ASSALAMU'ALAIKUM

Introduction to Big Data

DR. RAHMAD KURNIAWAN, ST., MIT., (MTA., CISDV.)

Learning Objectives

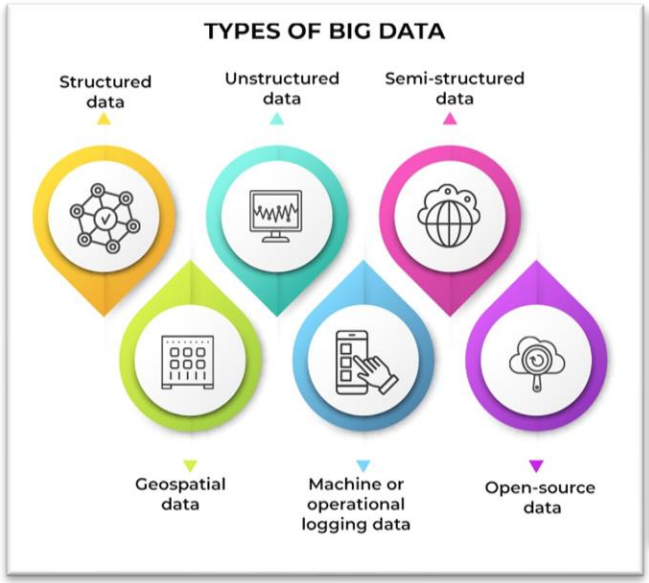
- **By the end of this session, students will be able to:**
  - Define Big Data and explain its key characteristics.
  - Identify major challenges in managing Big Data.
  - Understand the role of Big Data across various industries.

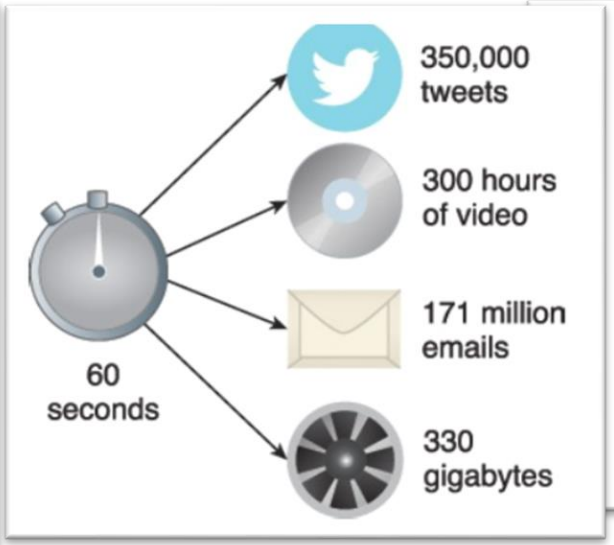
Outline

What is Big Data?

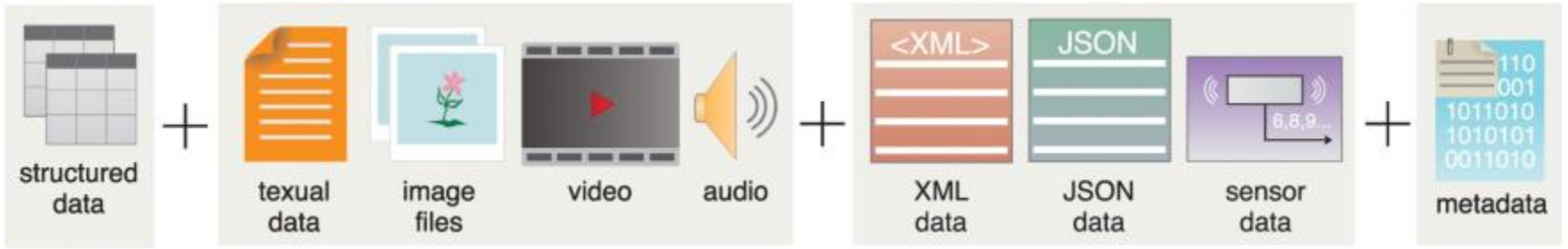
- **Big Data refers to large, complex datasets that traditional data processing software cannot handle effectively.**
- **Big Data addresses distinct requirements,**
  - such as the combining of multiple unrelated datasets, processing of large amounts of unstructured data and harvesting of hidden information in a time-sensitive manner.
- **Key Concepts:**
  - **Volume**: The massive amount of data generated every second.
  - **Velocity**: The speed at which new data is generated and processed.
  - **Variety**: The different types of data (structured, unstructured, semi-structured).

