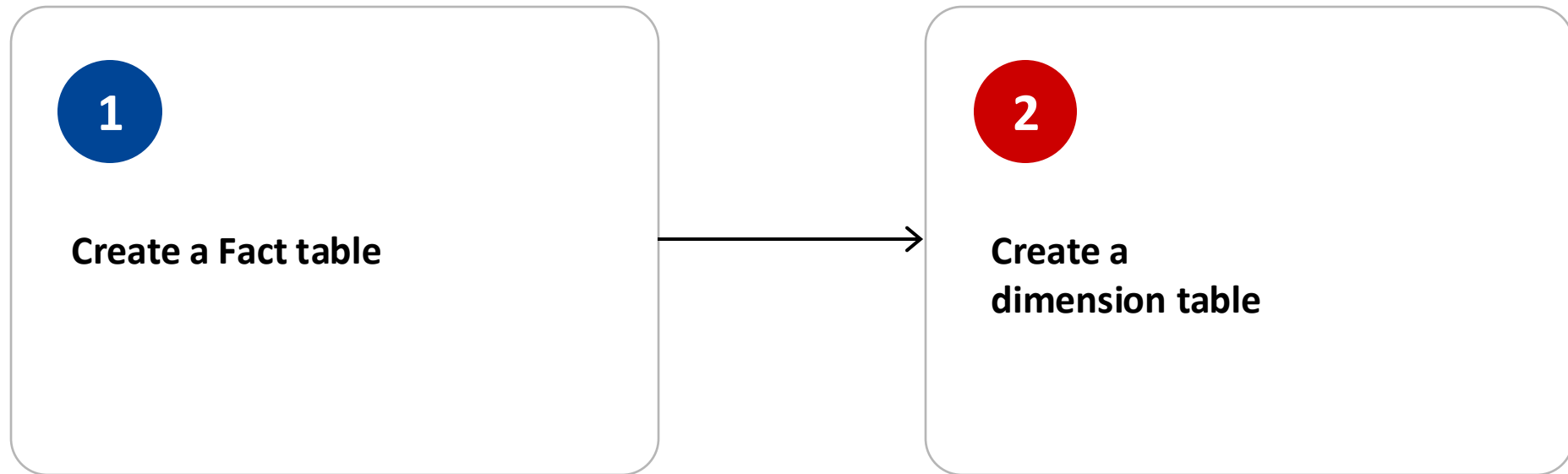
Spark natively uses a data structure called a **resilient distributed dataset (RDD)**.

Spark **dataframes** ~= Pandas dataframes in Python.

