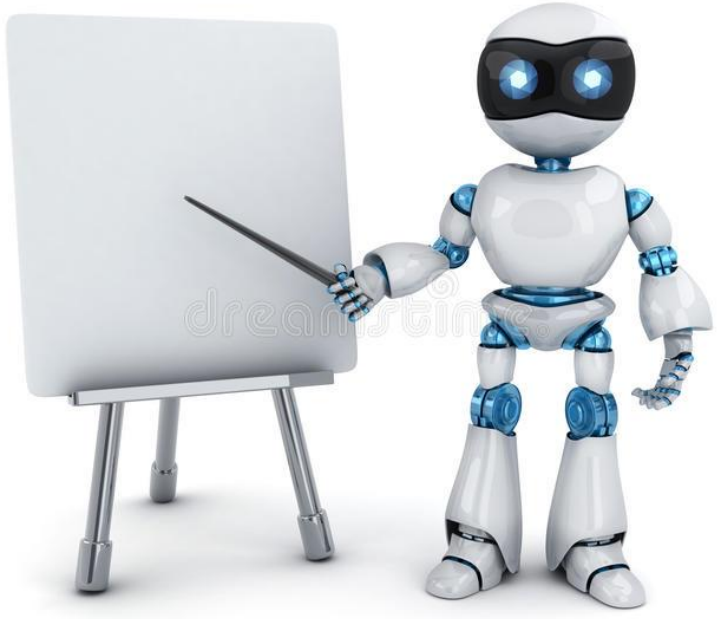


Big Data

Overview

Datamites

What is Big Data?

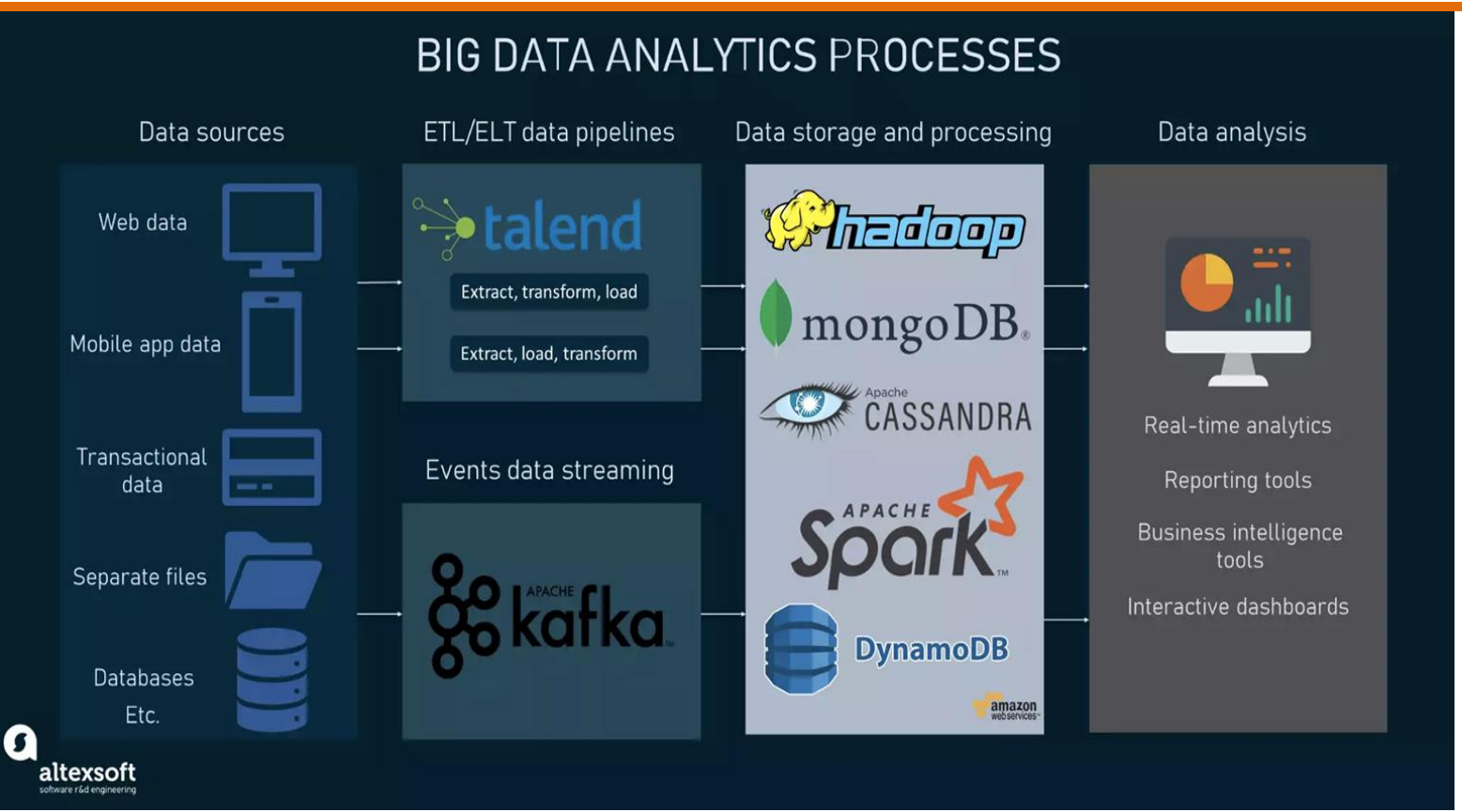
- Any data is big data when it meets below requirements.

