Enterprises are collecting vast amounts of big data and the volume of collected data continues to increase. Before insights can be extracted from this data, an important interim step of data cleaning and manipulation needs to take place. This step can often be time and resource consuming. Some estimate that such data cleansing and prep tasks constitute 50-80% of the development time and cost in data warehousing and analytics projects. Because a lot of time is spent on low-level data cleansing, actionable insights are not quickly extracted.

Trifacta offers a solution that has been certified with the Hortonworks Data Platform (HDP) to simplify the data cleansing and formatting so that business users are able to quickly extract information that they need. Trifacta has a unique Predictive Interactive™ Solution that allows users to create concise data transformation scripts that compile natively to Hadoop and are designed to scale.

Predictive Interaction™ tools replace tedious manual coding of data transformation scripts with a lightweight interaction for the user, grounded by concrete data transformation scripts generated by the software. This new interactive technology transforms low-level programming tasks into high-level visual interaction.

What Business Challenges are Solved?

Increased Productivity

IT Teams and analysts can build and manage data products and transformation scripts more effectively. The time for data cleanup and manipulation is considerably reduced resulting in faster access to relevant insights.

Improved Data Usability

Accelerating the time to clean and manipulate data allows domain experts to work with more data and to explore deeper relationships between data sets.

Streamlined Business Process

The goal of data analysis is not simple insight but rather an improved business process. Successful analytics lead to product or operational changes that drive value for the organization.

www.hortonworks.com
Trifacta in the Modern Data Architecture

Trifacta provides solutions that enable data transformation and streamline development and operations.

Features and Benefits of the Combined Solution

- Certified solution with HDP
- Native scripts that can be run on Hadoop to format data so that analytics can be done quickly
- No tedious hand-coding of scripts
- Predictive Interaction™ allows users to visualize, predict and interact directly with their data

Trifacta is a Certified Technology Partner

The Hortonworks Certified Technology Program reviews and certifies technologies for architectural best practices, validated against a comprehensive suite of integration test cases, benchmarked for scale under varied workloads and comprehensively documented.

Hortonworks. We do Hadoop.

Hortonworks is a leading commercial vendor of Apache Hadoop, the open source platform for storing, managing and analyzing Big Data. Hortonworks Data Platform, our distribution of Apache Hadoop, provides an open and stable foundation for enterprises and a growing ecosystem to build and deploy Big Data solutions.

Hortonworks is the trusted source for information on Hadoop, and together with the Apache community, Hortonworks is making Hadoop an enterprise data platform. Hortonworks provides unmatched technical support, training and certification programs for enterprises, systems integrators and technology vendors.

For additional questions, contact:

- Trifacta  
  www.trifacta.com  
  (415) 429-7570  
- Hortonworks  
  www.hortonworks.com  
  (855) 8-HORTON