*Optimized for Spark's distributed processing*

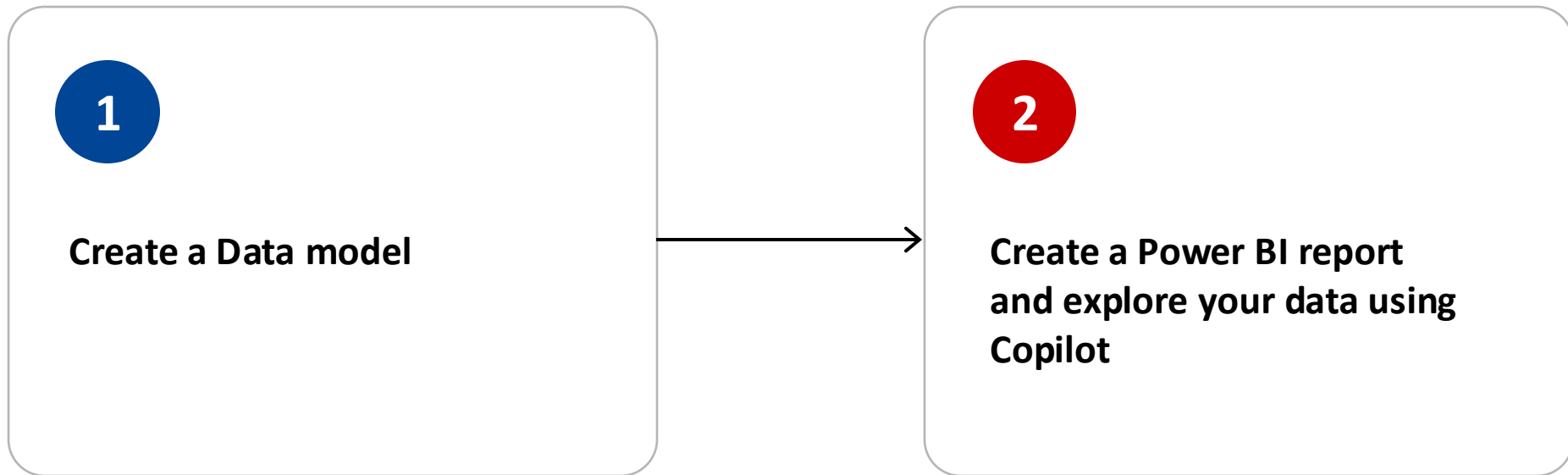


# Lab 2: Data Transformation



# Break

HITACHI



# Real Time intelligence

HITACHI





# Microsoft Fabric

The unified data platform for the era of AI



Data  
Factory



Synapse Data  
Engineering



Synapse Data  
Science



Synapse Data  
Warehousing



Real-Time  
Intelligence



Power BI



Industry Data  
Solutions



Real-Time Hub



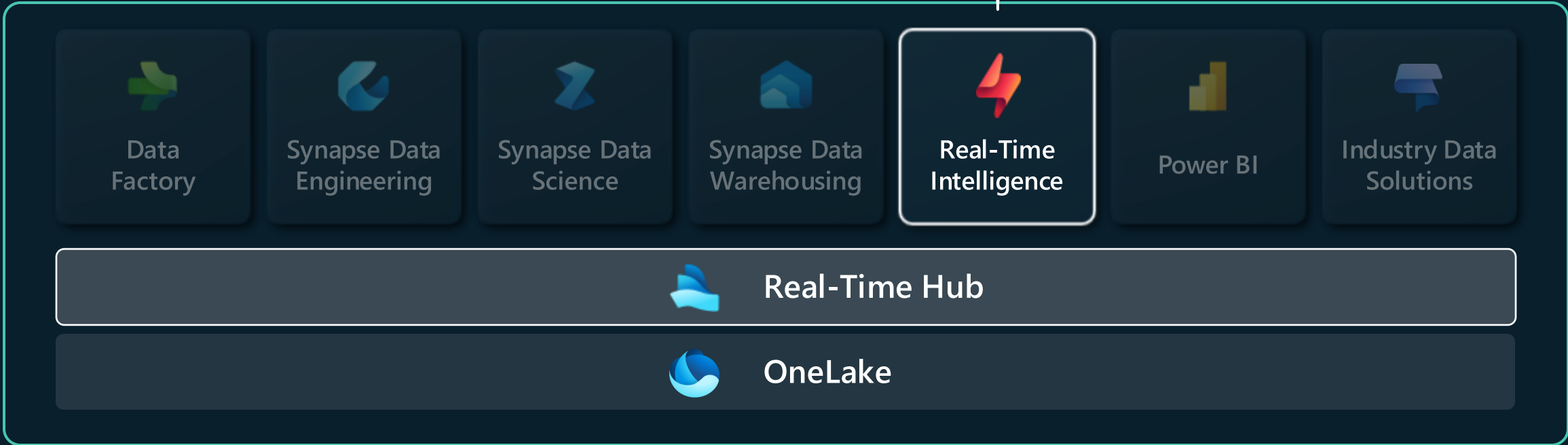
OneLake



Purview

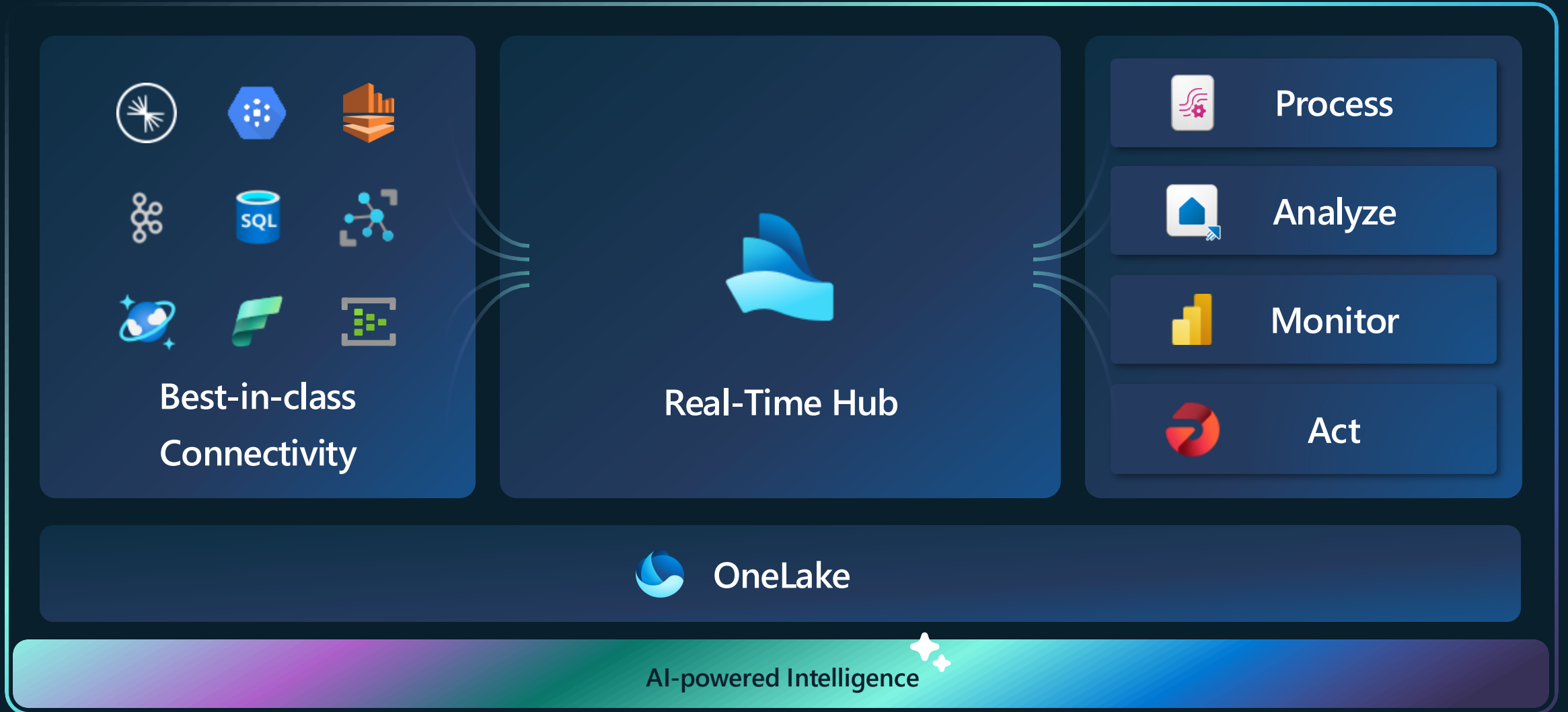


# Real-Time Intelligence in Microsoft Fabric





# Real-Time Intelligence in Microsoft Fabric





# Real-Time Intelligence in Microsoft Fabric



## Enterprise real-time data platforms

Azure Event Hub  
Azure Event Grid  
Azure Stream Analytics  
Azure Data Explorer



