

Microsoft Fabric Modernization Pathways in Action

datAvail





Tobin Thankacen

Global Practice Leader,
Data Management,
@ Datavail

- ✓ Proficient Data Analytics & AI Leader, Data Strategist
- ✓ Cloud, Big Data, Large scale Lakehouse, Data lake, and Traditional Data Warehouse expertise
- ✓ Cross-functional project leader using advanced data modeling and analysis techniques
- ✓ Enhancing organizational performance through the evaluation of data warehouse design and resolution of data quality issues across ETL and analytics systems



<https://www.linkedin.com/in/tobin-t-832b6213/>

Put the Fabric Pieces Together

Fill out your session evaluation form for a chance to win a LEGO set.



Why Datavail & Microsoft Azure?

Solutions Partner

- ✓ Data & AI
- ✓ Digital & App Innovation
- ✓ Infrastructure

Microsoft Preferred Partner

- ✓ Microsoft Fabric Featured Partner
- ✓ Oracle to Azure Migration & Modernization
- ✓ PostgreSQL Azure Migrations, Apps, & AI

Microsoft Specializations

- ✓ Migrate Enterprise Apps to Microsoft Azure
- ✓ Analytics on Azure
- ✓ Infra and Database Migration to Azure (2025)
- ✓ Data Warehouse Migration (2025)
- ✓ Kubernetes on Azure (Future)
- ✓ DevOps with GitHub on Azure (Future)

Azure Marketplace Offerings

- ✓ [Oracle on Azure Assessment](#)
- ✓ [Azure Data Lake Roadmap & Implementation](#)
- ✓ [Azure Synapse Implementation](#)
- ✓ [SQL Server Migrations to Azure](#)
- ✓ [Cassandra to Cosmos DB Assessment](#)
- ✓ [Oracle to Azure Migration](#)
- ✓ [MongoDB to Azure Cosmos DB Migration](#)
- ✓ MySQL [Assessment](#) & [Implementation](#)
- ✓ [Guidewire CDA Connector for Azure](#)

Managed Partner (Growth)

- ✓ Funding Programs - ECIF and AMM & Azure Innovate Eligible
- ✓ ISD (MCS) Approved Sub-Contractor – Commercial & Federal
- ✓ Co-sell Ready Program

We've Done this Before...

1000+

Consultants focused on Cloud deployments and managed services

17+

Years of Data Platform & AppDev Experience

10+

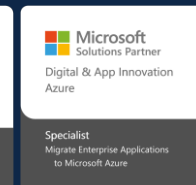
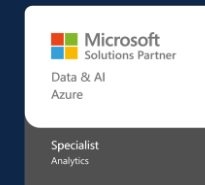
Years Cloud Experience

500+

Customers with average lifespan of 7+ years

550+

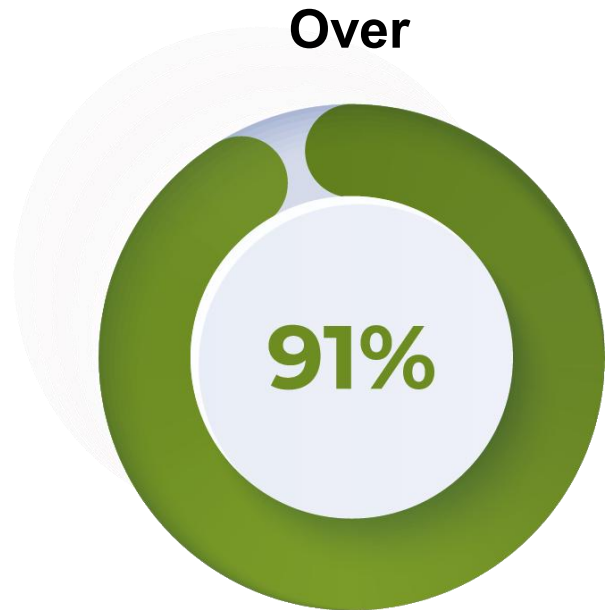
Microsoft consultant certifications



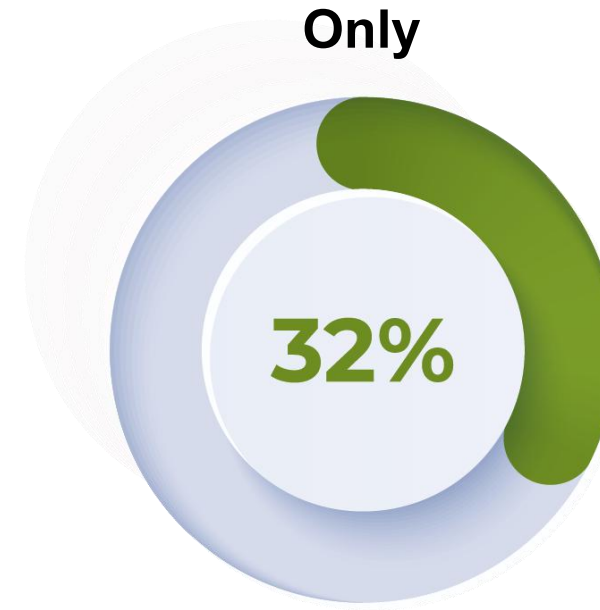
What are Organizations Striving for Today?

- ✓ Becoming Data Driven
- ✓ Leveraging Data as a Strategic Asset

But achieving these benefits comes with challenges



of leaders report their organizations have increased investment in data and analytics



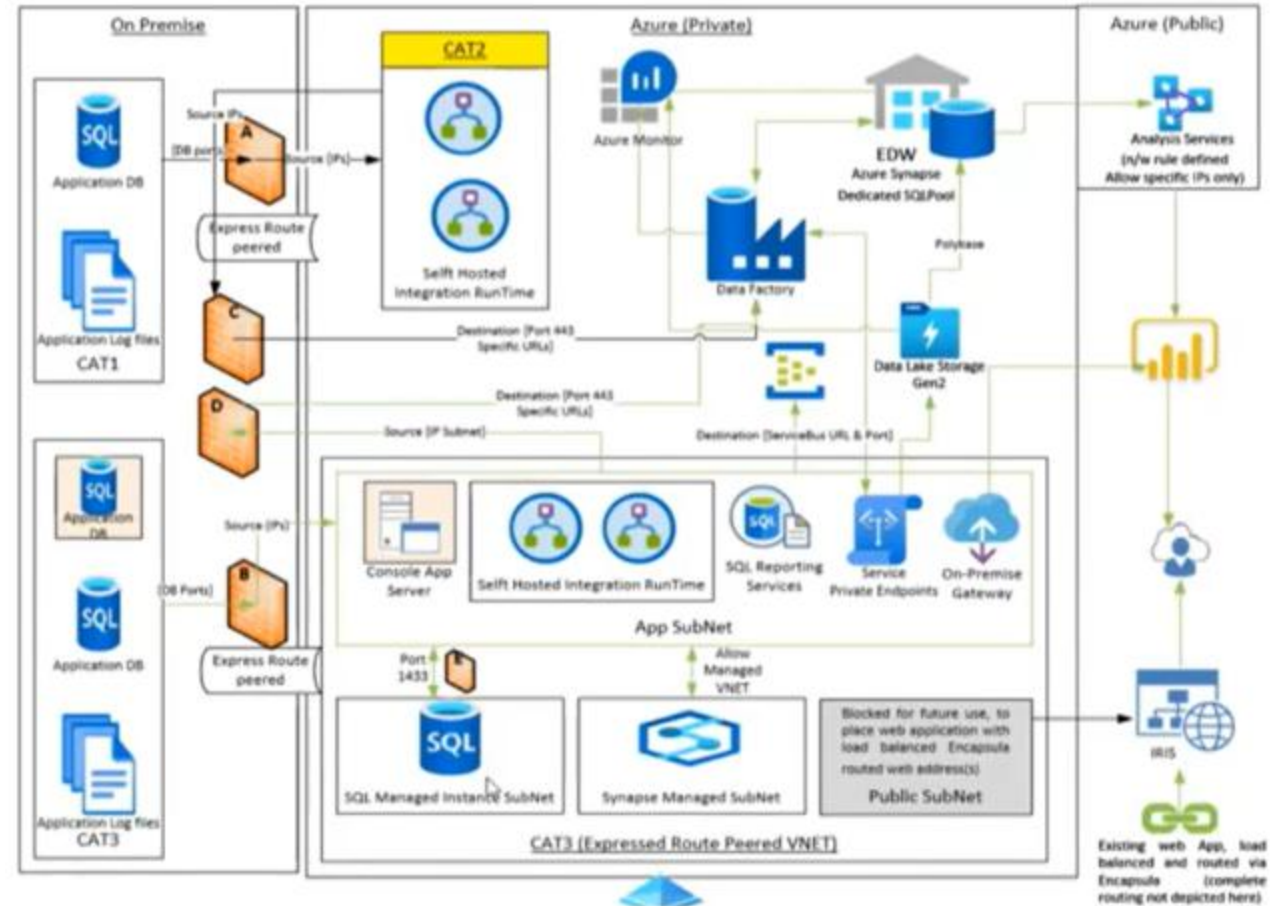
of organizations report being able to realize tangible value from data and analytics

