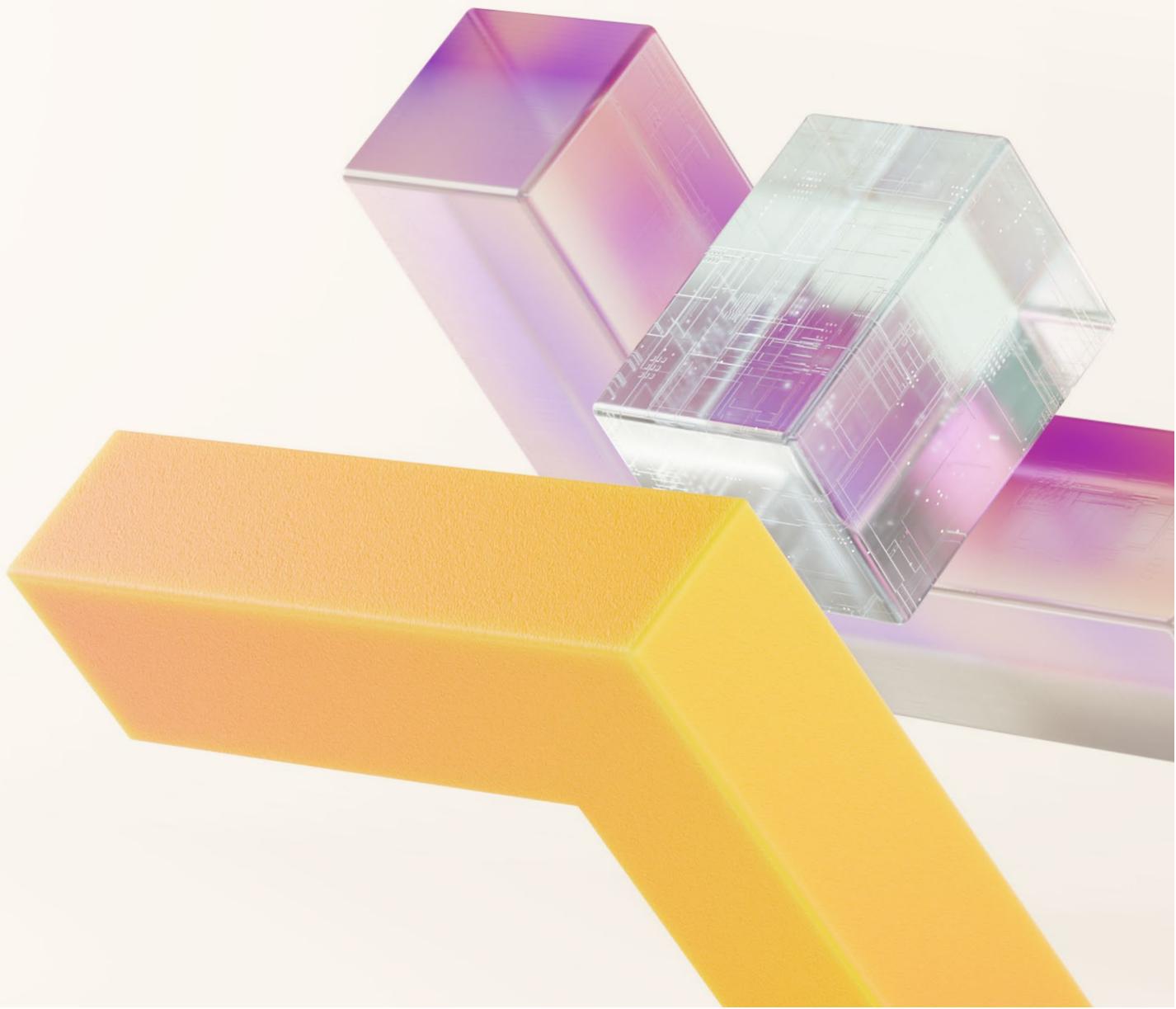




# The **Azure** Databricks Advantage: Benchmark Report Highlights

Discover how Azure outpaces other clouds for Databricks workloads



# Contents

Why data and AI are best on an open, unified, and governed intelligent platform	3
Success by the numbers	4
Key findings	5
The power of partnership	6
<b>Unlocking the power of Azure Databricks with an integrated ecosystem</b>	7
<b>Seamless, real-time access to Databricks data</b>	8
<b>Democratized data intelligence</b>	9
Customer use cases and success stories	10
Next steps	12

Tests conducted by Principled Technologies showed Azure Databricks outperforming Databricks on AWS in both single-query and concurrent-query scenarios. Explore the findings in detail and discover what makes Azure Databricks the ideal platform for faster data-driven insights.