## Self-service reporting and activation

OneLake  
Activator  
Power BI



## Real-Time Intelligence in Microsoft Fabric

Fully integrated,  
no/low-code real-time SaaS  
data platform



# Real-Time Intelligence in Microsoft Fabric

## Complete SaaS solution

Everything, unified

---

SaaS experience

---

Lightweight modeling

## Single data estate & open

Events out-of-the-box

---

Open data format

---

Data rests in OneLake

## Rapid app development

Reuse of models

---

Multi-tenant security

---

Event pub-sub

## Real-time AI insights

Non-obvious insights

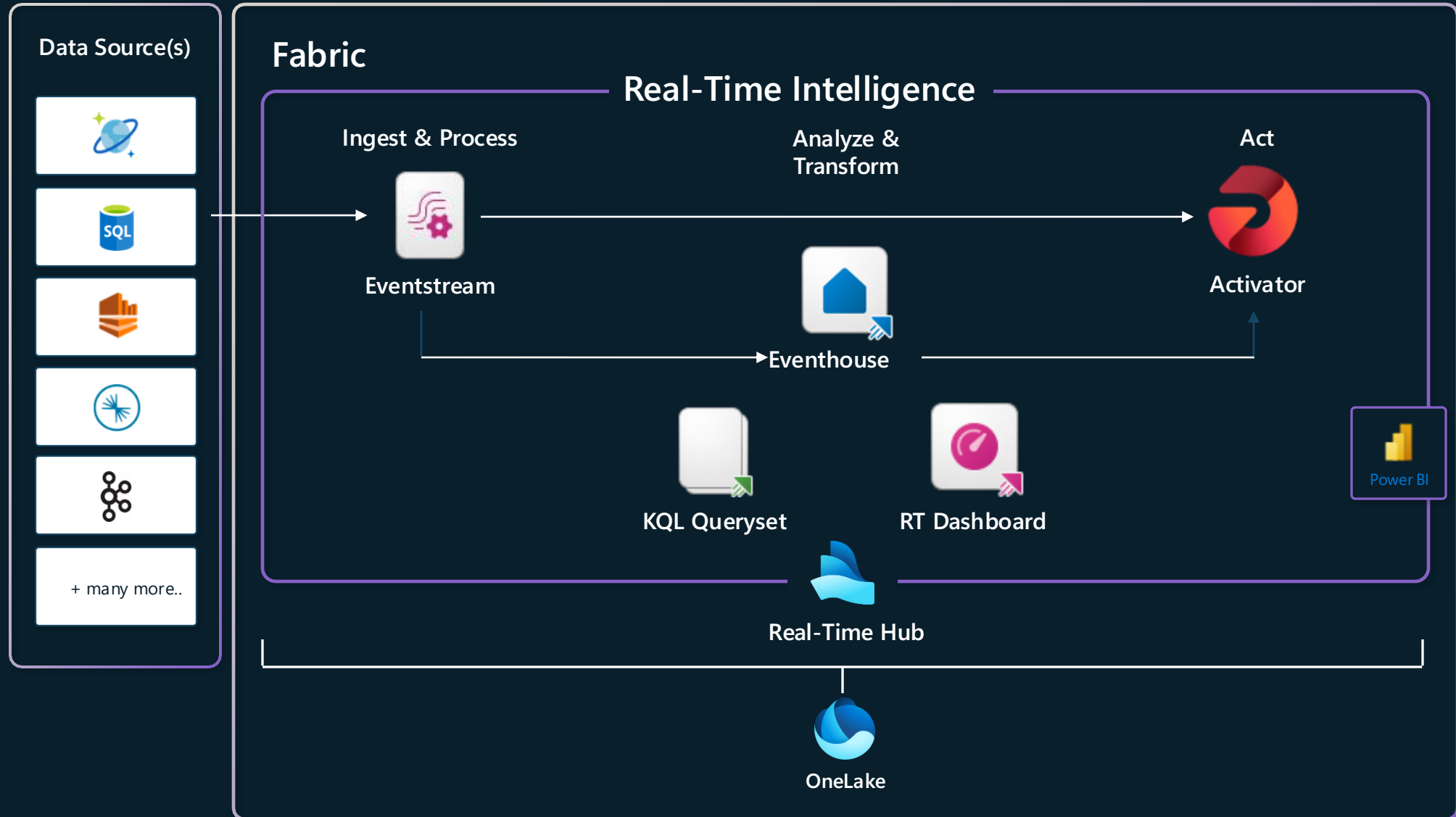
---

Automatic learning

---

Copilot acceleration

# Components of Fabric's Real-Time Intelligence





# Real-Time Intelligence in Microsoft Fabric



## Ingest & Process

- **Streaming sources** including AMQP, Kafka, Azure Event Hubs and more
- **No-code experience** including event processing
- **Event routing** to Eventhouse, Reflex and other Fabric entities



## Analyze

- **Timeseries database** with unlimited scale (query, ingestion, storage)
- **Transformation on read** of complicated data structure
- **Copilot** for generating queries



## Act

- **Take actions automatically** when patterns or conditions are detected
- **Light-weight modeling** empowering the business user



## Real-Time Hub

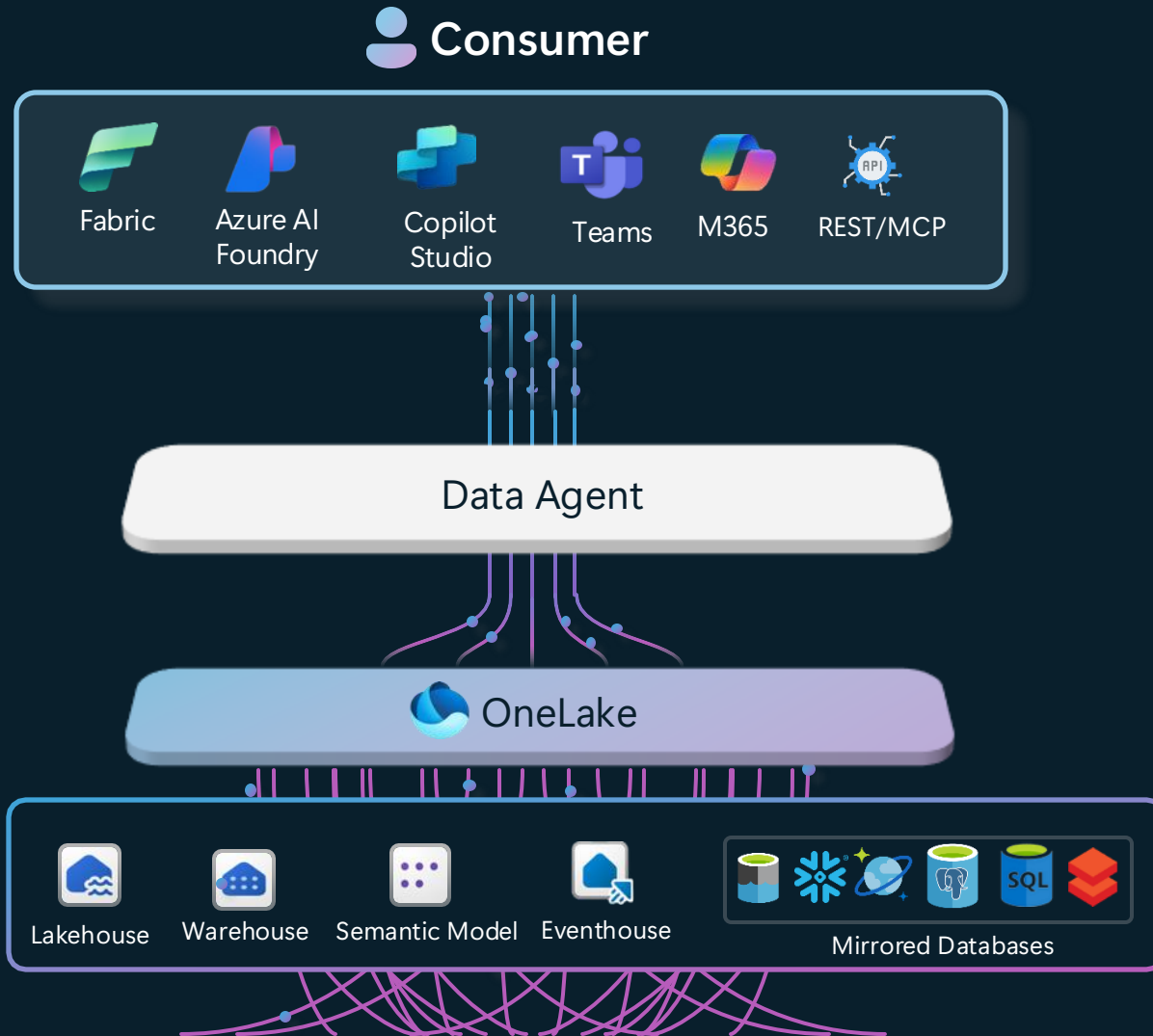
Central place to discover data and launch all related tasks

