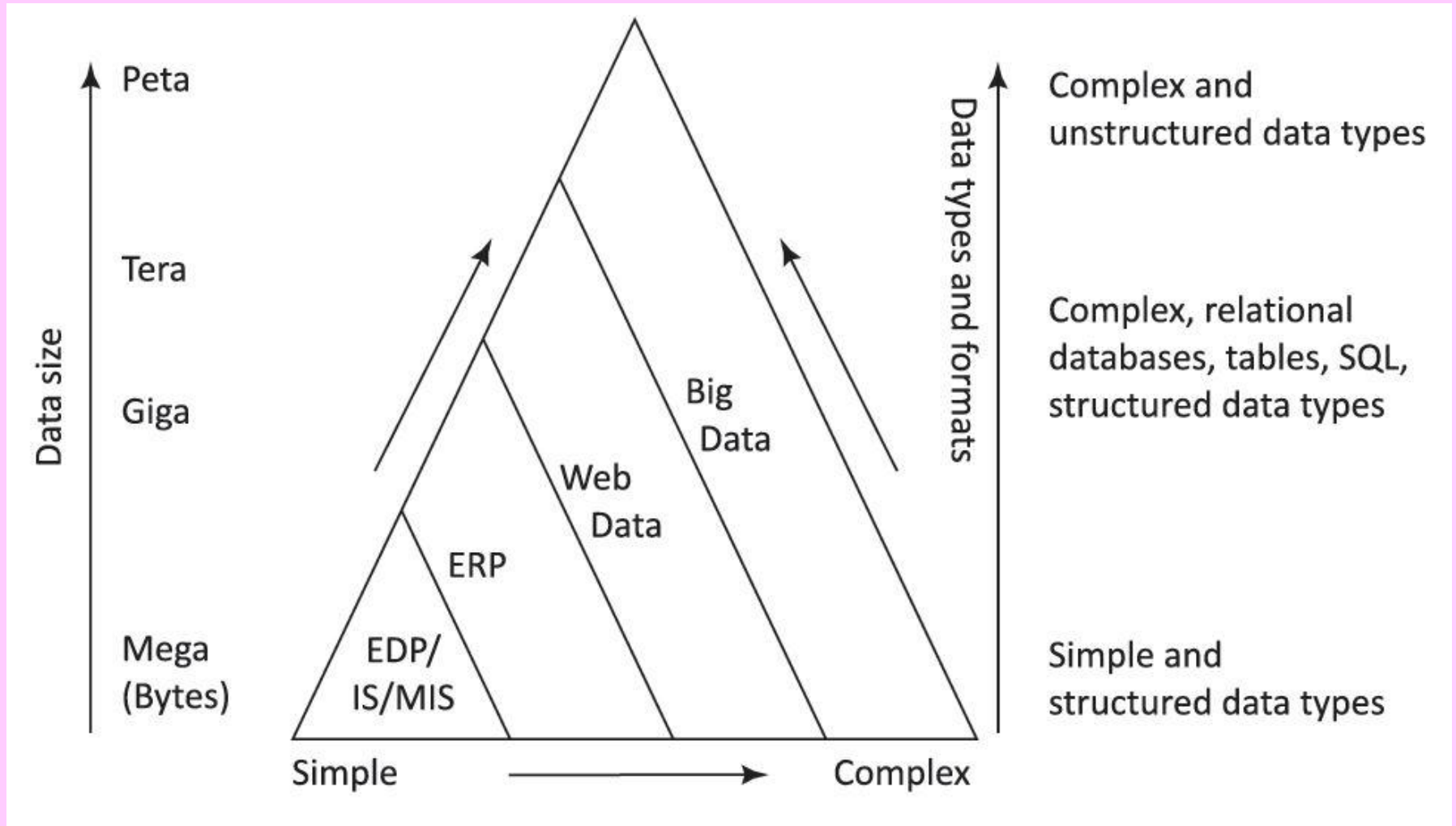


Lesson 1

Big Data Characteristics, Types and Classifications

Figure 1.1 Evolution of Big Data and their characteristics



Big Data Definition

- Big Data is high-volume, high-velocity and/or high-variety information asset that requires new forms of processing for enhanced decision making, insight discovery and process optimization (Gartner¹ 2012)
- ¹ [http:// www.gartner.com/it-glossary/big-data](http://www.gartner.com/it-glossary/big-data)

