



ClearStory Data and Hadoop

Fast, Scalable Analysis on Diverse Data



Partner Brief



“ClearStory Data places big data analytics in the hands of users on the front lines of business. Users of any skillset can easily tap into their Hadoop data store using ClearStory’s fast data access model, create Interactive Data Stories on-the-fly, speed exploration of data and insights and uncover insights from Hadoop and Hadoop-based Data Lakes. “

Vaibhav Nivargi
Technical Co-founder
ClearStory Data

Today, competitiveness on the front lines of business is dictated by the speed of data access, increased data diversity, data exploration capabilities in the hands of users and and timeliness of decision-making.

Hortonworks offers a scalable, fault-tolerant, and cost-effective platform for the volume of data that drive an enterprise’s big data analytics needs. ClearStory Data, together with the Hortonworks Data Platform (HDP), enables fast-cycle analysis across disparate data stored in Hadoop. ClearStory’s solution enables fast blending of derived data from Hadoop and supports direct data connections to other internal data platforms and external sources. Users do not require any specialized skillset, but rather can quickly access data from HDP, reduce the lengthy process of “data wrangling” and pre-modeling via ClearStory’s data inference and harmonization capabilities and reach visual insights immediately. The insights generated via ClearStory allows for rapid data exploration and fast diagnostic and discovery-based analysis.

ClearStory’s underlying data processing is built on *Apache Spark* for scalable, fast analysis. ClearStory’s integrated User Application is designed for data stewards, data producers, and everyday data consumers. ClearStory’s solution also includes all the necessary ‘data governance’ and ‘user governance’ capabilities required for data stored and accessed from HDP. The solution also lets users collaborate in real-time on the data workflow and on interactive insights to drive conclusions and enable fast, data-driven decisions.

Fast-Cycle Analysis Across Diverse Data, Fast Discovery

Easily connect to Hadoop, and other internal or external data sources. ClearStory infers what’s in your data to speed data preparation and converge disparate data on the fly. It’s a new, easy way to speed access to more sources, and reduce the dependency on costly IT cycles.

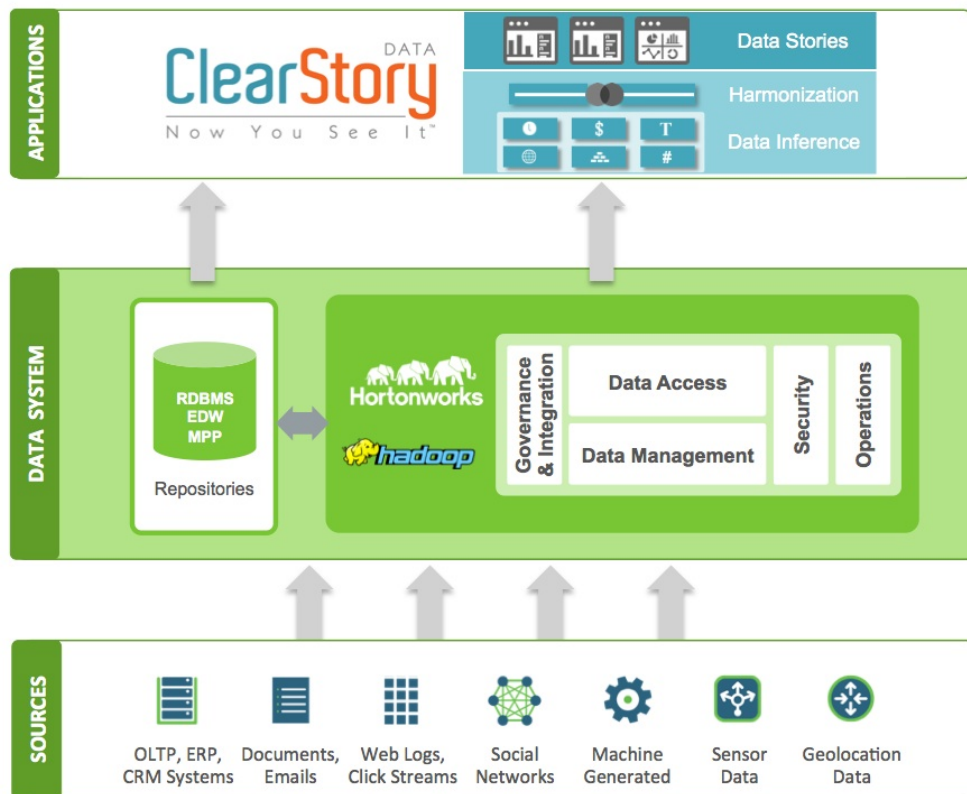
Making Big Data Work for the Enterprise and for ‘Data Lake’ Initiatives

Enable line-of-business users to securely access data sets in a Hadoop environment and easily uncover visual insights, with the user and data governance that is crucial to the success of your Big Data initiative.

Continuous, Diagnostic and Exploratory Analysis

Accelerate analysis and reach holistic, interactive answers faster. Arrive at deeper insights via fast intra-day, daily, and weekly analysis across more data.

ClearStory Data in the Modern Data Architecture



Fast Holistic Insights on HDP and other direct data connections.

- *Intelligently match data with common dimensions*, making it easy to blend disparate data stored in Hadoop. Blend in, on the fly, other data sources if required for immediate insight and *eliminate tedious data prep or data modeling*.
- *Track data lineage* and data updates through ClearStory's integrated *metadata management* model.
- Leverage ClearStory's *User Governance and Data Governance* capabilities to govern how data is accessed, who sees it, and maintain an audit trail for every analysis and user.
- Let line-of-business 'data consumers' *deeply explore* data in Hadoop, *safely iterate* on insights, and *collaborate in-context* insights

See It Today

ClearStory Data connects directly to Hadoop and other internal and external data sources. Register for a demo at www.clearstorydata.com/demo and experience the fastest way to analyze diverse data and see blended insights. See how ClearStory eases the data consumption and insights for data stewards, data producers and everyday data consumers.

ClearStory Data 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:

- ClearStory Data
www.clearstorydata.com
(650) 322-2408
- Hortonworks
www.hortonworks.com
(855) 8-HORTON