



WEBINAR

Data Observability with Microsoft Fabric

Agenda

- Introductions and Opening Remarks
- Data Observability & Governance
- Data Observability Built into Microsoft Fabric
- Microsoft Fabric Integrations Enhance Data Observability
- Power BI Integration Deep Dive: Direct Lake, Direct Query, & Data Import
- Microsoft Fabric Business Benefits & Timeline
- Q & A with our Expert Panel
- Ways to Get Started



Expert Panel



Walt De Petris

Senior Vice President
Modern Apps – Data & Analytics Practice



Gjnana Duvvuri

Sr Technical Architect
Data & Analytics Practice



Raja Manzoor

Principal Technical Specialist
Data Intelligence



Mohini Varma

Sr. Technical Specialist
Business Intelligence



Divya Paduvalli

Solution Architect
Sr. Data & AI Cloud



Peter Shand

Principal Cloud Solution Architect
Data & AI



Ramesh Kalava

Data & Analytics Leader,
Data & AI and Data Governance



netwoven.com



info@netwoven.com



+1 877 638 9683

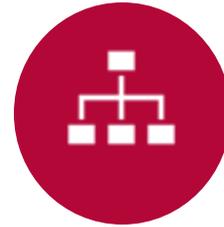
65%

Analysts at Gartner have predicted that 65% of the world's population in 2023 will be covered by laws similar to GDPR. This reality means that as different stakeholders acquire datasets, they must comply with various data and privacy regulations.

83%

Of organizations experience more than one data breach in their lifetime. Hence the importance of Data Governance place a key role in digitalization

Data Governance Implementation Challenges



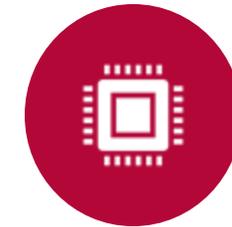
Lack of alignment



Data Silos



Data Security & Classification



Integrated Security

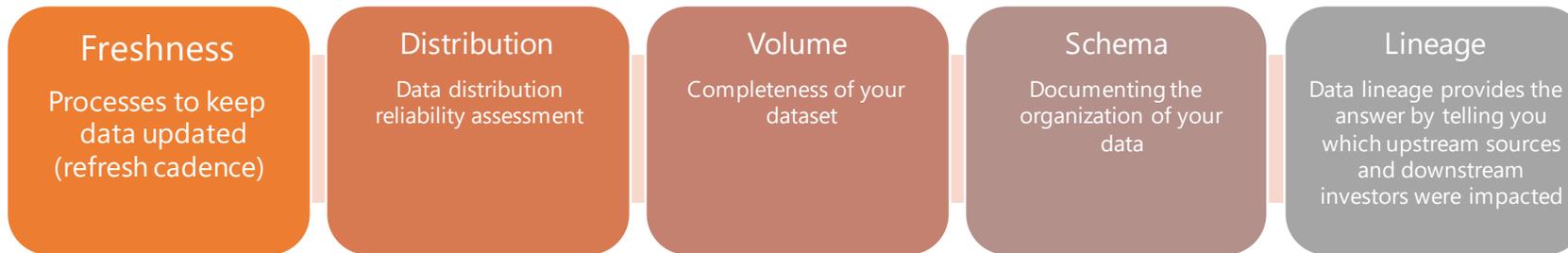
More details in our free [Data Observability eBook](#)

Data Observability

What is Data Observability

“Data observability is the ability of an organization to have a broad visibility of its data landscape and multilayer data dependencies (like data pipelines, data infrastructure, data applications)...” [2023 Top Tech Trends Data Observability](#)

Main Pillars



Better Data Quality



Why It Matters



Azure Cloud Adoption Framework Supports Data Observability

Azure Services Support Data Observability

Azure enables Data Observability & Governance with the following services:

- Data Platform Service Monitoring
- Data Pipeline Performance Monitoring
- Data Quality Monitoring
- Data Lineage
- Data Discovery

Modern Data Maturity Model

Cloud Adoption Framework (CAF) gives a 5 Stage maturity rating to help assess where your organization is on the journey of data observability

- Stage 1 (Learning)
- Stage 2 (Planning)
- Stage 3 (Evolving)
- Stage 4 (Advanced)
- Stage 5 (Highly Advanced)

Data Governance and Observability Tightly Connected



Data governance helps set the framework for data observability, setting guidelines for what data to monitor and how frequently while also establishing necessary thresholds. Once these guidelines are set, data observability helps handle these activities through continuous identification, troubleshooting, and resolution of problems outlined in the data governance strategy.



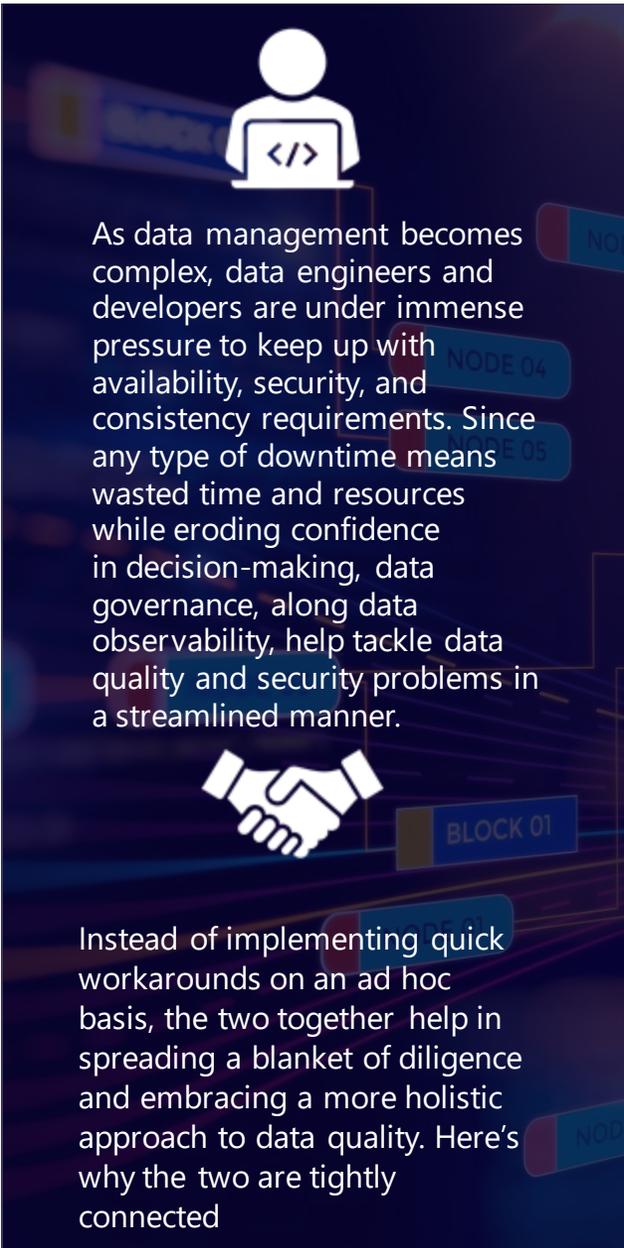
Data governance and data observability form the foundation of a comprehensive data pipeline. Together, they streamline data management and make maintenance easy through continuous monitoring of data metrics, logs, and traces to predict anomalies and abnormal behavior.



