NorthBay

How to Solve
The Data Quality Conundrum in Financial Services

At a Glance:

Adding new data sources to legacy data warehouses is a complex and time-consuming process (six-plus months, on average, for each new source). Purpose-built for Amazon Web Services analytics environments, the NorthBay Data Quality Engine streamlines the data ingestion process by:

- Automatically auditing existing datasets
- Implementing custom rules for data consistency
- Detecting anomalies in real-time
- Identifying relationships between different datasets
- Ensuring easy integration with AWS analytics

Data Quality in Financial Services – Why We Need a Fresh Approach

Financial services organizations want to leverage data analytics and machine learning algorithms such as generative AI to acquire and retain customers, increase customer lifetime value, and develop new products in order to maintain their competitive edge. But legacy data warehouses need to be modernized and moved to the cloud in order to make it happen. What's more, legitimate concerns about data quality often paralyze IT leaders from getting started or moving forward.

Rather than trying to integrate offerings into legacy core banking or insurance systems, adopting cloud-native solutions allows financial services organizations to scale core operations, and opens up access to a vast ecosystem of valuable services to drive meaningful innovation.

AWS is a pioneer at the intersection of financial services and technology, and thousands of financial services organizations — from established enterprises to fast-growing fintechs — are redefining their future on AWS.

The Data Quality Dilemma

Integrating additional data sources into legacy data warehouses is a complex and time-consuming process. IT teams and business stakeholders must manually correct, complete, reformat, and clean the data every time (think about the mistakes made by sales reps using customer relationship management systems, or consumers filling out web forms), which often takes six or more months per data source.

And, while most IT teams are well aware that input errors and omissions from users are perpetual problems, a lack of data governance standards, users who don't adhere to standards, and mismatches in data structure from multiple sources amplify data quality challenges. The result? Inaccurate, incomplete, and unstructured data that can introduce bias, hinder machine learning model accuracy, and impede progress toward the desired data-driven insights.

aws PARTNER Premier Tier Services

NorthBay Solutions Data Quality Engine — northbaysolutions.com

---

NorthBay

How Data Quality Engines Help

Even as financial services organizations embrace AWS for agility and scalability, data quality has still remained a critical concern. Enter data quality engines.

Data quality engines clean and structure data, identify errors and help correct them, streamline input checks, apply transformations, and ensure high-quality output. Incorporating a data quality engine into a centralized data platform on AWS ensures rapid and precise data ingestion and enables significantly faster access to data-driven insights and decisions. An effective data quality engine provides the following benefits:

Early issue detection

Automated checks find and help correct issues during the ingestion process by scanning, identifying, correcting, and flagging errors. This proactive approach performs some of the heavy lifting of analyzing a new data source and preparing it for inclusion in the analytics platform by reducing data quality problems early on.

Efficient data standardization

With built-in transformation functions, the data quality engine standardizes data more quickly and aligns the data more closely with desired governance. It ensures consistency across your ingested datasets and makes it easier to analyze and visualize meaningful results.

Streamlined data processing

The data quality engine eliminates the need to directly load source data into your analysis/visualization tool, avoiding downstream corrections, de-duplications, or other data processing tasks. This saves your team massive time and effort.

Find and Fix Data Quality Issues with the NorthBay Data Quality Engine

NorthBay Data Platform Engines and Accelerators are processes and best practices purpose-built to help financial services organizations speed up time-to-value while minimizing risk. Our Data Quality Engine automates a multitude of validation rules (based on business logic) to help reduce errors and streamline the integration of additional data sources. It's an extremely valuable accelerator for any financial services organization on the path toward data platform modernization.