# Demo Real-Time Intelligence

# Fabric data agents

Virtual analyst for deeper data insights

# Fabric data agents



This conversational data agent allows users to interact with a virtual analyst.



Tooling for creators from UI and SDK, with new debugging capabilities, making it easier to understand and refine responses.

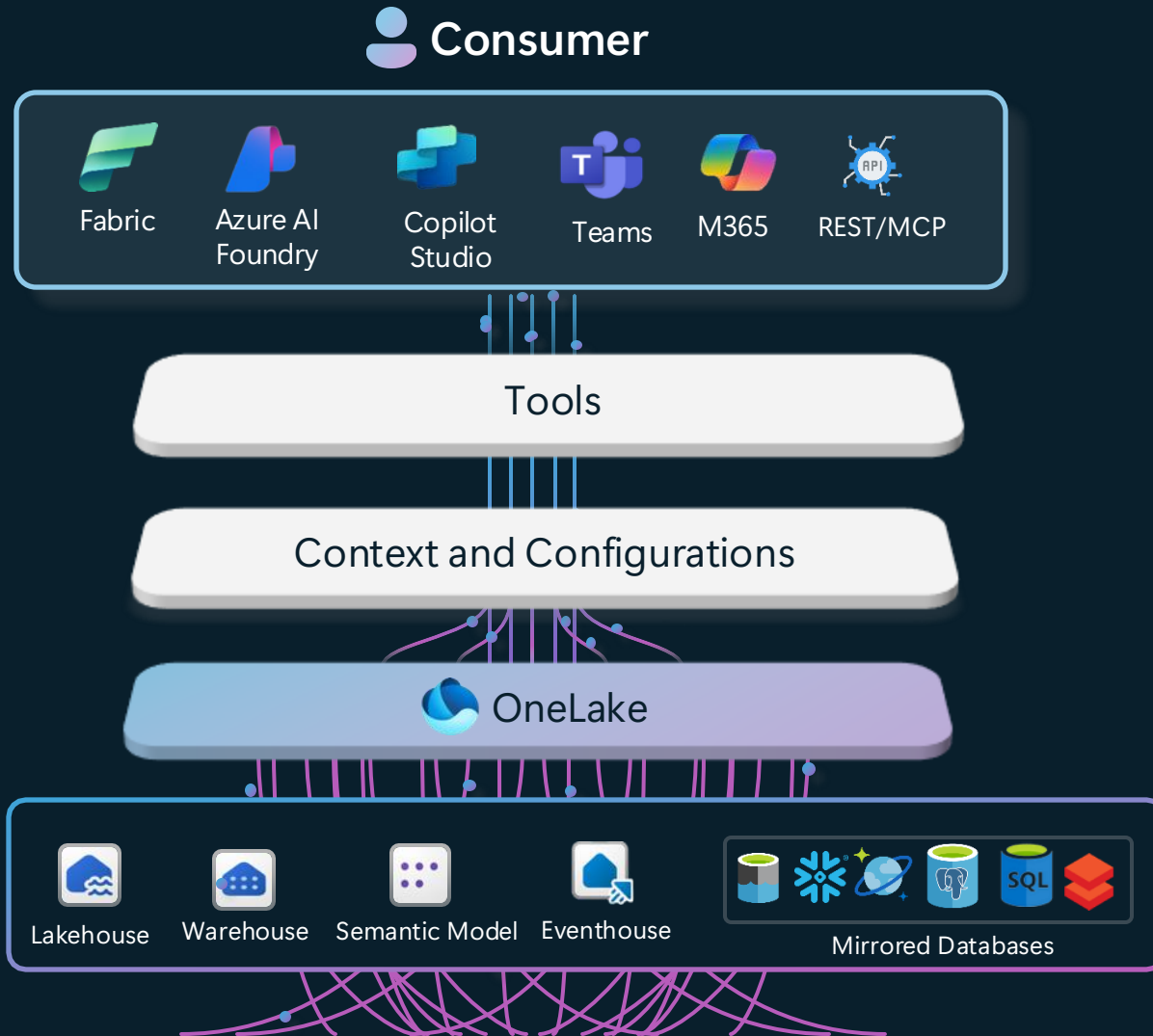


Seamlessly reason over data in OneLake, including shortcuts and mirrored databases to reason over tables, files and real-time events, to create a powerful data expert tailored to your data domain.



Your Data Agents can be consumed inside and outside of Fabric. Stay tuned for upcoming integrations with your own custom applications.

# Fabric data agents



This conversational data agent allows users to interact with a virtual analyst.



Tooling for creators from UI and SDK, with new debugging capabilities, making it easier to understand and refine responses.



