Microsoft Azure

Maximize Data Value

Transform your data estate for AI innovation

Contents

From proof of concept to AI transformation 3

Tackle data complexities and challenges 4

From unified data to AI innovation 6

Start unifying your data for AI 11

From proof of concept to AI transformation

A unified data and AI foundation

99.1% **of companies surveyed say that investment in data and AI is a top organizational priority[1]**

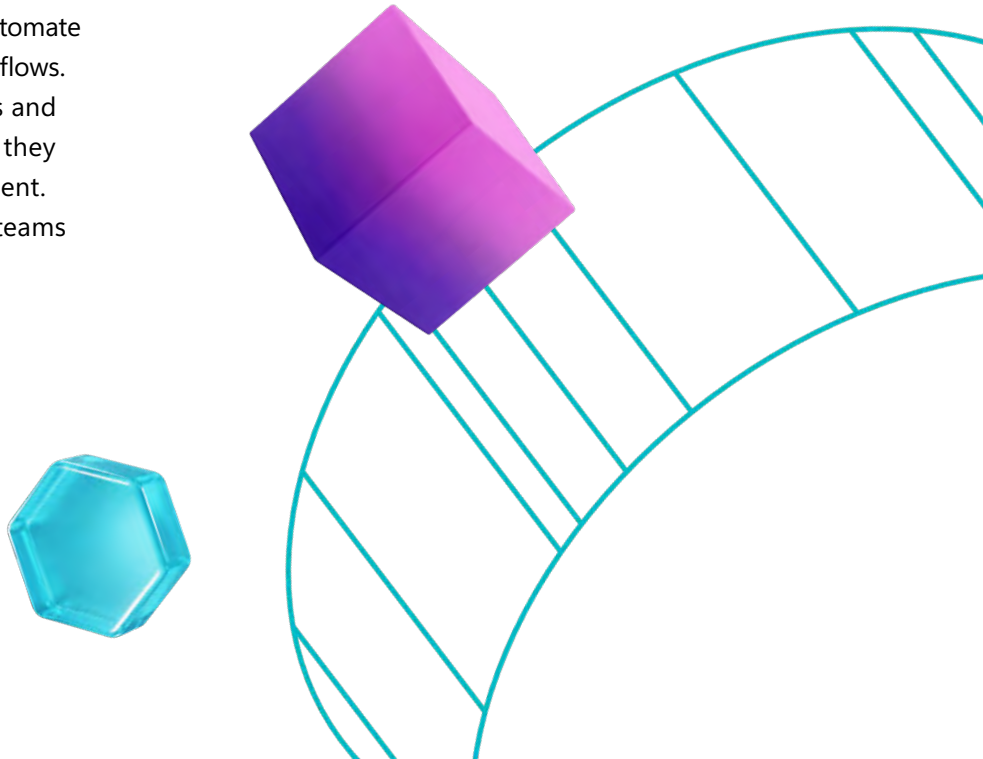
Reaching the AI frontier requires a unified data estate—with built-in governance, security, and compliance—to power analytics and insights across the organization. An integrated AI platform—paired with data—can help unlock innovation and provide your teams with real-time insights. The foundation is a unified data lake that brings your organization's data together, allowing teams to reduce duplication and sprawl, enforce consistent governance, and offer open standards.

With unified, governed data, you can power a system of intelligence based on shared business vocabulary. You can customize agents that help teams explore, build, and act using natural language, or build agents that work alongside people to monitor conditions, automate tasks, connect tools, and drive event-based workflows. Real-time insights and automation keep copilots and agents grounded in the latest trusted data, so they can recommend or initiate actions in the moment. AI-powered tools built into the platform help teams get more done in less time.

About this e-book

This e-book is for AI data leaders and technical decision makers who want to modernize their data estate and prepare it for scalable analytics, AI apps, and agents. It helps you maximize the value of your data by:

- **Unifying your data before scaling AI.**
- **Using a unified data lake to simplify complexity.**
- **Building and scaling an AI-ready platform.**



Tackle data complexities and challenges

It's tempting to push AI innovation forward as fast as possible. You will get better outcomes when you start with a data foundation that is deeply integrated, well-governed, and secure. That's because many data estates were optimized for reporting and transactions, not continuous, cross-domain AI use. As a result, you may be managing growing complexity as you build and scale models and analytics across disconnected systems. However, you can start addressing core data challenges with two steps: unifying data and reducing duplication.