[Source: Accenture, Closing the Data Value Gap](#)

[Source: Gartner, Modernize Data and Analytics Capabilities](#)

Analytics is complex and fragmented

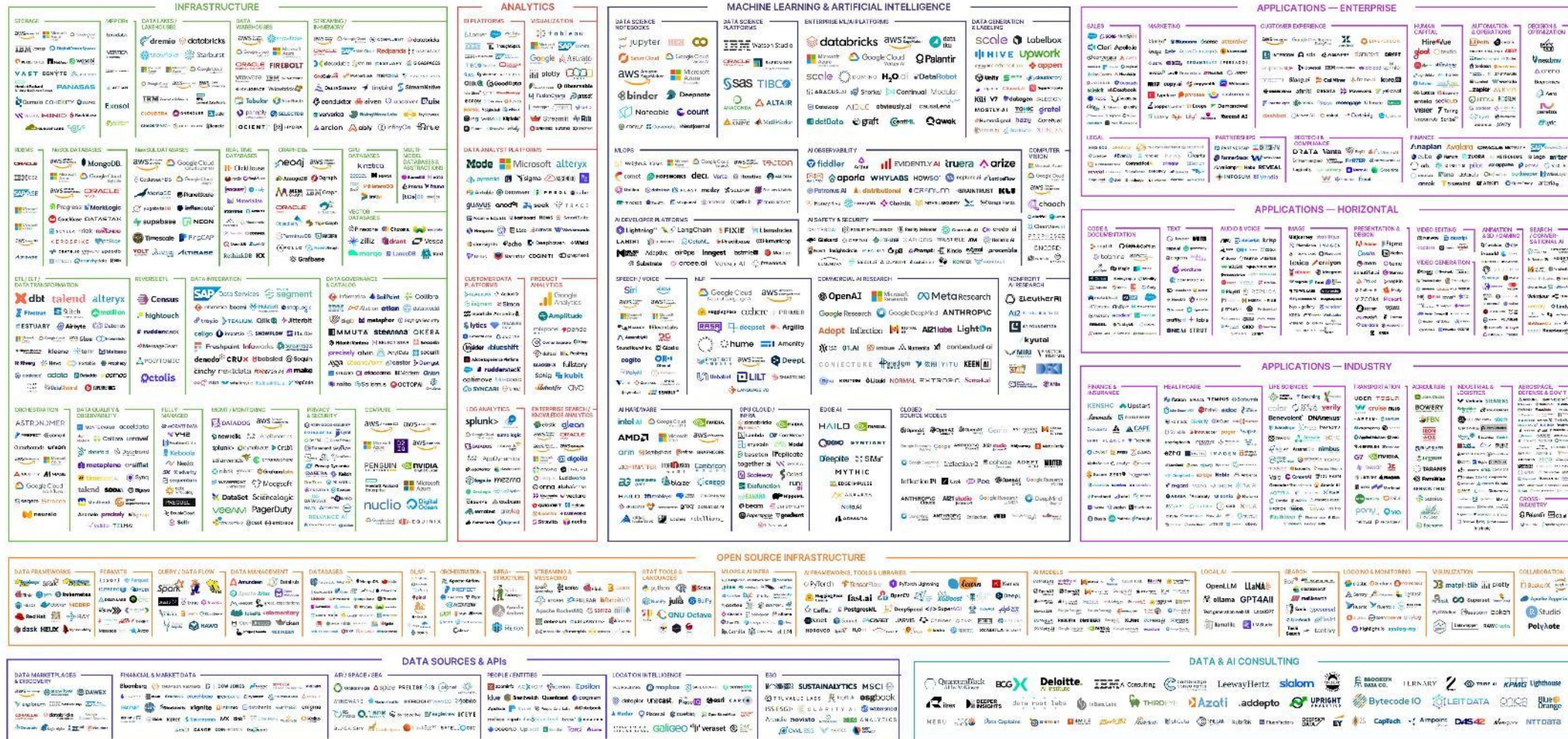
- ✓ Every project has many subsystems
- ✓ Every Subsystem need a different class of Product
- ✓ Product often comes from multiple vendors
- ✓ Integration is complex, fragile and expensive



There is massive fragmentation of the modern tech stack

datAvail

THE 2024 MAD (MACHINE LEARNING, ARTIFICIAL INTELLIGENCE & DATA) LANDSCAPE



Version 1.0 - March 2024

© Matt Turck (@mattturck), Aman Kaboor (@AmanKaboor1) & FirstMark (@firstmarkcap)

Blog post: mattturck.com/MAD2024

Interactive version: MAD.firstmarkcap.com

Comments? Email MAD2024@firstmarkcap.com

FIRSTMARK
EARLY STAGE VENTURE CAPITAL



“

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

Every CDO, Every Enterprise

datAvail

What we're hearing from our customers



How do I unify all disparate data sources cost-effectively?



How do I minimize security breaches and risks?



How do I improve analytical agility for my organization?



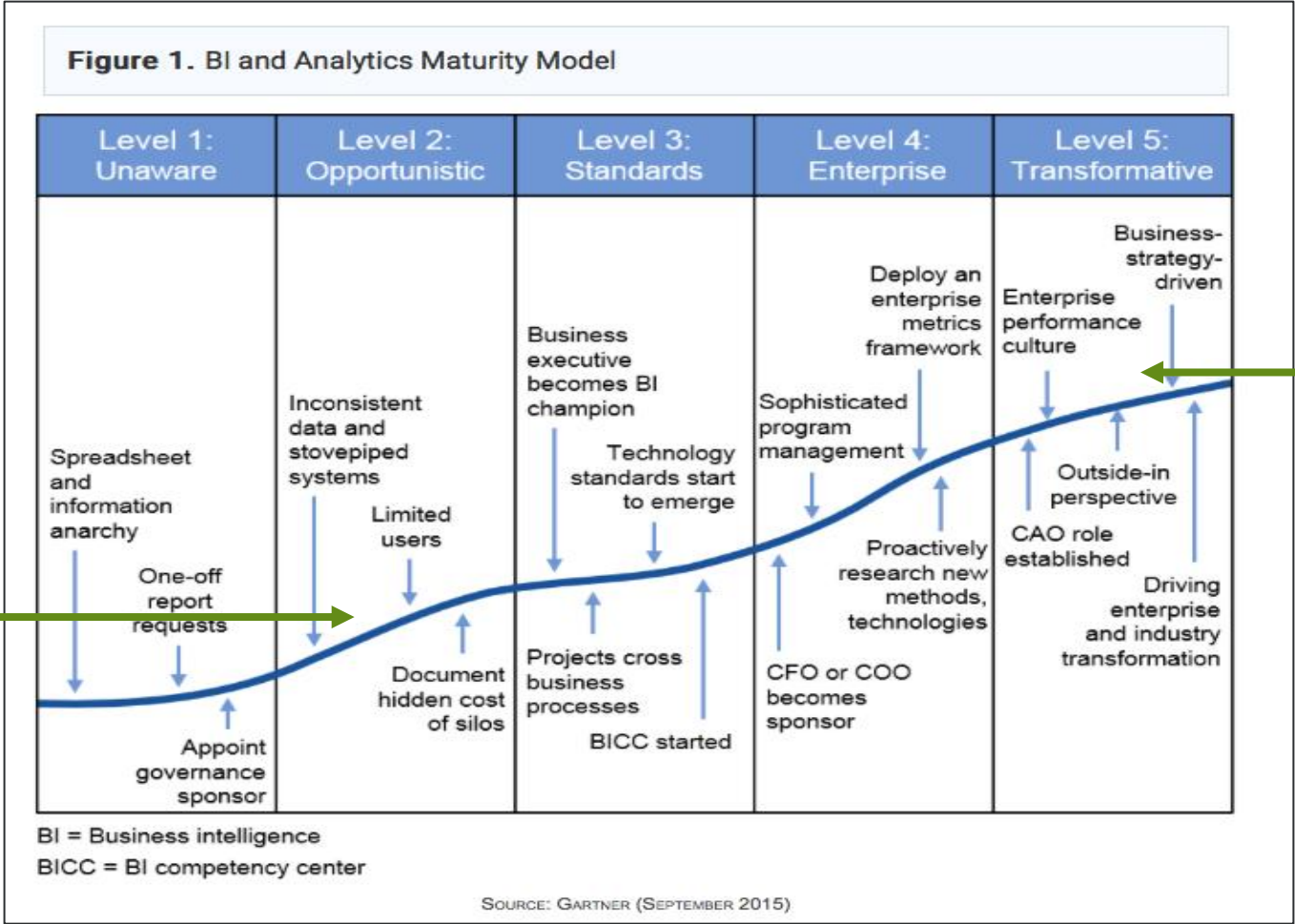
How do I ensure business users use the insights to get value from my investments?