Data observability also helps in bridging the gaps that exist in data governance. Since most data governance policies are owned and managed by IT/dev teams that do not have an end-to-end business perspective, data observability ensures a well-rounded, comprehensive, and contextual approach to resolving bottlenecks and driving results.



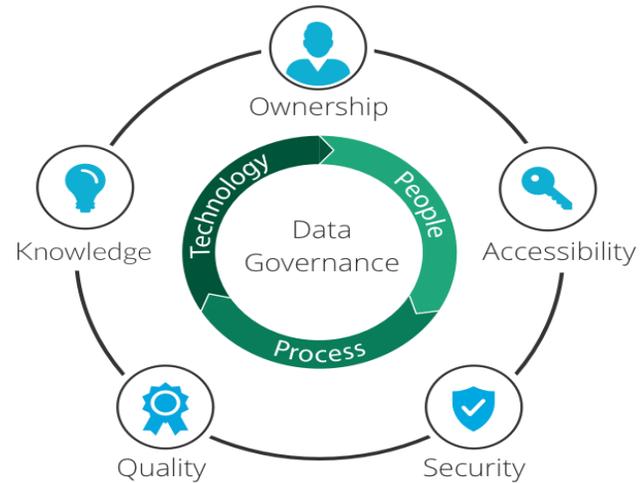
When used together, data governance and observability also help organizations check if performance drops below the minimum threshold and alert engineers in advance to implement proactive fixes.



As data management becomes complex, data engineers and developers are under immense pressure to keep up with availability, security, and consistency requirements. Since any type of downtime means wasted time and resources while eroding confidence in decision-making, data governance, along with data observability, help tackle data quality and security problems in a streamlined manner.

Instead of implementing quick workarounds on an ad hoc basis, the two together help in spreading a blanket of diligence and embracing a more holistic approach to data quality. Here's why the two are tightly connected

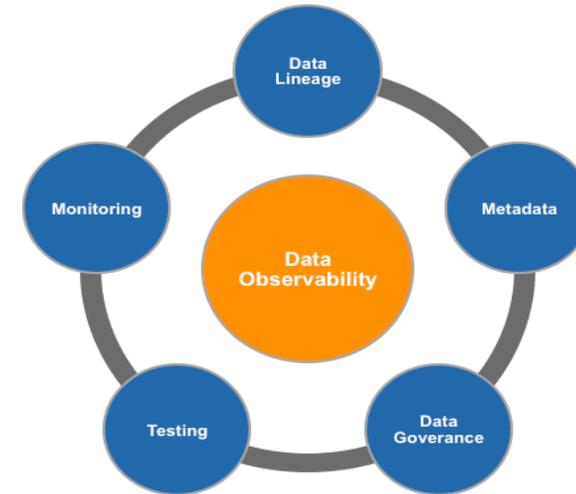
Difference between Data Observability and Data Governance?



Data Governance

Data governance helps in setting the right data management policies and procedures.

Although data governance helps in establishing the right set of data management policies and procedures, current data stacks are growing beyond boundaries. With data sets now scaling with more data sources, more tables, and more complexity, there is a pressing need to maintain a constant pulse on the health of these systems. Since any amount of downtime can lead to partial, erroneous, missing, or otherwise inaccurate data, organizations need to do better than just implementing a handful of policies.



Data Observability

Data observability aids in the constant monitoring of data to identify, resolve, and prevent issues, thus strengthening the overall data architecture.

Data observability enables organizations to cater to these increasingly complex data systems and support an endless ecosystem of data sources and formats. By providing a real-time view of the health and state of data across the enterprise, it empowers them to identify and resolve issues and go far beyond just describing the problem. Observability provides much-needed context to the issue, paving the way for a quick resolution while also ensuring it doesn't transpire again.

Data Observability Built into Microsoft Fabric

- Microsoft Fabric Integrated Solution Software as a Service
- Fabric Brings Data Observability Industry Framework to Enterprises

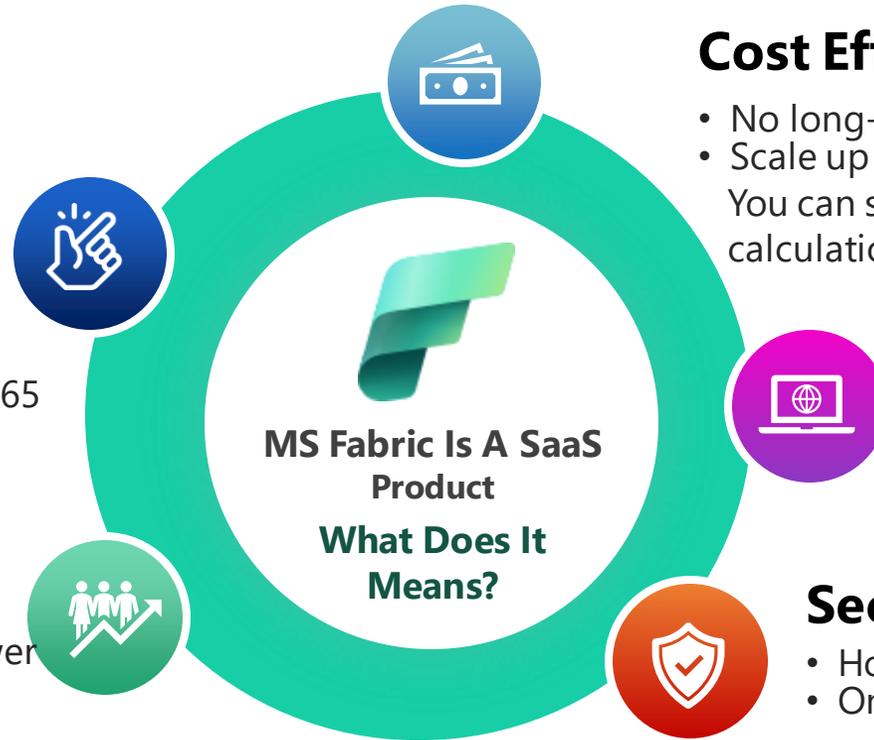
Microsoft Fabric Integrated Solution Software as a Service

Ease of Use

- No infrastructure to manage
- Start in minutes, without IT knowledge
- Same UI as Power BI, Microsoft 365

Scalability

- Add or remove users as you wish
- Scale up or down processing power as needed
- Bursting and smoothing help you pass the Monday 9 am rush



Cost Effective

- No long-term commitment
 - Scale up and down when you need
- You can scale up for a few hours – for end of period calculations – and then scale down

