

IBM Unleashes the Power of Machine Learning with Watson-enabled Data Platform

Delivers world's fastest data ingestion engine and machine learning as a service to accelerate AI for business

LAS VEGAS - 25 Oct 2016: IBM (NYSE: [IBM](#)) today announced IBM Watson Data Platform to help companies gain more valuable insights from data. The platform delivers the world's fastest data ingestion engine and cognitive-powered decision-making to data professionals, allowing them to collaborate in the IBM Cloud, with the services they prefer. IBM is also making IBM Watson Machine Learning Service available – making machine learning simple with an intuitive, self-service interface.

“Machine learning is incredibly powerful, but many of today's data professionals lack the skills to fully exploit it for business and the ability to effectively collaborate on datasets,” said Bob Picciano, Senior Vice President, IBM Analytics. “Watson Data Platform applies cognitive assistance for creating machine learning models, making it far faster to get from data to insight. It also, provides one place to access machine learning services and languages, so that anyone, from an app developer to the Chief Data Officer, can collaborate seamlessly to make sense of data, ask better questions, and more effectively operationalize insight.”

Built on Apache Spark, Watson Machine Learning intelligently and automatically builds models from structured and unstructured data and open machine learning libraries, while accelerating model deployment into business operations. Its patented Cognitive Assistance for Data Science technology scores each machine-learning algorithm against the data provided to recommend the best match for the need. It also includes the most comprehensive set of algorithms in the industry.

Expanding AI's business impact by enabling collaboration

According to a recent Harvard Business Review survey, 80 percent of organizations believe the inability for teams to work together on common data is an inhibitor to an organization's ability to achieve business objectives in a timely fashion.¹ Data professionals work in silos, are trained on different languages, lack a consistent view of data, and spend too much time on data collection and cleansing. Watson Data Platform helps address all of these issues.

Watson Data Platform enables a high level of collaboration across data professionals – such as data scientists, data engineers, business analysts, and developers – by allowing them to work together with a data set, applying the languages, services and tools they want to use. Additionally, the platform enables data professionals to easily visualize and share insights across the enterprise.

Watson Data Platform capitalizes on IBM's investment in Apache Spark, The IBM Cloud, cognitive computing, and The Weather Company – and takes advantage of several technologies developed by IBM Research. Now data professionals have the ability to:

- Ingest large volumes of diverse data into the cloud at record speeds – in more than 100 gigabytes per second
- Cleanse, edit and shape data for easier modeling
- Add and remove collaborators as needed while maintaining version control
- Drag and drop services into analytic notebooks for better productivity and time management.

Through the IBM Cloud-based Watson Data Platform, companies can combine their own data along with external data leveraging built-in governance to address process, privacy and regulatory requirements, while maintaining control of their data.

The Watson Data Platform is building an expanding ecosystem of technology service providers, allowing data professionals to use the language they are comfortable with and the services they prefer. IBM has incorporated SQL, Python, R, Java, and Scala into Watson Data Platform as well as more than 20 ecosystem partners to extend the platform services, including:

- Qubole® – Enables users of the IBM Data Science Experience to process data using Spark on their choice of public cloud infrastructure;
- RStudio – enables the development of R packages and integrates existing tools for R, including Shiny and the new R interface for Apache Spark, sparklyr, to quickly bring data science workflows into production;
- Keen IO – provides a set of powerful APIs that allow data scientists to collect, analyze, and visualize events from anything connected to the internet.

"Data professionals have a variety of tools and languages that they can choose from when working with data. Whether it is writing in R or python or Scala – or using different machine learning or visualization services, they want to use what they know when working with data," Dan Vesset, Group Vice President, Analytics and Information Management, IDC. "IBM Watson Data Platform gives data professionals an open and robust ecosystem of partners that provides them with the services they know in an easily consumable platform that encourages collaboration within an organization."

Availability

Watson Data Platform brings to market "Project DataWorks" and is immediately available on the IBM Cloud through both self-service and enterprise plans that allow data professionals to use the tools they need, when they need them.

IBM Watson Machine Learning Service will begin with Apache SparkML with additional algorithms included in the future and can be accessed through Watson Data Platform, as an API on IBM Bluemix or on z/OS.

<https://uk.newsroom.ibm.com/2016-Oct-25-IBM-Unleashes-the-Power-of-Machine-Learning-with-Watson-enabled-Data-Platform>