Data Management Maturity Level

As a starting point, we leverage an enterprise analytics maturity model to define where your organization is today and the objectives that you want to achieve

Run the Business

Data Driven



Modernizing your data & analytics platform is critical for resilient business transformation



Variety of siloed solutions and data



Integrated, SaaS-based suite that can integrate any data source

Data with security risks



Industry-leading, built-in security, compliance, and governance

Demand for data stewards



Easy-to-use analytics and center of enablement for increased business agility

Costly integration and ongoing maintenance of legacy systems



Cost-transparent suite with reduced management overhead

Introducing Microsoft Fabric

Bring all these analytics services together in a single end-to-end analytics platform that can go from the data lake to the business user



**Data
Integration**

Data Factory



**Data
Engineering**

Synapse



**Data
Warehouse**

Synapse



**Data
Science**

Synapse



**Real Time
Analytics**

Synapse



**Business
Intelligence**

Power BI



**Partner &
Industry
Workloads**



Copilot in Fabric



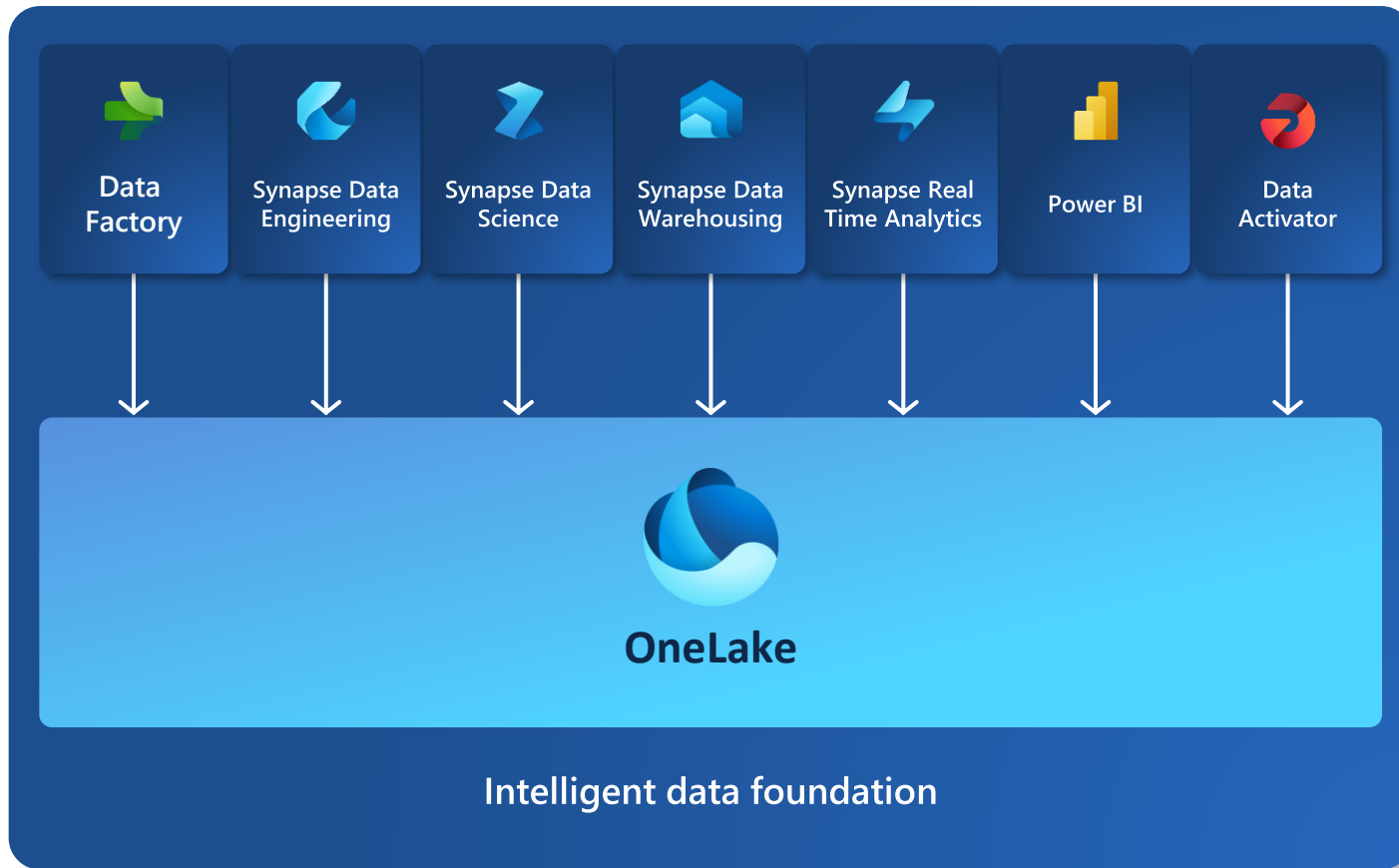
OneLake



Microsoft Purview

OneLake for All Data

“The OneDrive for Data”



A single SaaS lake for the whole organization

Provisioned automatically with the tenant

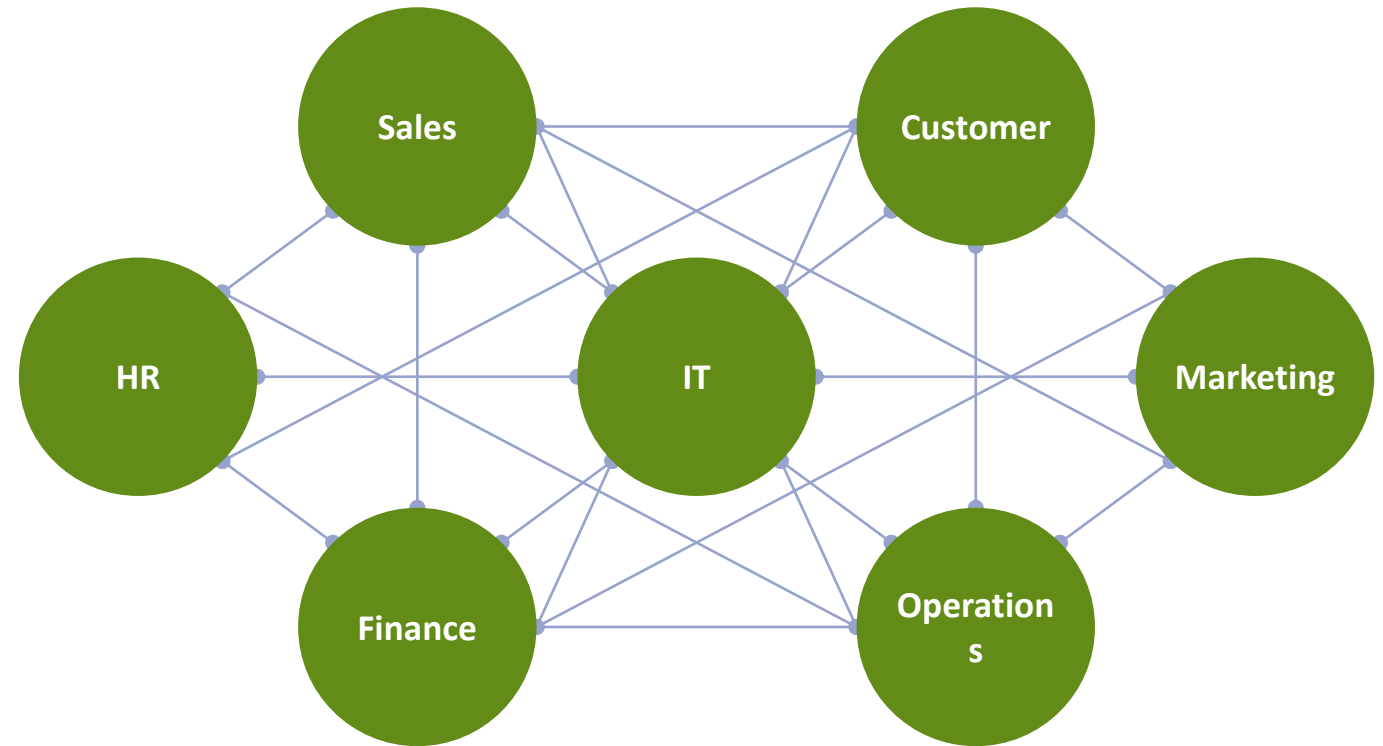
All workloads automatically store their data in the OneLake workspace folders

