

PRODUCT SHEET

Trillium DQ for Big Data

Maximize the business value of Big Data and Cloud with the power of highly-scalable data profiling and data quality

Overview

Organizations are gathering larger volumes, and greater variety, of data to achieve more insights and make better data-driven business decisions. Yet surveys consistently show executives lack trust in their data. To have confidence in decision-making, regulatory compliance and more, enterprises require data quality tools that can handle these growing and complex data sets.

Trillium DQ for Big Data provides industry-leading data profiling and data quality at scale, designed specifically to meet the challenges presented by today's data environments, so you can drive successful data governance, advanced analytics, AI, machine learning, and focused business insights.

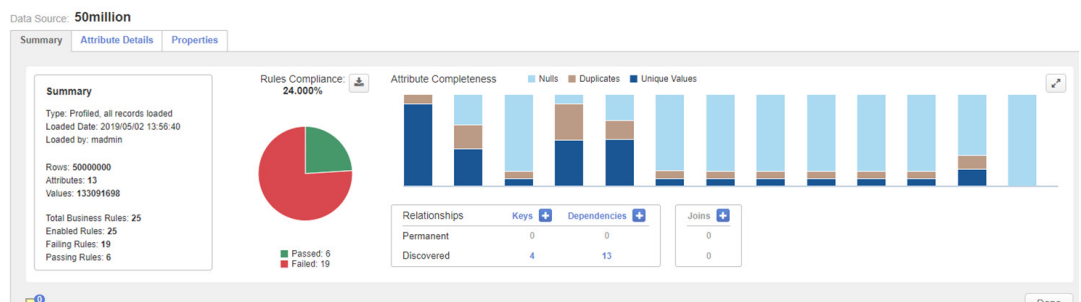
Trillium DQ for Big Data empowers your users to focus on understanding and addressing critical data quality issues and requirements. Quickly and natively connect to Big Data sources to execute data profiling tasks, and visually create and test data quality processes which can be deployed and run directly within Big Data execution frameworks on premises or in leading cloud platforms.

Unlike other enterprise data quality tools, Trillium DQ for Big Data automatically manages the technical aspects of executing data profiling and data quality jobs at run time, including dynamic performance optimization based on available system resources and your chosen compute framework. You can design once and deploy anywhere, with no tuning or re-coding, even if you change frameworks.

With native high-performance data profiling, you can:

Visually assess the quality of the data sources in your data lake and assess their completeness, consistency, validity, and accuracy. Trillium DQ for Big Data includes:

- Time-tested, robust data profiling capabilities built from decades of industry-leading data quality expertise. Users select, connect, and run data profiling against Big Data sources in a few easy steps, with no user coding skills or Big Data expertise required.
- A business user interface to broadly explore your data with rich out-of-the-box functionality to uncover data defects or outliers, and evaluate data relationships across sources, drill down to any detail, and annotate your findings.
- Powerful yet straightforward business rules to focus in on key validation criteria, whether simple, complex, or compound conditions for measuring data quality, and evaluate data sources according to thresholds you set.
- Native connectivity to Big Data sources, Intelligent Execution for highly scalable processing, and extensible storage so you can scale to the data volumes you need to address.



Using industry-leading data quality processes, you can:

Visually build and locally test data quality projects using Trillium DQ for Big Data's industry-leading advanced data quality capabilities, and deploy to your fast-growing Big Data environment, including:

- Entity resolution to link any data entity – customers, suppliers, products, etc. – into a trusted single view to support a broad array of business-critical use cases (e.g. Customer 360, fraud detection, AML).
- Multi-domain deduplication and matching with the most comprehensive set of match comparisons available, including fuzzy matching, distance comparisons, and more.
- Parsing and standardization for complex multi-domain data, extended with enrichment and verification of critical address and geolocation data – all leveraging out-of-the-box templates.
- “Design once, deploy anywhere” approach that lets you build data quality business logic while the product automatically and dynamically handles the technical aspects of framework execution – with no coding or tuning required.
- Outcomes from data quality process flows enabling you to adjust quality and integration processes as needed, for continuous data quality improvement.
- Optimized native processing to leverage the high-performance compute power of distributed Big Data frameworks including Hadoop MapReduce and Spark to process high volumes within targeted time windows to meet critical Service Level Agreements (SLA's).

Key Benefits

- Maximize the value and time-to-ROI of your Big Data investments. Get your new machine learning and advanced analytics initiatives off to a fast, successful start with assessment of overall data quality and delivery of reliable, complete, fit-for-purpose data, trusted by data scientists, analysts and business users alike while removing technical barriers.
- Understand and govern the content of data within your data lake whether on-premises or in the cloud. Address the need to gain insight into the quality of data intended for better business decisions while improving compliance and governance over that data.
- Build a true 360-degree view of your global customers, products, assets and more, using Trillium DQ's industry-leading data matching, standardization and other advanced cleansing functionality. With comprehensive, unified records in place, your data lake becomes a trusted source for a true 360-degree view of any data entity and new, breakthrough business insights.
- Optimize operational efficiency. Perform data quality faster to help meet critical Service Level Agreements. Minimize time spent on downstream data remediation efforts caused by incomplete, inaccurate, or “dirty” information in the data lake.