Accessibility

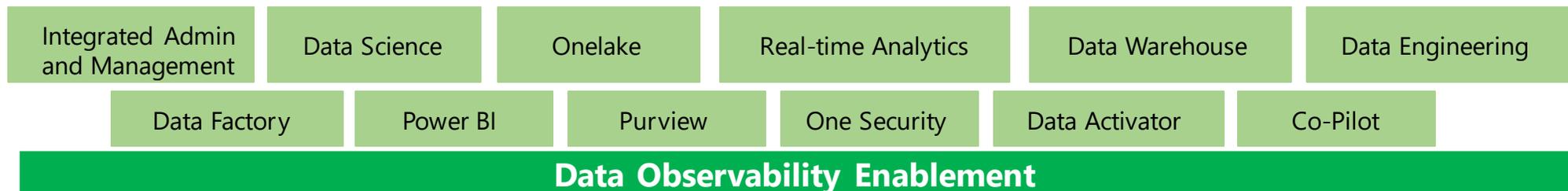
- Works on any PC or Tablet
- You just need a web browser

Security

- Hosted on Microsoft cloud infrastructure
- One security: same rights across tables, lake files

Software As A Service

**You Rent The Software + The Infrastructure
Pay Only For What You Use**



Microsoft Fabric Licensing*

No cost during preview & 60 Day Trial available after GA

Infrastructure



Processing Capacity

Shared across all analytics workloads can be resized up, down, and even paused attached to a specific region. Sold as capacity unit (CU), from 2 to 2048
\$0.18/CU/Hour, minimum 2 CU (Only \$9/Day!)



Additional Storage

You pay for the storage you consume. Prices vary from region to region
\$0.023/GB/Month for US West 2



Network

Cross-region data transfer charges, egress charges for multi-cloud shortcuts



Individual License

Each Fabric user needs a license. There are 3 types of licenses: Free, Pro, and Premium

FREE

Access Fabric web app
Create Fabric workspaces
Create non-Power BI items in Fabric workspaces
Share non -Power BI items in Fabric workspaces
View shared Power BI items if workspace has a Fabric Capacity => 64CU

PRO
\$10/ User/
Month

View shared content
Create Power BI items in all workspace

PREMIUM Enable Power BI premium features
\$20/ User/
Month

* As Of 10/10/23, subject to Microsoft licensing terms at General Availability(GA)

Fabric brings Data Observability industry framework to enterprises

Time to value

1

Microsoft Fabric connects to your existing stack quickly and seamlessly and not require modifying your pipelines, writing new code, or using a particular programming language.

Security-first architecture

2

Microsoft Fabric provides robust identity and access management, encryption at rest and in transit, and threat detection capabilities and governance.

Minimal Configurations

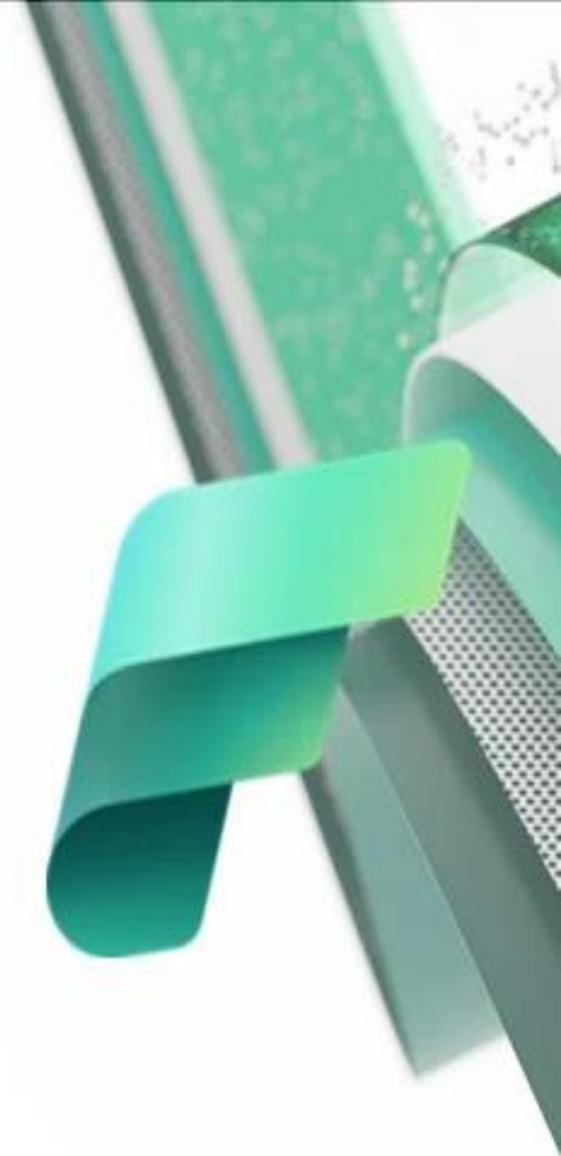
3

Microsoft Fabric requires minimal configuration on your end to get up and running and practically no threshold-setting? A great data observability platform uses ML models to automatically learn your environment and your data.

[Data Observability using Microsoft Fabric\(netwoven.com\)](https://netwoven.com)

Data Observability Integrated in Microsoft Fabric

- Data Activator helps Data Exploration
- Purview helps go beyond compliance
- Graph Data Connect helps integrate Microsoft 365 data
- Dataverse “View in Microsoft Fabric” simplifies data reporting
- Data Wrangler improves Data Observability
- Power BI & Direct Lake Improves performance



Data Activator (Public Preview 10/10/23)

Applied Data Observability

How Data Activator works

Simple 3-step Process to Get Started:

- 1 **Connect to Your Data**
- 2 **Detect Actionable Conditions**
- 3 **Trigger Action**

Putting Data Activator to Use



Sales

Alert a sales manager if a particular customer is in arrears with their payments



Inventory

Check whether inventory levels for a particular product are sufficient and notify an operations manager if not.



