

REPORT REPRINT

Informatica highlights its depth and breadth with Big Data Management update

MATT ASLETT

21 MAR 2016

Enterprises now face a disparate array of emerging data management vendors. With its 10.1 update, Informatica is focusing on the benefits of running a single, cohesive Big Data Management suite.

THIS REPORT, LICENSED EXCLUSIVELY TO INFORMATICA, DEVELOPED AND AS PROVIDED BY 451 RESEARCH, LLC, SHALL BE OWNED IN ITS ENTIRETY BY 451 RESEARCH, LLC. THIS REPORT IS SOLELY INTENDED FOR USE BY THE RECIPIENT AND MAY NOT BE REPRODUCED OR REPOSTED, IN WHOLE OR IN PART, BY THE RECIPIENT, WITHOUT EXPRESS PERMISSION FROM 451 RESEARCH.



©2016 451 Research, LLC | WWW.451RESEARCH.COM

Few vendors in the data management space can match the depth and breadth of Informatica. While the company faces growing competition from an array of emerging startups, with version 10.1 of its Big Data Management suite, Informatica is focusing attention on the benefits of running a single, cohesive suite rather than integrating a disparate collection of point products. In that vein, the company has also introduced a new offering for self-service data discovery and preparation that is based on the technology behind its Rev cloud service.

THE 451 TAKE

As we've previously noted, the ability to integrate, govern and secure data in multiple platforms will be critical to turning the concept of the 'data lake' (into which raw data flows from multiple sources and is available as trusted data assets to multiple users for multiple purposes) into reality. As such, it makes sense for Informatica to highlight the potential benefits of using a suite of big-data management products from a single vendor. Many of Informatica's emerging rivals are arguably more closely associated with emerging data platforms and approaches to data integration and analysis, but the company can boast strength in depth, while the introduction of Intelligent Data Lake illustrates how it is rising to the competitive challenge.

CONTEXT

The last time we checked in with Informatica last November, it had just completed the v10 launch of its data management portfolio with the introduction of Big Data Management, a new offering positioned as a single platform for integration, quality, governance and security of big data in Apache Hadoop, and Hadoop-based 'data lakes,' as well as relational and non-relational databases. The company has now updated that platform with the launch of Informatica Big Data Management v10.1, which includes the Intelligent Data Lake, a self-service offering for data discovery and preparation (previously known as Project Sonoma). Informatica is looking to strike back against the multitude of data management startups that have emerged to take advantage of Hadoop and other emerging data platforms.

The rise of these new platforms, and the ecosystem of startups they have inspired, means that Informatica arguably faces a greater number of competitors than ever. However, these are largely niche providers (see Competition section below) that are tackling one of several aspects of the data management landscape (data ingestion, data transformation, data blending, data cleansing, data security, master data management and data governance). The key message with Big Data Management v10.1 is that rather than try to integrate a variety of tools from a variety of vendors, enterprises can look to Informatica to deliver a single, cohesive suite of offerings that share common metadata and in doing so, provide greater flexibility and autonomy.

With Informatica Big Data Management v10, the company already delivered a single platform for integration, quality, governance, metadata management and security of big data, based on Blaze, a YARN-native high-performance execution engine for complex batch processing of data in Apache Hadoop. Version 10.1 introduces a variety of improvements over v10, including expanded support for Apache Spark, real-time matching and linking, and higher-performance data quality. The company has also extended the portfolio to include on-premises self-service data preparation with the launch of Intelligent Data Lake – previously known as Project Sonoma.

Informatica already responded to the growing trend for self-service data preparation with the launch of Rev (formerly known as Project Springbok) in late 2014. Although Rev is a free, cloud-based data-preparation service, it is now teamed with IT monitoring and governance capabilities and integrated into the larger Big Data Management suite. The resulting Intelligent Data Lake is designed to offer business users the ability to search, discover, prepare and publish data in Hadoop, as well as provide IT users the ability to track and monitor usage, and manage data lineage and security.

COMPETITION

As noted above, Informatica competes with a variety of vendors across the data management landscape in areas such as data ingestion, data transformation, data blending, data cleansing, data security, master data management and data governance. In terms of comparable breadth, IBM perhaps comes closest. SAS Institute offers a broad range of data management products, as does Talend, which is steadily expanding its portfolio beyond its data-integration roots.

For data ingestion, the competition is likely to come from the likes of Talend, SnapLogic and Syncsort, with open source tools such as Sqoop, Flume and Kafka (for real-time integration) popular in Hadoop environments, and emerging startups including StreamSets and Striim. Talend and Syncsort also have offerings for data transformation and data cleansing, along with companies such as Hitachi Data Systems' Pentaho, Diyotta, SnapLogic and Harte-Hanks' Trillium Software.

Data security in Hadoop environments is a specialist area for a number of vendors including Centrify, Dataguise, Thales's Vormetric, Protegrity and Hewlett Packard Enterprise's Voltage, while the Hadoop distributors have maturing security offerings of their own. For master data management, Informatica cites Tamr as a competitor, but Talend also has offerings in this space, and Hadoop-based data-governance players include Zaloni, Waterline Data, Alation, Ataccama and Collibra.

The new Intelligent Data Lake offering will compete with data blending and preparation providers such as Trifacta and Paxata, as well as Alteryx and Zaloni. This is a particularly crowded space at the moment because IBM, Oracle, SAS Institute and Hitachi (Pentaho) have entered the market, along with discovery and visualization specialists such as Platfora and Datameer, and Talend also recently introduced its free Talend Data Preparation desktop application.

SWOT ANALYSIS

STRENGTHS

Informatica has a broad customer base and undoubted expertise in data integration, data quality and data management, making it a trusted provider for many enterprises.

WEAKNESSES

The company has trod more carefully than some of its rivals with regard to embracing Hadoop, but is now getting into its stride with Informatica Big Data Management.

OPPORTUNITIES

The breadth and depth of Big Data Management gives the company an edge over its rivals, many of which are peddling point offerings that only serve one aspect of overall management of big data.

THREATS

Informatica is competing with more nimble and agile startups that are more closely associated with emerging data platforms and approaches to data integration and analysis.