All the data is organized in an intuitive hierarchical namespace

The data in OneLake is automatically indexed for discovery, MIP labels, lineage, PII scans, sharing, governance and compliance

OneLake for all domains

A true **hub & Spoke** data mesh across organizational data domains.



Workspaces and artifacts for different data domains, contribute to building the same Data Lake.

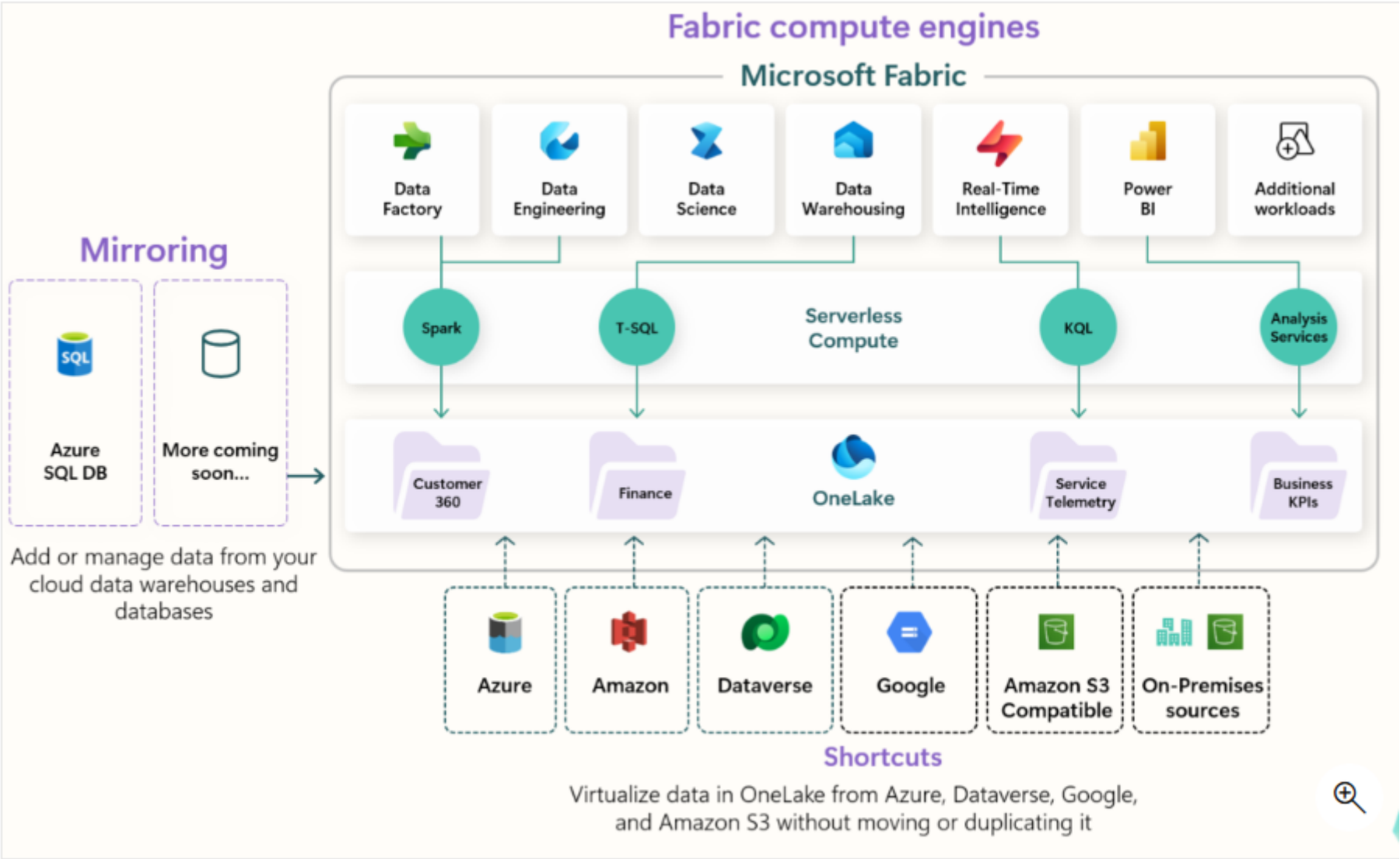
Without data movement, data from different domains can be analyzed, blended and transformed together

Data is secured and governed in one place while remaining easily discoverable and accessible to all who should have access across the organization.

Data can be certified by domain experts to enabling trust for data which is discovered

Fabric Compute Engines

All Microsoft Fabric compute experiences are **preconfigured** to use **OneLake** as their default storage.



The simplicity of the Microsoft Fabric business model

Rather than provisioning and managing separate compute for each workload, with Microsoft Fabric, your bill is determined by two variables: The amount of compute you provision and the amount of storage you use.

Compute

- ✓ A shared pool of capacity that powers all capabilities in Microsoft Fabric, from data modeling and data warehousing to business intelligence.
- ✓ Pay-as-you-go (per sec billing with one minute minimum).

Storage

- ✓ A single place to store all data.
- ✓ Pay-as-you-go (\$ per GB / month).

What Does Fabric Mean for My Data & AI Teams?

datAvail



Fabric addresses the top pain points of every analytics role

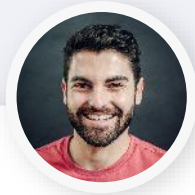


Data Engineers

Complexity. Implementing ETL can require considerable coding, especially when loading only modified rows.

Cost. Implementing ETL can require an exorbitant cost of purchasing additional tools and licenses.

Data latency. Implementing ETL adds a time delay for running the analytics and prevents fresh insights.



Data Scientists

Siloed development. Isolated collaboration between Data and ML teams leads to low quality data pipelines and unproven ML ROI. Difficulty in stream-lining decision-making processes.

Complex operationalization. ML PoC purgatory. High cost of management, scale, reliability, RAI. Reduced trust disconnected from data source.



Data Analysts

Bad data. Lack of access to the right data leads to bad insights.

Limited capacity and speed. Struggle to keep up with the number of reports needed to be created and maintained, slowed by dependencies on other data teams.

Embedding. Finding ways to integrate and share reports where stakeholders work.



Business users

Finding insights. Digging into insights requires accessing reports from multiple applications and systems.

Time waste. Identifying critical and action-orientated insights from BI reports is time intensive.



Data Stewards

Data sprawl. Lack of visibility and strong governance tools leads to sprawling data, excessive replication of data, and security risks.

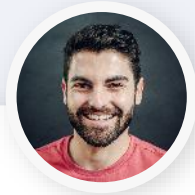
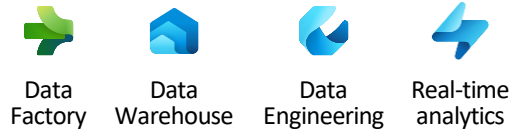
Empower users with role-specific workloads on a single platform



Data Engineers

Data Engineers can easily transform large volumes of data from various sources and make it available for analysis.