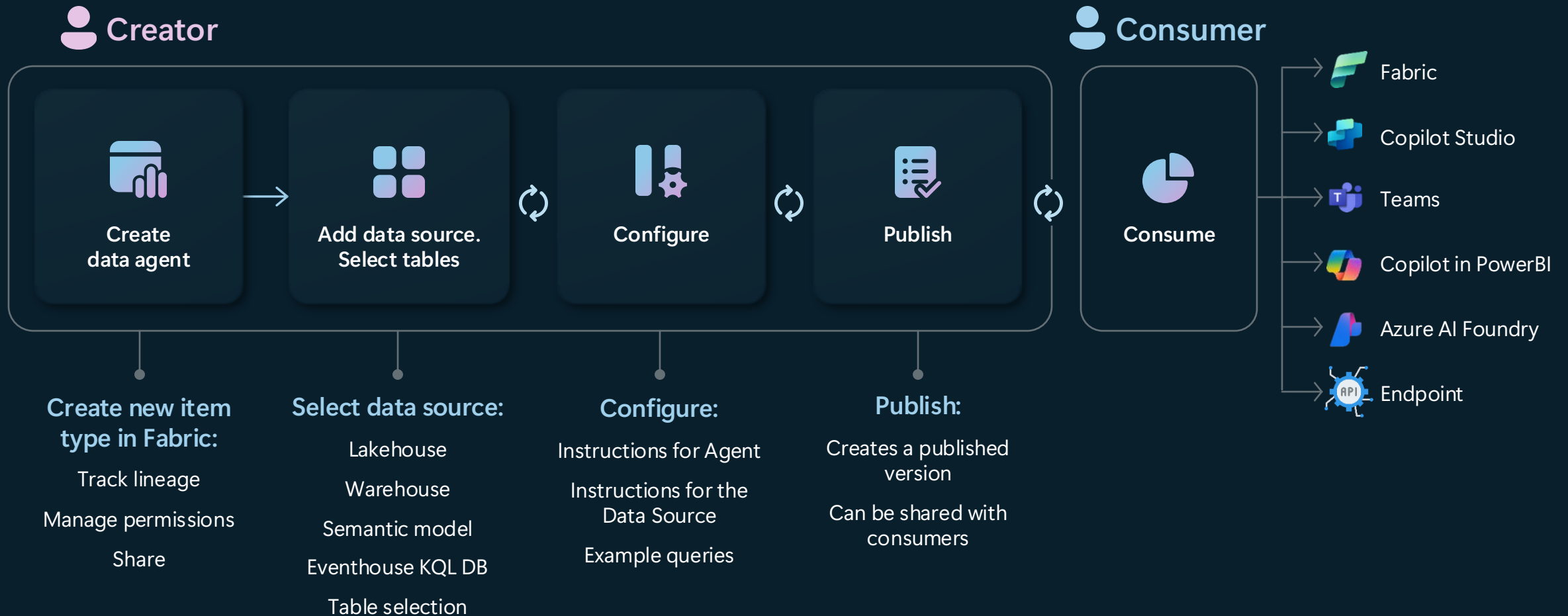
Seamlessly reason over data in OneLake, including shortcuts and mirrored databases to reason over tables, files and real-time events, to create a powerful data expert tailored to your data domain.



Your Data Agents can be consumed inside and outside of Fabric. Stay tuned for upcoming integrations with your own custom applications.

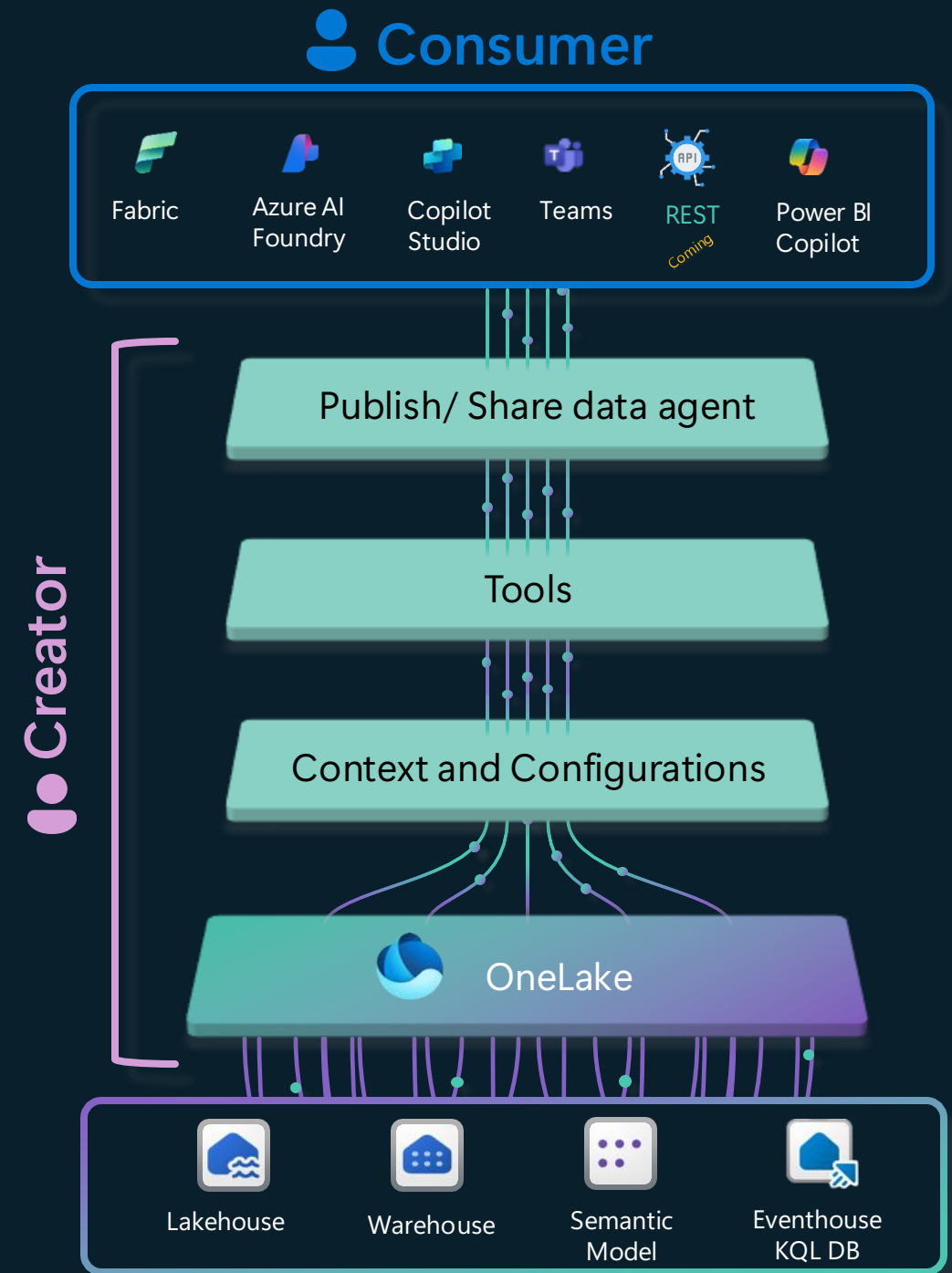
# Data agent end-to-end flow

End-to-end scenario



# Data agents roadmap

- Enhanced tooling for creators
- Data Agent REST Endpoint
- ALM and Git support
- Support Mirrored Databases
- Unstructured data support
- Visualizations
- Feedback, monitoring & evaluation experiences
- Data agent actions