# Why data and AI are best on an **open, unified, and governed intelligent platform**

With so much data at their fingertips, organizations today have a chance to get ahead by delivering real-time insights, personalized experiences, and intelligent automation—all powered by AI. Databricks helps them bring those opportunities to life by unifying data engineering, data science, analytics, and AI on one platform. But to realize even greater value from their Databricks, those organizations also need an intelligent data and AI platform to support it.

Choosing the right cloud for Databricks is key to maximizing its value. The right platform streamlines data ingestion, accelerates processing, and puts machine learning frameworks within easy reach. It also dictates how smoothly you can scale workloads, manage costs, and maintain governance. In short, the cloud isn't just where Databricks runs—it's what helps turn your data into decisions, actions, and innovation.

This is where Azure Databricks truly shines. As a first-party Microsoft service, it offers deep integration with Azure's security, networking, and data services—enabling seamless access to tools while ensuring enterprise-grade compliance, scalability, and performance.

But what's the real benefit of running Databricks natively on Azure? Principled Technologies decided to see if there were any advantages, conducting cross-cloud tests and publishing a benchmark report that confirmed Azure's edge in speed and performance.

## What do enterprises need to succeed in a data-rich future?

- **A scalable, high-performance data and AI architecture** that can handle today's exploding data volumes while staying flexible for tomorrow's demands.
- **Seamless collaboration across data roles** so engineers, scientists, and business users can work together without friction.
- **Robust governance and compliance frameworks** that protect sensitive information and maintain trust across the organization.
- **Strong ROI from cloud and AI** by ensuring every workload is optimized for cost and performance.

### Key benchmark findings confirm the Azure advantage for Databricks

- **Up to 21.1% less time** to complete a stream of queries
- **Saved over 9 minutes** while running four concurrent streams



# Success by the numbers

## Azure Databricks demonstrates superior performance

Not all Databricks cloud solutions perform the same. To see how Azure measured up, Principled Technologies ran a series of tests to see which could deliver faster insights for data teams working with large datasets.

### To explore this, the testers compared:

#### Azure Databricks

First-party, co-engineered solution on Microsoft Azure

versus

#### Databricks on AWS

Third-party deployment on Amazon Web Services

### The tests focused on two real-world scenarios:

The performance of a single user running queries and the performance of multiple users simultaneously. These scenarios reflect the day-to-day realities of data teams relying on platforms that are decision support systems (DSSs).

To ensure a fair comparison, the testers created clusters on each platform with similar computing resources, thereby verifying that both could handle a 10TB dataset for a DSS workload. Each platform ran the same version of Databricks with performance features enabled, so the results reflected differences in cloud infrastructure rather than software.

### Testing Azure Databricks versus Databricks on AWS

- **Single-user tests:** These measured how quickly one person could complete a stream of queries from start to finish. Think of an analyst running a detailed report or dashboard—this test shows how much time they would save with faster query execution.
- **Multi-user tests:** These tests measured performance when four users ran queries simultaneously. In many organizations, multiple analysts or teams work with the same data simultaneously, so understanding how the platform handles concurrent workloads is crucial.

Running these tests revealed how Azure and AWS perform under both focused, individual use and more complex, collaborative scenarios. This setup enabled testers to compare the platforms fairly, providing a realistic picture of how each would perform in day-to-day enterprise operations.



# Key findings

The results of these tests clearly demonstrated that Azure Databricks outperformed Databricks on AWS across both single-query and concurrent-query scenarios.

**Here's a closer look at the numbers and their implications:**

## Faster query stream processing

- **Single query stream (power test):** Azure Databricks completed a sequential stream of 99 queries, 13.4%–21.1% faster than Databricks on AWS.
- **Time saved:** In absolute terms, this meant nearly 9 minutes faster for a single session without Autoscale.

