

## IBM Unveils New Data Prep Tool Designed to Help Speed DataOps - New InfoSphere Advanced Data Preparation designed to automate data prep to help streamline data operations and accelerate AI adoption

ARMONK, N.Y., June 28, 2019 /PRNewswire/ -- IBM (NYSE: [IBM](#)) today announced a new data preparation solution designed to help clients improve their dataops processes to get their data ready for AI quickly and efficiently.

Data preparation is an integral step in building machine learning and predictive models, but it's also one of the most cumbersome and time-consuming, leading many data scientists to devote up to 80 percent of their time to it.<sup>1</sup> And while the quality of the data remains a critical factor in producing accurate models – and more accurate insights – the time-intensive process can stall AI projects.

To ease this process, IBM introduced today InfoSphere Advanced Data Preparation, a new solution designed to help clients transform raw datasets by formatting, structuring and enriching the datasets for analytic processing and standard reporting. Jointly developed with data prep software provider, [Trifacta](#), the new InfoSphere solution is engineered to work in conjunction with clients' existing data environments, including data lakes.

Among its many features, the new InfoSphere solution includes an intuitive dashboard for visualizing the data prep process, including the progress of tracking data quality and lineage (where the data originated, and where it's been). With the resulting cleaned datasets, clients can move them into the business analytics tool of their choice.

InfoSphere Advanced Data Preparation resides on top of a client's data lake or data warehouse and provides automated transformation capabilities. Through the solution's self-service user interface, business users, as well as data scientists, can access, explore, prepare and enrich datasets for analytics. In addition to data prep, the tool is designed to empower users of all levels of technical expertise to generate business-ready data insights.

"When you have accurate datasets ready for AI, it serves as a launchpad for all sorts of new business capabilities," said Dumisani Mthimkhulu, Head of Data Asset Management Platforms, Standard Bank of South Africa Limited. "We can start making strategic decisions because our data is curated and it's trusted, and our data scientists can use it to build some really interesting and valuable models."

"Organizations across the board are looking to leverage data for strategic decision making. At the same time, we've seen analytics, machine learning and AI initiatives throttled by poor data quality, inefficient data preparation processes, and a lack of governance," said Adam Wilson, CEO, Trifacta. "We're excited to bring Trifacta's self-service approach to data preparation to an innovative platform like IBM Infosphere and Watson to empower a broad base of business users in IBM's ecosystem. This collaboration will empower organizations to accelerate data preparation for self-service analytics in a governed and centrally managed environment."

"The new InfoSphere solution adds to our growing stable of dataops services and capabilities that are designed to help organizations automate much of the cumbersome preparation work and get to the business of conducting data science and building AI models fast," said Daniel G. Hernandez, Vice President, IBM Data and AI.

### **About IBM Data and AI**

For more information go to <https://www.ibm.com/analytics/>.

### **Contact**

Michael Zimmerman  
IBM Media Relations  
[mrzimmerman@us.ibm.com](mailto:mrzimmerman@us.ibm.com)

<sup>1</sup> Figure Eight, Inc., Data Science Survey

SOURCE IBM

---