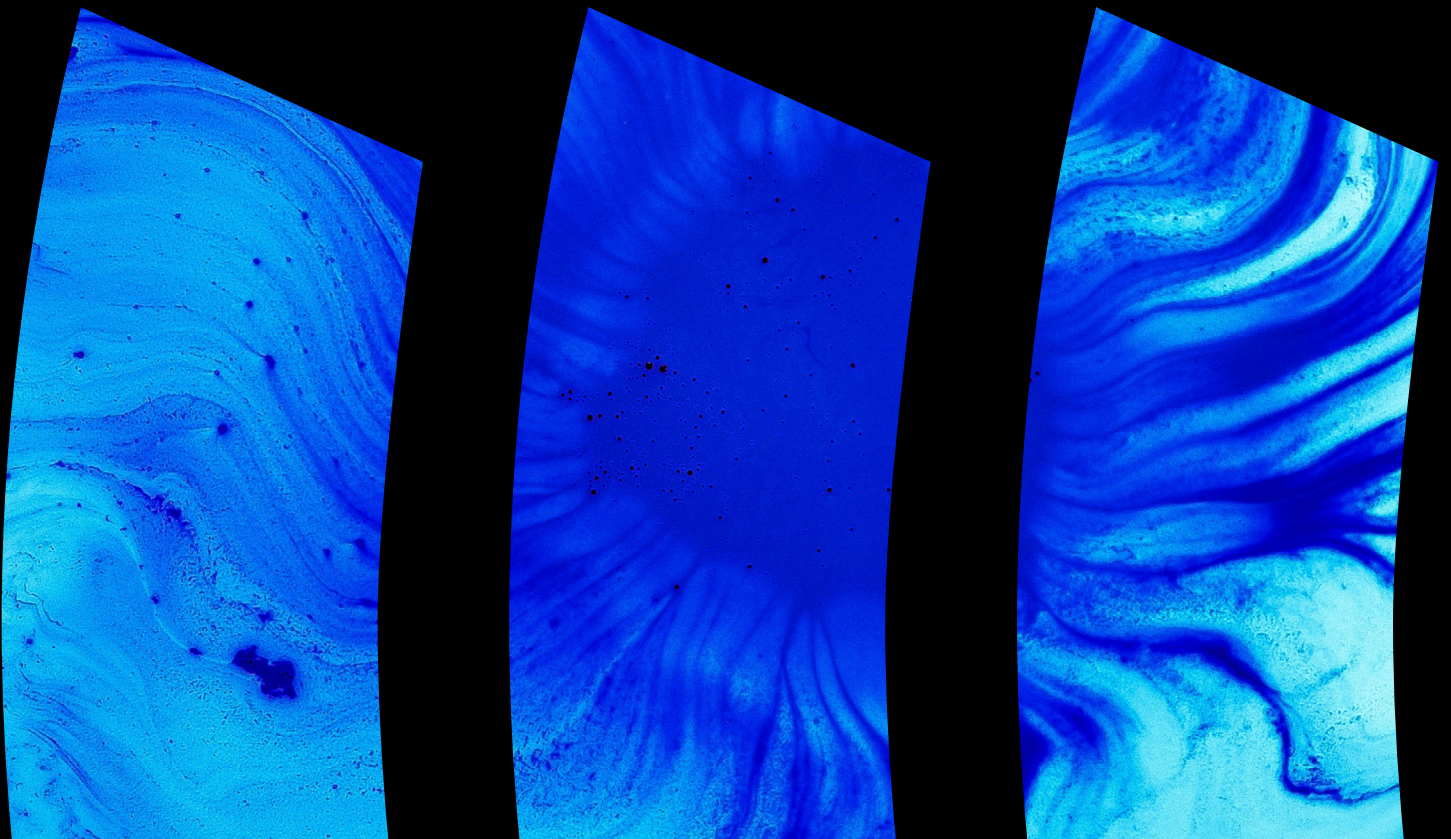


bjss

Microsoft

Fabric at Microsoft Ignite 2024



As a Microsoft Partner actively delivering Fabric to multiple enterprises, many of the announcements at Ignite support our view that Fabric is ready for production at scale.

The continued delivery of new features and innovations demonstrates that Fabric's evolution is advancing rapidly, from adding must-have capabilities to providing business benefits not seen in many equivalent products.

Framing Fabric's goals

One of the messages that Ignite has helped with, is the reframing of Fabric's goals and where it fits into a modern data-driven organization.

1. **An AI-powered data platform**
2. **An open and AI-ready data lake**
3. **AI-enabled business users**
4. **A mission-critical foundation**

Given the size and scale of functionality in Fabric we've broken down the key announcements to align with the Fabric four goals described above.



Goal 1: An AI-powered data platform

There were significant updates across the Fabric Real-Time Intelligence, Fabric Data Factory, Fabric Data Engineering, Power BI, and Industry Data Solutions workloads in Fabric.

Fabric databases

Microsoft announced a new class of database, allowing transactional and analytic workloads from within the Fabric experience. The first database, SQL database in Fabric, is now in preview.

