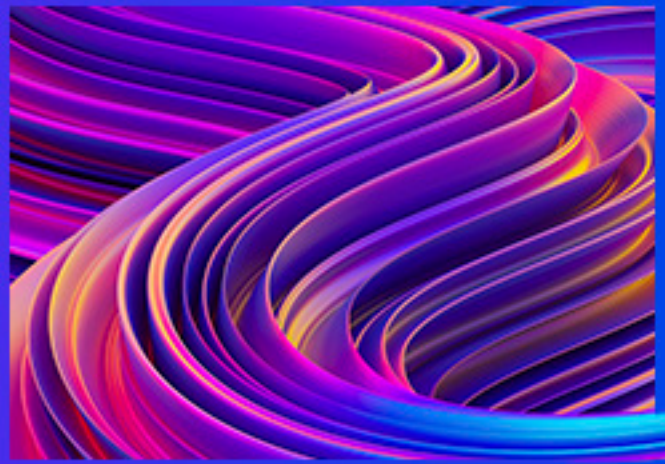# Unify your data, analytics and AI on one platform

The business world is swimming in data. Around 2.5 quintillion bytes of new data are generated each day.[1] Data warehouses and data lakes can help ingest, manage, and analyze large amounts and different types of business data, but both platforms are limited in their ability to support modern analytics in a unified data repository across disparate systems.

Traditional data warehouses are designed for structured, relational tables of data and specific analyses such as ad hoc queries and reporting. They often rely on batch processing, which is appropriate for aggregating totals but is not suited for real-time queries and processing. However, their data models can complicate data topic on-boarding efforts, product limitations can restrict real-time analysis, and the overall architecture can incur high ownership costs.

Data lakes deal with massive amounts of unstructured data and various analyses, including data exploration, predictive modeling, and automated decision-making. They also support streaming data, machine learning (ML) capabilities, and a high degree of scalability. But with data lakes, flexible storage can degrade into data swamps, flexible standards can erode data trust, and flexible standards can lack analytical data optimization.

Data lakehouse—the best of both worlds

The Databricks Lakehouse Platform is an open data management architecture that combines the best of data warehouses and data lakes. The Databricks Lakehouse Platform delivers the reliability, strong governance, and performance of data warehouses with the openness, flexibility, and ML support of data lakes.

Three essential qualities can describe the benefits of a Databricks Lakehouse Platform:

Simple:

Organizations can unify their data warehousing and artificial intelligence (AI) use cases on a single platform. Databricks Unity Catalog provides governance for data and AI assets in the lakehouse, giving organizations much better control, management, and security across clouds. This approach eliminates data silos that traditionally separate analytics, business intelligence, data science, and ML.

Open:

Delta Lake forms the open foundation of the lakehouse by providing reliability and record-setting performance directly on data in the data lake. Organizations can easily share data and build their modern data stack with unrestricted access to the ecosystem of open-source data projects.

Multicloud:

Databricks takes a cloud-agnostic approach, avoiding vendor lock-in and providing a consistent management, security, and governance experience across all cloud platforms. Organizations don't need to invest in reinventing processes for every cloud platform that they use to support their data and AI efforts.
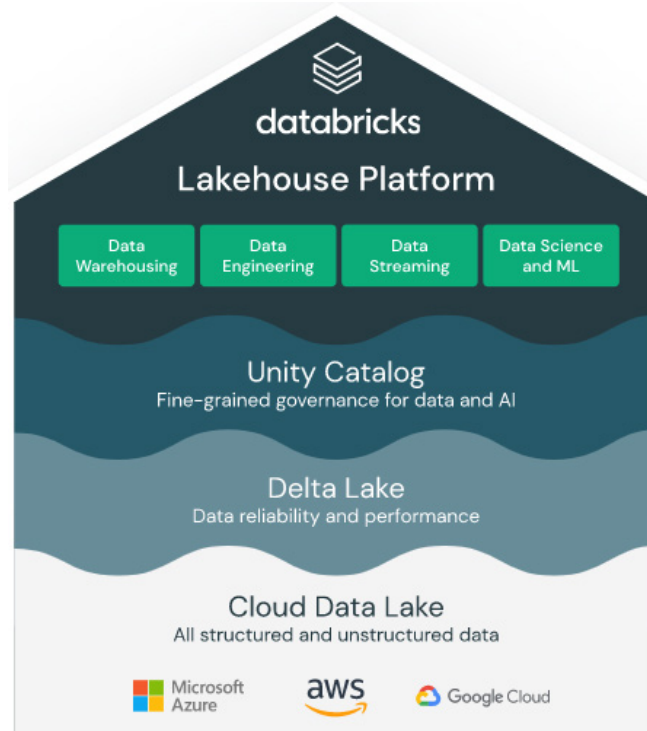
Your value-add with KPMG

KPMG LLP (KPMG) and Databricks have established a strategic alliance and our technology specialists work closely with Databricks daily to provide our joint clients with data, analytics, and AI solutions using the Databricks Lakehouse Platform. KPMG draws upon extensive industry knowledge, deep domain experience, a broad understanding of business processes, and technology know-how at every stage of the implementation process.

KPMG and Databricks professionals can help develop a future-state roadmap aligned with the overall business vision of the customer, establish pipelines from existing data sources and data lakes to build the Data Lakehouse Platform, foster collaborative experiences between data scientists and engineers, and implement data policies that support data privacy regulations.

Additionally, KPMG offers a number of accelerators, frameworks, and tested methodologies to help optimize the strategy, design, implementation, and support for the Data Lakehouse:

KPMG Modern Data Platform (MDP) is a scalable, cloud-native platform and approach designed to streamline data infrastructures and accelerate advanced data analytics initiatives. MDP integrates and harmonizes disparate data in a common data model, simplifying and democratizing access to data so business and data professionals can rapidly innovate and collaborate on their own, with less dependence on information technology.

KPMG Ignite is our patented, easy-to-use AI solution that combines a feature-rich AI development platform with a robust portfolio of prebuilt, industry-tested, AI-enabled capabilities. The solution ingests various types of data components from different sources and applies AI-based automation patterns to create intelligent workflows, helping to solve business problems specific to data strategy.

KPMG Ambient Data Management uses advanced AI and ML algorithms to monitor key elements of the data—such as quality, lineage, metadata, and master reference data. This provides organizations with the ability to identify and work with the highest quality of data possible to enhance analytic insights and accelerate problem resolution.

Contact us

For more information on how the KPMG and Databricks alliance team can help unify all of your data, analytics and AI on one platform.

Dan Fisher
Principal, Advisory
Lighthouse | KPMG LLP
T: 214-840-2180
E: danielfisher@kpmg.com

Barry Raghunathan
Managing Director, Advisory
Lighthouse | KPMG LLP
T: 732-501-7073
E: braghunathan@kpmg.com

Jeff Gilmore
Senior Director, Databricks
Alliance Lead | KPMG LLP
T: 312-231-3938
E: jeffgilmore@kpmg.com

Some or all of the services described herein may not be permissible for KPMG audit clients and their affiliates or related entities.

kpmg.com/socialmedia

January 2023