Announcing



Fabric IQ



# Microsoft Fabric

The unified data platform for AI transformation



Data  
Factory



Analytics



Databases



Real-Time  
Intelligence



Power BI

Fabric Platform



Copilot

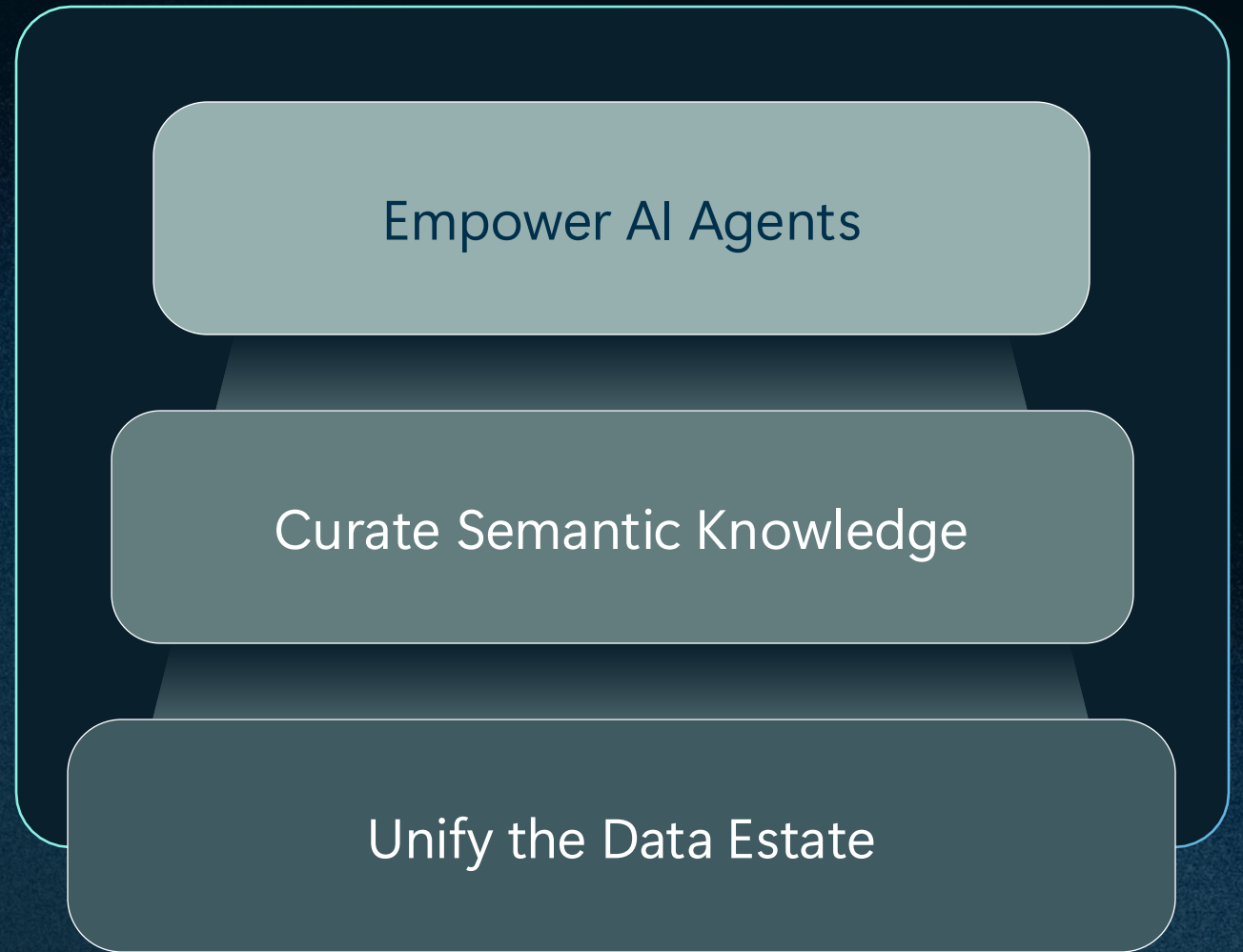


OneLake

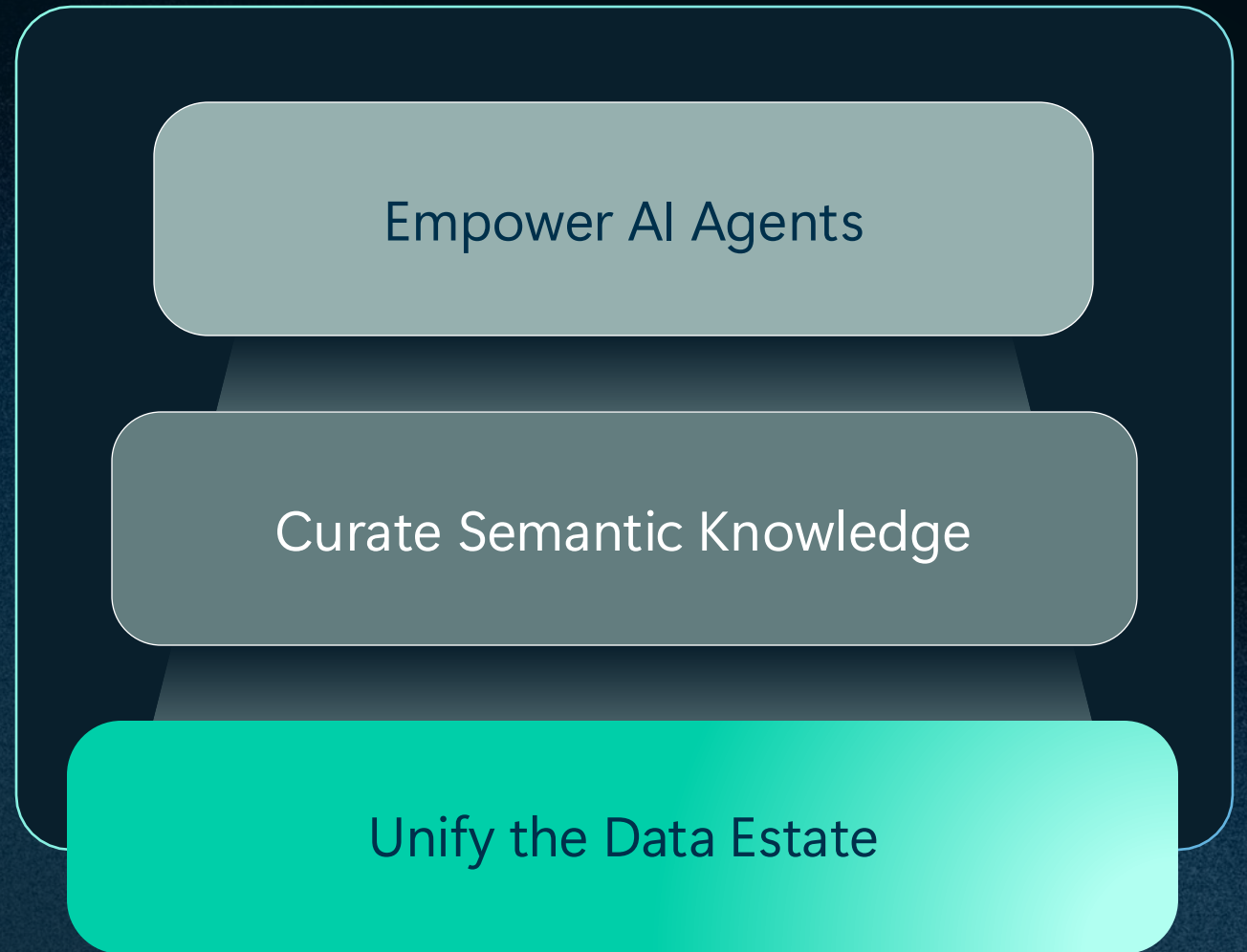


Governance

# The Data Estate for the Age of AI



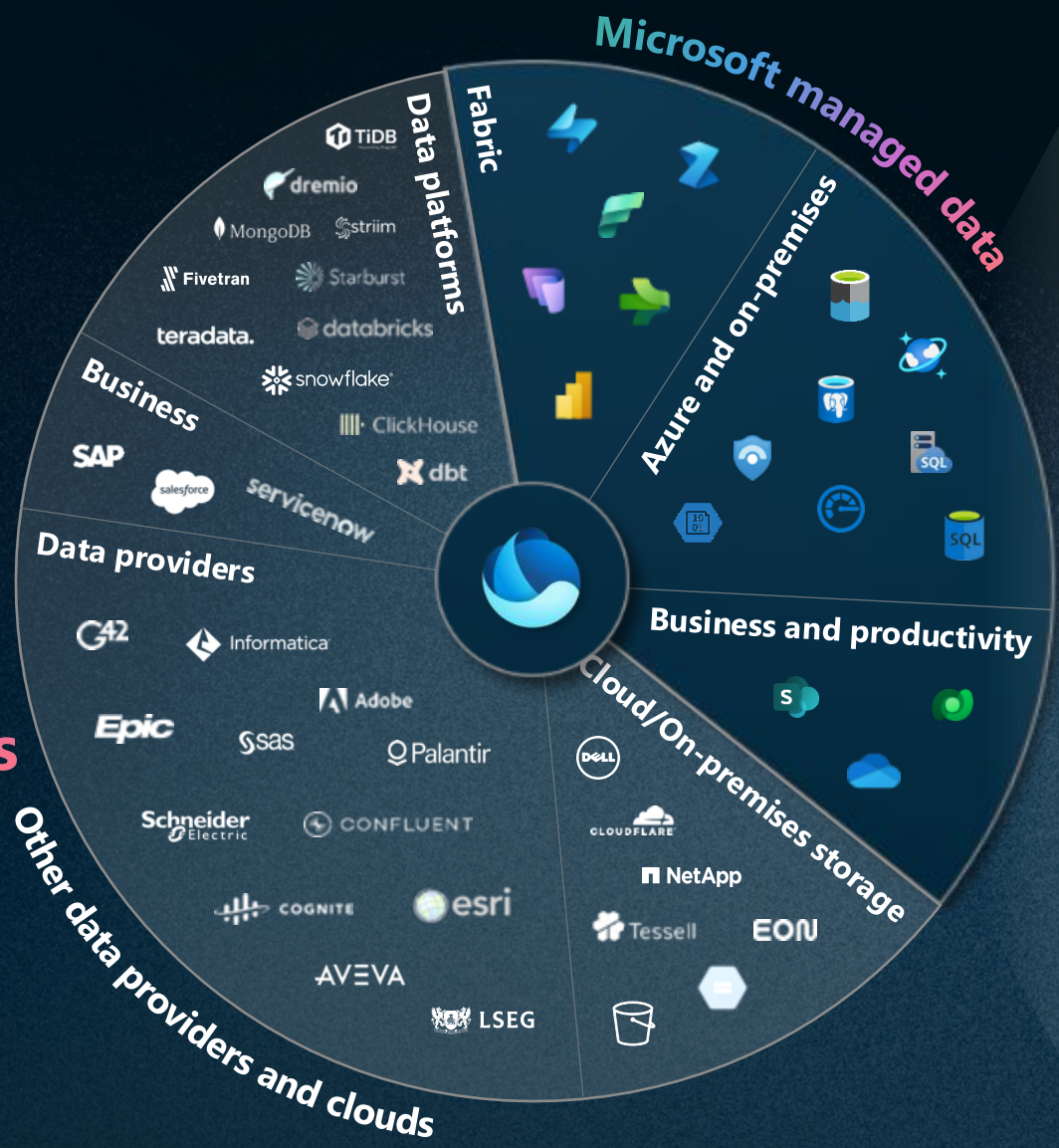
# The Data Estate for the Age of AI





# OneLake unifies the world's data

All clouds, on prem, all data bases, apps and files



## Announcing

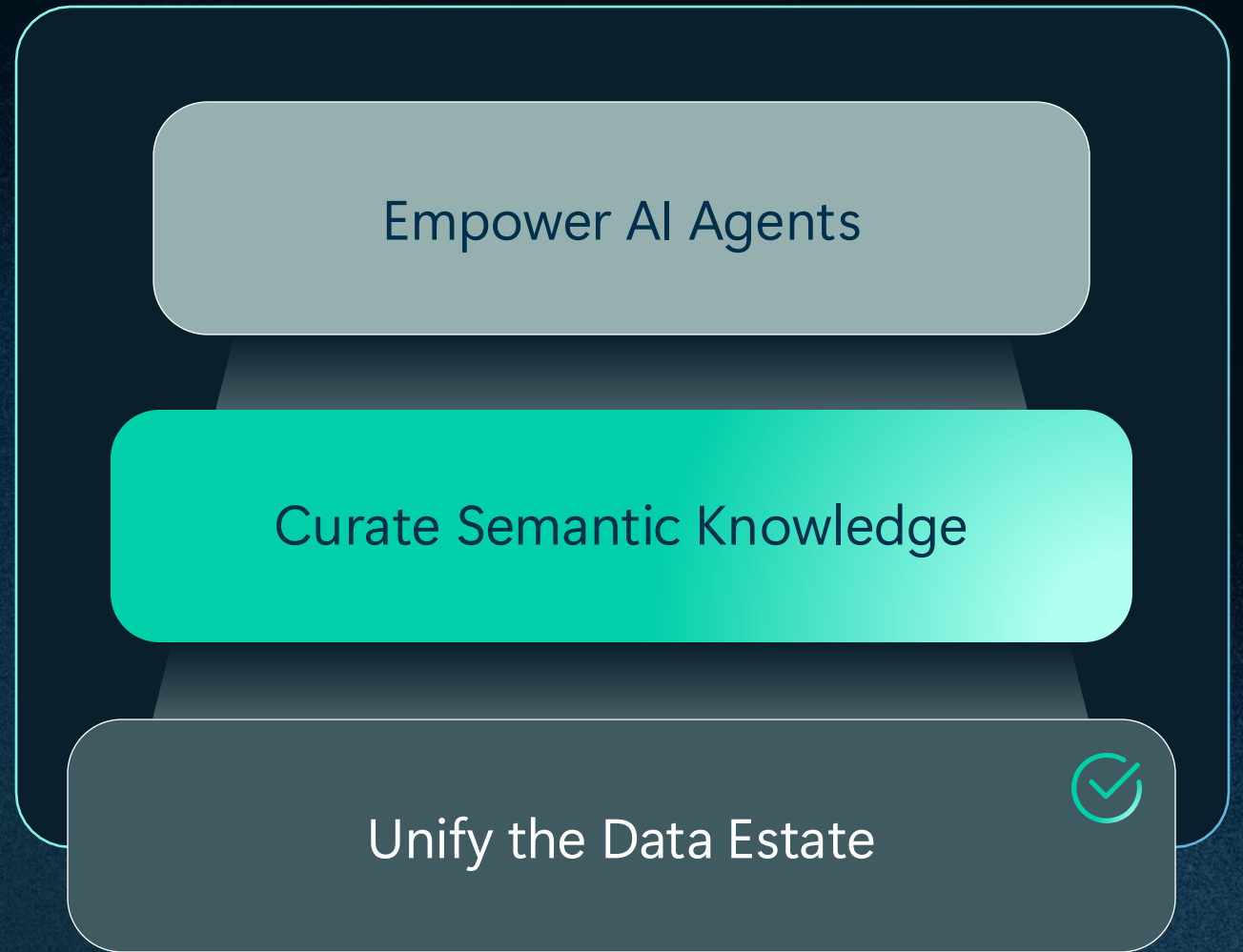
## Generally available

- Azure SQL MI
- Azure Blob Storage

## Public preview

- Oracle DB
- SharePoint/OneDrive
- Google BigQuery
- SAP Datasphere

# The Data Estate for the Age of AI





# Microsoft Fabric

The unified data platform for AI transformation



Data  
Factory



Analytics



Databases



Real-Time  
Intelligence



Power BI

Fabric Platform



Copilot



OneLake



Governance



# Microsoft Fabric

The unified data platform for AI transformation



Data Factory



Analytics



Databases



Real-Time Intelligence



Fabric IQ



Power BI

## Fabric Platform



Copilot



OneLake



Governance



# Microsoft Fabric

The unified data platform for AI transformation



Fabric IQ



Semantic Models



Ontology Models



Digital Twins



Graph



Operations Agents



Data Agents

Fabric Platform



Copilot



OneLake