The NorthBay Data Quality Engine provides a range of features, including:

- Tracking and notifying schema evolution, checking and validating the number of columns and their data types, and checking the business logic of dimension table columns.
- Calculating data quality metrics such as size, completeness, compliance, mean, and uniqueness with rules configurations that are based on the latest trends in data lake technology, and enables no-code configuration of new rules.
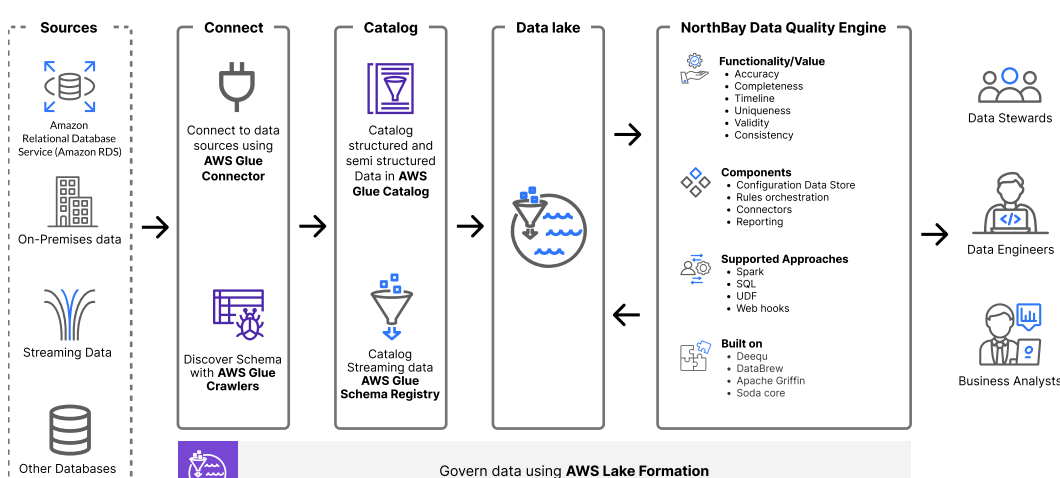
Data quality checks

Data quality rules replace manual scripts to ensure accuracy, consistency, completeness, and validity. New rules can be configured without extensive and time-consuming coding by IT.

A wide-range of data quality checks help organizations to quickly find and fix data quality issues before they impact business operations. The engine also allows businesses to perform data quality checks on new datasets quickly and easily, reducing the risk of data quality issues arising.

Data cleansing and transformation

Comprehensive cleansing and transformation tools and processes ensure proper data preparation.

- Cleansing tools include capitalizing and lowercasing text, formatting dates and phone numbers, adding suffixes, and removing characters.
- Transformation operations include changing data types, removing columns, duplicating columns, and renaming columns.

NorthBay Solutions Data Quality Engine — northbaysolutions.com

---

NorthBay

NorthBay Data Quality Engine Example

Key capabilities of the NorthBay Data Quality Engine

Automated Data Profiling

Automating data quality checks is essential for quickly identifying any issues with the data and taking corrective action. Automated data profiling captures detailed information about your data sources, allowing you to catch possible errors before they become serious problems.

Business Rules-Based Filtering and Monitoring

A rules-based filtering solution continuously monitors your data for accuracy and quality, both in real-time and in batch processing. This ensures that only accurate data is used in your analysis, and it automatically flags any suspicious data or outliers in the system.

Data quality Data Cleansing and Checks Standardization

In addition to identifying errors and detecting outliers, an effective data cleansing tool ensures that all incoming data is formatted consistently. This makes it easier to query, analyze, and visualize the results accurately— allowing more timely decisions based on reliable facts.

Automated Transformation

The engine should automate the transformation of incoming data based on predefined rules to cleanse and reformat it before loading it into the warehouse.

AI-Powered Data Analysis

A data quality engine can use artificial intelligence to analyze datasets for outliers and inconsistent data, reporting potential issues and allowing for proactive corrections to maintain data quality.

Learn More About Data Platform Engines and Accelerators

For more information on it and NorthBay's other Data Platform Accelerators, visit nortbaysolutions.com/financialservices or check out the 3-minute Data Platform Accelerators video.

About NorthBay Solutions

NorthBay Solutions is a U.S.-based organization focused on delivering technology driven business solutions. With over a decade of successfully delivering solutions around AI/ML, big data and data lakes-based analytics, DevOps, applications modernization and database migrations, NorthBay is geared to help customers solve their business challenges.

NorthBay Solutions Data Quality Engine — northbaysolutions.com