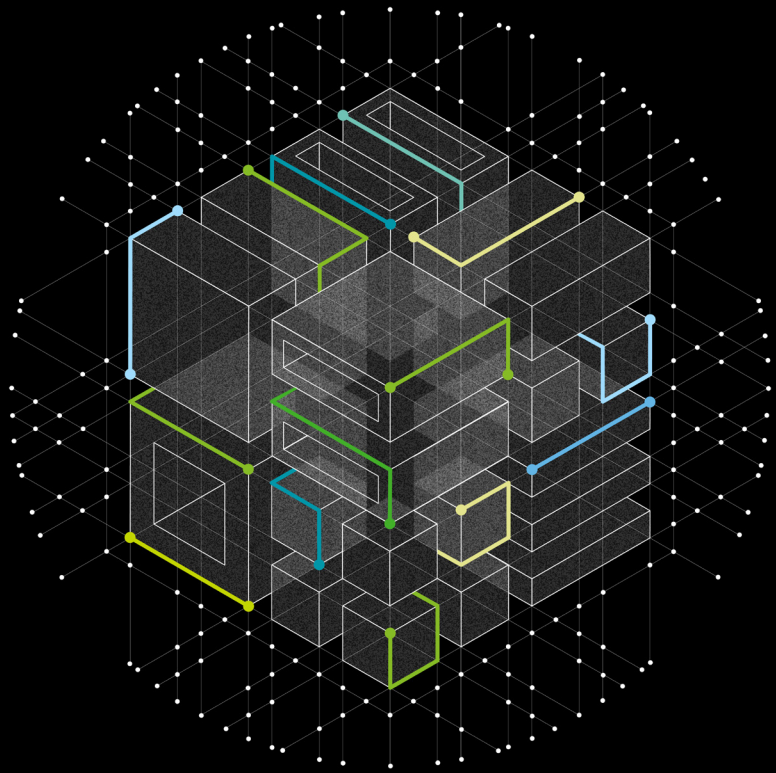# Deloitte.

databricks

A financial services leader's guide to building a modern governed data platform

Unlock key business outcomes, reduce costs, and harness the full potential of your data with Deloitte and Databricks.
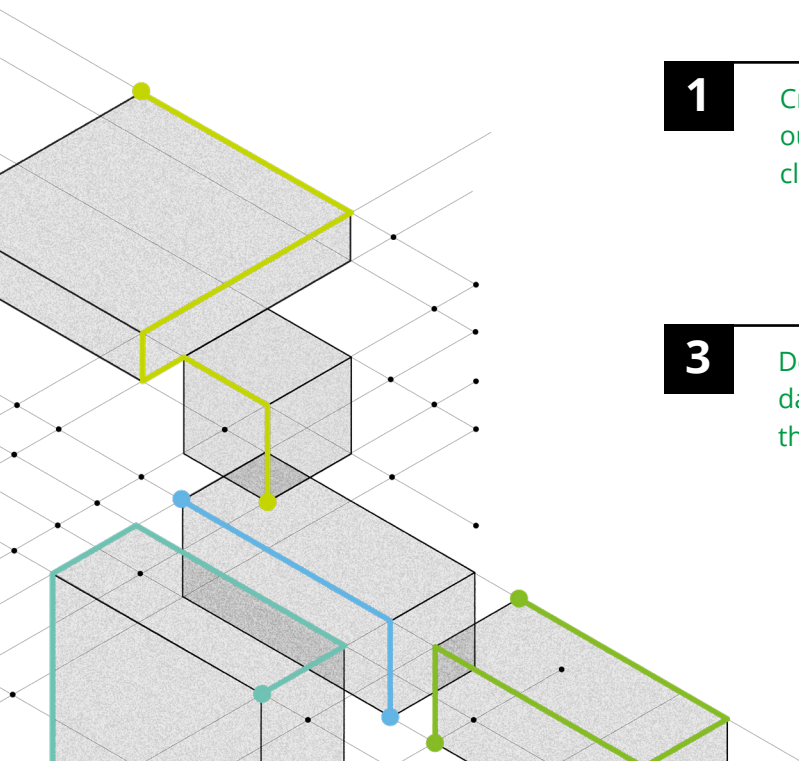
Today, as organizations seek to capitalize on rapid advancements in technology such as artificial intelligence (AI), machine learning, and hyper-personalization, the need for a modern governed data platform is more relevant than ever. This is particularly true in the financial services industry, an ever-evolving space which depends on quality and traceable lineage to understand, analyze, and cater to customer needs.