**What this means:** For analysts running detailed reports one at a time, Azure Databricks can shave minutes off each session. Across multiple reports or repeated analyses, these savings add up, enabling faster decision-making and more agile responses to business needs.

## Improved performance for concurrent queries

- **Four concurrent query streams (throughput test):** Azure Databricks handled multiple simultaneous workloads 7.3%–9.4% faster than Databricks on AWS.
- **Time saved:** When Autoscale was disabled, the Azure cluster finished 9 minutes and 14 seconds sooner than the AWS equivalent.

**What this means:** In collaborative, multi-user environments—common in medium and large enterprises—Azure Databricks allows teams to run parallel analyses more efficiently. Users experience fewer delays when querying shared datasets, which supports better collaboration and faster cross-departmental insights.

## Unified lakehouse benefits

Both platforms leverage Databricks' unified lakehouse architecture, which streamlines the entire data lifecycle from data engineering and ETL processes to machine learning, AI, and business intelligence. Structured, semi-structured, and unstructured data can all flow seamlessly through a single repository, supporting complex analytics without requiring data to be moved between systems.

While AWS and Azure share this architecture, the study highlighted that the underlying cloud infrastructure, integration, and support models materially affect performance. Azure's co-engineered, first-party service provides **tighter integration** and **more consistent performance**, which directly translates into faster insights.

## Why these findings matter

When it comes to AI and analytics, speed directly drives impact. Faster query processing and more efficient handling of concurrent workloads mean insights are delivered sooner, enabling teams to explore data, test hypotheses, and refine models without delay. The time saved per query allows analysts to respond quickly to emerging trends, adjust strategies based on fresh data, and scale analyses without performance bottlenecks.

Optimized performance also ensures resources are fully utilized—more queries run at the same time, and processes scale reliably. Ultimately, faster, more dependable analytics empower organizations to act on AI-generated insights when they matter most, supporting real-time decision-making and enhancing outcomes across operations, marketing, and customer experiences.

# The power of partnership

## How Microsoft and Databricks enable AI-enhanced analytics

Databricks is a powerful, intelligent data and AI platform built on the lakehouse architecture, capable of handling massive volumes of structured, semi-structured, and unstructured data. It brings the entire data and AI lifecycle together—from data engineering and ETL (extract, transform, load) to data science, machine learning, AI, and business intelligence—into a single, centralized platform.

While Databricks can run on multiple major cloud platforms, the way it's delivered makes a real difference. As a first-party service, Azure Databricks offers tighter integration, more optimized performance, and smoother management compared with running Databricks as a third-party service on other clouds. And when you're using AI to drive analytics, every bit of speed and efficiency counts.

**What does this mean for teams and their organizations?** The partnership between Microsoft and Databricks gives teams a unified, secure platform to harness data and AI more efficiently. "Together, we are empowering our customers with Azure Databricks to help them harness the full potential of their data, build the next generation of autonomous systems, and accelerate AI transformation," says Judson Althoff, Executive Vice President and Chief Commercial Officer at Microsoft.

"This extended partnership with Microsoft shows our long-term joint commitment to helping organizations innovate faster with a unified, open, and governed platform for data and AI on Azure Databricks. Together, we're laying the foundation for the next generation of AI-powered enterprises."

**Ali Ghodsi,**  
Co-founder and CEO of Databricks



## Unlocking the power of Azure Databricks with an integrated ecosystem

Co-engineered by Microsoft and Databricks to run natively on Azure infrastructure, Azure Databricks integrates tightly with core Microsoft and Azure services. These integrations help organizations secure their data, accelerate innovation, and bring analytics into everyday business decision-making—creating built-in resonance with the Microsoft solution ecosystem that makes it easier to unlock value across the entire data and AI landscape.

