



CASE STUDY

DATA GOVERNANCE USING BIG DATA LAKE FOR A LEADING MULTINATIONAL BANK

Reduced technology costs by 30% and enhanced data governance

Client Background

Our client is a leading multinational investment bank and financial services corporation. Their financial services encompassed asset management, banking and consumer finance, credit and charge cards, insurance, investment banking, and trading. The client's critical inventory and infrastructure data was spread across multiple applications. This diffused nature of business resulted in the decentralization of data and its duplication. In other words, they experienced data integrity issues, lacked a single source of trusted data with timely insights and business value.

The key objectives included:

- Development of data models (both AS-IS and TO-BE) and data dictionary with complete profiling details for governance.
- Enhancement of data quality, governance, and integrity by building ingestion utility and business rules engine to port the inventory and infrastructure data in big-data Data Lake.

Xoriant Solution | Key Contributions

Xoriant partnered with the client to contemplate a robust data governance strategy using big data lake. The data governance solution was developed using new-age technologies. Our key contributions included:

- Development of a data dictionary to profile data, ensuring consolidated metadata for each application including AS-IS data model and creation of automated scripts for data analysis. This allowed the client to get insights into their existing data.

KEY BENEFITS

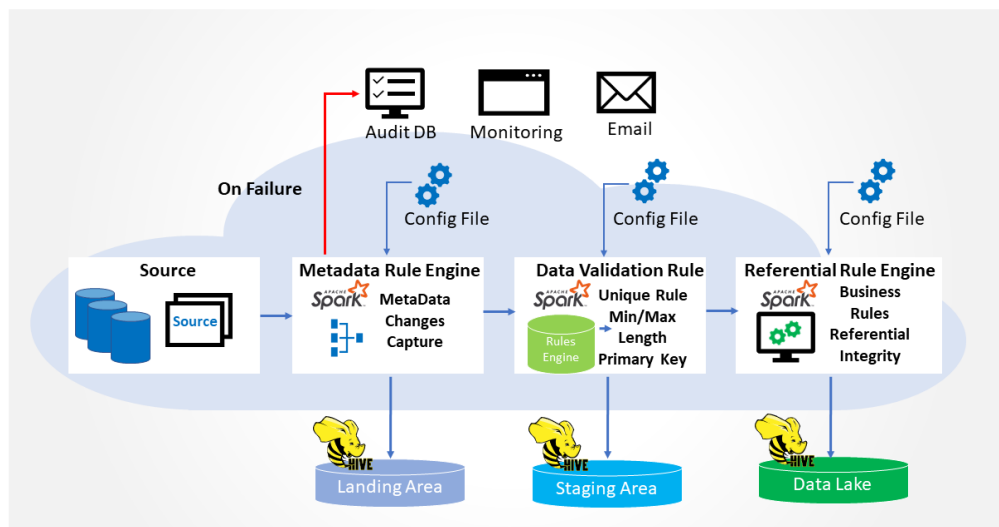
- Gained deep insights into data landscape with enhanced quality and integrity for governance
- Experienced effective data collection and management in big data lakes using ingestion utility and rules engine
- Reused the solution components for multiple client applications with enhanced extensibility
- Reduced technology costs by 30% using big data open source ecosystems – Apache Spark and Hive

- Design and development of an ingestion utility in Apache Spark to ingest data from disparate sources in the data lake on the Hadoop system.
- Defining a comprehensive and elaborate integrity rule engine configured for each table.
- Ensuring comprehensive auditing reports at each stage of the data ingestion to provide feedback to stakeholders and apprise them with issues in data.

Client Testimonial

“Xoriant team’s data management acumen and engineering expertise streamlined our data landscape. They ensured a single source of trusted data, enhanced data quality and governance with timely insights and business value.”

Architecture Diagram



Technology Stack:

Apache Spark | Java | Hive



Xoriant is a product engineering, software development and technology services Company, serving technology startups as well as mid-size to large corporations. We offer a flexible blend of onsite, offsite and offshore services from our eight global delivery centers with over 3600 software professionals. Xoriant has deep client relationships spanning over 30 years with various clients ranging from startups to Fortune 100 companies.