

# Big Data Overview

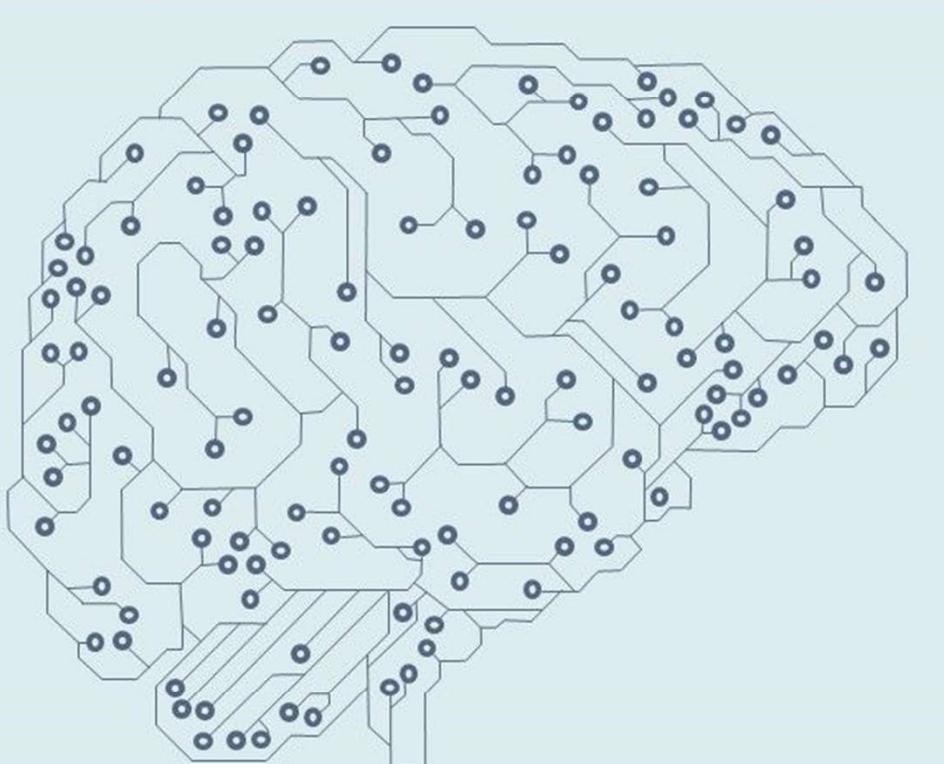


Businesses need to understand that Big Data **is not just about technology**—it is also about how these technologies can push an **organization forward**.



