

Big Data is meaningless to your organization if it is inaccurate, incomplete and inaccessible. You can't make informed business decisions if your Big Data analytics tools don't have access to reliable data. You won't gain valuable customer insights if you're unable to match data points from internal and external data sources to single customer records. Overall, you won't maximize the return on investment in your Big Data initiatives if the information underlying them is not accurate.

### Power Big Data with the Right Data

To power your Big Data initiatives with valuable business information, you need an innovative data quality solution that can manage the full scope of your organization's data today and in the future.

Trillium Big Data empowers you to rapidly derive business value from high-volume data sets and uncover new business insights by deploying a scalable, multi-domain data quality solution across global data in your Hadoop environment in 90 days or less. This solution runs natively in Hadoop to ensure all of your business information is integrated, fit-for-purpose and accessible across the enterprise. Trillium Big Data has at its foundation the innovation and expertise that has led the data quality market for over 20 years.

Using Trillium Big Data, you can:

- Profile multi-domain data to provide better understanding of data in Hadoop
- Easily build and deploy data quality projects using Trillium's graphical interface, without the need for MapReduce programming
- Deploy data quality workflows as native, parallel MapReduce functions for optimal efficiency
- Integrate, parse, standardize, cleanse and match new and legacy customer and business data from multiple disparate sources
- Analyze data quality workflow metrics, and adjust workflow rules and processes to meet business requirements

### Key Benefits

#### Uncover new data insights

Profile and analyze multi-domain data in Hadoop to understand the full breadth of your business information and identify the value of previously inaccessible data sets

#### Build the best view of your global customer base

Create more comprehensive single customer views to better detect fraud, create personalized customer experiences and deploy targeted marketing campaigns

#### Maximize the value of your Big Data investments

Power machine learning initiatives and analytics platforms with reliable, fit-for-purpose data that supports timely, accurate business decisions

#### Optimize operational efficiency

Minimize time spent on downstream data remediation efforts and maintain high-performance processing as volumes and variety of data increase

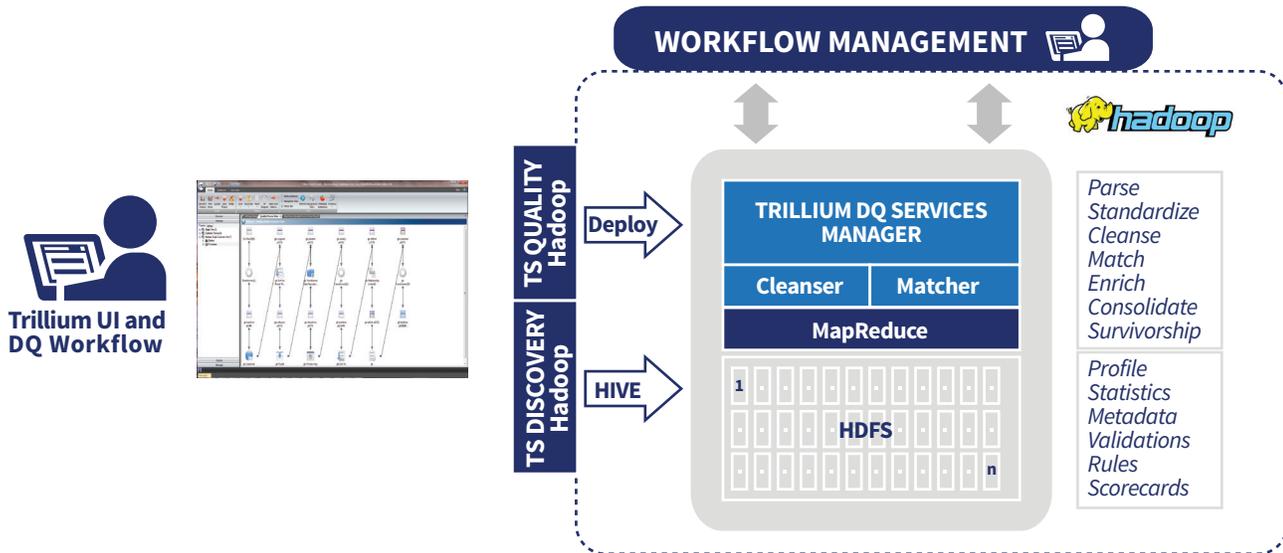
#### Seamlessly extend existing investments in Trillium Software

Repurpose existing data quality processes, workflows and business rules to maintain enterprise data quality standards



## Trillium Big Data Solution Overview

Trillium Big Data encompasses a full suite of data quality capabilities, architected to meet the high-performance and scalability demands of Hadoop. Organizations can apply traditional data quality principles – including profiling, cleansing and matching – to their growing Hadoop environment to monitor the condition and improve the value of Big Data.



Trillium Big Data deploys data quality workflows as native MapReduce functions that run in parallel across all nodes in your Hadoop cluster, ensuring maximum speed and operational efficiency.

## Easily Create Hadoop Data Quality Projects

Trillium customers can graphically build workflows and test and modify data quality processes prior to deployment into an operational environment. Prebuilt workflows are available to configure and customize to meet specific business requirements. Trillium customers can easily extend existing data quality workflows to Hadoop, ensuring consistency in data standards across the enterprise.

## Rapidly Improve the Quality of Big Data

Once a data quality workflow is ready for operational use, it is easily invoked by Trillium customers using standard Hadoop job management tools such as Hue or Oozie, ensuring consistency with existing operational procedures. Trillium Big Data configures the workflows as MapReduce steps that run in parallel across all nodes in your Hadoop cluster, ensuring maximum speed and operational efficiency.

Trillium Big Data's cleansing and matching capabilities automatically:

- Parse data values to their correct fields
- Verify global addresses, and match and enrich postal data, using global postal reference sources
- Match like records and eliminate duplicates
- Match data from external, third-party sources to internal records to create comprehensive single "golden" records
- Identify records that belong to the same domain (i.e., household or business)

## Analyze Your Results

Trillium Big Data enables data analysts and stewards to profile and analyze Big Data prior to and following data quality processes so they can clearly understand the condition of data in Hadoop. Users can analyze data sets down to individual data rows to examine specific data anomalies. After viewing the results of a data quality process, users can determine whether changes to data quality rules are needed to ensure cleansed data consistently meets business requirements.

### Corporate Headquarters

978-436-8900  
trinfo@trilliumsoftware.com  
[www.trilliumsoftware.com](http://www.trilliumsoftware.com)

### EMEA

trillium.uk@trilliumsoftware.com  
+44-0-118-940-7600

### Central Europe

+49-7031-714-756  
trillium\_DE@trilliumsoftware.com

**TRILLIUM  
SOFTWARE**

A Harte Hanks Company