Governance

# The Data Estate for the Age of AI

Empower AI Agents

Curate Semantic Knowledge 

Unify the Data Estate 

# AI Agents in Fabric

Org-wide intelligence that **understands, reasons, and acts**



## Chat with Your Data Agents

Provides insights based on Power BI dashboards



## Data Agents

Virtual analysts for deeper data insights



## Operations Agents

Virtual teammates: always observing, always acting



# The Fabric IQ advantage



Layered on top of a unified data foundation



Millions of Power BI models to jumpstart ontology



Native real-time platform for live business context



Microsoft 365 & Microsoft Foundry integration



Business agility from visual no-code modeling tools



Licensing that accelerates adoption, not cost

# Microsoft Fabric IQ + Microsoft Foundry IQ



## Fabric

Unified business semantics across data, models, and systems

### Unify analytical, time-oriented, and geospatial data

Provide a live, connected view of the enterprise using a semantic model of business entities and their relationships to create org-wide business context

### Improve decision-making with context-rich insights

Leverage a connected semantic foundation and context-aware experiences to accelerate insight discovery and optimize operational processes

### Enhance how AI agents access and interpret information

Develop AI agents and applications that can search, reason, and learn continuously as they're deployed across diverse, unstructured sources



## Foundry

Build AI agents that can search, reason, and learn across diverse sources



OneLake serves as the connective tissue to build custom AI apps powered by your data

# Fabric IQ resources

## Engage

Ask questions on the forum

[aka.ms/fabric-iq-forum](https://aka.ms/fabric-iq-forum)

Submit ideas and vote

[aka.ms/fabric-ideas](https://aka.ms/fabric-ideas)

## Learn more

Use the docs

[aka.ms/fabric-iq-overview](https://aka.ms/fabric-iq-overview)

Complete a tutorial

[aka.ms/ontology-tutorial](https://aka.ms/ontology-tutorial)

Get certified

[aka.ms/dp700](https://aka.ms/dp700)

Find customer success stories

[aka.ms/fabric-customer-success](https://aka.ms/fabric-customer-success)

## Stay updated

Read the blog

[aka.ms/fabric-iq-blogs](https://aka.ms/fabric-iq-blogs)

Watch on YouTube

[aka.ms/fabricyoutube](https://aka.ms/fabricyoutube)

Check the release plan

[aka.ms/fabricreleaseplan](https://aka.ms/fabricreleaseplan)

# How did we do?



# Take the next step on your fabric journey

Fabric Trial Capacity



Microsoft Digital  
Briefings



Fabric Capacity  
Calculator



# Contact



**Graeme Gord**  
Senior Architect  
gord@hitachisolutions.com



**Anthony Sheldrake**  
Architect  
asheldrake@hitachisolutions.com



**Matt Holmes**  
Business Development Director, Commercial  
mholmes@hitachisolutions.com

Hitachi Solutions UK