# What is Big Data

---



## Big Data

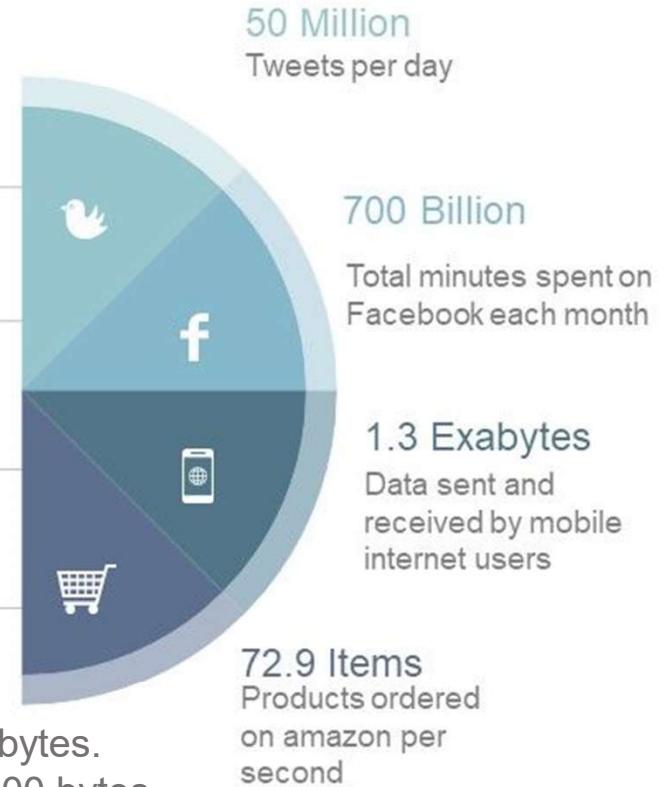
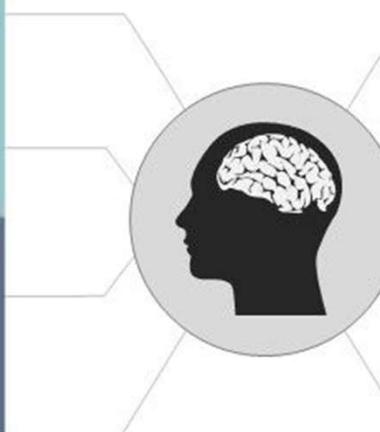
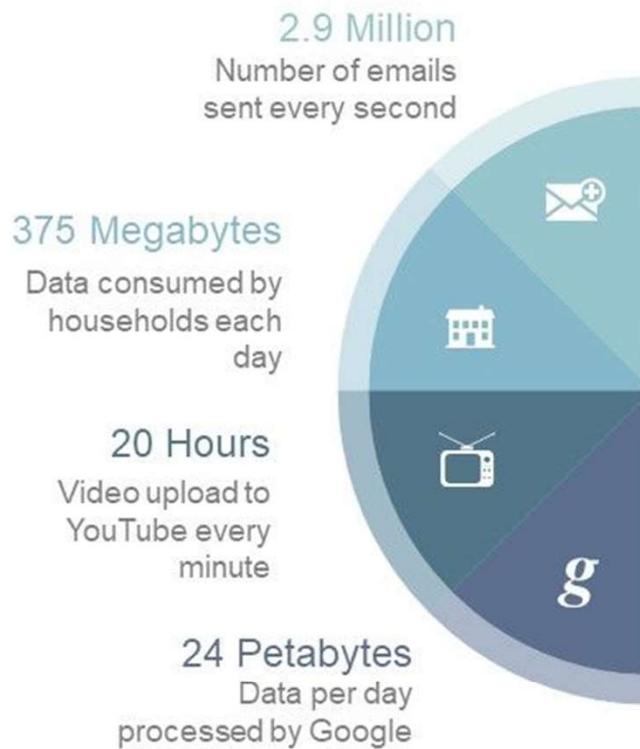
---

Big data is characterized by large volumes of different types of data (e.g. Social, web, transaction, etc.) That builds very quickly.

It exceeds the reach of commonly used hardware environments and software tools to capture, manage and process in a timely manner for its users.