### Data access and storage: [Azure Data Lake Storage \(ADLS\)](#)

- Provides seamless access to massive datasets for advanced analytics
- Delivers optimized I/O performance for scalable data and AI workflows
- Ensures secure data management with fine-grained access controls

### Business Intelligence and visualization: [Power BI](#)

- Connects directly to Azure Databricks for real-time dashboards and reports
- Makes insights easy to share across business teams
- Supports large-scale data models powered by Azure Databricks

### Identity and security: [Microsoft Entra ID](#)

- Simplifies authentication, supports single sign-on (SSO), and enforces role-based access control (RBAC)
- Integrates with Conditional Access, enables multi-factor authentication (MFA), and logs identity events for auditing
- Automatic Entra ID management in Azure Databricks enables secure and seamless workflows
- Connects with Microsoft 365 and other apps, enabling federated identity across hybrid environments

### Low-code innovation: [Power Platform \(Power Apps, Power Automate, Power Virtual Agents, Copilot Studio\)](#)

- Allows business users to build low-code apps that tap into Azure Databricks data
- Automates workflows leveraging big data access
- Empowers citizen developers to create intelligent data-driven solutions

### Monitoring and optimization: [Azure Monitor](#)

- Tracks metrics, logs, and system health across Databricks clusters
- Correlates logs and integrates with Log Analytics for deep insights
- Triggers automated actions and helps optimize performance and cost

### Governance and compliance: [Microsoft Purview](#)

- Enables unified data governance across Azure Databricks notebooks and pipelines
- Automates data discovery and lineage tracking for transparency
- Offers compliance-ready metadata management for sensitive datasets

These integrations make it easier to keep data flowing smoothly and unlock richer analytics and AI use cases. They also bring governance and security under one roof, making compliance much more manageable. And by simplifying management and cutting costs, they help free teams to spend more time collaborating and less time wrestling with complexity.

In addition, two key integrations push the boundaries even further—laying the foundation for unified data and AI innovation: Mirrored Azure Databricks Unity Catalog in Microsoft Fabric and Azure AI Foundry.

## Seamless, real-time access to Databricks

[Mirrored Azure Databricks Unity Catalog in Microsoft Fabric](#) makes it effortless to bring your Databricks-managed data into Fabric, all without duplicating it or building complex ETL pipelines. With just a few clicks, you can mirror entire catalogs, schemas, or individual tables into Microsoft OneLake, where they're instantly available for SQL queries, analytics, and Power BI Direct Lake reports. Updates in Databricks automatically sync to Fabric, so your teams are always working with the latest data, giving your organization a single, trusted source of truth that spans your entire data estate—secure, governed, and ready to use.

This integration makes data easier to access and transforms how teams work. Mirrored Unity Catalog tables behave like native Fabric tables, meaning analysts, data engineers, and business users can immediately leverage them in analytics, AI models, and applications. It reduces storage costs, eliminates unnecessary duplication, and keeps your data governance strong through OneLake's security framework. The result is faster insights, smoother collaboration, and more time spent turning data into action, rather than managing pipelines.

## Azure AI Foundry integration: Powering rapid AI innovation

[Azure AI Foundry](#) brings the next layer of intelligence to Azure Databricks by streamlining the deployment and management of generative AI use cases. Offering a native connector, Foundry can directly tap into Databricks AI/BI Genie, enabling AI agents to reason over real-time enterprise data.

### This creates powerful new possibilities:

- **Domain-specific context:** AI agents access real-time analytics, not just static snapshots.
- **Governance and security:** Unity Catalog ensures every interaction is compliant and auditable.
- **Enterprise-grade AI agents:** Developers can create agents that deliver accurate, business-aware insights and deploy them across apps like Microsoft Teams.
- **Natural language intelligence:** Users can query Databricks data using plain language, democratizing access to analytics.
- **Responsible AI:** Foundry enforces fairness checks, transparency, and auditability by design—critical for regulated industries.
- **Multi-agent workflows:** Foundry supports orchestrating multiple agents to handle complex business processes.
- **Cross-team collaboration:** Business users, data scientists, and developers can all interact with the same governed data in the interface that fits their needs.

Together, Azure Databricks and Azure AI Foundry enable organizations to build intelligent, secure, and collaborative AI ecosystems where innovation happens faster, insights are more accessible, and AI operates responsibly at enterprise scale.



## Democratized data intelligence

"We're entering a world where natural language becomes the interface to data. That's a fundamental shift—and it starts with having the right platform underneath."