> *To stay on the vanguard of change, organizations require a data, analytics, and AI platform that equips them with advanced tools and capabilities that drive informed, data-driven decisions— all while putting in place controls to ascertain the efficacy of their data.*

Databricks, the world's first cloud-based lakehouse architecture, addresses these needs with its unified and interoperable system that covers everything from data input, preparation, and curation to in-depth analytics, machine learning, and AI. But implementing a data, analytics, and AI technology solution is just the first step in the process. To achieve the expected results, financial services organizations should implement an integrated approach that effectively navigates both the intricate landscape of industry-specific obstacles and individual organizational needs.

In this report, we'll explore four key challenges to modernization efforts in the financial services industry—and how Deloitte's experience paired with Databricks capabilities can:

1 Create an outcome-oriented cloud strategy

2 Align stringent data controls with key business outcomes

3 Develop a new data model for the cloud

4 Configure data governance, data quality, lineage, and security in the cloud

Raising the stakes: Under the regulatory lens

Organizations across sectors are facing unprecedented levels of scrutiny and accountability from customers, investors, regulators, and governments alike, and the financial services industry is no exception. Accelerated timelines for regulatory reporting in the EU under MiFID II is just one recent example.

While regulatory control and governance are deeply ingrained in the industry's DNA, the current climate requires an even greater level of vigilance to navigate complex data ecosystems. In this era of generative AI, there is exponential growth and utilization of data, which correlates to more compliance requirements for how data is used to train AI models. To meet the demands of both internal audits and regulatory requirements, teams must be able to prioritize effective operational execution of data governance while ensuring data stewardship and data quality, providing data lineage, and enabling control testing. Data unification, privacy, residency, security, and ethical use of data and analytics are not optional.

Key considerations by sector

Insurance

- **Regulatory:** The insurance industry is subject to a high degree of regulatory control and oversight to protect consumers and ensure fair competition.

- **Long Duration Targeted Improvements (LDTI):** A modified accounting model standard for insurance contracts.

- **The California Consumer Privacy Act and California Privacy Rights Act (CCPA + CRPA):** Two acts to ensure and protect the privacy rights of California consumers. (In the not-so-distant future, this will likely expand to other states.)

- **The General Data Protection Regulation (GDPR):** A law established by the European Union which enforces data protection and privacy for individuals within the EU and the European Economic Area.

- **Customer 360/Customer 720:** By constructing a comprehensive customer profile, insurers can not only gain valuable insights into how to improve customer service, but they can also pinpoint profitable opportunities.

Banking

- **Regulatory and financial reporting:** There exist various regulations and requirements within the industry related to financial reporting such as capital adequacy, liquidity, and credit risk management. Regulatory guidelines such as Basel Committee on Banking's Supervision's standard number 239 (more commonly known as BCBS 239) have also put an emphasis on governed data and analytics.

- **Integrated Finance risk:** The integration of various finance and risk management functions, including treasury, liquidity management, and asset-liability management.

- **Customer 360/Customer 720:** Building a complete view of customers, such as their financial behavior and risk profile, to facilitate anti-money laundering measures, credit decision-making, and ongoing credit lifecycle management.

- **Ethical AI:** As organizations delve into AI's potential to enhance customer experiences, ethical considerations surrounding AI, such as detecting biases, are increasingly prominent.

Real Estate/Investment Management

- **Sales and distribution:** The process and regulations surrounding the marketing and selling of real estate or investment products.

- **Customer and channel insights:** Insights into customers and channels to enable investment managers to optimize distribution and grow their customer base.

- **Assets under management (AUM):** Sometimes referred to as funds under management, this is the total market value of assets that a financial institution manages on behalf of clients.

- **Trade lifecycle:** The process of executing and settling trades in financial markets—from order placement and trade execution to clearing and settlement.

At Deloitte, we have developed tailored solutions to help organizations in the financial services sector manage their data in a way that satisfies these requirements while also driving business outcomes.

CHALLENGE 1

Create an outcome-oriented cloud strategy

At the cornerstone of any successful cloud strategy is the capability to leverage data and analytics to prioritize effective business decision-making and achieve specific business outcomes. Cloud technology drives this critical prioritization, yet many organizations still make the mistake of migrating to the cloud without a clear understanding of how the technology can best enable the outcomes they seek. And even with a smooth migration, organizations often still fall short of achieving their goals, resulting in wasted time and resources.