92.7%

of firms report that interest in AI has led to a greater focus on data[1]

Streamline your data estate

To deliver reliable AI outcomes, organizations need clean, accurate, and consolidated data. When data is spread across systems, teams spend too much time extracting, moving, and reshaping it to make it usable. The result is slower time to insight, higher costs, inconsistent definitions, and outdated data by the time it reaches the people and processes that need it. It also becomes harder for business users to confidently manage and scale AI initiatives across the enterprise.

But a unified data estate isn't the end goal—it's the foundation for building, grounding, and scaling AI applications and agents. Unlike fragmented environments that duplicate data across tools, a connected estate establishes shared meaning, consistent governance, and reusable data across analytics and AI workloads.

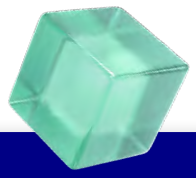
By reducing unnecessary data movement and fragmentation, organizations can improve consistency, governance, security, model performance, and operational efficiency—while lowering costs, reducing latency, and enabling teams to work from trusted, current data.

Reduce duplication

When the same data is stored in multiple places, duplicates accumulate as teams integrate information across systems. This creates confusion about what's current and correct, and it slows your ability to deliver reliable analytical insights and AI experiences because more time goes into reconciling records than generating insights.

A unified foundation reduces data duplication and its challenges, including:

- **Accuracy and reliability**
  Get a single, trusted view of data, analytics, and AI to produce more reliable results.
- **Storage efficiency and retrievals**
  Reducing unnecessary copies lowers storage costs and keeps queries fast by minimizing the amount of data that needs to be scanned and managed.
- **Operational overhead**
  A unified foundation simplifies ongoing data management by reducing exception handling, remediation cycles, and the coordination needed to reconcile datasets.



"Microsoft Fabric was a game changer because of its ability to create shortcuts without physically moving data from one place to another. Before, if I had to incorporate three sources, I had to create pipelines to bring in the data. That pipeline had a cost. The data movement had a cost. With Fabric, it's two clicks and that's it."

José Viegas
Senior Data Architect, IWG

Read the full story →

From unified data to AI innovation

True innovation requires a platform that can turn that foundation into action. Enter Microsoft OneLake and Foundry.

OneLake: the foundation for lakehouse architecture

A lakehouse is a modern data architecture that lets you store structured and unstructured data together and analyze it using both Spark and SQL—without moving data between systems. Implemented on OneLake, it reduces the friction of splitting data across separate platforms, making it easier to apply consistent governance and use the same underlying data for analytics and AI.

Without data duplication or movement, you can unify your multicloud and on-premises data sources. Through a broad ecosystem of connectors, this approach enables seamless access to data across platforms—including cloud data warehouses, enterprise databases, operational systems, and SaaS applications—such as Azure SQL Database, SQL Server, SAP on Azure, and Azure Databricks. This also extends to data managed across other platforms and open ecosystems like IBM Netezza, MongoDB, Oracle, Servicenow, and more, enabling bi-directional access and interoperability without requiring data migration. The result is the ability to work with data where it lives, breaking down silos while maintaining consistency, governance, and performance across the entire data estate.

Fabric mirroring continuously replicates operational data into OneLake as open Delta tables for analytics. That governed data becomes accessible across workloads—such as Data Factory, Data Engineering, Data Science, Data Warehouse, Real-Time Intelligence, and Power BI. Your data is further enriched by Fabric IQ, which empowers teams and AI agents to understand and act on live, context-rich insights, driving smarter decisions, and better business outcomes.

With OneLake as the shared foundation, you gain several advantages that help teams scale analytics and AI, including:

- **Scalable storage and processing**

  Lakehouses scale to support larger data volumes and more users over time. In Fabric, lakehouse data is stored in OneLake and can be processed with multiple engines, including Spark and SQL, as your needs grow.

- **Single source of truth**

  A lakehouse helps you standardize consistent, governed data across teams and tools. OneLake organizes lakehouse and warehouse data in a single, organization-wide data foundation.