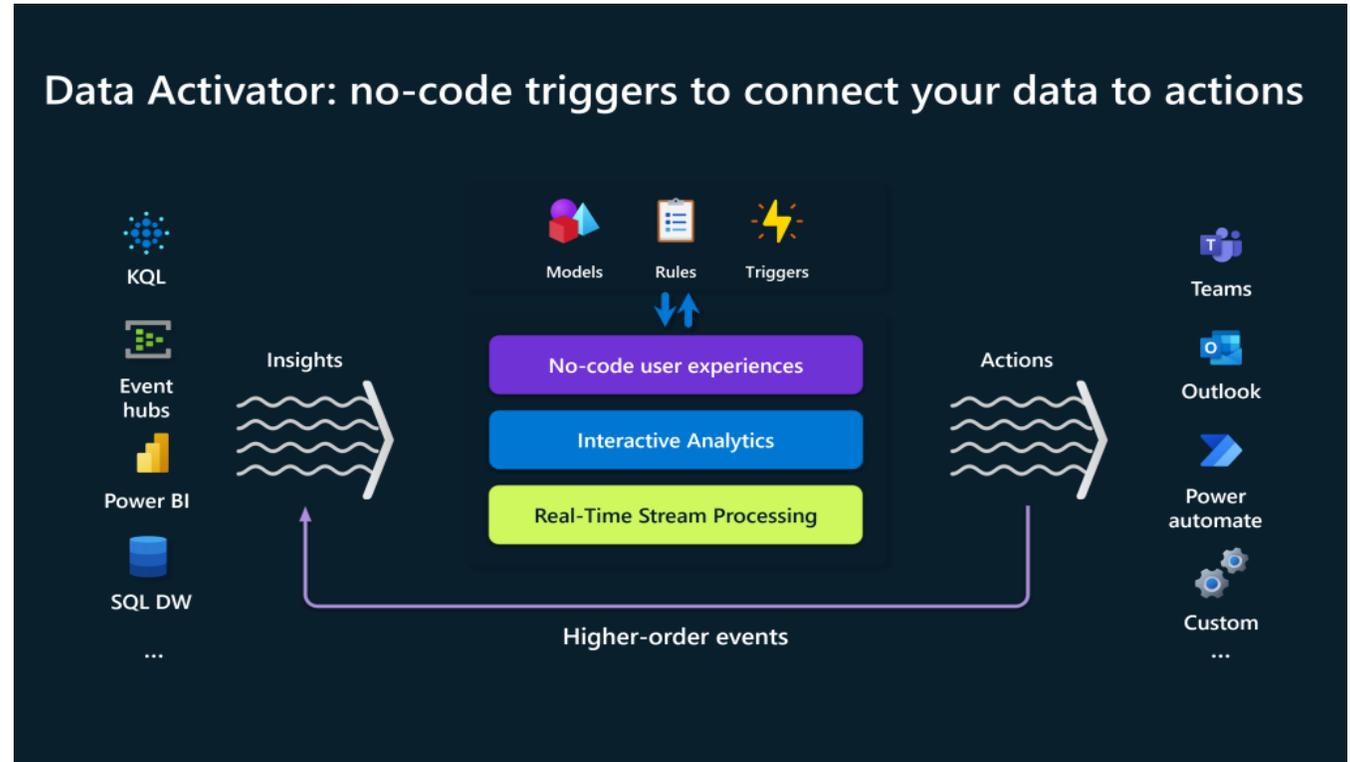
IT Operations

Automatically monitor data quality metrics and kick off remedial processes if those metrics are below target.



IoT

Automatically create an engineering support ticket if a refrigerator temperature is too high.



Go Beyond Compliance with Microsoft Purview

- 1. Microsoft Purview** represents a new way to approach your data strategy, with a unified view across multi-cloud, multiplatform environments.
- 2. Security and compliance leaders** can finally bring together the business users of data with the protectors of data
 - Data visibility and governance
 - Discover, understand, and govern data across clouds and platforms
 - Data loss prevention Safeguard data wherever it lives
 - Data risk management
 - Identify, manage, and reduce insider risk and other potential data vulnerabilities.
- 3. Purview(admin only) is already in Microsoft Fabric**

Power BI Microsoft Purview hub

Fabric Trial: 59 days left

Microsoft Purview hub (preview)

Explore, govern, and secure your data in Microsoft Fabric and beyond. [Learn more about Microsoft Purview](#)

Govern, secure, and manage your entire data estate

Get started with Microsoft Purview

Browse, search, and discover data from your org

Data Catalog

Classify and protect your data, wherever it lives

Information Protection

Microsoft Fabric data

Items Sensitivity

View key insights for your Microsoft Fabric data. [Learn more about Microsoft Fabric insights](#)

Workspace name	Certified	Promoted	None
Total	0	0	

Settings

Preferences

- General →
- Notifications →
- Item settings →
- Developer settings →

Resources and extensions

- Manage personal storage →
- Power BI settings →
- Manage connections and gateways →
- Manage embed codes →
- Azure Analysis Services migrations →

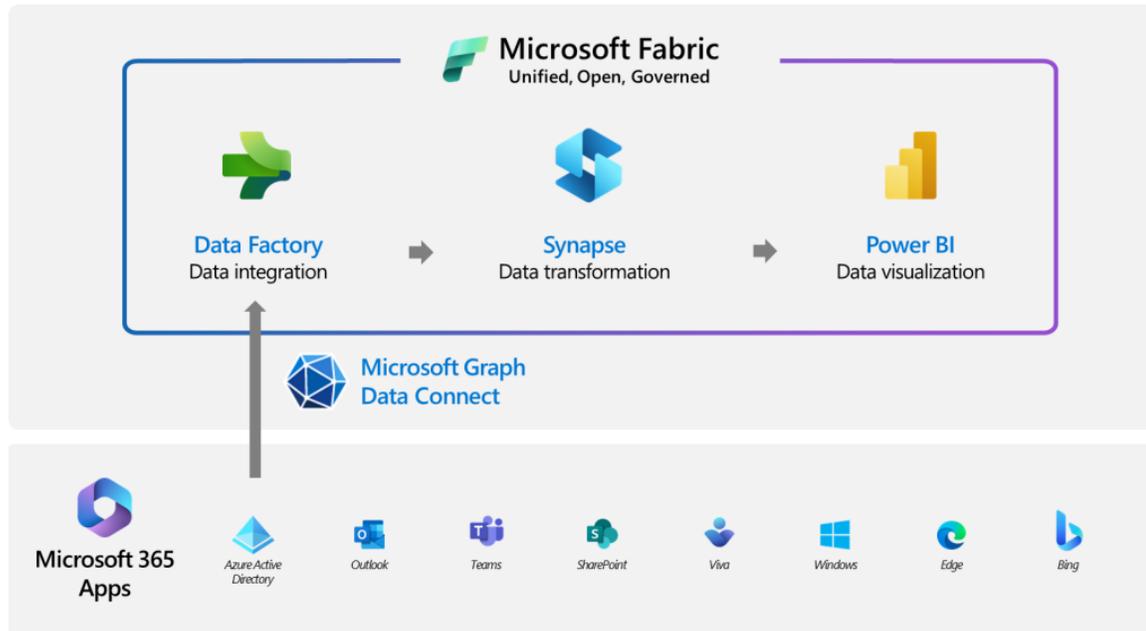
Governance and insights

- Admin portal →
- Microsoft Purview hub (preview) →

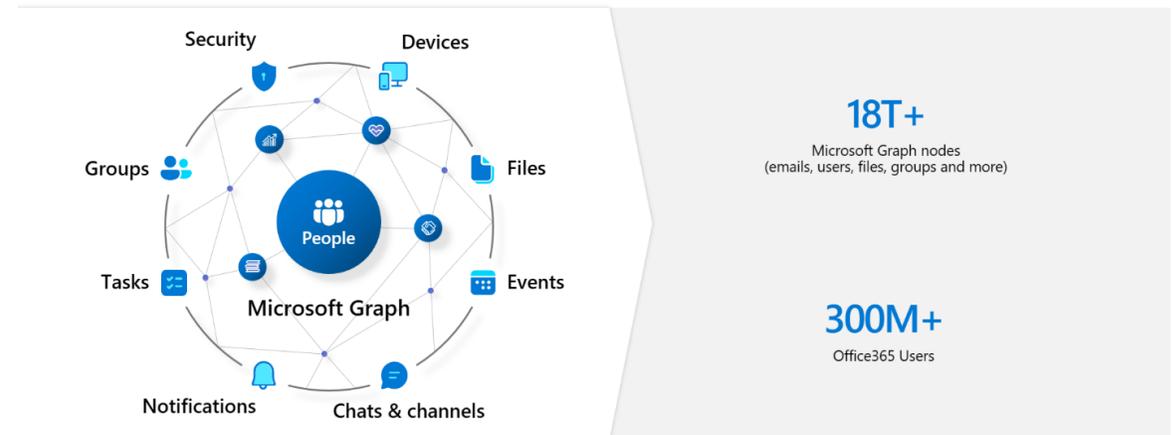
Collaboration and Communication Integration

Access all your Microsoft 365 data through Microsoft Graph Data Connect

To access your Microsoft 365 in the Microsoft Graph for enterprise analytics, we use Microsoft Graph Data Connect, a secure, scalable solution that enables you to integrate relevant Microsoft 365 datasets into Microsoft Fabric.