Supporting experiences



Data Scientists

Data scientists can democratize AI with complete end-to-end ML workflows, serving through their data warehouse with one-click flows and Power BI visuals.

Supporting experiences



Data Analysts

Analysts can explore data visually and bring insights to life with stunning, easy-to-use reports and AI-enhanced visuals.

Supporting experiences



Business users

Data citizens can explore Power BI reports in the apps they already use like Microsoft 365 to ensure every decision and action taken is data-driven.

Supporting experiences



Serve data via warehouse or lakehouse

Serve transformed data

Serve insights via embedding

Serve data via warehouse or lakehouse



Data Stewards

Data stewards can remain in complete control of their data while enabling teams across the organization to use the most accurate and trusted data

Understanding Your Upgrade Pathways to Microsoft Fabric

Key Considerations

dat~~Avail~~



Key Upgrade Concept: Shortcuts

No data movements or duplication: virtualize data across clouds, domains and accounts

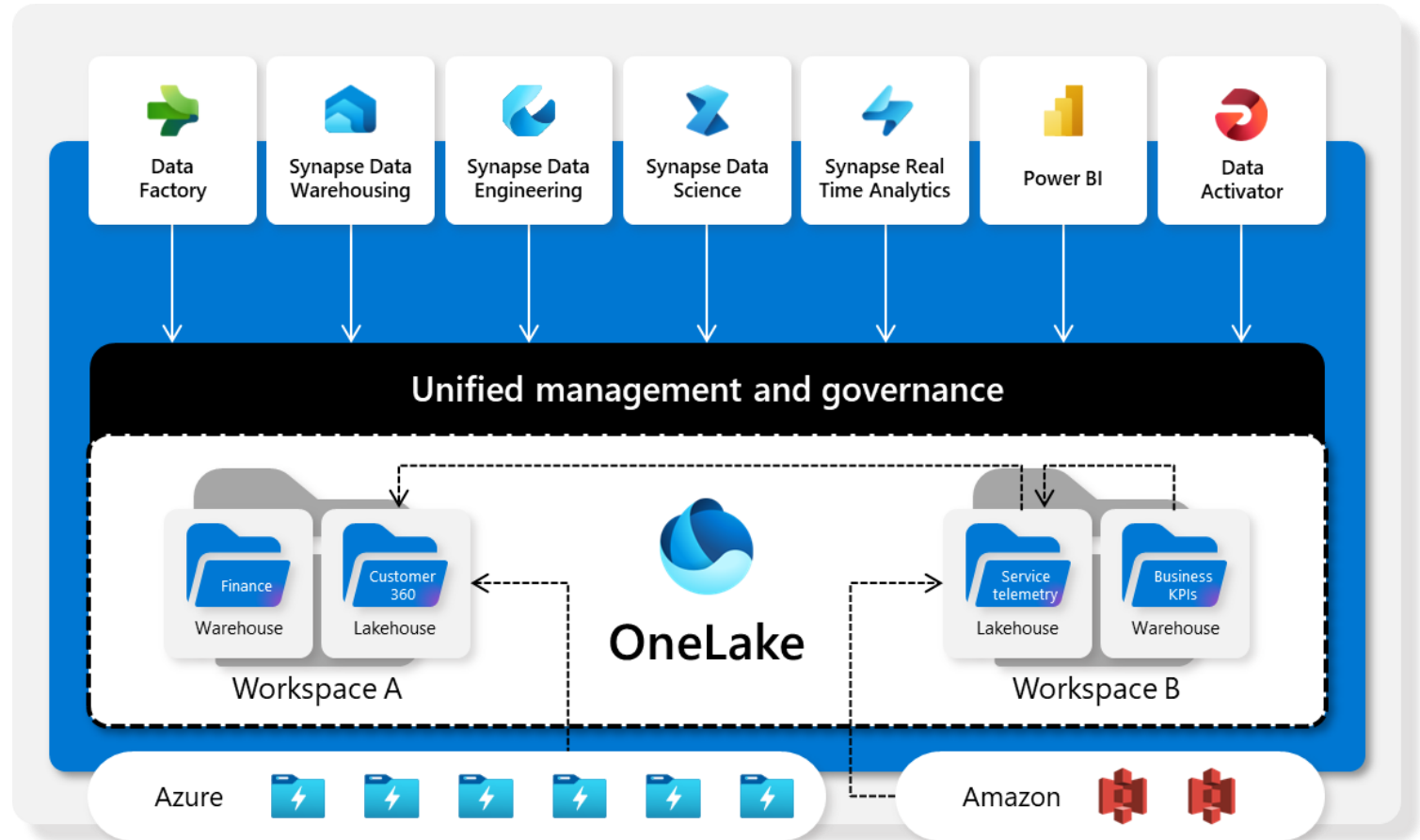
A Shortcut creates a live connection between OneLake and existing target data sources.

Create a Shortcut to virtualize existing file data in any format or tabular data in open delta format.

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.

Existing ADLS G2 storage accounts and Amazon S3 buckets can be managed externally to Fabric and Microsoft while still being virtualized into OneLake with Shortcuts.

All data is mapped to a unified namespace and can be accessed using the same APIs including the ADLS G2 DFS APIs.



Shortcuts produces one data copy for all computes

Real separation of compute and storage

Compute powers the applications and experiences in Fabric. The compute is separate from the storage.

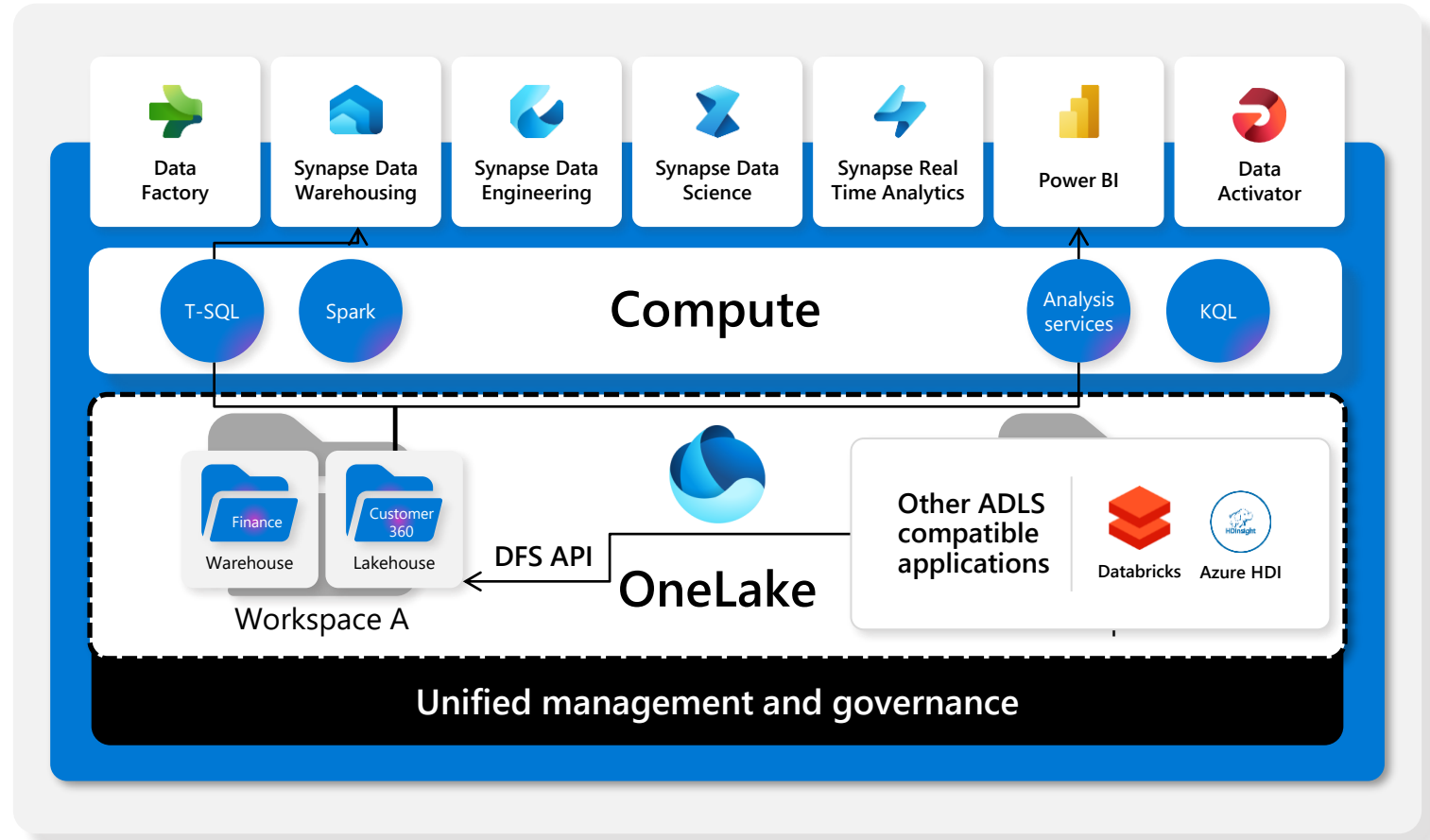
Multiple compute engines are available, and all engines can access the same data without needing to import or export it. You can choose the right engine for the right job.