- **Optimized for analytics and AI**

  The lakehouse architecture standardizes both analytical and operational data in open Delta tables that multiple engines can use from the same source, improving performance, consistency, and timeliness for analytics and AI workloads.

- **Zero-copy access with shortcuts**

  Shortcuts and mirroring let you connect to data where it already lives and make it available in your lakehouse without creating another copy. You can reduce duplication and staging delays while keeping access governed through the platform's existing permissions and credentials.

- **Third-party data integration**

  With more than 150 connectors, you can easily bring in data from external and third-party sources—including SaaS applications, partner systems, and multicloud platforms—without complex integration work.

- **Low query latency**

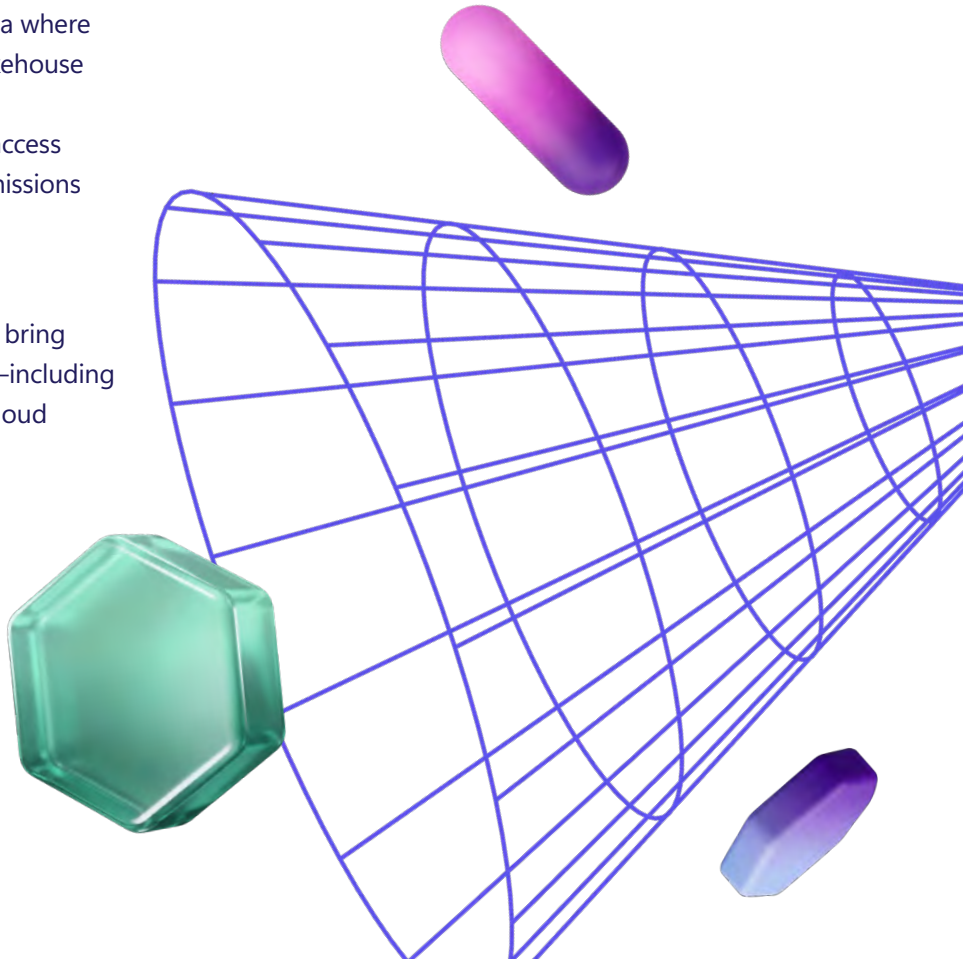
  Teams can retrieve insights quickly for interactive reporting and AI workloads. OneLake reduces latency from duplicate copies and handoffs, enabling fast access paths such as Direct Lake over Delta tables in the lake.

- **Data freshness**

  Lakehouses keep data fresh through streaming and frequent updates from operational systems. OneLake makes that data available across teams and workloads, through mirroring for near real-time replication and shortcuts to access data in place.