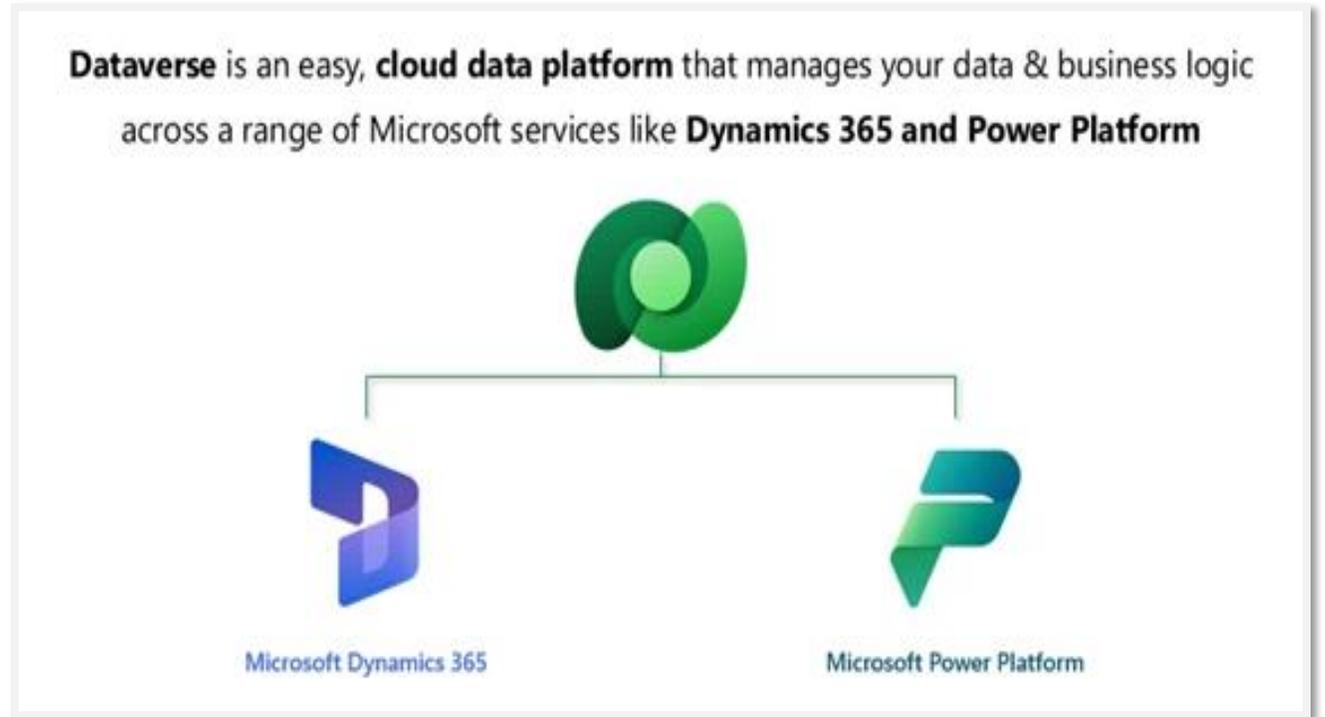


Collaboration and communication activities generate a massive, rich amount of data in M365



Dynamics 365 and Microsoft Fabric Seamless Integration

- **Organizations want to make better decisions that are based on business application data.** Dataverse integrates Dynamics 365 business application data with Fabric, making it easily discoverable and accessible directly.
- **Direct access reduce the risk of data silos and enhances data observability.** This integration helps data deliver actionable insights and reporting for better decision-making.
- **“View in Microsoft Fabric” for Dataverse.** Coming later this year to every Dynamics 365 customer, “View in Microsoft Fabric” automatically makes all your Dynamics 365 data available for analysis in Microsoft Fabric without having to copy data, build ETL pipelines, or use third-party integration tools. With just one click, you’ll get more insights from your business data stored in Dataverse.



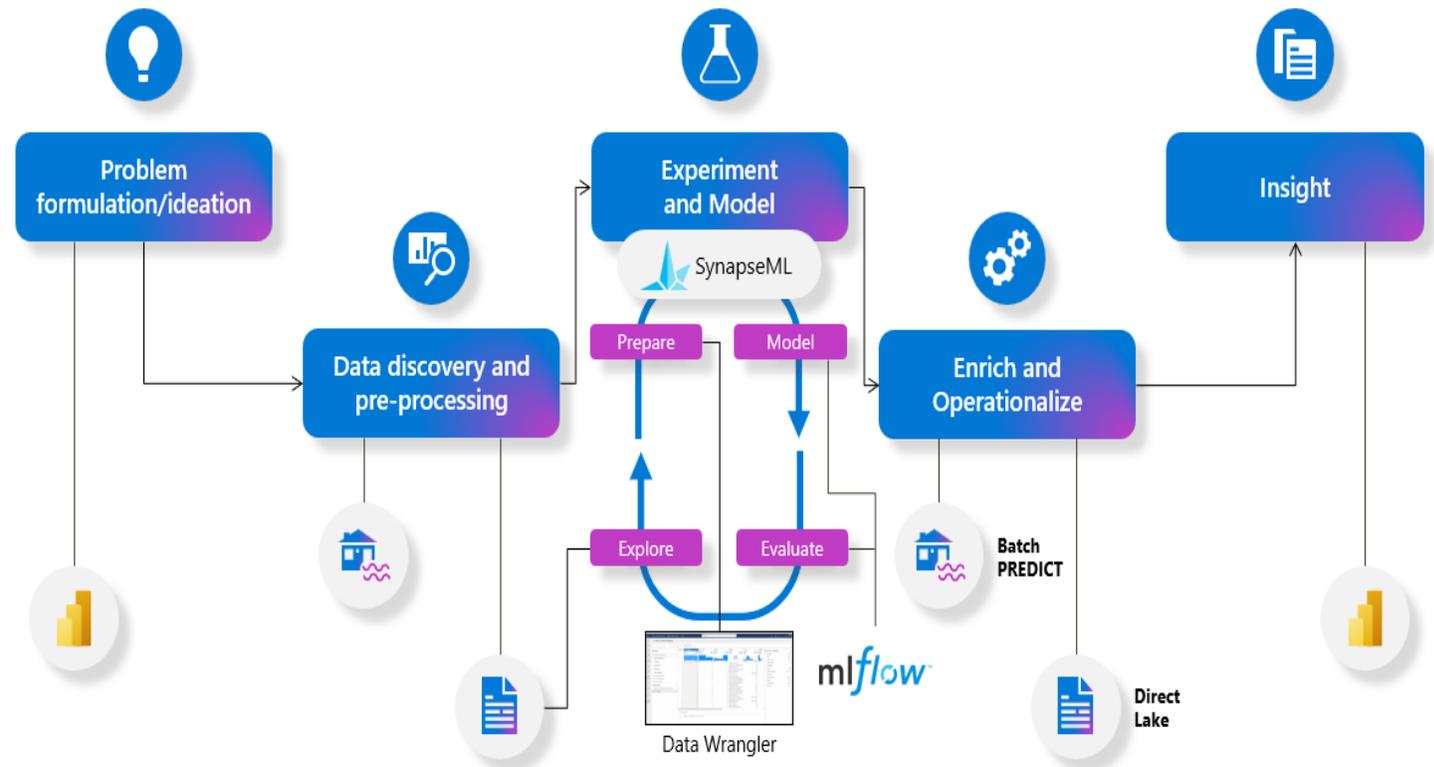
Data Science Integrated into the Analytics Workflow