  Volume:-Measure describing size of data.
  Velocity:-Speed at which applications are generating data.
  Variety:-How many variety of data like structured,unstructured and semi structured data.
  Veracity:-The data should be safe and trustable.
  Value:-The data should bring out some business values.

- In general term if any system not able to handle a specific amount of data,such data becomes big data for that organization.

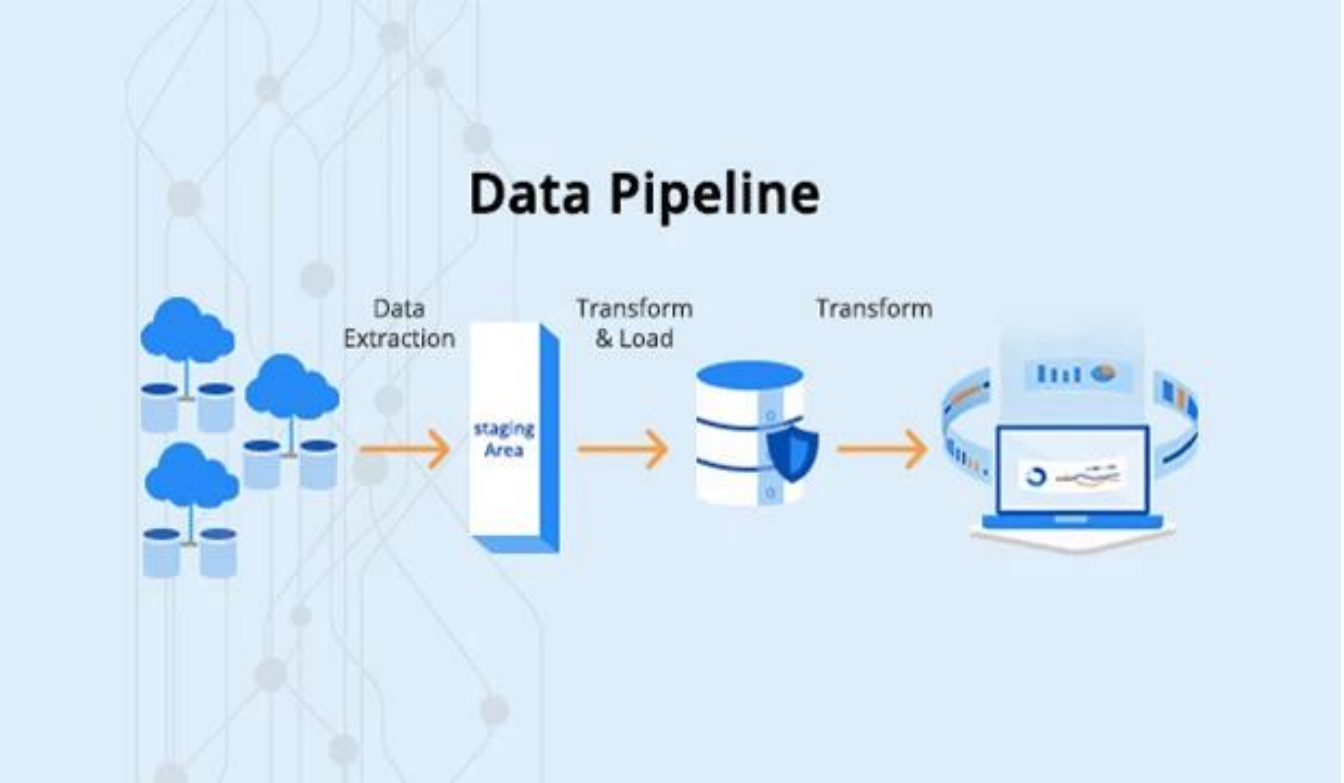
Datamites

What is Big Data?



Datamites

What is Big Data?



Datamites

What is Hadoop?