**Ali Ghodsi,**  
Co-founder and CEO of Databricks

At the heart of Azure Databricks is the lakehouse—a single, secure, and governed copy of your data managed through Unity Catalog. This foundation enables teams to run data warehousing, ETL, and business intelligence tools all on the same dataset, eliminating the need for duplicate resources. And with Azure's fully serverless architecture, organizations don't have to worry about managing infrastructure—everything scales automatically.

But that's just one of the major advantages. When you layer AI and generative AI on top of this lakehouse, you can unlock true data intelligence for everyone in your organization. Leaders and decision-makers no longer need to rely solely on data teams or dashboards. Instead, they can use natural language queries to ask questions of their data and get actionable insights without heavily depending on SQL, Python, or other technical skills.

Democratizing AI isn't only about helping business users interact with their data. It also empowers data and data science teams to build their own AI models tailored to the organization's data. These models go beyond general AI, providing data intelligence that understands and answers questions specific to your enterprise. The result is a unified platform where analytics, AI, and business intelligence converge, enabling teams to transform data into actionable insights across the organization—quickly.

### Unified insights start here

Azure Databricks makes it easy to tap into your existing data sources—whether it's Azure SQL Database, Synapse Analytics, or Cosmos DB. With built-in connectors and support for JDBC, you can ingest, transform, and analyze data from managed databases directly within your lakehouse environment.



# Customer use cases and success stories

Across industries, organizations are using Azure Databricks to move faster and get more out of their data. From slashing processing times from hours to minutes, to tripling the speed of data science cycles, to streamlining secure access for millions of users, customers are seeing real performance gains that translate into stronger business outcomes.

## Speeding up data processing with Azure Databricks

**Parker Aerospace** transformed its data analytics by adopting Azure Databricks, cutting nightly data processing time **from 14 hours to under 2**. This dramatic improvement enables faster insights into aircraft part performance, helping reduce maintenance delays and improve flight safety.

"It's been amazing. We're new to cloud computing environments at this scale, and Microsoft staff were always there to help us out every step of the way."

**Jae Lim,**  
Senior Principal Data Scientist,  
Parker Aerospace

## AT&T drives ROI and agility with Azure Databricks

**AT&T** modernized its massive on-premises data estate by migrating to Azure Databricks, retiring costly datacenters, and eliminating more than 80 schemas. The move delivered a **300% ROI** in five years, **tripled** the speed of data science cycles, and unified nearly 90,000 employees on a single source of truth. With Azure Databricks powering its lakehouse, AT&T is reducing costs, breaking down silos, and driving faster innovation for millions of customers.

"Moving to Azure Databricks has transformed the data culture at AT&T. Instead of people analyzing data on their own laptops, they're all coming to the cloud to collaborate in one place."

**Praveen Vemulapalli,**  
Director of Data & Gen AI Architecture,  
AT&T



## CPFB modernizes member services with Azure Databricks

Singapore's [Central Provident Fund Board \(CPFB\)](#) modernized its data management with a Unified Data Platform built on Azure Databricks, improving secure access to member data and enabling more sophisticated analytics. This streamlined reporting, automated manual tasks, and allowed teams to collaborate in real time—helping CPFB better serve over 4 million members through more efficient policy formulation and service delivery.

"Users now have access to broader datasets securely, and we can automate reporting and model development on a single platform. It's transformed how we work with data."