- **Data security and governance**

  OneLake provides built-in, centralized security alongside unified governance, enforcing permissions at both the management and data levels. The OneLake catalog centralizes security management. You can apply role-based access down to folders or tables, including row-level and column-level restrictions, ensuring consistent protection across your entire data estate.



From foundation to AI-ready capabilities

OneLake provides the shared data foundation, but preparing data for AI also requires a set of essential data capabilities to reduce movement, limit duplication, and prepare data for AI. Built on top of OneLake, these capabilities work together on the same governed data, rather than creating new silos.

Data integration

When you streamline data integration, you can reduce storage costs, simplify engineering, and improve trust in shared data. You can work with near-real-time replicas from sources such as Azure SQL, Azure Database for PostgreSQL, and Azure Cosmos DB for easy access to all of your data in OneLake.

In addition, Fabric can work with Azure Databricks Unity Catalog tables, mirroring catalog metadata and accessing the underlying data through OneLake shortcuts to avoid data movement and duplication.

Data transformation

You can transform data with low-code dataflows and automate the full process with pipelines. Dataflows in Fabric let you apply step-by-step cleaning and shaping, and pipelines schedule and coordinate end-to-end workflows. Fabric lakehouses also support Spark and SQL over the same data, so you can publish curated tables once and reuse them across workloads.

Data streaming

To act on what's happening now, you need to turn incoming signals into timely insights and automated responses. This is essential for agentic workflows, event-driven decisions, and Internet of Things scenarios, where systems detect patterns like thresholds, anomalies, or state changes and respond immediately.

Fabric Real-Time Intelligence ingests, transforms, and routes real-time events without code. Fabric Activator continuously monitors streams with subsecond latency to trigger downstream actions such as notifications or Power Automate flows when conditions are met.

Data querying

Querying helps you pull the right slice of data—features, records, and time windows—so analytics and AI models are trained and validated on relevant, high-quality inputs. You can query warehouses and SQL endpoints with T-SQL, lakehouses with Spark and SQL, or real-time and log data with KQL. You can then operationalize those results through semantic models and Direct Lake for fast, governed analysis.

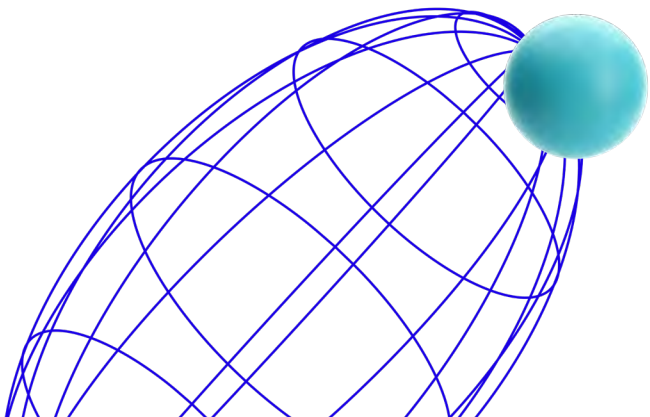
Data visualization and collaboration

When governed data becomes accessible to all, your teams can track outcomes, spot changes, and make decisions faster. Fabric includes Power BI as a first-class workload for interactive reporting on data in the platform. Direct Lake supports interactive analysis over large Delta tables in the lake, reducing refresh overhead compared to traditional import patterns.

Additionally, Microsoft Copilot in Fabric is available across workloads to help you interact with your data using natural language, transform and analyze it, generate insights, and create visualizations and reports across Fabric and Power BI.