# How Big is Big Data

---

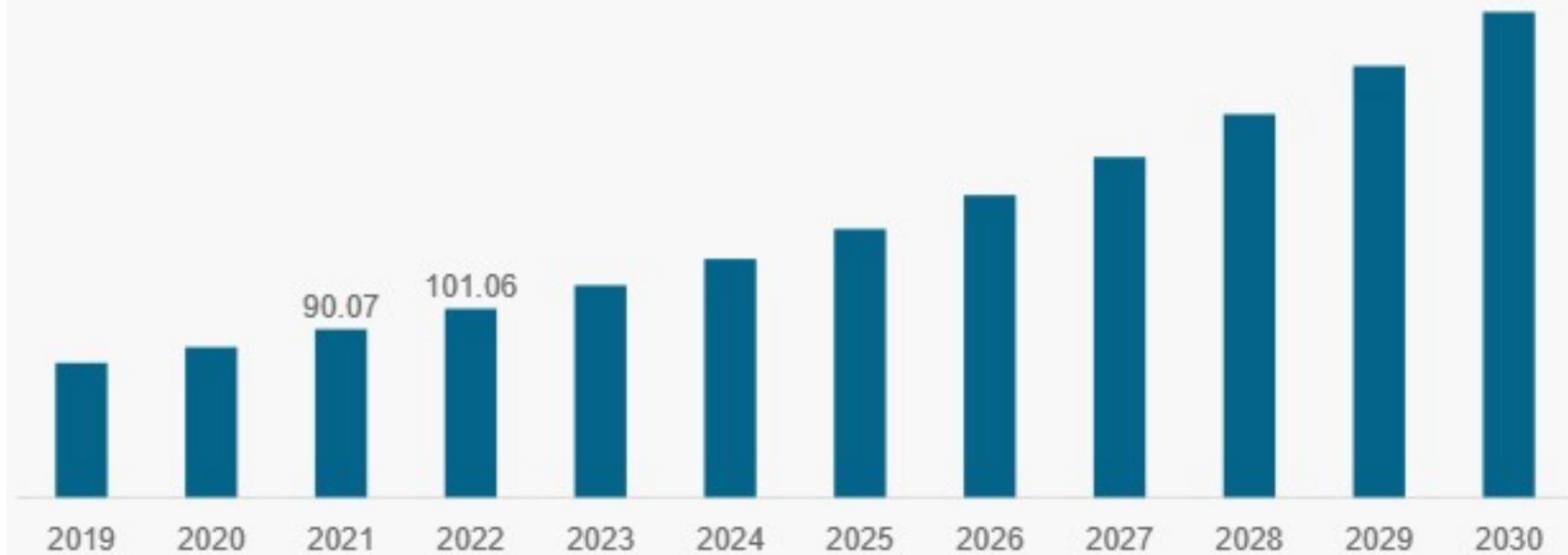


- Byte: 8 bits.
- Kilobyte (KB): 1,000 bytes.
- Megabyte (MB): 1,000,000 bytes.
- Gigabyte (GB): 1,000,000,000 bytes.
- Terabyte (TB): 1,000,000,000,000 bytes.
- Petabyte (PB): 1,000,000,000,000,000 bytes.
- Exabyte (EB): 1,000,000,000,000,000,000 bytes.

# Big Data Market Forecast

---

North America Big Data Analytics Market Size, 2019-2030 (USD Billion)

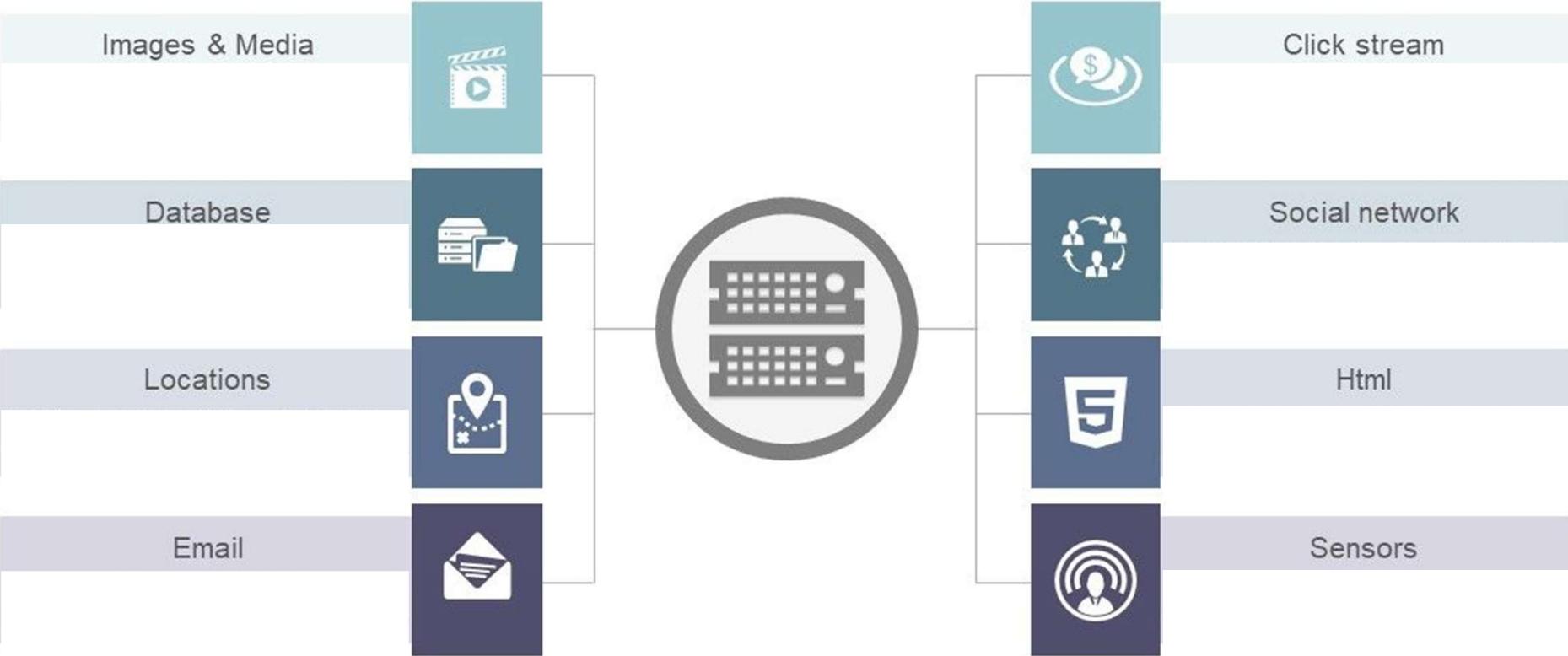


www.fortunebusinessinsights.com

The global big data analytics market size was valued at \$271.83 billion in 2022 & is projected to grow from \$307.52 billion in 2023 to \$745.15 billion by 2030