To effectively tackle this challenge, teams must first delineate how their cloud strategy will help them reach the necessary benchmarks by answering three critical questions:

How can business decisions be better enabled and supported by data, analytics, and AI?

Is the resulting strategy risk-managed and compliant?

How will this strategy help reduce costs and improve operational efficiency?

The solution

To aid organizations in their modernization journey, Deloitte and Databricks take an iterative approach that delivers tangible value within months. The process begins with a streamlined methodology for organizations to modernize and move to the cloud. Equally crucial to this migration process is evaluating each organization's unique approach to reengineering, which will be key to unlocking advanced features like data sharing and machine learning.

During this reengineering phase, Deloitte and Databricks help directly map business priorities to data, breaking down silos and rearchitecting operations to identify specific data domains and the analytical and AI use cases required to execute key business objectives. This phase also includes integrating risk profiles to improve compliance and reduce operational costs as well as organizing data for reuse to achieve greater efficiencies and further cost reductions.

Compliance starts with quality

High-quality and accurate data is essential for informed decision-making and regulatory compliance. Today, organizations face data silos across data lakes and data warehouses leading to a fragmented view of data across clouds and platforms. Governance poses a significant challenge for organizations, with potential consequences such as internal audits, regulatory issues, and error prone policy management. But it also provides a framework for teams to uphold their bottom line and prevent decisions based on unreliable data. This solution helps organizations implement governance across their data platform through data registration, governance, data quality, and lineage controls.

By taking a holistic approach to address these challenges, Deloitte and Databricks provide organizations with a comprehensive modernization solution—combining cloud data lake, database, and analytics capabilities—which drives the business outcomes integral to any cloud migration strategy.

Ensuring data accuracy is a multifaceted necessity that impacts various business outcomes. However, it has also historically been treated as a mere box to be checked, which can result in regulatory intervention.

A shift in mindset is therefore required to view data controls as intrinsic to the health and success of organizations. To achieve this, organizations must first understand how legacy systems and existing applications align with current and future business needs and outcomes. Controls are essential to prevent bad data from entering the system, resulting in inaccurate data output. The curation of clean, controlled, and governed data requires a systematic approach to receiving and organizing the different pieces of data from operational source systems to create reusable datasets, data products, and data views.

**Ensure peace of mind across the business**

Through this process of data curation, organizations can attain data accuracy and trustworthiness. In turn, this enables chief financial officers to report accurate financial information, risk managers to make informed risk decisions, and other business leaders to optimize value delivery, such as curating specific data for specific customers.

The solution

Databricks provides modern data analytics components that enable organizations to streamline their cloud modernization efforts, adhere to regulatory requirements, and expand into the AI ecosystem. With the Databricks Data Intelligence Platform, data processing involves four key stages:

Ingestion > *Curation* > *Consumption* > *Analytics*

...all supported by a unique medallion architecture that can help ensure traceability throughout.
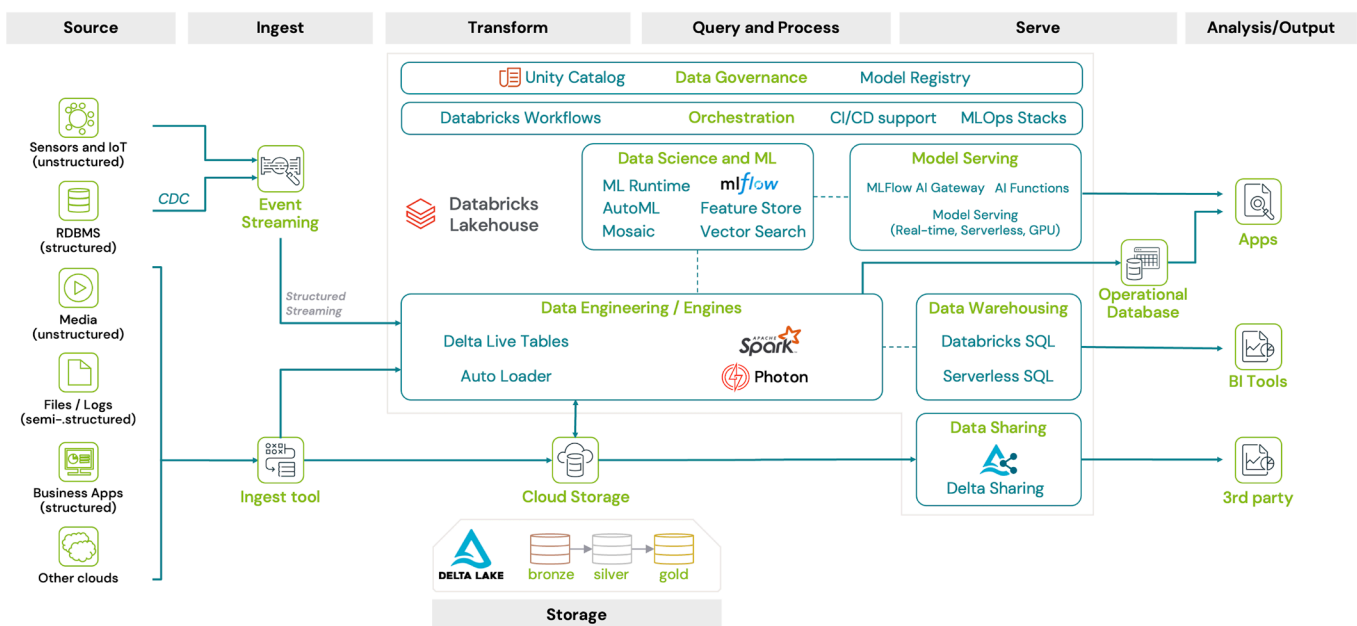
Databricks lakehouse architecture offers several unique features that set Databricks apart. For example, unlike many available options, the Databricks Data Intelligence Platform supports the entire data processing pipeline from raw data to cleansed and conformed data. It also coexists seamlessly with other tools, including cloud databases, allowing users to access data once it's ready to be used for analytics and more.