The integration of Azure AI, vector search and RAG support, as well as the management of the service from within Fabric may be a good option to build some AI offerings, around your data, without the need to stand up additional databases that are hosted and managed elsewhere.

Ultimately this is a drive to make consuming AI on your data easier and more democratised.

Enhancements to Real-Time Intelligence (RTI)

RTI is now generally available in Fabric. You can use both low-code and no-code tools to ingest high-volume streaming data with high granularity, dynamically transform streaming data, query data in real-time for instant insights, and trigger automated actions based on the data. RTI is supported by the Real-time Hub which acts as the central place to discover and manage streaming data.

Some of the other improvements, all currently in preview for RTI, included new Fabric events, enhancements to Eventstreams and Eventhouses, and easier real-time dashboard sharing.

For Eventhouses, Microsoft announced a migration experience from Synapse Data Explorer (SDX), now in preview, to move your SDX clusters to Eventhouse with minimal disruption.

Announcements regarding integrating Fabric CI/CD tools including REST APIs, Git integration, and deployment pipelines we also made.

Fabric Data Factory

Enhancements to data pipelines in Fabric Data Factory include export/import capabilities to support the Apache Iceberg format, and a new Copilot in Fabric Data Pipeline experience, now in preview, to aid in developing data integration solutions via natural language.

As with RTI Fabric is introducing CI/CD support, into preview, for Dataflow Gen2 and Copy Job, including GIT integration, deployment pipelines, and REST APIs. This could be very significant for organizations that want to focus on repeatability and automation.

The support for formats like Iceberg, continue to show that Microsoft understand that you are more likely to bring your data to Fabric if you can integrate and consume via native pipeline support for third-party sources.

Fabric Data Science

AI functions are coming soon in preview to Fabric. AI functions provide a simplified API in Fabric notebooks for common AI enrichments for text like summarization, translation, classification, sentiment analysis, grammar correction. With a few lines of code you can leverage these functions without the need to wire-up your own solution.

Fabric Data Engineering

Fabric notebooks now support a pure Python experience, helping developers use Fabric notebooks for smaller datasets with a Python runtime. The Fabric notebook Python experience offers multiple built-in Python kernels, easy data integration with Fabric lakehouses and a quick setup with a five-second spin-up time. Live versioning for Fabric notebooks, enabling you to track change history, compare and restore previous versions, is now included. This, like many other features announced, are going to help manage adoption and effect more controlled usage.

New features to support improved geographic information system (GIS) mapping systems has been added, enabling users to run spatial analytics and visualize geographic data directly within Fabric Data Engineering and Data Science workflows.

The API for GraphQL is now generally available, enabling developers to access data from multiple sources in Fabric with a single query API. Introduced were several new capabilities including support for new data sources like Azure SQL and Fabric SQL databases. A dashboard has been added to simplify monitoring performance and API activity.

Power BI

A new Tabular Model Definition Language (TMDL) view in Power BI Desktop is coming soon. Providing a semantic-modelling-as-code scripting environment utilising TMDL to modify the semantic model within a code editor. Also announced was the ability to include Copilot-generated reports and page summaries in the emails of a Power BI subscription.

Sustainability industry solution is now GA

Announced at Ignite was the general availability of sustainability data solutions in Microsoft Fabric. Environmental, Social and Governance (ESG) data is supported to aid removing of data silos, standardization, enrichment, benchmarking and governing the data for integrations with business and management systems.

Workload development kit is now GA

The Fabric Workload Development Kit announced last spring is now generally available. It was created to help developers design, build, and interoperate applications within Fabric. Applications built using this kit appear as a native workload within Fabric, providing a consistent experience for users directly in their Fabric environment. These applications become available in the workload hub, providing a more marketplace, centralized resource for intra-organizational reuse, and consumption of third-party solutions.

Goal 2: An open and AI-ready data lake

OneLake, Fabric's unified, multi-cloud data lake, is designed to span an entire organization, connect to data, and reduce data duplication. Two key announcements have caught our eye, the OneLake catalog, and Open Mirroring.

OneLake catalog

The OneLake catalog is a solution to explore, manage, and govern your Fabric data estate. It's an evolution of the OneLake data hub experience with enhancements to help users discover and manage trusted data. It also provides governance for data owners with insights, and recommended actions. The OneLake catalog comes covers two areas: explore and govern.

Explore allows users to explore and manage the Fabric items they have access to in a central location with filters for domains, item type, owner, endorsement, tags. Users can select a Fabric item to take actions or explore description, data lineage, permissions, and activity.

Govern allows data owners to see an overview of their data with insights on endorsements, data labelling, and data type, getting recommended actions to enhance their data for improved quality and compliance. Explore tab is now available and the Govern tab will be coming soon in preview.

You can extend the capabilities of the OneLake catalog by connecting it with the Microsoft Purview Data Governance solution. Purview's catalog offering has capabilities for data discovery, curation, and quality.

Announced were deeper data quality support for sources like OneLake, Azure Databricks Unity Catalog, and Snowflake Polaris.