Big Data Definition

- “Data of a very large size, typically to the extent that its manipulation and management present significant logistical challenges.”
- [Oxford English Dictionary (traditional database of authoritative definitions)]

Big Data Definition

- “A collection of data sets so large or complex that traditional data processing applications are inadequate.”—
Wikipedia

Big Data Characteristics

- Industry analyst Doug Laney described the ‘3Vs’, i.e. volume, variety and/or velocity as the key “data management challenges” for enterprises
- Analytics also describe the ‘4Vs’, i.e. volume, velocity, variety and veracity as the characteristics

Big Data Volume

- Term **big** relates to size of the data and hence the characteristic
- Size defines the amount or quantity of data, which is generated from an application(s)
- The size determines the processing considerations needed for handling that data

Big Data Velocity

- Term velocity refers to the speed of generation of data
- Velocity is a measure of how fast the data generates and processes

Big Data Variety

- Term refers to a variety of data, due to the availability of a large number of heterogeneous platforms in the industry, multiple sources in a system
- Variety introduces ‘complexity’.
- Variety refers to data consisting of various forms and formats

Big Data Veracity

- Can also considered an important characteristic to take into account the quality of data captured, which can vary greatly, affecting its accurate analysis

4Vs (i.e. volume, velocity, variety and veracity) Tools

- For mining, discovering patterns, business intelligence, artificial intelligence (AI), machine learning (ML), text analytics, descriptive and predictive analytics, and the data visualization

Big Data Types

- A team from University classified Big Data Types ²

² <https://statswiki.unece.org/display/bigdata/Classification+of+Types+of+Big+Data>

Big Data Types

- Another team from IBM developed a classification of Big Data types.³

³ <https://www.ibm.com/developerworks/library/bd-archpatterns1/>

Big Data Suggested Types

1. Social networks and web data, such as Facebook, Twitter, e-mails, blogs and YouTube.
2. Transactions data and Business Processes (BPs) data, such as credit card transactions, flight bookings, etc. and public agencies data such as medical records, insurance business data etc.

Big Data Suggested Types

3. Customer master data, such as data for facial recognition and for the name, date of birth, marriage anniversary, gender, location and income category,
4. Machine-generated data, such as machine-to-machine or Internet of Things data, Computer, sensors, trackers, web logs, ...

Big Data Suggested Types

5. Human-generated data such as biometrics data, human–machine interaction data, e-mail records with a mail server and MySQL database of student grades

Big Data Examples

- (i) Chocolate Marketing Company with large number of installed Automatic Chocolate Vending Machines (ACVMs)

Big Data Examples

(ii) Automotive Components and Predictive Automotive Maintenance Services (ACPAMS) rendering customer services for maintenance and servicing of (Internet) connected cars and its components

Big Data Examples

- (iii) Weather data Recording, Monitoring and Prediction (WRMP) Organization
- (iv) A toy company optimizing the services offered, products and schedules, devise ways and using Big Data processing and storing for descriptive, predictive and prescriptive analytics

Basis of Big Data Classification

- Big Data sources
- Big Data formats
- Data Stores structure
- Processing data rates
- Processing Big Data rates
- Analysis types

Big Data Classification

- Big Data processing methods
- Data analysis methods
- Data usages

Table 1.1 for Details

Summary

We learnt

- Evolution of Big Data
- Big Data Definitions
- Big Data Characteristics
- Big Data Types
- Basis of Big Data Classifications

End of Lesson 1 on
**Big Data Characteristics, Types and
Classifications**