Read More at:-<https://www.fortunebusinessinsights.com/big-data-analytics-market-106179>

# Sources of Big Data



The diagram illustrates the various sources of big data. At the center is a grey circle containing two server racks. Four lines extend from this central icon to four groups of icons on the left and right sides. The left side features four categories: 'Images & Media' (teal box with video camera icon), 'Database' (dark teal box with database icon), 'Locations' (dark blue box with location pin icon), and 'Email' (dark navy box with envelope icon). The right side features four categories: 'Click stream' (teal box with speech bubble and dollar sign icon), 'Social network' (dark teal box with people icon), 'Html' (dark blue box with HTML5 logo icon), and 'Sensors' (dark navy box with person icon inside a circular signal pattern). Each category group is separated by a thin vertical line.

Images & Media



Database



Locations



Email



Click stream



Html



Sensors



# Sources of Big Data

---

## Media

Media and communication outlets (articles, podcasts, audio, video, email, blogs)



## Social

Digital material created by social media (text, photos, videos, tweets)



## Machine

Data generated by computers and machines generally without human intervention (business process logs, sensors, phone calls)



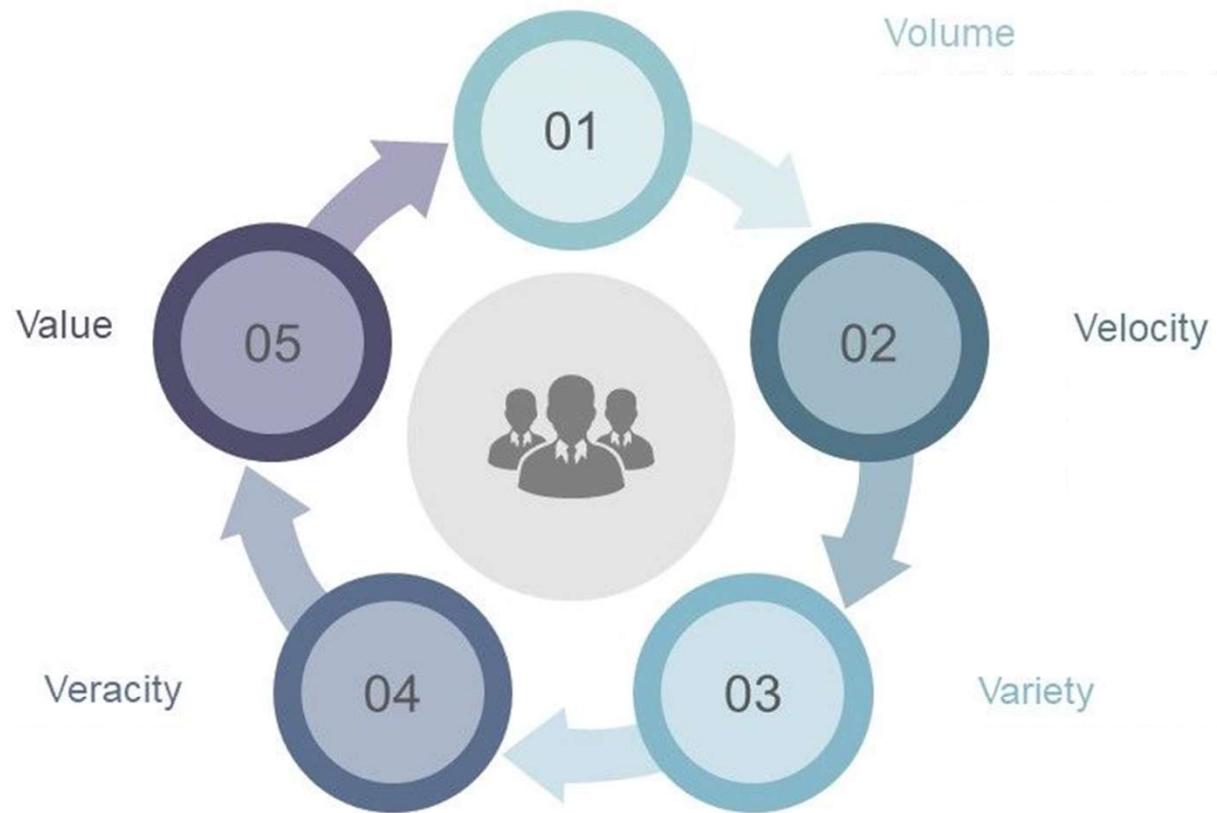
## Historical

Data about our environment (weather, traffic, census) and archived documents, forms or records



