

Modernizing data governance at scale with Databricks

Unified governance and agentic AI help Healthcare and Life Sciences (HLS) organizations strengthen compliance, reduce cost, and accelerate innovation.



HLS organizations must elevate data governance into a unified, AI-driven strategic advantage. In our work on Databricks, we help clients build unified governance with agentic AI, strengthening compliance, accelerating clinical innovation, and improving patient outcomes.

Navigating HLS data governance opportunities with Databricks

- Fragmented data ecosystems make regulatory compliance with HIPAA and GDPR increasingly complex and resource-intensive.
- Databricks Unity Catalog establishes a unified governance foundation that strengthens security and reduces operational overhead.
- Agentic AI automates discovery, modeling, and compliance workflows, minimizing manual intervention and improving data reliability.
- Unifying governance and operationalizing compliance accelerate decision-making and enable trusted AI adoption at scale.
- In many HLS environments, unified governance on Databricks can reduce data access timelines from weeks to hours, significantly accelerating clinical and operational decision-making.

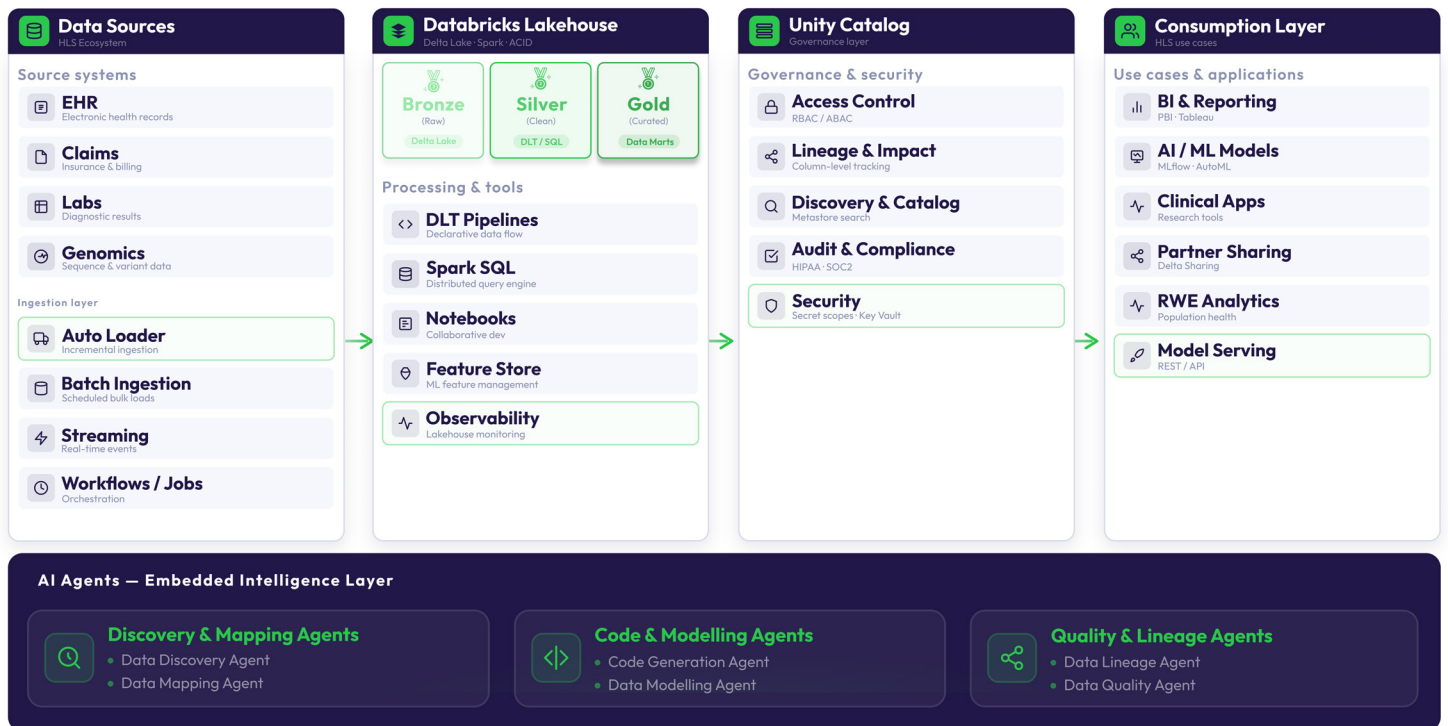
Why data governance is increasingly complex in HLS

HLS organizations manage some of the most complex and sensitive data in the world, spanning EHRs, clinical trials, and genomics. While these assets hold immense potential for growth and innovation, fragmented systems and an increasingly complex data ecosystem make regulatory compliance with bodies like HIPAA and GDPR increasingly difficult. Reframing data governance is thus critical. Databricks addresses this challenge through a unified governance solution, Unity Catalog, which provides a centralized framework for managing access, lineage, and compliance. Complementing this, its enterprise agent capabilities enable intelligent automation across governance workflows. Together, these capabilities shift governance from a control function to a strategic advantage that drives cost efficiency, innovation, and improved patient outcomes.

How a unified governance layer reduces risk and fragmentation

Healthcare and Life Sciences are high-stakes industries. Managing dozens of disconnected repositories with separate rules creates significant operational friction. A unified governance model directly reduces this operational complexity. Databricks Unity Catalog introduces a common governance layer for managing fine-grained access, end-to-end data lineage, and compliance tracking across enterprise systems. Establishing a single source of truth strengthens security, minimizes manual oversight, and aligns the data ecosystem with key growth targets. Consolidation directly impacts the bottom line by reducing duplication, simplifying infrastructure, and lowering operational costs.