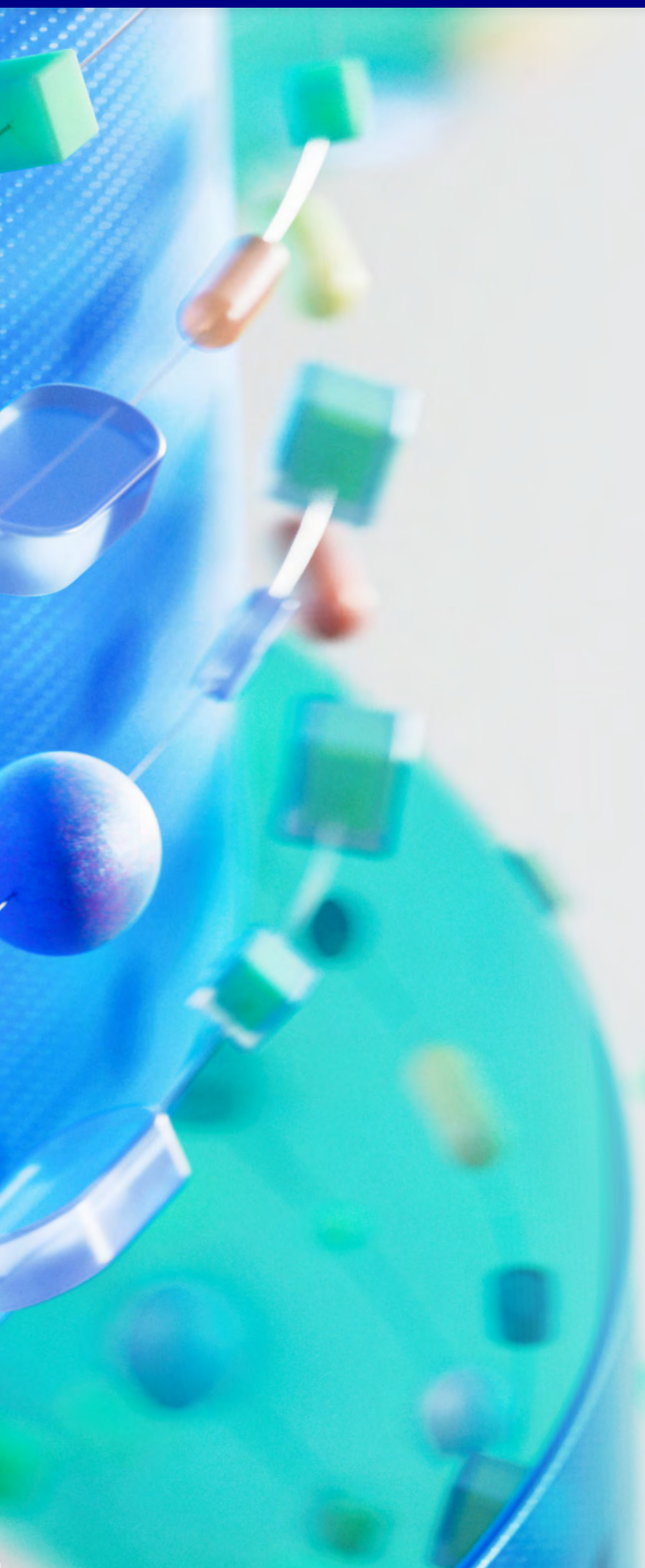
Data and operations agents

With unified, governed data, you can enable AI-powered agents that reason over your data and take action across systems. Data agents help you interact with data using natural language to discover insights and accelerate analysis, while operations agents monitor signals, trigger workflows, and automate processes in real time. These agents connect data, context, and action to drive more intelligent, adaptive operations.



Data intelligence with Fabric IQ

Fabric IQ adds a semantic intelligence layer to your data foundation—helping unify not just where data lives, but what it means by organizing data around business entities, relationships, and rules. This creates a consistent, governed understanding of concepts like customers, products, and operations across your environment, enabling analytics, AI agents, and applications to interpret data with the same meaning—reducing inconsistencies, improving trust, and accelerating insight with a unified semantic model that supports more intelligent, decision-ready outcomes at scale.



Azure Databases: Deliver AI apps and agents faster, with built-in capabilities for generative AI and LLM workloads

To make operational data usable for AI, it also needs seamless bridges—sometimes directly into AI, other times through analytics layers where data is transformed into insights that make AI more impactful. Historically, this required significant custom plumbing and armies of data engineers. With Azure Data, those integrations are built in. Every database connects natively into OneLake as the unified data foundation for analytics and AI, all the way through to the AI stack and AI Foundry.

Azure offers fully managed services for SQL, NoSQL, PostgreSQL, and MySQL, providing the foundation for any size business-critical workload your organization uses to achieve AI transformation and business innovation. Mirror your databases for easy access to all of your data in OneLake.

