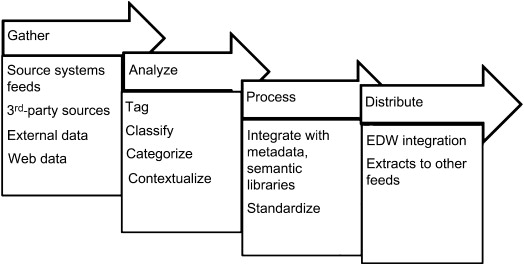
Big Data analysis with IBM Cloud Databases

Data processing:

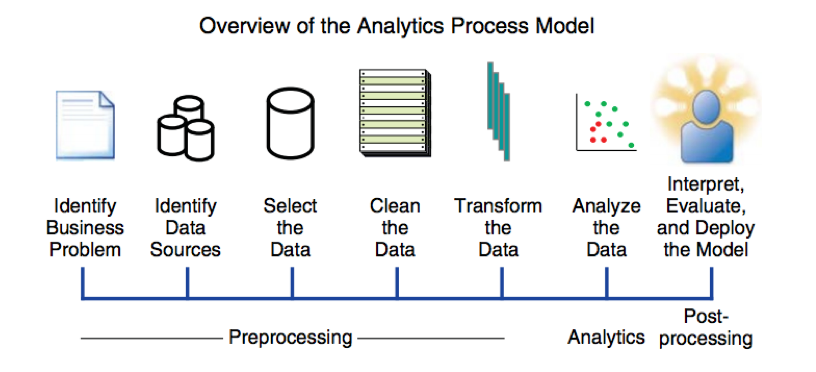
Processing of Big data

Big [Data processing](https://www.sciencedirect.com/topics/computer-science/data-analysis) involves steps very similar to processing data in the transactional or [data warehouse](https://www.sciencedirect.com/topics/computer-science/data-warehouse) environments. Figure 11.5 shows the different stages involved in the processing of Big Data; the approach to processing Big Data is:



**Six stages of data processing**

* Data collection. Collecting data is the first step in data processing.
* Data preparation. Once the data is collected, it then enters the data preparation stage.
* Data input.
* Processing.
* Data output/interpretation.
* Data storage.

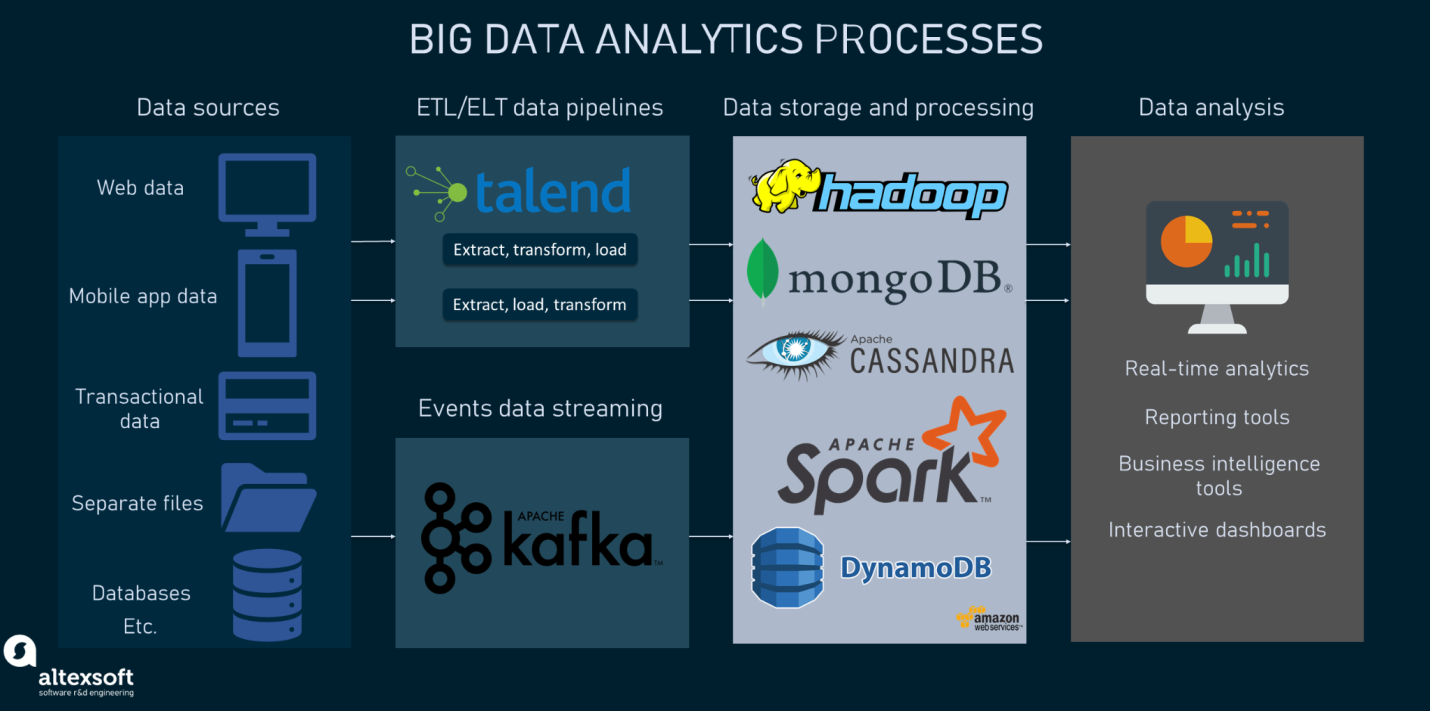


## Keys to an effective big data strategy

In an organization, [developing a big data strategy](https://www.techtarget.com/searchdatamanagement/feature/How-to-build-an-enterprise-big-data-strategy-in-4-steps) requires an understanding of business goals and the data that's currently available to use, plus an assessment of the need for additional data to help meet the objectives. The next steps to take include the following:

* prioritizing planned use cases and applications;
* identifying new systems and tools that are needed;
* creating a deployment roadmap; and
* evaluating internal skills to see if retraining or hiring are required.

To ensure that sets of big data are clean, consistent and used properly, a [data governance](https://www.techtarget.com/searchdatamanagement/definition/data-governance) program and associated [data quality management processes](https://www.techtarget.com/searchdatamanagement/feature/Data-quality-for-big-data-Why-its-a-must-and-how-to-improve-it) also must be priorities. Other [best practices for managing and analyzing big data](https://www.techtarget.com/searchbusinessanalytics/tip/6-essential-big-data-best-practices-for-businesses) include focusing on business needs for information over the available technologies and using data visualization to aid in data discovery and analysis.



Data analysis starts with identifying a problem that can be solved with data. Once you've identified this problem, you can collect, clean, process, and analyze data. The purpose of analyzing this data is to identify trends, patterns, and meaningful insights, with the ultimate goal of solving the original problem.

* Understanding the business problem.
* Analyze data requirements.
* Data understanding and collection.
* Data Preparation.
* Data visualization.
* Data analysis.
* Deployment.

Data analytics is the science of analyzing raw datasets in order to derive a conclusion regarding the information they hold. It enables us to discover patterns in the raw data and draw valuable information from them.

Data analytics provides businesses with deeper insight into their clients, helping them to customize customer experience to their needs, offer more customization, and create better relationships with them.

**CODING:**

