



Talend Introduces the First Apache Beam Powered Big Data Preparation Solution

March 6, 2017

Designed to help Future-proof Companies by Bringing Portability to Big Data Preparation

REDWOOD CITY, Calif.--(BUSINESS WIRE)--Mar. 6, 2017-- [Talend \(NASDAQ: TLND\)](#), a global leader in [cloud](#) and [big data](#) integration software, today announced the first Apache Beam-powered solution for self-service, big data preparation. Now a [top-level Apache project](#), Apache Beam is a unified programming model for executing both batch and streaming data processing pipelines that are portable across a variety of runtime platforms. [Talend Data Preparation](#) is a self-service solution to enable more employees to access, cleanse and analyze large data sets. Ultimately, the combination of Talend Data Preparation and Apache Beam is designed to help companies speed the time to insight by enabling more users to build data projects that can be run anywhere using the latest processing innovation.

"Modern businesses need better access to clean actionable data, in order to support real-time insight across their organization," said Laurent Bride, chief technology officer, Talend. "However, given the current rate of technology innovation, IT leaders often worry the investments made today, too quickly become obsolete and an obstacle to advancement tomorrow. We believe Apache Beam represents the future because it mitigates the need to re-write applications as new innovations are introduced, systems are moved to the cloud, or integration styles need to be alternated. Talend's use of Beam for Data Preparation will eventually allow customers to build their preparations once and run them anywhere, which is the ultimate in data agility."

Talend Data Preparation powered by Apache Beam was first introduced back in January, as part of the Winter '17 release of Talend's integration platform and signals Talend's continued commitment to this cutting-edge data processing technology. Talend has been collaborating on the development of Apache Beam [with Google and others since 2015](#), having made several contributions to the Beam community over the last two years. Moving forward, Apache Beam will become a key element of the [Talend Data Fabric](#) integration stack.

Empowering Employees with Qualified, Trusted Data

Self-service data preparation capabilities make it easier for line-of-business users to incorporate valuable information about customers, suppliers, products and partners into their daily workflows so they can more quickly respond to evolving business requirements and emerging market needs. Gartner research says the "failure of BI leaders to embrace self-service data preparation will leave slower-to-respond companies at a competitive disadvantage through an inability to fully exploit relevant data sources."¹

Turning data into insight is a team effort. Thus, companies need to fundamentally change how they enable access to and share data across their organization to orchestrate data collaboration if they want to advance their digitalization efforts. The Winter '17 release of [Talend Data Fabric](#) empowers IT to enable business users with access to corporate data lakes so they can expedite data preparation and cleansing activities. Talend data preparation capabilities allow customers to:

- Access any data source—whether it's housed in Hadoop, the cloud or traditional databases—and share it across users and groups to encourage collaboration
- Utilize a pre-configured data dictionary to auto-recognize the meaning of the raw data from the data lake, as well as augment the dictionary with their own vocabulary, such as product codes or names
- Crowdsourcing new data definitions from open data and/or the Talend Community

To learn more about the Apache Beam-powered version of Talend Data Preparation, or the broader set of capabilities and benefits of Talend Winter '17, visit <https://www.talend.com/products/talend-6/>. For further background on Apache Beam and Talend's involvement with the project, visit the [Talend blog](#).

Like this story? Tweet this: Future-proof #BigData projects with first #DataPreparation tool powered by #ApacheBeam <http://bit.ly/1NFMQ5I>

About Talend

Talend ([NASDAQ: TLND](#)) is a next generation leader in cloud and big data integration software that helps companies become data driven by making data more accessible, improving its quality and quickly moving data where it's needed for real-time decision making. By simplifying big data through these steps, Talend enables companies to act with insight based on accurate, real-time information about their business, customers, and industry. Talend's innovative open-source solutions quickly and efficiently collect, prepare and combine data from a wide variety of sources allowing companies to optimize it for virtually any aspect of their business. Talend is headquartered in Redwood City, CA. For more information, please visit www.talend.com and follow us on Twitter: [@Talend](#).

¹ "Embrace Self-Service Data Preparation Tools for Agility, but Govern to Avoid Data Chaos," by Cindi Howson, Gartner Research, May 2016.

View source version on businesswire.com: <http://www.businesswire.com/news/home/20170306005731/en/>

Source: Talend

Talend
Chris Taylor, 408-674-1238

VP, Corp. Communications

ctaylor@talend.com

or

Siobhan Lyons, 202-431-9411

Sr. Manager, Corp. Communications

slyons@talend.com