Non-Fabric engines can also read/write to the same copy of data using the ADLS APIs or added through shortcuts.

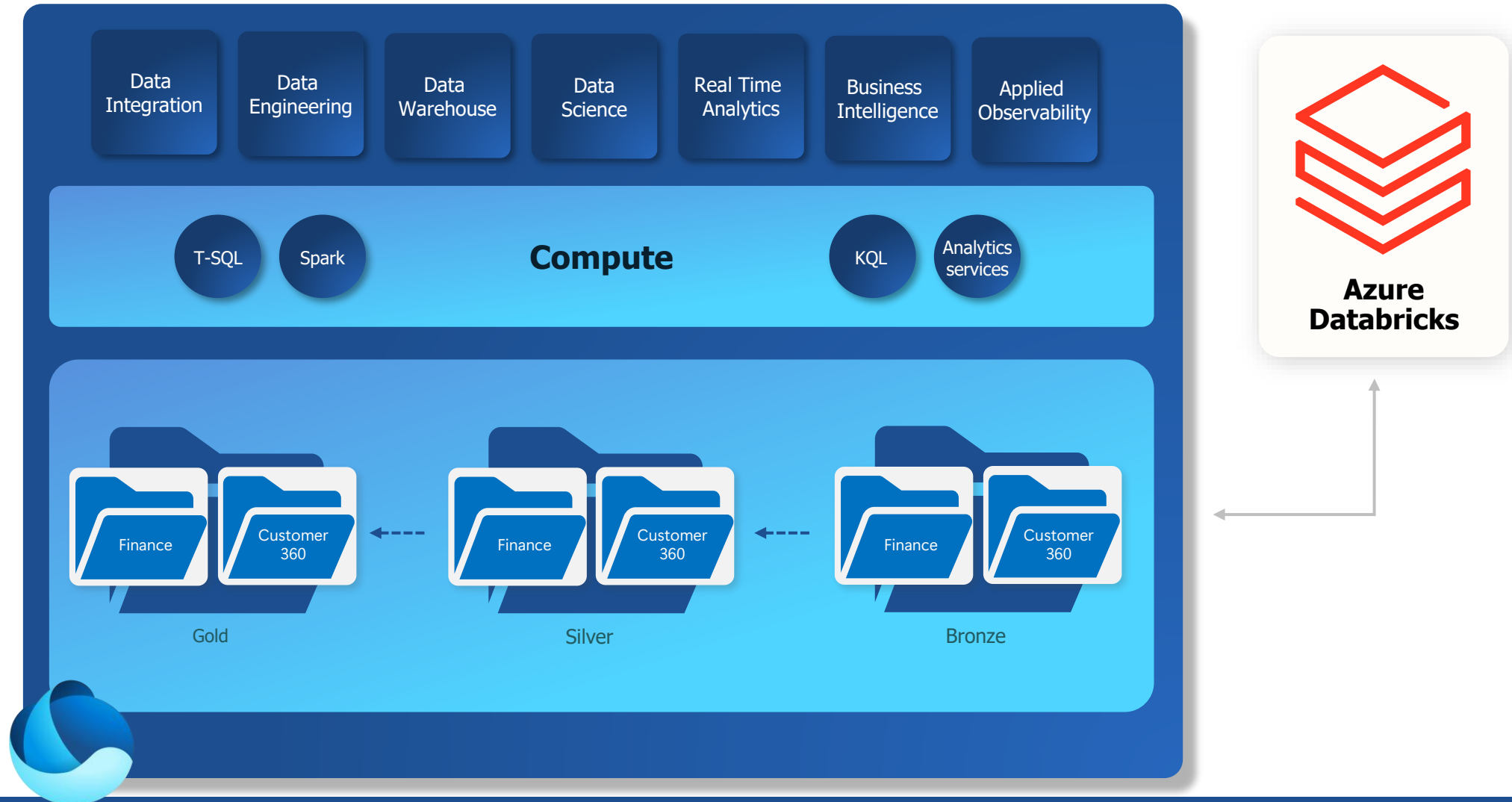
A data engineer can load data using a Databricks Notebook.

A data analyst can query the data using T-SQL.

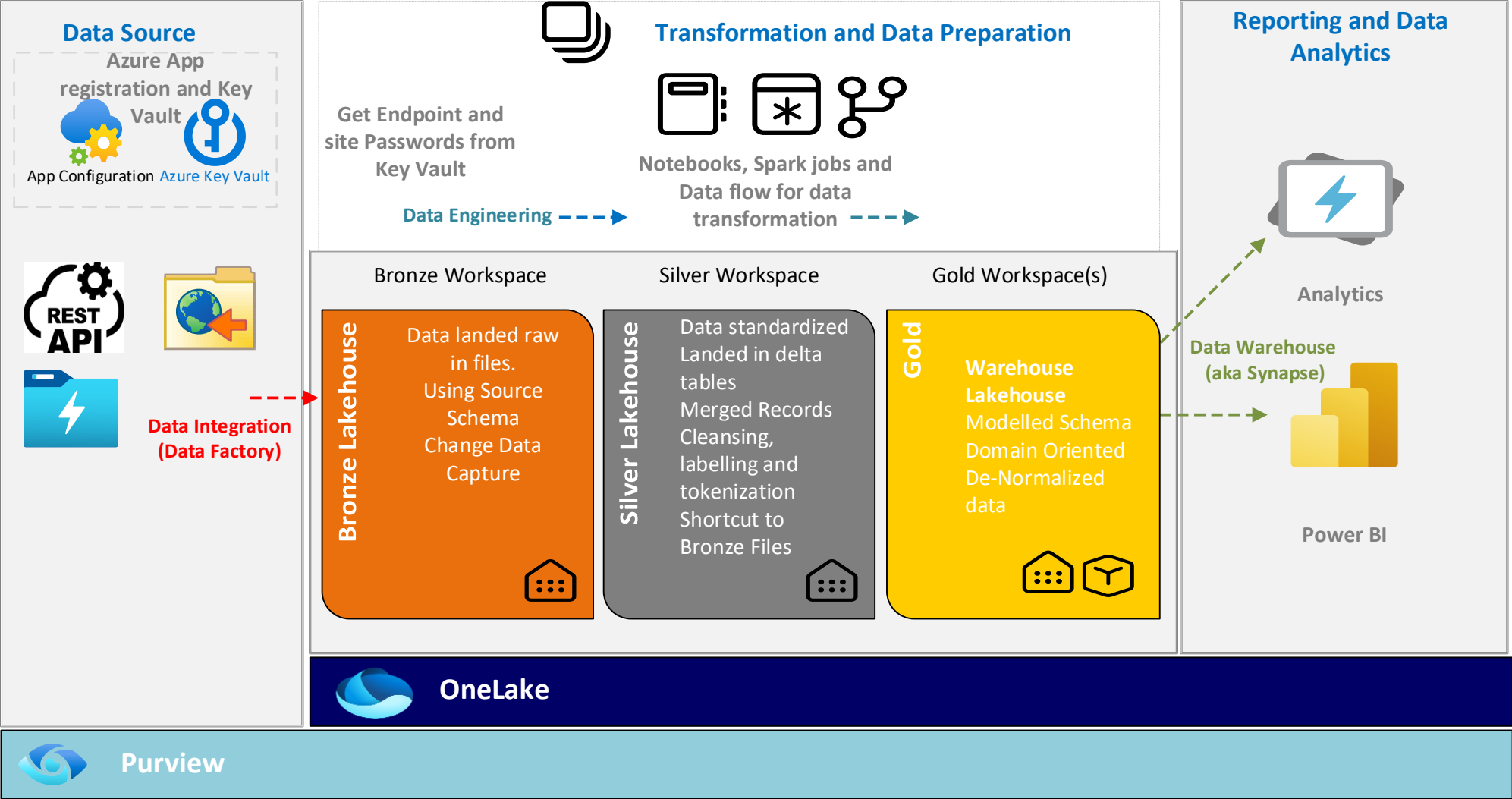
A business user can view the same data in a Power BI report using Direct Lake.