CHALLENGE 3 — Develop a new data model for the cloud

Adopting new technologies and tools can be a game-changer for modernizing organizational operations. Even so, success here still requires an integrative approach. And without a well-designed data model, organizations risk inefficiencies or performance failures— from workflow silos and data inaccuracies to increased risk due to unnecessarily repetitive processes and data copies.

But every organization naturally possesses a unique architecture and business plan— and therefore, a range of data modeling approaches that can be implemented based on various factors is key to laying the foundation for a dynamic new model for the cloud.

Deloitte's data model frameworks

More often than not, the data model used to curate and ingest the raw data will shape how data is organized as well as the final data domains and products—and Deloitte has developed two proprietary data model frameworks to help clients through this process.

- **Data as a Service (DaaS) for Insurance framework:** designed for the insurance industry.

- **Data as a Service (DaaS) for Banking framework:** designed for the banking industry.

Both frameworks allow organizations to accelerate the deployment of Databricks lakehouse to deliver curated data products to enable business outcomes

The solution

In the past, organizations have modeled their data directly into normalized data warehouses, resulting in inefficiencies and performance issues. With Databricks, the entire data architecture is underpinned by Delta Lake, an ACID compliant format which can drastically reduce total cost of ownership for data engineering workloads and analytics workloads. Teams are empowered to apply a tailored data-modeling approach at each layer of the process, from ingestion and curation to consumption, such as write-optimized data models (e.g., 3NF, Data Vault) for ingestion and curation layers and read-optimized data models (e.g., Star-Schema, Dimensional) for consumption layers. Moreover, every stage of data curation is automatically captured in Unity Catalog, Databricks' unified governance solution for data and AI.

Databricks enables the curation of data into data products (e.g., customer, transaction, payment, policy, billing). It also empowers aggregated views of data (such as Customer 360, Finance 360, Regulatory 360, or Channel 360) and provide this data as a data lake view, Database, API, or as a graph. Deloitte is poised to help accelerate and personalize this process for every organization's individual needs, and frameworks such as Data as a Service for Banking and Data as a Service for Insurance reduce risk in implementing such solutions.
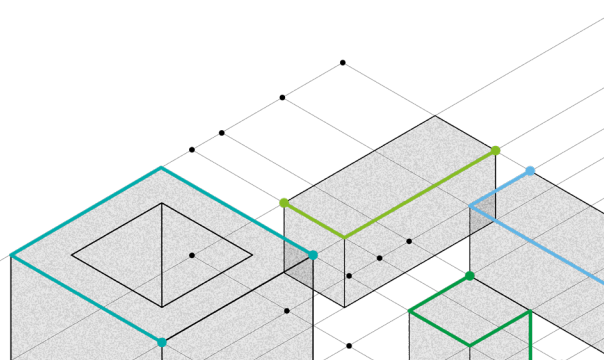
Configure data governance, data quality, lineage, and security in the cloud

As financial services organizations move to the cloud, configuring data governance, quality, lineage, and security becomes a critical challenge. Foundational data management is essential to operationalize data governance and data stewardship, ensuring that data is attributed to the appropriate ownership and managed accordingly. On top of this, the cloud environment itself presents new data governance and security challenges, prompting organizations to act with even more haste.

