News Release

An Industry First: Teradata Debuts Open Source Kylo™ to Quickly Build, Manage Data Pipelines

2017-03-08

Companies will benefit from simple, economical, accelerated data lake development; can focus talent on delivering high-impact business outcomes

Teradata (NYSE: **TDC**), a leading analytics solutions company, today announced a new and important contribution to the open source community that will deliver unprecedented efficiencies for companies creating data lakes. Teradata is introducing Kylo™, a data lake management software platform built using the latest open source capabilities such as Apache® Hadoop®, Apache Spark™ and Apache NiFi™. Kylo is a Teradata sponsored, open-source project that is offered under the Apache 2.0 license. Kylo evolved from code harvested from proven data lake engagements led by Think Big Analytics, a Teradata company, which will provide services and support for Kylo™.

With substantive experience creating business value from data lakes, Teradata is contributing Kylo™ to help organizations address the most common challenges they face in data lake implementation efforts. These include the central problem that data lakes simply take too long to build, and in the average 6-12 month build cycle, users find that use cases can become out of date and less relevant to quickly evolving businesses. Second, despite the lower cost of software, engineering costs quickly mount. Finally, a data lake, once created, may fail to attract users who find it difficult to explore, and so little value is realized.

Derived and developed from data lake deployments across industries, Kylo can easily help resolve these challenges, because it integrates and simplifies pipeline development and common data management tasks, resulting in faster time to value, greater user adoption and developer productivity. With Kylo, no coding is required, and its intuitive user interface for self-service data ingest and wrangling helps accelerate the development process. Kylo also leverages reusable templates to increase productivity.

“Many organizations find that implementing big data solutions on the Hadoop stack is a complex endeavor. Big data technologies are heavily oriented to software engineering, developers and system administrators,” said Nik Rouda, senior analyst with Enterprise Strategy Group (ESG). “Our research found 28 percent of organizations still struggle to staff teams with enough BI and analytics talent, much less big data and open source solution expertise. 77 percent of those surveyed say new big data initiatives will take between seven months and three years to show significant business value. It doesn’t have to be that way. I commend Teradata for open sourcing Kylo™ – an innovative and meaningful contribution.”

Encapsulating extensive experience from over 150 data lake engagements, Kylo™ helps organizations address the most common challenges they face in data lake implementation efforts, including:

- Skill shortage for experienced software engineers and administrators.
- Learning and implementing best practices around data lake governance.
- Driving data lake adoption beyond engineers.

When these challenges are overcome, high impact business outcomes are realized. In fact, Teradata has already helped many organizations save money and create new revenue streams from data lakes:

- A semiconductor manufacturer increased the yield quality of wafers; reducing waste, saving time, boosting output and thus increasing value to the business.
- An industrial equipment manufacturer enabled new service models, service-level agreements, intervention processes, and, notably, new revenue streams.
- A world-renowned research hospital reduced patient prep times, allowing doctors to treat more patients.

“Kylo is an exciting first in open source data lake management, and perfectly represents Teradata’s vision around big data, analytics, and open source software,” said **Oliver Ratzesberger**, Executive Vice President and Chief Product Officer, Teradata. “Teradata has a rich history in the development of many open source projects, including Presto and Covalent. We know how commercial and open source should work together. So we engineer the best of both worlds, and we pioneer new approaches to open source software as part of our customer-choice strategy, improving the commercial and open source landscape for everyone.”

Teradata’s vision for the blend of commercial and open source is recognized by customers, who continue to use Teradata to unleash their potential.

“At **Discover® Financial Services**, we are focused on leveraging leading-edge technology that helps us quickly bring

products to market while providing exceptional customer service. Kylo™ has a unique framework that has the potential to accelerate development and value on new data sources that leverage Apache NiFi," said Ka Tang, Director, Enterprise Data Architecture, Discover. "Kylo™ may provide an opportunity to leverage open source innovations while allowing the opportunity to give back to the open source community."

"Open source software has an appeal to users seeking independence, cooperative learning, experimentation, and flexibility for customized deployments," said Rick Farnell, President of Think Big, a Teradata company. "Our contribution is all about helping companies build a scalable data lake foundation that can continuously evolve with their business, technology data and analytical goals. We are removing impediments to use data to solve complex business problems and encouraging analytical users to contribute to the growing Kylo™ community. Going forward, our primary focus as a company is to help our customers create business value through analytics, rather than commodity capabilities. Kylo, along with our Teradata Everywhere approach to software and services, is a great example of our innovative strategy for the future."

To this point, a major telecommunications company implemented Kylo™ after a large team of 30 data engineers spent months hand-coding data ingestion pipelines. Using Kylo, one single individual was able to ingest, cleanse, profile, and validate the same data in less than a week. Kylo™ not only improved data process efficiencies, it allowed those additional engineers to focus on multiple major business priorities.

Kylo software, documentation and tutorials are available now, via the Kylo™ project website: **www.kylo.io** - or the GitHub web site: -- **https://github.com/Teradata/kylo**.

On request, Think Big offers these optional services, if required:

- Kylo™ support
- Kylo™ implementation services
- Kylo™ training
- Kylo™ managed services

"Kylo provides tooling on top of Apache NiFi to make it faster and easier to get data into your data lake," said Scott Gnau, Chief Technology Officer, Hortonworks. "Hortonworks is pleased to announce Kylo's certification with Hortonworks DataFlow and our expanded joint support relationship for NiFi."

Teradata will play a leadership role in the governance, stewardship and community-building around open-source Kylo™.

Relevant Links

- Join the community, download Kylo today: **CLICK** to download
- Obtain Kylo services and support: The Think Big Analytics Kylo web page.
- Think Big **Expands Capabilities for Building Data Lakes with Apache Spark**
- Teradata **Acquires Big Data Partnership Consultancy**, Expands Open Source Analytics Services
- Insights and Outcomes: **How Teradata is Helping its Customers Capitalize on Data and Analytics**

About Teradata

At Teradata, we believe that people thrive when empowered with trusted information. We offer the most complete cloud analytics and data platform for AI. By delivering harmonized data and trusted AI, we enable more confident decision-making, unlock faster innovation, and drive the impactful business results organizations need most. See how at **Teradata.com**.

Media Contact
Jennifer Donahue

Teradata

Jennifer.Donahue@Teradata.com