## Talend

* Talend simplifies and automates big data integration with graphical tools and wizards that generate native code. This means your team can start working with Apache Hadoop, Apache Spark, Spark Streaming and NoSQL databases today.
* Talend Big Data Integration platform delivers high-scale, in-memory fast data processing, as part of the Talend Data Fabric solution, so your enterprise can turn more data into real-time decisions.
* Blazing fast speed and scale with Spark and Hadoop
* Let anyone access and cleanse big data while governing its use
* Optimize big data performance in the cloud
* Protect your investments with a future-proof architecture
* Used for Data Governance & Meta Data management

## Trifacta

* Trifacta is a platform for preparing data for analysis. Trifacta is compliant in Hadoop using Spark or MapReduce
* Used for Data profiling & data wrangling
* Trifacta is a data preparation application that claims to enable users to transform complex data into structured formats for analysis. Their data wrangling experience is made up of six steps as outlined by Data Scientist Tye Rattenbury,[10] discovering, structuring, cleaning, enriching, validating and publishing. User can select elements to prompt transformation suggestions and use predictive transformation to preview a transform, intelligent execution for MapReduce, Spark and Javascript compilation, and collaborative data governance.[11]
* Trifacta has two products, Trifacta Wrangler and Trifacta Wrangler Enterprise. Trifacta Wrangler is a connected desktop application that allows users to transform their data for a variety of downstream analytic and visualization uses such as Tableau.[12] and Trifacta Wrangler Enterprise is claimed to improve an enterprise's big data by enabling users to turn raw, complex data into more structured formats.[13]

## Tableau

* A BI tool
* Tableau Software is a software company headquartered in Seattle, United States which produces interactive data visualization products[3] focused on business intelligence.[4] It initially began in order to commercialize research which had been conducted at Stanford University's Department of Computer Science between 1999 and 2002.[5] It was founded in Mountain View, California in January, 2003 by Chris Stolte, who specialized in visualization techniques for exploring and analyzing relational databases and data cubes.[6] The product queries relational databases, OLAP cubes, cloud databases, and spreadsheets and then generates a number of graph types.
* Tableau has a mapping functionality,[7][8] and is able to plot latitude and longitude coordinates.[9] It has been criticized for being overly US-centric.[10] They also offer custom geocoding,[11] as well as five ways to access their products: Desktop (both professional and personal editions), Server, Online (which scales to support thousands of users), Reader and Public, with the last two free to use.[12] Vizable, a consumer data visualization mobile app, was released in 2015.[13]

## Zaloni

## Zaloni, Inc. is a privately owned, software, and services company headquartered in Durham, North Carolina. Zaloni provides data management software and solutions for big data scale-out architectures, such as Apache Hadoop, Amazon S3. The company focus on management of data lakes with 2 products: Bedrock and Mica

## Pentaho

* Pentaho is a Business Intelligence (BI) software company that offers Pentaho Business Analytics, a suite of open source products which provide data integration, OLAP services, reporting, dashboarding, data mining and ETL capabilities. Pentaho was founded in 2004 by five founders.[2][3] and is headquartered in Orlando, FL, USA. Pentaho was acquired by Hitachi Data Systems in 2015

## Waterline Data

* Waterline Data is all about bringing self-service to finding, understanding, and governing Hadoop data
* Used for real-time tagging

## Arcadia

* A big data business intelligence visualization tool

## Collibra

* A Data governance platform

## Collibra

* Oracle GoldenGate 12c enables the continuous, real-time capture, routing, transformation, and delivery of transactional data across heterogeneous environments. As new or updated data is committed at the source system, it is continuously captured and applied to one or more target systems with low latency

## MarkLogic

* MarkLogic Corporation is an American software business that develops and provides an enterprise NoSQL database, also named MarkLogic. The company was founded in 2001 and is based in San Carlos, California. MarkLogic is a privately held company with over 500 employees and has offices throughout the United States, Europe, Asia, and Australia.
* MarkLogic has over 550 customers, including Aetna, BBC, Boeing, Broadridge Financial Solutions, Centers for Medicare & Medicaid Services, Condé Nast, Dow Jones, McGraw Hill Financial, NBC, Wiley, U.S. Army, U.S. Navy. Also, six of the top ten global banks..
* According to Wikibon, based on 2014 revenue, MarkLogic was at that time the leader among NoSQL databases,[1] and appeared in the Gartner Leaders Quadrant in the Magic Quadrant for Operational Database Management Systems.

## R Studio

* RStudio is a free and open-source integrated development environment (IDE) for R, a programming language for statistical computing and graphics. RStudio was founded by JJ Allaire, creator of the programming language ColdFusion. Hadley Wickham is the Chief Scientist at RStudio.