The solution

Ensuring the accuracy and trustworthiness of key business data requires operationalizing data controls. However, implementing these controls on legacy systems such as a mainframe can be challenging due to platform constraints. To establish sustainable (and compliant) data quality and security rules, the trick is then to implement these controls as close to the source system as possible through a data platform such as the Databricks Data Intelligence Platform.

With Databricks Unity Catalog, organizations can seamlessly govern their structured and unstructured data, machine learning models, notebooks, dashboards and files on any cloud or platform. Data scientists, analysts and engineers can use Unity Catalog to securely discover, access and collaborate on trusted data and AI assets, leveraging AI to boost productivity and unlock the full potential of the lakehouse environment. This unified approach to governance accelerates data and AI initiatives while helping to ensure regulatory compliance in a simplified manner.

For example:

- When raw data is extracted from the mainframe and imported into the Databricks Data Intelligence Platform, it is initially placed in a "raw zone." At this stage, critical data elements can be identified, data quality rules defined, stewardship defined, data owners assigned, and data classified and tagged. Finally, the data is registered into a metadata repository or data catalog.

- Data quality rules are then executed to validate the data.

- Subsequently, this raw data is cleaned up and moved into a "clean data zone". An Authorized Provisioning Point can be created for clean data. This data can then be further organized into reusable data products in the curated data zone on the Databricks Data Intelligence Platform.

- Once the data is curated, a data catalog can be further enriched with curated data products, creating a central repository for customer or business data domains that can be utilized again and again, by various users.

- Powered by Unity Catalog, Databricks Lakehouse Monitoring enables high quality, accurate, and reliable traceability and lineage of data and AI assets flowing through its platform, from source to raw, clean, and curated. The single, unified approach to monitoring enabled by lakehouse architecture makes it simple to diagnose errors, perform root cause analysis, and find solutions.

- The platform additionally facilitates appropriate data modeling techniques to enable this repository of clean, governed, and traceable information.

Before implementing this architecture, organizations should consider designating data stewards who understand the data at the raw and cleansed stages. These stewards can manage the issue resolution processes to operationalize data ownership and take accountability to facilitate cleaning up data.

Ultimately, Databricks helps organizations systematically organize raw data, creating various consumption layers that ensure quick accessibility to various users. As an example, Databricks incorporates AI into governance workflows to auto-classify PII data. This in turn enables teams across the organization to break free from previous data silos to discover, access, and utilize this cleaned and organized information for a range of business processes and outcomes—from transactions and billing to product customization and personalization.

CONCLUSION

Explore the art of the possible

In today's business landscape, organizations face numerous obstacles in realizing the full potential of data analytics and AI. Establishing an efficient, effective, and governed platform is therefore crucial to drive business outcomes and maintain a competitive edge. As the world's leading consulting firm, Deloitte is uniquely positioned to guide and advise your team as you journey toward a modernized data system, providing expert insight and implementation know-how to help you overcome any roadblock.

Databricks, meanwhile, equips organizations with advanced analytics, AI, machine learning, and other emerging data technologies within its easy-to-use, compliant platform. Databricks also provides the latest innovations focused on Generative AI and Large Language Models, such as MosaicML and Lakehouse AI, which can be tightly integrated with the Databricks Data Intelligence Platform. In a few simple keystrokes, organizations can quickly set up a Databricks environment and avoid intensive build processes. And the platform's pricing model further enables teams to accelerate their migration journey without upfront investments in infrastructure and maintenance costs, allowing them to deliver actionable insights and drive meaningful business outcomes. This is an end-to-end solution, which fosters reuse, efficiency, cost reduction, and multi-cloud capabilities—powering modernization and growth for teams across industries.

Contact us today to learn more and unlock the full potential of your data.

Jojy Mathew
Principal
Deloitte Consulting LLP
jojymathew@deloitte.com

Cindy MacFarlane
Managing Director
Deloitte Consulting LLP
cmacfarlane@deloitte.com

Dave Hurlbrink
Senior Manager | Channel Sales
Deloitte Consulting LLP
dhurlbrink@deloitte.com