- Problem formulation and ideation
- Data discovery and pre-processing
- Experiment and build ML models
- Enrich and operationalize
- Gain insights

Data prep and code generation with Data Wrangler

- **Data Wrangler**, a powerful, intuitive tool for data wrangling and preparation. Data Wrangler makes data cleansing and preparation easier than ever before, while still allowing users to take advantage of the power of coding and reproducibility of Python.
- The dynamic data display, built-in statistics and chart-rendering capabilities, and the ability to get started with **Panda's** data in just a few clicks make this tool easily accessible to a range of experience levels, from novice developers to seasoned professionals.
- Future updates will include support for Spark and a natural language processing "to code" functionality via Azure OpenAI.

Data science for business insights



Power BI is Part of Fabric

- Fabric helps Power BI as an integrated service
- Direct lake and different methods for power BI refresh methods
- Power BI benchmarks with different method



Microsoft Fabric Helps Power BI with Data Observability



**Seamless
integration and
immediately
accessible**



**Unified data
foundation with
OneLake and
Direct Lake mode**



**Enterprise-grade
collaboration with
Git integration**

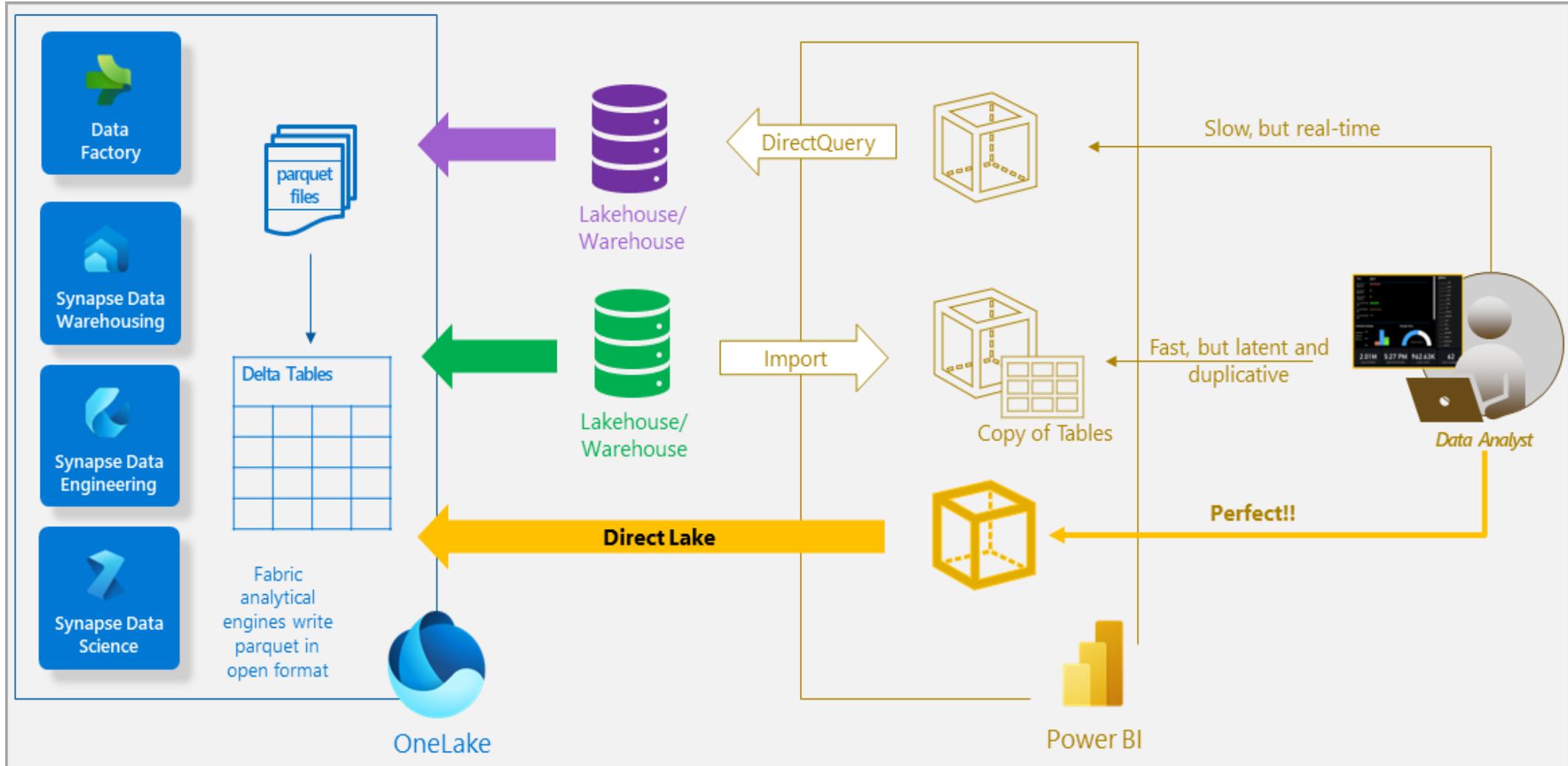


**Governance,
security and AI
monitoring in one
place**



**Next generation of
AI with Copilot in
Power BI & other
tools**

Direct Lake, Direct Query, & Import Methods to Power BI



Direct Lake – Power BI Benchmarks

Type of Report	Data Import	Direct Lake	Direct Lake	Notes
		V-Order On (default)	V-Order Off	
Simple Visual Report (SVR)	126-178ms	90-413ms	90-1910ms	Good performance for real time data
SVR with Date Slicer or filter	101-283ms	143-672ms	Did not complete	import performed little better
SVR with Average Top 20	110-224ms	95-125ms	111-1324ms	very good performance direct lake
Table Columns with Granular Data	23841-24443ms	25142-26120ms	25197-51949ms	even direct lake performed equally good

Data set details:
 - 500 k records
 This testing was conducted in Fabric preview environment with F64 CU

Observations

- Through tests, Power BI's Direct Lake mode showed promise, particularly when boosted with V-Order optimization, outperforming the traditional Import mode in certain scenarios. Its key feature is the live data connection, continuously providing the latest data, a convenience not present in the Import mode. However, this comes at a cost: maintaining the Fabric capacity continuously operational increases compute demands and, consequently, the cost. Hence, we caution, that if you don't need real-time time use import. Use semantic layer in direct lake mode. Use Gen2 dataflows
- Importantly, our tests revealed that Direct Lake doesn't consistently outperform Import mode without V-Order optimization. Therefore, **the choice between the two largely depends on your specific needs** . Now V-order is default
- An added complexity is Direct Lake's fallback to Direct Query under certain conditions, bringing along its own set of limitations, such as those related to DAX.
- In conclusion, Direct Lake presents an exciting potential for Power BI users, especially with V-Order optimization. It can handle large data sources and improve performance in specific scenarios. However, it's still a work in progress and we see improvements in performance and new updates that will further enhance this promising feature.

