Bringing AI to the data lakehouse

At Databricks, we're building a **data intelligence platform** on top of our lakehouse. We're excited about the potential of AI in data platforms and continue to enhance our features. Our open data lakehouse is unique, offering (1) unified governance across data and AI, (2) an open, unified storage layer and (3) a unified query engine for ETL, SQL, ML, AI and BI. Mosaic AI powers our Data Intelligence Engine, driving intelligence throughout our platform.

Data intelligence integrates across Databricks, enabling:

- **Data democratization with AI/BI:** Easily creates spaces for business teams to self-serve insights not already answered by their dashboards. AI/BI Genie lets people just converse with their own data, without having to go through others to build dashboards. Genie leverages Unity Catalog so the insights continuously learn your specific business context and semantics while including controls for guidance and security oversight.

- **Enhanced governance:** Data intelligence improves Unity Catalog by auto-generating descriptions and tags for all data assets such as tables and columns, enabling better semantic search, AI assistant quality and governance across the platform.

- **Platform optimization:** Automatically adjusts settings like column indexing and partition layout, strengthening the lakehouse foundation for better performance and lower TCO.

- **AI assistant:** Enhances Python and SQL code generation for text-to-SQL and text-to-Python capabilities.

- **Query performance:** Boosts query speed by using data predictions for optimal query planning that provides extremely fast query performance at a low cost.

- **Efficient scaling:** Optimizes ETL and orchestration by predicting workload needs for optimal autoscaling and cost reduction.

databricks

Building enterprise AI on the foundation of data management

Data intelligence makes building AI solutions easier. Mosaic AI integrates seamlessly to help enterprises create and deploy production-grade ML and AI applications.

- **Production-quality AI:** Delivers accurate outputs tailored to enterprise data, with reinforcement learning by business users. Models can be easily swapped for better accuracy as needed.

- **Unified governance:** Ensures oversight across data and AI assets, managing risk, privacy and accountability from data to models and applications.

- **Cost efficiency at scale:** Mosaic AI optimizes AI deployments, making enterprise-level AI affordable.

Companies are now building adaptable, high-performing and trustworthy AI by leveraging AI systems — integrating multiple models to boost adaptability, customization and transparency.

Databricks, powered by Mosaic AI, is the unified platform for developing and managing AI systems:

1. **Data prep:** Use tools like LakeFlow to ingest and prepare data. With the AI built on the lakehouse, there is no need to duplicate data — instead, you can automatically generate vector indexes and ML features from your production data.

2. **Build agents:** Choose from existing models, train new ones or serve models using Mosaic AI's tools for model training and serving.

3. **Deploy agents:** Deploy models securely at scale using MLflow and Mosaic AI Agent Framework.

4. **Evaluate agents:** Use human and machine evaluations to ensure quality, leveraging Mosaic AI Agent Evaluation and Lakehouse Monitoring.

Throughout, Unified Governance keeps data and AI assets secure and compliant, using tools like Unity Catalog and Mosaic AI Gateway for centralized control.

databricks

Conclusion

As we move forward and transition to new ways of working, adopt new technologies and scale operations, investing in effective data management is critical to removing the bottleneck in modernization. With the Databricks Data Intelligence Platform, you can manage your data from ingestion to analytics and truly unify data, analytics and AI.

Learn more about data management on Databricks: Watch now

Visit our Demo Hub: Watch demos

databricks

About Databricks

Databricks is the data and AI company. More than 10,000 organizations worldwide — including Block, Comcast, Condé Nast, Rivian, Shell and over 60% of the Fortune 500 — rely on the Databricks Data Intelligence Platform to take control of their data and put it to work with AI. Databricks is headquartered in San Francisco, with offices around the globe, and was founded by the original creators of Lakehouse, Apache Spark™, Delta Lake and MLflow. To learn more, follow Databricks on LinkedIn, X and Facebook.