With Azure Databases, companies get up to 5x faster performance, intelligent threat detection, and up to 58% lower total cost of ownership. New apps, AI agents, and data-driven solutions benefit from GitHub Copilot-assisted development, zero-touch recovery, single-digit millisecond latency, advanced vector search capabilities, and integrations with the most used developer tools, analytics platforms, and AI technologies.

With Azure Databases, your data becomes a strategic asset, helping turn your AI ambitions into real business results close to seamlessly with OneLake.

Foundry: The enterprise AI platform for innovation

Foundry is an end-to-end AI platform for enterprises looking to move beyond experimentation and scale AI across their business. It brings together a large model catalog, a unified control plane, deep integrations across the Microsoft technology stack, and robust observability features—giving teams everything they need to build, deploy, and manage AI solutions with confidence. Rather than assembling a patchwork of tools, organizations can use a single platform designed to carry AI use cases from development through to production.

Foundry IQ: Putting enterprise knowledge to work

Foundry IQ provides a context engineering platform for Microsoft's intelligence layer, unlocking knowledge for your agents wherever they live. Foundry IQ automatically handles how agents connect to, process and action enterprise information. With built-in enterprise security and agentic retrieval, agents deliver better results from the right context.

Foundry and Foundry IQ form the foundation on which teams can move from fragmented data to intelligent, scalable AI experiences—building and deploying the agents and applications that drive real business outcomes.

How Foundry IQ helps unlock data for enhanced agentic capabilities

- **No data left behind**
  Deep integrations across SharePoint, Fabric IQ, Logic App workflows, and 1,000+ data connectors mean every data source—structured or unstructured—is within reach.
- **Build once, use everywhere**
  Reusable knowledge bases eliminate the need to re-connect and re-process the same data for each new application, so your data works more efficiently across more use cases.
- **Data made AI-ready automatically**
  Pipelines handle extraction, parsing, chunking, enrichment, and embedding, so raw enterprise data becomes usable knowledge without manual preparation.

Get more from your data at retrieval time

- **Retrieve what actually matters**
  Automated, best-in-market RAG with declarative retrieval ensures agents surface the most relevant data—not just the most recent or most accessible.
- **Quality you can rely on**
  The agentic RAG engine is backed by rigorous benchmarks, so the data reaching your agents is consistently accurate and complete.
- **Data flows across your full stack**
  Integrations with Foundry Agent Service, Azure OpenAI, M365 Copilot, and frameworks including LlamaIndex, LangChain, and Semantic Kernel mean your data is available wherever your teams build.

Maximize data use without compromising control

- **Governance at scale**
  Microsoft Entra ID and OneLake Catalog help ensure that as more data is activated for AI, compliance and access control keep pace automatically.
- **Permissions travel with the data**
  Document-level ACLs are inherited directly from source systems, so expanding data access never means expanding risk.
- **Full visibility into data usage**
  Audit trails via Azure Monitor and Log Analytics give teams a clear picture of how data is being accessed and used across every agent and application.

Start unifying your data for AI

For AI to deliver real business value, start with a unified data foundation that every team can trust and reuse. The fastest path to impact is reducing friction—fewer silos, fewer copies, and less time spent reconciling which version is right.

An end-to-end platform gives every team what it needs to move from raw data to AI-powered intelligence. Across roles and departments, your users can more efficiently explore data, generate insights, and build AI apps and agents, powering your organization's AI transformation with a complete data platform.

[1] Bean, Randy and Davenport, Thomas H. Survey: How Executives Are Thinking About AI in 2026. Harvard Business Review. January 12, 2026.

© 2026 Microsoft Corporation. All rights reserved. This document is provided "as-is." Information and views expressed in this document, including URL and other Internet website references, may change without notice. You bear the risk of using it. This document does not provide you with any legal rights to any intellectual property in any Microsoft product. You may copy and use this document for your internal, reference purposes.

Take the next steps

Get started with a free trial of Microsoft Fabric →

Get end-to-end expert guidance and investments with Azure Accelerate →

Establish your AI platform for scale with Microsoft Foundry →