Microsoft Fabric Business Benefits & Timeline

- Business Benefits
- Microsoft Fabric Known Issues & Release Plans
- Journey to General Availability

Microsoft Fabric Business Benefits

1

Significantly reduce costs, improve collaboration, and simplify purchasing

Fabric provides deeply integrated, consistent, role-specific experiences for data engineers, data warehousing professionals, data scientists, data analysts, and business users. It makes collaboration easier, as data artifacts for all roles are stored in a single shared workspace. A single compute pool shared across all analytics workloads also simplifies purchasing and resource management and further reduces cost.

2

Avoid data silos, data duplication, and vendor lock-in

Fabric offers a SaaS, multi-cloud data lake called OneLake, which is built on ADLS Gen 2 and automatically available via existing APIs on a dedicated tenant to every customer. Data from ADLS Gen 2, AWS S3, and Google Storage (coming soon) can be directly linked via a virtualization capability called "Shortcuts." A single copy of the data and a shared open data format (Delta/Parquet) in OneLake powers all workloads, and a universal security model to set policy across all engines will ship soon.

3

Empower their business users and create a data culture

Fabric integrates seamlessly with Microsoft 365 apps. Users can easily access and analyze Onelake data from Excel, Teams, PowerPoint, and SharePoint. Fabric also connects with Microsoft Graph Data Connect to provide insights on customer relationships, business processes, security and compliance, and people productivity.

4

Accelerate time-to-market by harnessing AI

Copilot experiences powered by Azure OpenAI Service are built into every layer of Fabric to provide generative AI solutions for customers against their own data. Copilot in Fabric is coming soon and will enable all analytics personas, from the data engineer to the business user, to use conversational AI to speed up tasks. Copilot will also adhere to the organization's security, compliance, and privacy policies; Microsoft's Azure Open AI base models will not be improved using customer tenant data.



Fabric

Known Issues

Product experience

Administration & Management	Data Engineering	Data Factory	Data Science	Data Warehouse	OneLake	Power BI	Real-time Analytics
-----------------------------	------------------	--------------	--------------	----------------	---------	----------	---------------------

This is a list of known issues for Fabric features. Before submitting a support request, review this list to see if the issue that you are experiencing is already known and being addressed. Fixed issues are removed after 46 days. Click on the **URL** to redirect to the corresponding known issue documentation.

Issue ID	URL	Title	Issue publish date	Status	Fixed date
449	🔗	Pipeline isn't loading if user deployed with update app via public API Description: When the pipeline was deployed via public API with the 'update app' option (/rest/api/power-bi/pipelines/deploy-all#pipelineupdateappsettings), opening the pipeline page gets stuck on loading	7/5/2023	Fixed	8/24/2023
420	🔗	Moving files from outside of OneLake to OneLake with file explorer doesn't sync files	6/8/2023	Fixed	7/31/2023

Microsoft Fabric release plan documentation

The Microsoft Fabric release plan documentation announces the latest updates and timelines to customers as features are prepared for future releases.

Please follow
[Microsoft Fabric Known Issues](#) , aka.ms/FabricRoadmap and [Limitations](#)

Microsoft Fabric Journey to General Availability

From the date of the announcement of the Fabric preview we have seen improvements every month and regular introductions of new functionality. Here is our assessment of the progress:

- New features are being consolidated and integrated in each minor release
- Major kernel changes are settling down with improved performance and expanded functionality
- Customer adoption during preview is free for pre-pilot testing and can be upgraded to a cost-based enterprise version without migration when general availability occurs.
- Data governance and observability enhanced with AI & purview in recent upgrade
- A few things cannot work region to region like move, data warehouse and power BI. There are a few issues with shortcuts but we have found workarounds to help mitigate short-term needs
- Fabric licensing is integrated one model that lowers costs, is easier to administer and plan
- All-inclusive and integrated product services give better data observability.
- Check out the Microsoft Fabric open issues & release plan documentation sites often

Expert Panel



Walt De Petris

Senior Vice President
Modern Apps – Data & Analytics Practice



Gjnana Duvvuri

Sr Technical Architect
Data & Analytics



Raja Manzoor

Principal Technical Specialist
Data Intelligence



Mohini Varma

Sr. Technical Specialist
Business Intelligence



Divya Paduvalli

Solution Architect
Sr. Data & AI Cloud



Peter Shand

Principal Cloud Solution Architect
Data & AI



Ramesh Kalava

Data & Analytics Leader ,
Data & AI and Data Governance



netwoven.com



info@netwoven.com



+1 877 638 9683

3 Complimentary Ways to Get Started



Data Discovery & Advanced Analytics

Comprehensive overview of the modern cloud and hybrid data landscape to help accelerate time-to-value on your data:

- Data Optimization
- Data Scalability
- Data Performance
- Reporting, Real-Time, Predictive, Prescriptive Analytics



Modern Data Warehouse Demo

Modern Data and Analytics demos based on deployed solutions for real-world industry scenarios, including

- Retail
- Manufacturing
- Financial Services
- Other Sectors



Data Governance & Microsoft Purview

Comprehensive overview of the data ecosystem, and the benefits of:

- Data Transparency
- Compliance Assurance
- Data Collaboration
- Data Security

Wait. There's more!

- Additional Office Hours with Export QA
 - Oct 18 Wed, 9a PT/Noon ET
- Free eBook for Data Observability
 - Look for follow up email with links



Netwoven

Nick Simas

Customer Success Director
nsimas@netwoven.com

[Let's Meet! – My Calendar](#)



Abraham Ortiz (Abe)

Customer Development Manager
abraham.ortiz@netwoven.com

[Let's Meet! – My Calendar](#)





Netwoven Inc.

4000 Pimlico Drive
Suite 114-103 Pleasanton,
CA 94588, United States



+1 877 638 9683



info@netwoven.com



netwoven.com

Appendix



Netwoven Inc.

