

Data Governance with Unity Catalog

Secure, manage, and audit your data assets in one place using Databricks Unity Catalog

Business Challenges

Organizations face increasing challenges managing and governing their data assets as data landscapes expand, become more fragmented, and as data volumes grow. Furthermore, organizations also need to maintain a solid grasp on data access controls and regulatory compliance to minimize risk exposures from data misuse and data breaches. Finally, data governance is crucial for maintaining control over data assets. Effective governance includes establishing clear roles and responsibilities, ensuring data quality, and maintaining audit trails.

Solution Summary

Unity Catalog is a powerful tool for managing and securing data, providing centralized access control, auditing, lineage tracking, and data discovery capabilities. Imagine a large library with books, magazines, and digital content spread across multiple rooms and floors. Unity Catalog is like a smart cataloging system for this library that:

- ✓ Organizes all the content in a structured way
- ✓ Helps you easily find what you need
- ✓ Controls who can access which materials
- ✓ Keeps track of who viewed what and when
- ✓ Shows how different pieces of information are connected

Infinite's Databricks Unity Catalog Implementation Solution offers organizations a comprehensive approach to putting its capabilities to optimal use.



Features & Benefits of Unity Catalog

- **Define Once, Secure Everywhere:** Provides a unified platform for administering data access policies across workspaces and user personas.
- **Metadata Management:** Centralizes metadata management, enabling standardization of data definitions, structures, and formats.
- **Data Discovery:** Facilitates data discovery through tagging, documentation, and an intuitive search interface for easy access to relevant data assets.
- **Data Lineage:** Automatically captures data lineage for tables, view, and columns across the data lifecycle, from source ingest to downstream consumption.
- **Auditing and Logging:** Automatically captures detailed audit logs to track data access and usage.



Databricks Unity Catalog

Discovery

Access Controls

Lineage

Monitoring

Auditing

Data Sharing

Tables



Files



Models



Notebooks



Dashboards



We can do your Unity Catalog migration if you are...

1 Not currently using Databricks or Unity Catalog

- For organizations looking for a complete migration to Databricks and Unity Catalog
- For organizations using other governance solutions that don't have the full suite of centralized data governance capabilities Unity Catalog offers

2 Using Databricks but not fully using Unity Catalog

- For organizations who want to complete a previously started migration to Unity Catalog
- For organizations already on Databricks that have assets in Hive Metastore or AWS Glue
- For organizations using basic elements of Unity Catalog who wish to use more features such as Delta Sharing

Migration Approach

We tailor our Unity Catalog migration approach to clients' unique needs, adapting it according to the organization's technical environment and data governance requirements.



Conduct discovery to gather technical and functional requirements. Run pre-migration assessment. Review catalogs, workspaces and access control patterns. Tailor the migration plan.

Setup and configure the Unity Catalog at the Account level. Document the current state, objects to be migrated, and design the Unity Catalog architecture. Create metastore.

Create catalog and schemas. Sync legacy metadata, and upgrade tables and views, then create access controls. Start with one workspace and validate, then repeat for other workspaces.

Test thoroughly to validate successful object migration. Test access controls. Create go live plan, checklist, and rollback options.

Perform the migration in production. Finalize and review documentation. Do knowledge transfer to ensure a smooth transition and handoff.