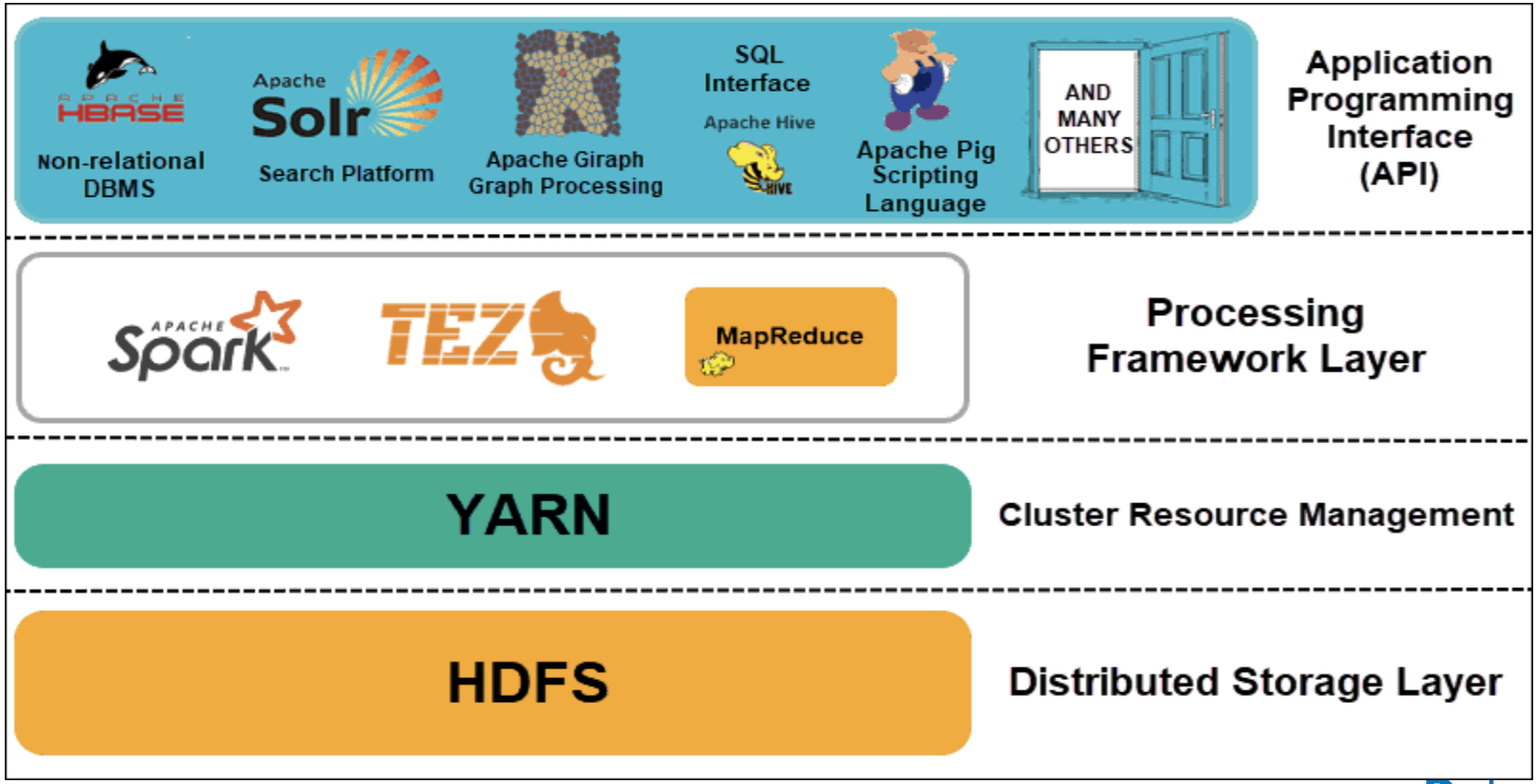
Hadoop is a framework that uses distributed storage and parallel processing to store and manage big data.

There are three components of Hadoop:

1. Hadoop HDFS - Hadoop Distributed File System (HDFS) is the storage unit.
2. Hadoop MapReduce - Hadoop MapReduce is the processing unit.
3. Hadoop YARN - Yet Another Resource Negotiator (YARN) is a resource management unit.

Datamites

Components of Hadoop



Big Data Analytics

Big Data analytics is the process of finding patterns, trends, and relationships in massive datasets that can't be discovered with traditional data management techniques and tools.

There are four key types of Big Data analytics singled out.

1. Descriptive analytics is a common kind of analytics that allows you to find out what happened and when.
2. Diagnostic analytics explains why and how something happened by identifying patterns and relationships in available data.
3. Predictive analytics uses historical data to uncover patterns and make predictions on what's likely to happen in the future.
4. Prescriptive analytics provides specific recommendations on what should be done better.

Datamites

key Processes

- Data ingestion
- Data storage and processing
- Data cleaning
- Data analysis

Real-life Big Data analytics use cases

- Healthcare: Ginger app uses Big Data for mental health monitoring.
- eCommerce: Amazon analyzes Big Data to enhance its recommender engine.
- Hospitality: Marriott makes decisions based on Big Data analytics.
- Machinery: GE uses advanced Big Data analytics to optimize a wind farm.
- Fintech: American Express leverages Big Data to detect fraudulent activity

Datamites