



DataHub



databricks

Get unified AI & data context across your entire Databricks ecosystem and beyond



Real-time platform to manage data and AI assets

25% savings in data storage costs

Cut cloud costs instantly by identifying and removing unused or duplicate data across your Databricks ecosystem with real-time metadata intelligence. Save money instantly. Process changes in real time at any scale.

Launch AI models faster with unified platform

Leverage the only modern metadata platform with REST and GraphQL APIs that optimize performance for complex queries to accelerate AI deployment. Ensure models train on quality data with complete lineage tracking and integration.







Boost productivity of data users with enterprise-grade extensibility

DataHub automatically generates comprehensive data lineage extending across your Databricks ecosystem and beyond. Skip extensive manual setup and instantly identify data sources, users and dependencies for an in-depth understanding of your data supply chain.

Cut development costs with business-adaptive architecture

Adapt to business changes without the cost of custom development. Configure, don't code, to match your specific needs and save 80% while ensuring your platform evolves as your business does.

With DataHub, **trusted AI-ready data** becomes a reality

-  **Discover relevant datasets** using powerful search and AI-based natural language interfaces
-  **Uncover insights with AI-assisted summaries** of transformation logic and popular queries
-  **Understand business context in real-time** through AI-generated documentation and automated context propagation
-  **Monitor data quality** through automated checks, AI anomaly detection, and assertions
-  **Track lineage across your data ecosystem** to reveal hidden dependencies and make changes with confidence through impact analysis.
-  **Implement dynamic governance** with compliance forms, certification workflows, and enforced standards

Complete data context beyond Unity Catalog

NEW

Bi-Directional Metadata Sync

Cross-platform integration

End-to-end lineage

Business context enrichment

AI-powered data discovery






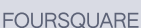




FOURSQUARE

DataHub serves as central control panel for Foursquare geospatial datasets in Databricks and beyond.

“Our teams reported a reduction in data discovery and access time from the order of **days to minutes**.”

Vikram Gundeti, Chief Technology Officer

Trusted by 6300 + innovative data teams with largest deployments

 Figma  T-Mobile  NETFLIX  Notion  coursera  FOURSQUARE  slack  Optum  myob  airtel

Download the Guide:
7 Reasons to Rethink Your Data Catalog