4000 Pimlico Drive
Suite 114-103 Pleasanton,
CA 94588, United States



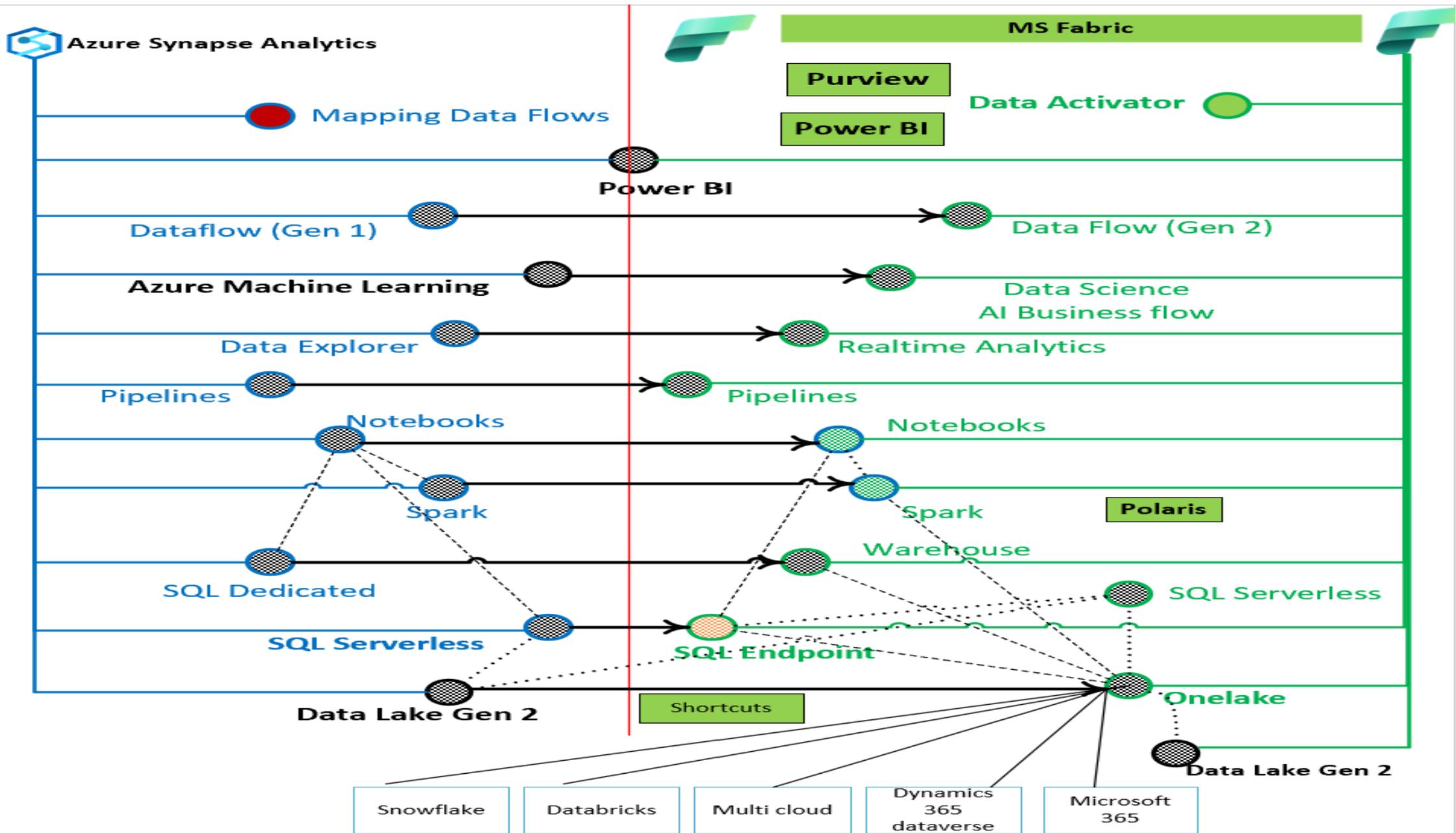
+1 877 638 9683



info@netwoven.com



netwoven.com



Azure CAF Data observability Maturity Explained

Stage 1 Learning	Every critical platform service (provider-managed and self-managed) is monitored in the data landscape.	Pipeline monitoring is minimal. Failures trigger alerts, but have no insights into any possible cause.	Data quality tests can be run from the pipeline, but no metric is measured or tracked.	Data lineage doesn't exist.	Data discoverability doesn't exist.	Stage 2 (Planning)	An initial draft of SLO, SLI, and SLA covers the most critical components needed for data observability. Platform monitoring data is centralized and there is a unified view of the entire data environment. All data incident management is manual.	Data pipeline performance metrics are defined and measured.	Data quality checks exist, but no standard metric is defined, measured, and visualized.	Data lineage is limited to single data product or isn't tracked.	Data discoverability is achieved but no sophisticated tools are used.
-------------------------	---	--	--	-----------------------------	-------------------------------------	---------------------------	--	---	---	--	---

Stage 3 (Evolving)	Well defined SLO, SLI, and SLA cover most critical almost all components for Data Observability. Data incidents are managed with specialized tools.	Platform monitoring data is correlated with data pipeline performance monitoring using some amount of automation.	Data quality checks are well defined and mapped to custom metrics.	Data lineage has matured to contain enough metadata needed for decision making.	Data discoverability is achieved using specialized data catalog tools.
---------------------------	---	---	--	---	--

Stage 4 (Advanced)	Dashboards track SLO, SLI, and SLA across the most critical data observability components. Platform monitoring data and pipeline performance monitoring data are correlated using automation.	Data incident tools monitor and measure TTD and TTR metrics for any incidents.	Data quality is maintained through a framework that's usable across multiple data products and tracked using dashboards.	Data lineage includes data quality tags and is connected to data discoverability.	Data lineage is now connected to data discoverability and includes data quality tags as well.
---------------------------	---	--	--	---	---

Stage	Data platform service monitoring	Data pipeline performance monitoring	Data quality monitoring	Data lineage	Data discovery
Stage 5 (Highly advanced)	Data is collected across all the data observability components from one or more data products in a unified view and is correlated using machine learning to find any anomalies. Dashboards track SLO, SLI, and SLA across all data observability components.	Data pipeline performance metrics are tracked across multiple data products. Root cause analysis is completed and driven by the system.	A high Level of trust in data quality is established. Data consumers can verify the reliability of data.	Data lineage is visually represented and is used in multiple ways, such as tracing root causes of pipeline failure, data quality analysis, and compliance.	Data consumers can easily find available data that they need.

Internal for panel members

1. The Growing Need for Data Governance and the interconnections with Observability

- Important to reduce data loss and comply with increased worldwide compliance requirements
- We spend a considerable amount of time on Data Governance and Observability.
- we have our own methodologies and Road Maps
- We work with other industry-leading organizations like Gartner and follow Azure CAF data observability recommendations. We also read a lot of research publications and blogs to improve our framework.
- The same we followed in our presentation. Hope the Microsoft team will not have any objections to supporting us in this subject because we took Azure the CAF reference from Microsoft in the presentation.
- We will not hold Microsoft on our focus on Data Observability on Microsoft Fabric. This is up to the Netwoven team to assert in the webinar presentation.

2. For The second part of the presentation, we expect the Microsoft team can help us which will give us variety and authenticity

- What does Microsoft Fabric Mean for Power BI through Data Observability?
- Direct lake with Power BI and other refresh methods
- Report Benchmarks comparisons
- New business opportunities with Microsoft Fabric – the data platform for the era of AI

3. The third part preview to GA progression we can take in QA or a brief presentation (Netwoven)