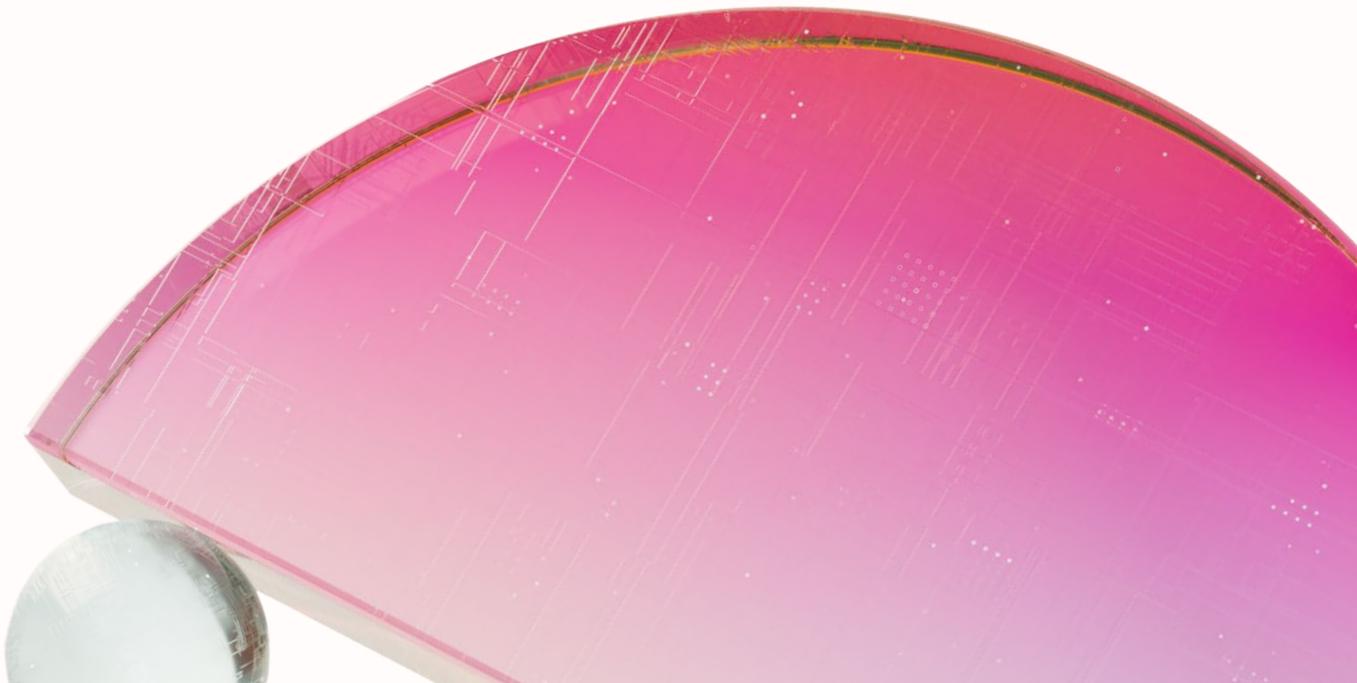
**Benedict Ho,**  
Senior Deputy Director, Data Science Accelerator  
CPFB

## JLL delivers faster, smarter real estate insights with Azure Databricks

[Jones Lang LaSalle \(JLL\)](#) transformed its commercial real estate analytics by adopting Azure Databricks, giving clients instant access to AI-driven insights from previously siloed data. With JLL Azara, their analytics platform, clients can ask targeted questions and uncover actionable intelligence in **minutes instead of months**—improving decision-making, cutting costs, and advancing sustainability goals.

"Azure Databricks has completely changed how our clients interact with data. They can explore insights freely, make faster decisions, and unlock value they didn't think was possible before."

**Sharad Rastogi,**  
CEO of Work Dynamics Technology  
JLL



# Next steps

The data doesn't lie: Azure is the superior cloud for running Databricks workloads. Azure Databricks removes the complexity of infrastructure management, letting your teams focus on what matters—turning data into insights. Meanwhile, integrated billing, support, and operations let you manage everything under one roof, without juggling multiple platforms.

Offering a secure, open lakehouse, Azure Databricks helps you unify analytics, data engineering, and AI on a single platform, breaking down silos and giving your teams more freedom to innovate. Plus, with built-in security backed by \$1 billion annual cybersecurity investments and Microsoft's 3,500 security experts, your teams can work with confidence knowing their data is protected.

Want unmatched performance, seamless integration, and robust governance for your Databricks workloads? Look no further than Azure Databricks.

Read the full comparative report from Principled Technologies.

[Read now](#) →

Gain Azure Databricks skills in our structured learning plan.

[Discover Azure Databricks](#) →

Check out Azure Databricks features, resources, and more.

[Start your learning journey](#) →

Learn more about how Azure Databricks works from the bottom up.

[Explore documentation](#) →

© 2025 Microsoft Corporation. All rights reserved. This document is provided "as is." Information and views expressed in this document, including URLs 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