Our approach: A governed data lake house on Databricks with agentic AI



Data sources: Intelligent discovery

Healthcare data from EHRs, claims, labs, clinical trials, and genomics systems is automatically discovered and classified by AI agents. These agents identify sensitive PHI and PII, generate relevant metadata, and recommend optimal ingestion

Databricks lakehouse: Automated data engineering

Data is ingested into bronze, silver, and gold layers within the Lakehouse architecture. AI agents automate schema mapping, data quality validation, and transformation logic, while supporting legacy ETL migration and modern data modeling. This enables faster delivery of trusted, analytics-ready datasets.

Governance with Unity Catalog: Autonomous governance

Databricks Unity Catalog provides centralized access control, lineage, discovery, and compliance. AI agents continuously classify data, recommend governance policies, identify compliance gaps, generate documentation, and monitor access patterns, strengthening security and trust across the data estate.

Consumption layer: AI-driven insights

Governed data is consumed through BI dashboards, AI and ML models, clinical applications, and APIs. AI agents assist users in discovering relevant datasets, generating insights, automating reporting, validating outputs, and accelerating decision-making, while reducing migration and testing effort.

How agentic AI extends governance across data and workflows

While unified governance establishes the foundation, scaling it across complex HLS ecosystems requires intelligent automation. Databricks agent capabilities extend governance by introducing autonomous, goal-driven workflows that improve efficiency, reduce risk, and accelerate decision-making.

- 1. Intelligent data discovery:** AI agents automatically scan datasets across domains, tag sensitive information like PHI and PII, and recommend certified datasets for use. This reduces manual cataloging efforts and ensures that executive dashboards reflect highly trusted, real-time insights.
- 2. Automated data modeling:** Agents analyze usage patterns to suggest optimized data models for advanced analytics. This rapid creation of analytics-ready datasets supports population health initiatives and predictive clinical modeling, reducing the heavy lifting typically required from data engineering teams.
- 3. Strategic modernization:** Legacy systems often stall business agility. Agentic AI helps convert legacy ETL processes into native pipelines. It optimizes queries for performance and cost while ensuring strict governance policies remain embedded by design, significantly reducing migration timelines.
- 4. Quality and risk automation:** Agents automatically generate test cases for data pipelines and continuously monitor data quality rules. This proactive risk management approach detects anomalies early, leading to higher data reliability and better patient safety outcomes.

Translating governance into business impact

Secure data sharing capabilities allow organizations to collaborate across internal departments and external partners without compromising strict privacy controls. For life sciences teams executing clinical trials, moving from fragmented data discovery to an automated, integrated access model can reduce data access time from weeks to hours. This securely accelerates drug development timelines and improves global team collaboration. Furthermore, embedded policy enforcement such as role-based access controls and automatic masking of sensitive patient data streamlines regulatory reporting. Compliance audits that traditionally took weeks can conclude in days, drastically lowering the risk of violations and reducing manual effort.

Our HLS use cases powered by Databricks and agentic AI

Intelligent data discovery

AI agents address disconnected data environments by automatically scanning datasets, tagging sensitive data like PHI and PII, and recommending certified datasets for use. This transforms discovery into an intelligent, continuous process.

MANUAL CATALOGING EFFORT

40%
reduction

Improved data usability through automated tagging and certification.

Code migration and modernization

Agentic AI enables automated conversion of legacy pipelines into Databricks-native architectures, optimizes queries for performance and cost, and embeds governance controls by design. This shifts modernization from a manual, fragmented effort to a more consistent and scalable process.

MIGRATION EFFORT

50%
reduction

Improved scalability through optimized, governance-aligned pipelines.

Maximizing value from Databricks unified governance and agentic AI

- Establish a unified governance foundation on Databricks early to eliminate scaling inefficiencies, reduce costs, and build a resilient data ecosystem.
- Operationalize compliance workflows with automated lineage and fine-grained access controls to proactively manage enterprise risk and reduce effort.
- Invest in intelligent data discovery to empower business leaders with trusted, real-time, decision-grade insights for significantly faster strategic planning.
- Reposition governance from a static control mechanism into a strategic value driver enabling trusted AI, innovation, and better outcomes.

About Brillio

Brillio is a digital technology services company that drives AI-first engineering and design-led experiences for global enterprises. Born digital in 2014, its consulting-led services span Customer Experience, Data & AI, Product Engineering, and Digital Infrastructure. With an industry-leading NPS of 71, Brillio accelerates time to market through its proprietary BrillioOne.ai platform, powered by AI-ready talent with deep domain expertise.

Brillio is the official Digital Transformation Partner and the official Data and AI Services Provider of Atlassian Williams Racing. Brillio partners with leading technology providers including Microsoft, AWS, Google Cloud, Salesforce, Adobe, Databricks, and Snowflake and operates with 6,000+ “Brillians” across 15 global delivery centers.

Consistently recognized as a Great Place to Work® since 2021, Brillio blends innovation, talent, and purpose to deliver measurable outcomes for clients and fulfilling careers for employees.



<https://www.brillio.com/>

Contact Us: info@brillio.com

brillio

The logo for Brillio, featuring the word "brillio" in a lowercase, white, sans-serif font. Below the letters "i" and "l" are two small, solid green circles.