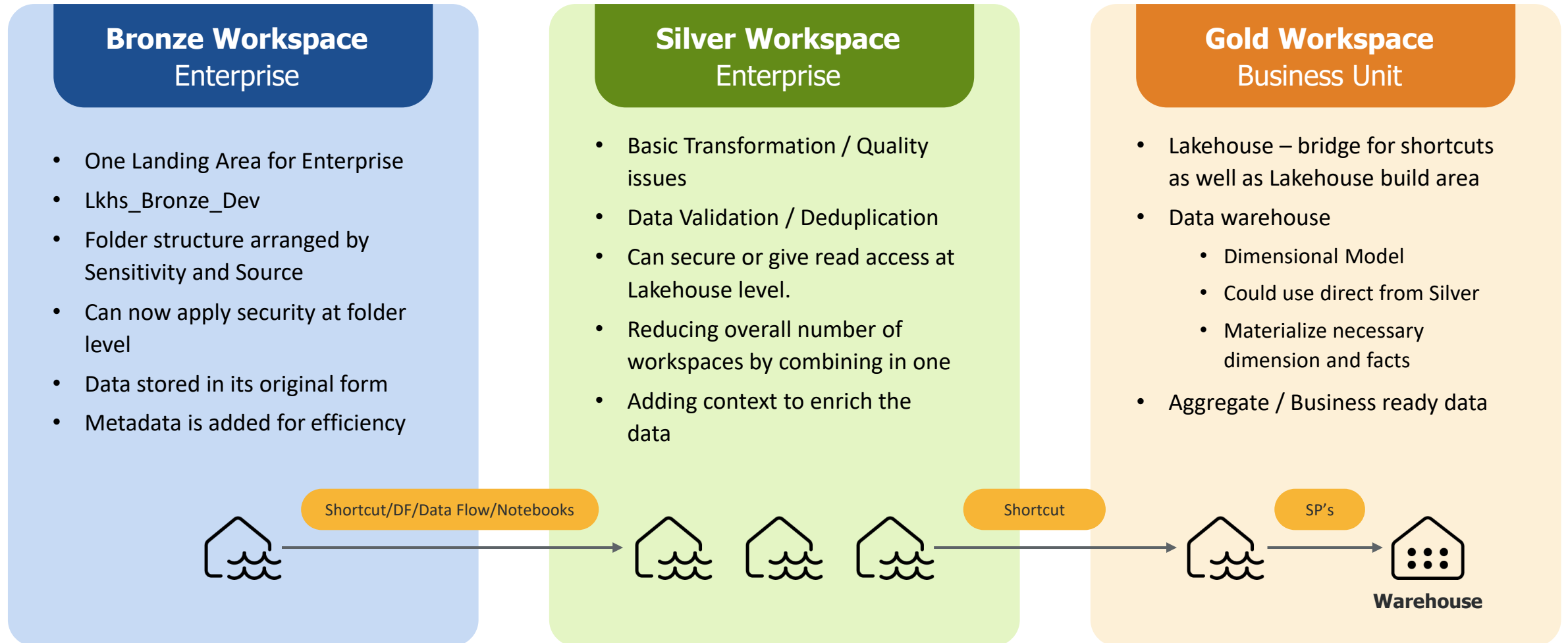
Medallion architecture built natively in OneLake



Microsoft Fabric Use Case



Medallion Architecture



Your Analytics upgrade path to Microsoft Fabric

Applicable for existing customers

*Solution applicable for investments in Synapse Spark.

Investment	Pattern	Solution
 #1 Azure Databricks only	Using Azure Databricks for entire implementation.*	Continue relying on Azure Databricks Lakehouse. Enable Microsoft Fabric with ADLS G2 Shortcuts to use the best of both worlds between ADB + Fabric.* Replatform notebooks and pipelines to Lakehouse in Fabric (OneLake, Spark, Notebooks, Delta)
 #2 Azure Databricks and Synapse	1. Using Azure Databricks and Spark and data engineering. 2. Synapse SQL Dedicated Pools for data warehouse and serving	1. ADLS G2 Shortcuts for data lake. Data Factory to move data into OneLake-based lakehouses or Fabric tables. 2. Large datasets: consider incremental copy or using ADLS → OneLake shortcuts to avoid full copy during migration. 3. Delta/Transactional data: preserve change data capture (CDC) or use incremental load patterns to avoid long outages.
 #3 Power BI	Using Power BI Dataflow for data preparation or other legacy reporting services.	1. Plan Fabric Capacity 2. Enable Fabric in Tenant 3. Optionally convert Power BI Dataflows to Data Factory Dataflows. 4. Validate and Optimize

Current Platform: Data Warehouse Appliances

Investment

Using a legacy data warehouse appliance, like Teradata, Oracle, Netezza, Greenplum, and others.

Current Upgrade Pathway



Datavail is here to help you explore Microsoft Fabric for new use cases and new workloads.

Where applicable, Datavail can help you modernize your data estate towards a Lakehouse pattern. Do a full upgrade to a Lakehouse today with Fabric. Use Fabric Shortcuts to a Delta Lakehouse in ADLS G2 to use the best engines seamlessly to Fabric.

Microsoft Fabric Advantages



Complete analytics platform

- ✓ Use a single product with a unified experience and architecture that provides role-specific capabilities for all data professionals.



Lake centric and open

- ✓ Connect the clouds and services you need to an open, scalable analytics solution, all standardized on Delta Parquet format.



Empower every business user

- ✓ Equip everyone with access to powerful, self-serve analytics to innovate faster, enabling real-time insights that unlock impact.



AI-powered experiences

- ✓ Seamlessly augment analytics systems with generative AI to reduce data estate fragmentation.

Current Platform: Azure Data Factory

Investment

Using ADF for integration to Azure or other services.

Current Upgrade Pathway



Datavail is here to help you explore Microsoft Fabric. For new use cases and new workloads, consider Data Factory in Microsoft Fabric. Plan for a full upgrade to Fabric for your Azure Data Factory workloads in 2025.

Fabric data pipelines will cover the majority of ADF activity capabilities. Notably, a few ADF activities will not be available in data pipelines (U-SQL (deprecated), AML v1, SSIS, and Mapping Data flows).

- MS actively working on enabling native SSIS package execution.
- For Mapping Data flows (MDF), the following will be available:
 - For smaller ADF usage, guidance docs and videos on how to upgrade your Mapping Data Flow (MDF) to Dataflow Gen2.
 - For large ADF usage, customers will be able to mount ADF pipelines in Fabric or convert their MDF to run as a Spark Job
 - In future, customers can mount their ADF, enabling them to run their ADF pipelines as it is in Fabric.
- There is the capability to mount data pipelines in a Synapse workspace.

Microsoft Fabric Advantages

Govern data across your organization

- ✓ CI/CD for Power BI.
- ✓ Foundational cataloging with Purview included.
- ✓ Security model provides source-to-serve data security at artifact level.

Unify your data estate

- ✓ Legacy Microsoft (SSIS, SSRS, SSAS, Power BI Report Server, AAS, Azure Time Series Insights) stack transitions easily to Fabric.
- ✓ Connect other storage accounts or clouds to OneLake.

Manage powerful AI models

- ✓ Build your own LLMs with Azure Open AI Service integration.
- ✓ Explore AI through BI use cases with PyTorch, sci-kit learn, and TensorFlow support.
- ✓ Adds Data Science capabilities not currently available in Synapse.

Empower everyone in your business

- ✓ ROI in cost savings and performance increases.
- ✓ Break data silos: Multiple personas access the same copy of data – single source of truth.