..Cont


TYPES OF BIG DATA
Structured data
Unstructured data
Semi-structured data
Geospatial data
Machine or operational logging data
Open-source data

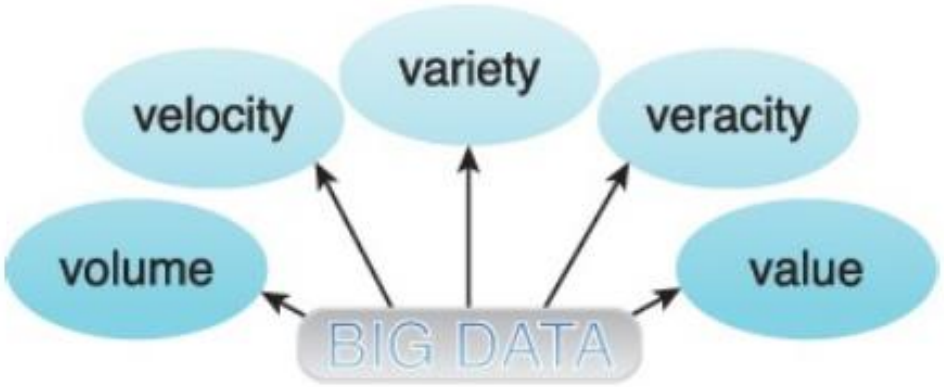

350,000 tweets
300 hours of video
171 million emails
330 gigabytes
60 seconds


BIG DATA
COLLECTION
STORAGE
NETWORK
RESEARCH
CLOUD TECHNOLOGY
ANALYSIS
VISUALIZATION
VOLUME


structured data
texual data
image files
video
audio
<XML>
JSON
6,8,9...
XML data
JSON data
sensor data
110
001
1011010
1010101
0011010
metadata

Beyond the 3V's: Veracity & Value

- **Veracity:**
  - The uncertainty or trustworthiness of data focusing on data quality.
- **Value:**
  - The worth derived from analyzing Big Data, turning data into actionable insights.



Challenges in Big Data Management

- **Scalability → Managing and scaling infrastructure to handle large datasets.**
- **Data Quality → Ensuring the accuracy and reliability of data.**
- **Data Integration → Combining data from various sources into a cohesive whole.**
- **Security and Privacy → Protecting sensitive data from unauthorized access and breaches.**

Opportunities in Big Data

- **Enhanced Decision-Makin ➔ Leveraging data to drive business strategy.**
- **Innovation ➔ Developing new products and services based on data insights.**
- **Personalization ➔ Tailoring experiences based on individual data.**

Sources of Big Data

❖ Examples of Big Data Sources:

- **Social Media ➔** Platforms like Twitter, Facebook, Instagram, etc
- **Sensor Data ➔** IoT devices collecting real-time environmental data.
- **Transaction Data ➔** Financial transactions, e-commerce purchases.

Summary and Review

❖ Key Takeaways:

- Big Data is defined by its Volume, Velocity, Variety, Veracity, and Value.
- Managing Big Data comes with significant challenges but offers vast opportunities.
- Various industries are leveraging Big Data to gain competitive advantages.

❖ Reflective Questions:

- How does Big Data impact the industry you are most interested in?
- What are the potential ethical concerns with Big Data usage?

❖ Thank you