

CASE STUDY

From manual governance to observable governance

How UnionBank of the Philippines embedded lineage, ownership, and trust into 40,000+ data assets with Collate on OpenMetadata

**90%**

Decrease in service-desk requests related to data discovery and observability

60%

Faster root-cause analysis and remediation (5 days → 2 days)

40,000+

Production data assets made discoverable, traceable, and governed

UnionBank of the Philippines is one of the country's largest and most digitally forward banks, operating in one of the world's most regulated industries. As its data ecosystem grew to more than 40,000 production data assets across hundreds of dashboards and interconnected pipelines, governance still depended on manual coordination, fragmented metadata, and "data historians" who carried critical lineage and ownership knowledge in their heads. By running **Collate's managed OpenMetadata service**, UnionBank made governance operationally visible, embedding lineage, ownership, and trust signals directly into how data operates, for both human and AI consumers.

INDUSTRY

Banking / Financial Services

HEADQUARTERS

Pasig City, Philippines

FOUNDATION

Collate's managed OpenMetadata service.
Observable data governance layer.

TECHNOLOGIES

Collate, Snowflake, AWS, SageMaker,
QuickSight

"In a highly regulated industry, data governance is not simply a documentation exercise. It impacts regulatory compliance, operational resilience, audit readiness, customer trust, and increasingly the organization's readiness for analytics and AI-driven operations."

Cirene Simon R. Simbahan, Head of Data Trust and Governance Office,
UnionBank of the Philippines

Governing a Deeply Interconnected Bank Without Operational Visibility

As UnionBank's enterprise data ecosystem grew, the team's visibility gaps and manual approaches to their governance posture were creating risk for their organization. More pipelines, more reports, more dashboards, and more interconnected data flows meant that in a highly regulated bank, these visibility gaps were not just inconveniences; they were control gaps.

OBSERVABILITY QUESTIONS FLOODING THE SERVICE DESK

11% of governance service-desk requests (98 of 886) were about observability and discoverability alone: where did this report come from, which upstream datasets changed, who owns this asset, which dashboards are impacted. These questions could not be answered reliably through tribal knowledge or disconnected spreadsheets.

DEPENDENCE ON DATA HISTORIANS

Critical governance knowledge lived in individuals, not systems: people who knew the details of where reports originated, which pipelines fed which dashboards, and why certain transformations existed. When those historians left, transferred, or retired, that knowledge left with them. In a regulated bank, that dependency on tribal knowledge was not sustainable.

SLOW, MANUAL LINEAGE INVESTIGATIONS

With more than 40,000 production data assets and 200+ dashboards consuming shared datasets, tracing lineage or identifying root cause could take several days. The ecosystem had outgrown manual governance approaches, with their existing spreadsheets and static policies not keeping pace with the business.

INTERCONNECTED RISK ACROSS REGULATED REPORTING

Banking ecosystems are deeply interconnected. One upstream issue can hit regulatory reporting, operational dashboards, analytics outputs, downstream decisions, and AI-driven use cases at the same time. That made visibility and traceability control requirements, not operational conveniences.

Observable Data Governance with Collate

UnionBank shifted its mindset from traditional, documentation-driven governance to **observable and automated data governance**, where metadata becomes operational intelligence and governance is embedded into how data operates, not just how it is audited. Collate was rolled out to connect lineage, ownership, and trust signals directly into operational workflows.

GOVERNANCE EMBEDDED IN THE DATA PRODUCT OPERATING MODEL

As UnionBank transitioned to a data product operating model, it embedded governance capabilities directly into that model: data product owners, data grading, lineage, quality standards, and usage context. That context is connected directly to the data assets themselves, so governance is not theoretical but an actionable part of people's workflows.

A FOUR-TIER DATA GRADING FRAMEWORK

To communicate trust and intended usage consistently, UnionBank built a data grading framework based on the medallion architecture with four grades: authoritative (regulatory and critical reporting), validated (business decision-making), curated (analytics and operational reporting), and exploratory (discovery and sandboxing). These trust signals were made operationally visible through metadata, so users could immediately see whether a dataset could be trusted and how it was intended to be used.

METADATA AS OPERATIONAL INTELLIGENCE

Instead of static policies and spreadsheet-based monitoring, metadata became a live intelligence layer. Lineage became searchable and traceable, ownership became discoverable, and impact analysis became faster, turning governance from a gatekeeping function into an enablement capability.

SEARCHABLE, SYSTEM-OF-RECORD LINEAGE AND OWNERSHIP

Governance knowledge moved out of individuals' heads and into a centralized, searchable, traceable system. Lineage and ownership became discoverable for both human and AI consumers, so governance continuity could survive organizational changes, transfers, and turnover.

Making Trust Operationally Scalable

As governance visibility became integrated into operational workflows, UnionBank saw measurable improvements to their data and business processes. Governance scales better when visibility is embedded directly into the system, and the team has further initiatives planned on their operational maturity journey.

90% FEWER DISCOVERY & OBSERVABILITY REQUESTS

Service-desk requests related to data discovery and observability dropped by 90%, freeing the Data Trust and Governance team to spend time on higher-value governance work instead of answering "where did this come from" questions.

ROOT-CAUSE INVESTIGATIONS: 5 DAYS TO 2 DAYS

Investigation timelines for complex root-cause analysis and remediation that previously took up to five days were cut to two. That 60% reduction gave the team time to reinvest in other governance and AI initiatives.

100% LINEAGE AND 100% QUALITY COVERAGE ON CRITICAL DATA

UnionBank achieved complete metadata and lineage visibility for dashboards consuming governed data products, and full data quality coverage on critical data assets, not just request-specific or domain-specific coverage.

DATA USAGE ALIGNED TO RISK TOLERANCE

With grades attached to assets and visible in metadata, teams get clear guidance on which data can be used for which purpose. That improved alignment between data usage and the bank's risk tolerance, and surfaced gaps in data quality and governance earlier.



"In a highly regulated industry, data governance is not simply a documentation exercise. It impacts regulatory compliance, operational resilience, audit readiness, customer trust, and increasingly the organization's readiness for analytics and AI-driven operations."

Cirene Simon R. Simbahan

Head of Data Trust and Governance Office, UnionBank of the Philippines



getcollate.io

OBSERVABLE DATA GOVERNANCE