Given its advanced features and its ability to ingest data from multiple catalogs, Microsoft Purview Data Catalog is being renamed to Microsoft Purview Unified Catalog to reflect its broader capabilities.

Mirroring

Mirroring support has been extended with new sources available for mirroring, a zero-ETL way of ingesting data in near-real time from existing databases or data warehouses.

Azure SQL DB Mirroring is now general availability and SQL Managed Instance mirroring is now in preview.

Microsoft also introduced a new capability called Open Mirroring. Now in preview this feature extends Fabric by allowing applications or data providers to bring their data estate directly into OneLake. By enabling apps to write change data directly into a mirrored database within Fabric, Open Mirroring simplifies the handling of complex data changes and the resulting mirror.

Open Mirroring could allow a new approach to syncing your data and be analytics-ready from within OneLake.

Goal 3: AI-enabled business users

The right insights can help organizations be more efficient and get more done, but unless those insights are available in the tools your business teams use everyday then there remains a gap in productivity.

AI skills

At the Microsoft Build event last May, Microsoft announced a new capability called AI skills, which is an AI-powered conversational Q&A experience against your data. Announced today were enhancements, coming soon, for the support of semantic models and Eventhouse KQL databases, and multiple data sources at the same time in AI skills.

AI skills are being integrated with the Azure AI Agent Service. This will allow developers to build AI apps using Fabric data as a core knowledge source. This could be via a low-code chatbot or custom generative AI apps in Azure AI Studio.

Goal 4: Mission-critical foundation

All our clients want to be assured that Fabric meets their requirements with regards to running their key line of business and mission critical workloads. It means business continuity and disaster recovery, compliance with standards across industry like GDPR, CCPA, and HIPPA, security and governance are non-negotiable.

Workspace monitoring

To support observability, workspace monitoring in Fabric was announced. It is designed for admins and developers to view detailed diagnostic logs and workload metrics about their workspaces to troubleshoot performance issues, capacity performance, and data downtime.

Fabric will provide a read-only Eventhouse KQL database of workspace logs that users can query ad-hoc, analyze for patterns and anomalies, or save drafted queries as query sets. You can use this database to conduct root-cause analysis for errors, long running queries, refresh failures, and other issues. This feature is important, aiding operational management. We've experienced what happens when you can't fully uncover what is happening in the workspace, so we are excited by this feature.

Microsoft Purview integration

There is extended integration with Microsoft Purview, Microsoft's unified data security, data governance, and compliance solution. Microsoft Purview Protection policies have been extended to automatically enforce access permissions to Fabric Databases alongside already supported items like semantic models, notebooks, and lakehouses. This is also coming soon to Power BI reports, datamarts, and warehouses.

Integration with Purview Data Loss Prevention (DLP) policies will enable security admins to detect the upload of sensitive data. You can now automatically restrict access to semantic models with sensitive data.

Billing for DLP, including for Power BI artifacts, will be part of a new Purview pay-as-you-go consumptive model that will begin charging in January.

Cross-tenant capabilities

Cross-tenant collaboration capabilities in Fabric are being extended, with external data sharing now being general availability. External data sharing allows you to directly share OneLake tables and folders with other Fabric tenants.

As this is built on OneLake's shortcut capabilities, a welcome benefit is that you can share data in place without copying the data. Coming soon is the ability to share multiple OneLake folders and tables in a single share as well as sharing directly from data warehouses and Eventhouses in Fabric.

Also announced was a new tenant switcher in Fabric. Users with access to more than one Fabric tenant can switch between tenants directly from the account manager in the top right corner of the Fabric portal.

Fabric billing and consumption updates

Three announcements stood out when it came to billing and consumption.

Organizations with multiple capacities can now direct Copilot in Fabric consumption and billing to a specific capacity, no matter where the Copilot in Fabric usage actually takes place.

Admins can assign specific members of their organization to the specified F64 or higher capacity for all their Copilot requests. These requests will be consumed and billed on that assigned F64+ capacity, ensuring Copilot in Fabric usage doesn't impact priority jobs while expanding Copilot access to any workspace regardless of its capacity.

Capacity admins now have more control over the Fabric jobs running in their capacities. Surge Protection, now in preview, helps protect capacities from unexpected surges in background workload consumption.

Admins can use surge protection to set a limit on background activity consumption which will prevent background jobs from starting when reached. Admins can configure different limits for each capacity in your organization.

So, what is next?

The reality with Fabric is with the rate of change, continuing maturity, and functionality will remain a challenge to keep on top of. The size and scale of Fabric, the breadth of personas and use cases it supports means that it will continue to require specialist support to maximize your investment in the product.

At BJSS we believe that Fabric will open up more data opportunities for organizations and help drive data-modernization and data-driven insights. If you want to know more about what Fabric could do for your organization, or are interested in the broader Cloud, Data & AI landscape then please reach out.

Get in touch

Daniel Amini

Head of Cloud, Data & AI Services

daniel.amini@bjss.com

