INFORMATICA IS THE DATA LAKE LIFEGUARD

NUCLEUS RESEARCH

ANALYST

Anne Moxie

THE BOTTOM LINE

As IT teams are faced with exponentially increasing volumes of data, they also have to meet the demand for data democratization. Nucleus has found that industry-specific solutions can help to reduce ongoing analytics maintenance while also shortening the payback period by 57 percent, on average. Informatica's Data Lake Management solution now provides industry-specific offerings and Nucleus expects that, for organizations that are drowning in a data lake, these new offerings will allow them to regain control.

•••

THE ANNOUNCEMENT

On September 21, Informatica announced its new out-of-the-box Data Lake Management solution, which is a centralized solution for finding, preparing, governing, and protecting data. The new solution packages together Informatica's existing suite of big data management products along with a number of new Informatica solutions including:

- Informatica Enterprise Information Catalog, which leverages metadata and automated scanners for discovery and visibility of data assets across the enterprise.

- Informatica Intelligent Streaming allows organizations to quickly capture and process data from sources such as machines, social media feeds, and website click stream using a visual development interface. In addition, it supports real-time events for IoT, marketing, and fraud detection.

- Informatica Blaze technology improves Hadoop processing performance using data pipelining, job partitioning, job recovery and scaling, supported by a cluster aware data processing engine that is integrated with YARN.

- Informatica Secure@Source provides functionality for monitoring data based on risk conditions and user access and activity.

- Informatica Big Data Management product can now be easily deployed in the cloud through the Microsoft Azure Marketplace.

- Informatica Cloud Microsoft Azure Data Lake Store Connector provides self-service data integration.

WHY IT MATTERS

Nucleus expects that Informatica's Data Lake Management solutions will provide users with a faster time to value, and will also allow them to maintain control over their growing volumes of data.

ACCELERATING TIME TO

Nucleus has found that offering industry-specific solutions in the analytics space accelerates the time to value by 57 percent on average (Nucleus Research, *p67 – Vertical solutions drive greater value in analytics,* March 2015). The significantly faster payback period is a result of prebuilt capabilities and integrations that reduce the need for custom code.

> ### Nucleus found that industry-specific solutions in the analytics space accelerate the time to value, on average, by 57 percent.

Having industry-specific solutions for data ingestion and management is especially important. This is because IT is often separated from business users and this can inadvertently cause long delays in delivering capabilities that, in theory, should be reusable. Therefore, they often struggle to decide what data they should provide, how they will organize it, how they will prepare it, and how they will ensure it is trustworthy. Informatica offers decades of experience in the space to provide IT with the tools to meet the unique requirements of their industry.

For example, the Marketing Data Lake, announced in May, allows users to find more sources of customer data in real time and enables users to quickly prepare and share the data. The new Fraud Data Lake, helps to identify transaction data in real time

and also facilitates fast preparation and distribution of the data. To support these two offerings, Informatica leverages Informatica data intelligence, which is powered by the Informatica Live Data Map technology, to automatically identify certain types of data. With these tools incorporated into the industry-specific data lakes, IT teams are able to take advantage of pre-built parsers and connectors, as well as policy and rules based processing, to quickly identify, prepare, and distribute the specific data that they need.

WITH DATA COMES RESPONSIBILITY

Faced with a growing amount of data, combined with an increasing demand for self-service data accessibility, companies need to be prepared to keep up with compliance and regulatory requirements. Informatica Enterprise Information Catalog is a key component in the Informatica Data Lake Management solution. It essentially offers users the ability to hold stewardship of their data using metadata. This metadata is a summary of the data that's available, giving the users visibility into their data assets. Informatica Secure@Source then provides the security measures necessary to alert users to any potential concerns. As a result, customers do not need to spend as many man-hours trying to manually meet compliance and regulatory requirements.

LOOKING AHEAD

Customers have struggled with the volume, variety, velocity, and veracity of data for years, which has prevented them from achieving scalable analysis. Today, customers are demanding more from analytics tools because they want to democratize data for ad-hoc use. At the same time, IT has to manage exponentially increasing volumes of structured and unstructured data from more and more sources.

Informatica is stepping in as the lifeguard for IT teams that are drowning in these growing data lakes. Nucleus expects that, based on the success we found in previous research, the new industry-specific solutions will help Informatica customers to experience a 57 percent faster time to value. Moving forward, Informatica's centralized Data Lake Management offering can continue to add value for customers by leveraging decades of experience to build additional vertically-focused solutions.