Capturing and processing big data isn't easy. That's why Apache products such as Spark, Kafka, Hadoop, and NiFi that scale, process, and manage immense data volumes are so popular. The drawback is that they do not make data flows easy for business users to access, optimize, and analyze. Kylo, from Think Big, overcomes that challenge. It lets businesses easily configure and monitor data pipelines in and through the data lake so users have constant access to highquality data. It also enhances data profiling and discovery with extensible metadata.

The Solution That Literally Means Data Flows

The name Kylo is the Greek phonetic name for flow. Kylo solves critical problems of data pipeline control across complex big data production environments. It links together world-class engineering and data science control frameworks to solve critical analytics operational challenges. For example, a telecommunications company that was using Think Big's Velocity™ Services implemented Kylo after a large team of 30 data engineers spent months hand-coding data ingestion pipelines. Using Kylo, one individual was able to ingest, cleanse, profile, and validate the same data in less than a week. Kylo not only improved efficiencies, it allowed those engineers to focus on other business priorities.

Kylo works with all flavors of Apache open-source technologies. This gives businesses the flexibility to choose the programming languages and tools that their analysts are most comfortable and familiar using, including Aster, R, Python, and Scala.

Business Benefits of Kylo

Speed to market
Kylo can accelerate your big data efforts, helping your program stay ahead of the competition.

Growth through innovation
Using Kylo, the prioritized use cases you select will help deliver business value and new opportunities across your company.
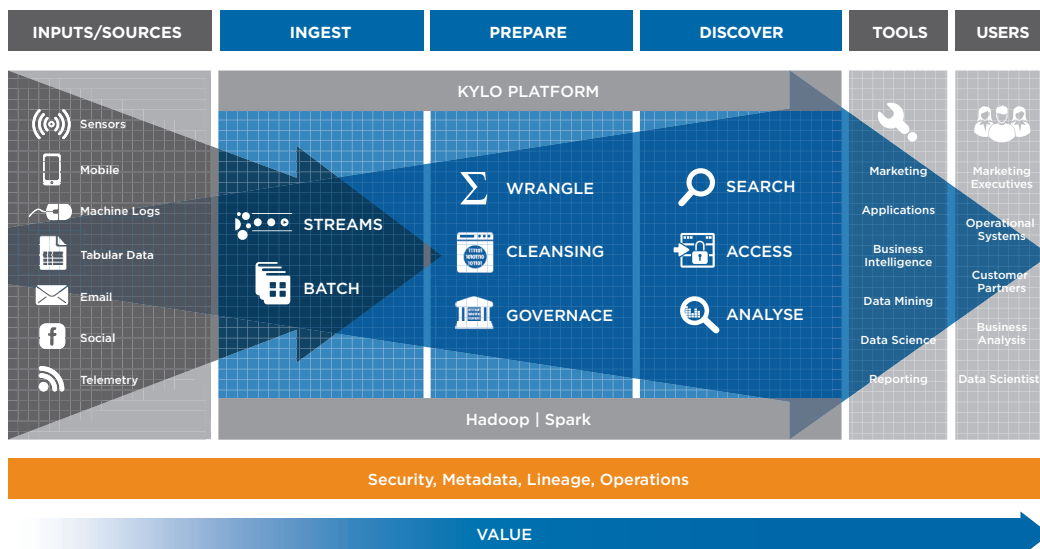
Improved quality, security and governance
Kylo can help increase the quality, security and governance of data, helping meet SLAs.

Cost reduction
Kylo can help your organization build custom engineered data lakes at a fraction of the typical cost.

Gain More Business Value From Your Data Lake

In the past, organizations had to choose between an "engineered" data lake using low-level open-source components or expensive legacy ETL tools with bolt-on support for Hadoop. Now there's a better way. Kylo makes it possible to leverage high-quality open-source technology for data lakes. By enabling orchestration, governance, and management, Kylo keeps the data lake from turning into an opaque swamp that delivers little corporate value. It turns data into an asset for the enterprise that leads to optimization of processes and better business planning and success.

Automating Processes Boosts Productivity

Kylo is not a traditional IT tool. It's for all kinds of data producers and consumers including analysts, data scientists, data stewards, and operations. Unlike other big data technologies, Kylo can be used to import and access data in a data lake. It offers self-service and data wrangling capabilities, including more than 100 transform functions and a visual SQL builder. Kylo can also automate many of the tasks associated with data lakes, such as ingesting, profiling, and monitoring data.

Typically, the process to implement a data feed and then manage, work with, and profile the data takes several weeks. With Kylo, it only takes hours. This gives analysts a huge boost in productivity and accelerates time-to-value. Kylo also lets users integrate new data sources in their data lake, perform transformations, and publish to target systems. In addition, it tracks and records lineage, providing traceability for data objects as they move through or are transformed in the data lake. Analysts can then quickly catalog, discover, and qualify data.

Take Advantage of Open Source

Kylo offers an integrated enterprise data lake platform built on open-source frameworks such as Apache Spark and NiFi. Analysts as well as developers benefit from the open-source technologies by being able to create and maintain data flows and processing in data lakes. They can also easily scale to accommodate large numbers of data feeds.

Open Source Availability

Starting in late first quarter 2017, Kylo will transition from private beta to open-source code base, available with an Apache license. Whether organizations are looking to setup and manage a data lake foundation or build out their current one, Kylo gives them the ability to create intelligent data flows and processing using open-source technologies.

Kylo framework simplifies creating and modifying data lakes with security built in. Businesses gain a turn-key data lake solution with an intuitive user interface for self-service data ingest and organization—with no coding required. This overcomes the common data lake challenge of finding people with big data skills to conduct custom development.

Operations and dashboard features offered by Kylo allow users to monitor the health of their data lake, read performance reports, and receive alerts about potential problems. These features help the business gain confidence in its data, provide the ability to observe and enforce SLAs, and make monitoring data feeds easier than ever.

Building Future and Growth with Full Lifecycle Velocity™ Services

Working in tandem with our Velocity™ Services, and using the **Kylo** as a key deliverable and on-going process within your programme, we will reduce overall risk and provide a measurable way to validate your strategic investment. Following our assessment, our full lifecycle Velocity™ approach takes your company through the process of creating a big data strategy to achieve your technical and enterprise goals.

Regardless of what stage your company is at on its big data journey, our **Kylo** and Velocity™ Services will enable you to accelerate speed to market and enhance efficiency while generating tangible returns. The battle-tested methodology enables your organisation to extract business value in the following ways.

+ **Strategy & Achitecture** - Prioritized implementation plan to maximize short and long-term returns.

+ **Data Lake Solutions** - Best practices and pre-built components to accelerate time to value.

+ **Analytic Solutions** - Building scalable big data infrastructures to meet your future needs.

+ **Data Science** - Boosting growth through innovation and modern analytics approaches.

+ **Managed Services** - 24/7 expert support empowering continuous improvement.

+ **Training** - A comprehensive selection of topical courses delivered by our certified experts.

THINK**BIG**

A TERADATA COMPANY

Teradata Velocity™ Services approach combines speed with the experience to proceed in the right direction. These services include: Strategy & Architecture, Data Lake, Analytic Solutions, Data Science, Managed Services, and Training to accelerate the time to value and increase investment returns. Teradata has also developed innovative technology frameworks through our production work with enterprise customers to speed deployment of big data solutions. With Teradata Velocity™, we deliver proven business outcomes in an expedient manner with minimal risk and maximum flexibility that supports your ongoing business strategy.

f Think Big Analytics   ▼ ThinkBigA   in ThinkBigAnalytics   ▶ Think Big Analytics