

Naveego Complete Data Accuracy Platform and Cloudera Data Flow Platform



“Naveego’s Complete Data Accuracy Platform™ harnesses the power of CDF® to rapidly scale and deploy data quality and master data management to make Big Data accurate and usable.”

Katie Horvath

CEO

Naveego, Inc.

Big Data mandates the need to automate data quality and scale data management to quickly transform data lakes into actionable business information.

Naveego’s Complete Data Accuracy Platform™ is used to fix and monitor data health with data quality checks and to keep data consistent from multiple disparate data sources across an enterprise with master data management (MDM). Naveego empowers organizations to proactively transform data into viable, actionable information assets across enterprise systems. The platform delivers exceptional data quality in the Hadoop world, without the gross complexity and high cost of customized installation and ongoing maintenance that go hand-in-hand with legacy solutions. Using API connections, Naveego’s solution nimbly connects with any type of data where the data resides, eliminating the need to make infrastructure changes.

By leveraging CDF technologies including Apache Kafka and Apache Spark, the Naveego platform is able to handle high volumes of data from any source, in real-time, regardless of its structure or schema. CDF is software designed to solve dataflow challenges. It provides scalability and real-time stream processing of data at high volume and high scale, ingestion from IoT devices, edge applications and streaming sources. Working together, Naveego and Cloudera provide data accuracy at high volume with real-time streaming from any data source.

Naveego brings data accuracy into the Hadoop eco-system by adding master data management and data quality to Cloudera data assurance management technologies.

Data Quality

Naveego adds a data quality solution to CDF by performing data quality checks inside of a customer security boundary to assess the data where it sits. Naveego provides data validity putting business rules first. Quality checks are based upon natural language that define accurate data. Quality checks include resolution instructions, empowering the business, who best knows the data content, to correct exceptions. Naveego provides quantitative data health assessments that may be tracked over time to view global data health trends for all data sources of an enterprise in one portal.

Master Data Management (MDM)

Naveego manages data relationships using master data management, including distributed processing of record linkage using matching and merging the data into a golden record that may be bidirectionally synced back to the data sources. The MDM system has been designed to handle probabilistic matching and merging at a massive scale across distributed environments. Leveraging Apache Kafka and Apache Spark enables the platform to manage today’s massive new load requirements.

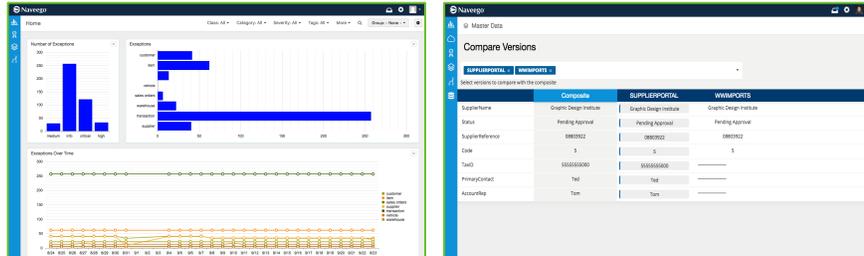
Data Governance

Naveego tracks, logs and normalizes data across CDF distributed systems. The platform’s Data Flow Log tracks all data over distributed systems with multiple processors, and deliver invaluable end-to-end visibility to users. This capability, unique to Naveego, identifies duplicate records, cleanses data and performs exact matching, in addition to tracking the lifecycle of all data across time and who or what changed it.

For integration projects, Naveego preserves and reuses system mappings enabling customers to nimbly adopt new systems and applications by freeing data from any one application.

NAVEEGO COMPLETE DATA ACCURACY

Connect to any data source in minutes using Naveego's custom NiFi processors. Add real-time data profiling, issue tracking and validation, workflow optimization, and quantitative data health analysis to the power of CDF.



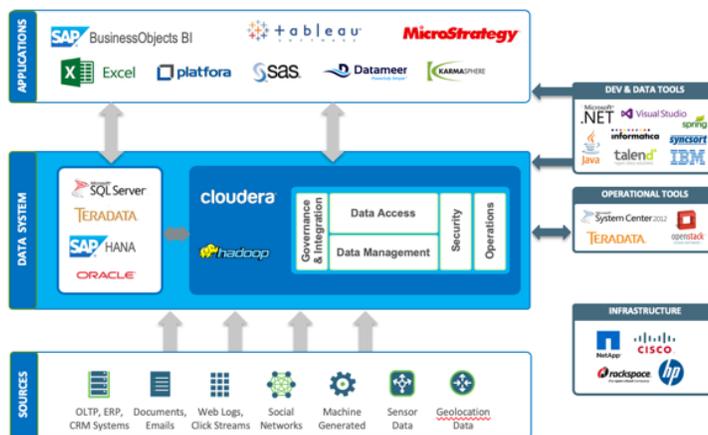
NAVEEGO INTEGRATION WITH CDF

Big Data and enterprise investment in data collection has created massive amounts of data that needs to be made accessible for use.

Working together, Naveego and Cloudera improve management, flow and usability of data, effectively transforming data into actionable business information ready for analytics and reporting.

Auto-Provisioning and Self-Service: The Naveego platform is fully self-service capable, allowing new partners and customers to add new tenants to the system. While Naveego's support staff is available to assist, users can perform all operations without customization, and achieve MDM implementation in hours.

Naveego in the Modern Data Architecture



MODERN ARCHITECTURE WITH NAVEEGO COMPLETE DATA ACCURACY

NAVEEGO HAS EARNED THE CLOUDERA CDF CERTIFICATION BADGE

Cloudera Certified Technologies have been fully tested and validated to comply with Cloudera's integration guidelines