Use cases

Use case 1 : A very small law firm with ten users who create and use the Power BI content

Size of Business	No. of Users	Operational Hours (Per Day)	Operational Days (per Month)	Total Hours (Per Month)	Peak Hours (Per Month)	Workload	Data Size (GB)	Total Estimation (PRO)	Total Estimation (PPU)
XXS	10	8	22	176	0	Low	50	\$164.51	\$264.51

- **Use Case -1** : A very small Law Firm with 10 users who create and use the Power BI content
 - Size of the Business : Extra Extra Small
 - Number of Users use Power BI: 10
 - Operational hours: 8 Hours per Day
 - Operational Days : 22 Days per Month
 - Workload: Low workloads
 - Data Size: 50 GB
- **Cost Estimation** :Based on the above attributes , it's a very small team with Low workloads , so the best possible option for this Law Firm is to acquire individual user licenses with basic Fabric Capacity to maintain the low workloads
 - Power BI Pro License (\$10) * Number of Users (10) => $\$10 * 10 \Rightarrow \100.00
 - F2 Capacity Per hour(\$0.36) * Total Number of Operational Hours (176) => $\$0.36 * 176 \Rightarrow \63.36
 - Onelake Storage Per GB (\$0.023) * Total Data Size in GB (50) => $\$0.023 * 50 \Rightarrow \1.15
 - Total Cost Estimation => \$164.51

To avail Premium features , they need Premium Per User (PPU) License

 - Premium Per User(PPU) License (\$20) * Number of Users (10) => $\$20 * 10 \Rightarrow \200.00
 - F2 Capacity Per hour(\$0.36) * Total Number of Operational Hours (176) => $\$0.36 * 176 \Rightarrow \63.36
 - Onelake Storage Per GB (\$0.023) * Total Data Size in GB (50) => $\$0.023 * 50 \Rightarrow \1.15
 - Total Cost Estimation => \$264.51

Use cases

Use case 2 : A medium team of sales with one hundred users who create and use the Power BI content

Size of Business	No. of Users	Operational Hours (Per Day)	Operational Days (per Month)	Total Hours (Per Month)	Peak Hours (Per Month)	Workload	Data Size (GB)	Total Estimation (PRO)	Total Estimation (PPU)
M	100	12	22	264	50	Medium	500	\$1106.54	\$2106.54

- **Use Case -2** : A Sales Team with 100 users who create and use the Power BI content
 - Size of the Business : Medium
 - Number of Users use Power BI: 100
 - Operational hours: 12 Hours per Day
 - Operational Days : 22 Days per Month
 - Workload: Medium workloads
 - Data Size: 500 GB
- **Cost Estimation** :Based on the above attributes , it's a medium sized team with moderate workloads , so the best possible option for this Sales team is to acquire individual user licenses with basic or decent Fabric Capacity to maintain the workloads, **during peak hours we can scaleup the CUs based on the need**
 - Power BI Pro License (\$10) * Number of Users (100) => $\$10 * 100 \Rightarrow \1000.00
 - F2 Capacity Per hour(\$0.36) * Total Number of Operational Hours (264) => $\$0.36 * 264 \Rightarrow \95.04
 - Onelake Storage Per GB (\$0.023) * Total Data Size in GB (500) => $\$0.023 * 500 \Rightarrow \11.50
 - Total Cost Estimation => \$1106.54

To avail Premium features , they need Premium Per User (PPU) License

 - Premium Per User(PPU) License (\$20) * Number of Users (10) => $\$20 * 100 \Rightarrow \2000.00
 - F2 Capacity Per hour(\$0.36) * Total Number of Operational Hours (264) => $\$0.36 * 264 \Rightarrow \95.04
 - Onelake Storage Per GB (\$0.023) * Total Data Size in GB (500) => $\$0.023 * 50 \Rightarrow \11.50
 - Total Cost Estimation => \$2106.54

Use cases

Use case 3 : A large team of health care professionals with two hundred and fifty users who create and use the Power BI content

Size of Business	No. of Users	Operational Hours (Per Day)	Operational Days (per Month)	Total Hours (Per Month)	Peak Hours (Per Month)	Workload	Data Size (GB)	Total Estimation (PRO)	Total Estimation (PPU)
L	250	24	30	720	150	Medium	2000	\$2805.20	\$5305.20

- **Use Case -3** : A large team of Healthcare Professionals with 250 users who create and use the Power BI content
 - Size of the Business : Large
 - Number of Users use Power BI: 250
 - Operational hours: 24 Hours per Day
 - Operational Days : 30 Days per Month
 - Workload: Medium workloads
 - Data Size: 2000 GB (2TB)
- **Cost Estimation** :Based on the above attributes , it's a large sized team with moderate workloads , so the best possible option for this team is to acquire individual user licenses with basic or decent Fabric Capacity to maintain the workloads, **during peak hours we can scaleup the CUs based on the need**
 - Power BI Pro License (\$10) * Number of Users (100) => $\$10 * 250 \Rightarrow \2500.00
 - F2 Capacity Per hour(\$0.36) * Total Number of Operational Hours (470) => $\$0.36 * 470 \Rightarrow \169.20
 - F8 Capacity Per hour(\$1.44) * Total Number of Peak Hours (150) => $\$1.44 * 150 \Rightarrow \216.00
 - Onelake Storage Per GB (\$0.023) * Total Data Size in GB (2000) =? $\$0.023 * 2000 \Rightarrow \46.00
 - Total Cost Estimation => \$2931. 20

To avail Premium features , they need Premium Per User (PPU) License

 - Premium Per User(PPU) License (\$20) * Number of Users (10) => $\$20 * 250 \Rightarrow \5000.00
 - F2 Capacity Per hour(\$0.36) * Total Number of Operational Hours (470) => $\$0.36 * 470 \Rightarrow \169.20
 - F8 Capacity Per hour(\$1.44) * Total Number of Peak Hours (150) => $\$1.44 * 150 \Rightarrow \216.00
 - Onelake Storage Per GB (\$0.023) * Total Data Size in GB (2000) =? $\$0.023 * 2000 \Rightarrow \46.00
 - Total Cost Estimation => \$5431. 20

Use cases

Use case 4 (Production Ready) : very large team with thousands of users who create and use the Power BI content

Size of Business	No. of Users	Operational Hours (Per Day)	Operational Days (per Month)	Total Hours (Per Month)	Peak Hours (Per Month)	Workload	Data Size (GB)	Total Estimation (PRO)	Total Estimation (PPU)
XL	500	24	30	720	250	High	5000	>=F64 - \$8,409.60(PAYG)	
XXL	1000	24	30	720	300	Very High	10000	>=F64 - \$8,409.60(PAYG)	
XXXL	1000+	24	30	720	350	VV High	10000	>=F64 - \$8,409.60(PAYG)	

- **Use Case -4** : A very Large Team with 1000s of users who create and use the Power BI content
 - Size of the Business : Extra Large
 - Number of Users use Power BI: 500+
 - Operational hours: 24 Hours per Day
 - Operational Days : 30 Days per Month
 - Workload: High Workloads - Very Very High Workloads
 - Data Size: 5000 GB
- **Cost Estimation** :For Large teams with hundreds of users and heavy workloads , its recommended to go with F64 and above

OneLogin

INC14526307 - RJO Access usin

Microsoft Fabric

Getting Started | Microsoft Fabr

app.fabric.microsoft.com/home

SQL Server Reportin... Knowledge_Base New folder Datavail HDFC Google Replication WMI_Events AG SSAS All Bookmarks

Microsoft Fabric Home

Home

Your free trial of paid Power BI features has ended. Upgrade your license to continue using those features. [Learn more](#)

Microsoft Fabric

All your data. In one location. Organize. Collaborate. Create.

Explore the experience

Power BI

Find insights, track progress, and make decisions faster using rich visualizations.

Data Factory

Solve the most complex data integration and ETL scenarios with cloud-scale data movement and data transformation services.

Data Activator

Monitor data to trigger alerts and automated actions so your organization adapts to changing conditions in real time.

Microsoft

Sign out

Tobin Thankachen
tobin.thankachen@datavail.com

License type:
Free account

Start trial

[Learn more](#)

[Buy Fabric now](#)

[Buy Pro now](#)

[View account](#)

www.datavail.com

34



Put the Fabric Pieces Together

Fill out your session evaluation form for a chance to win a LEGO set.



Thank You



Tobin Thankachen

Global Practice Leader, Data Management,

info@datavail.com

datAvail