# 5 Vs of Big Data

---



# Small Data vs Big Data

---

## Small Data



- Low Volumes
- Batch Velocities
- Structured Varieties

## Big Data



- Into Petabyte Volumes
- Real-time Velocities
- Multistructured Varieties

Vs

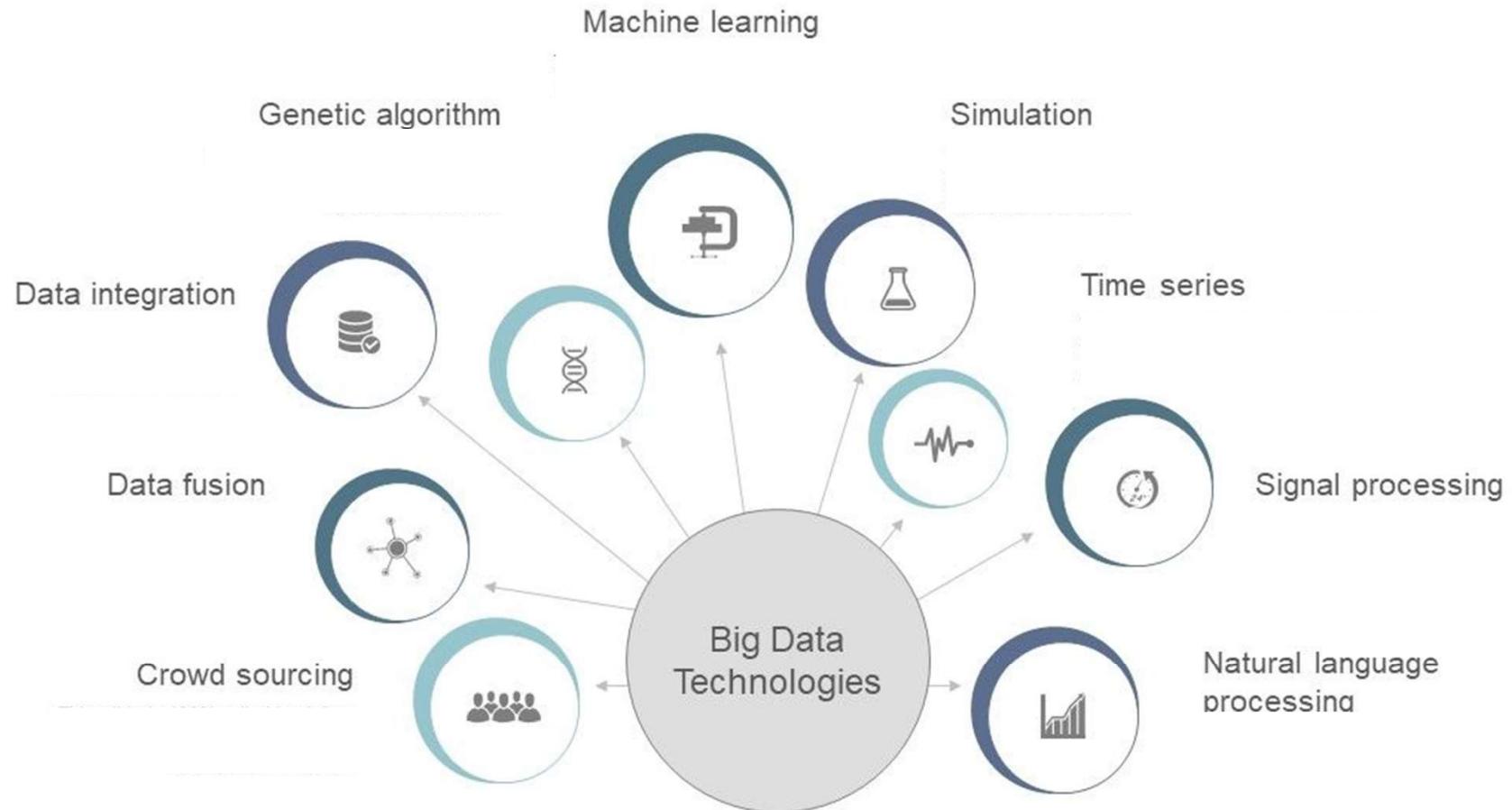
# Objective of Big Data

---



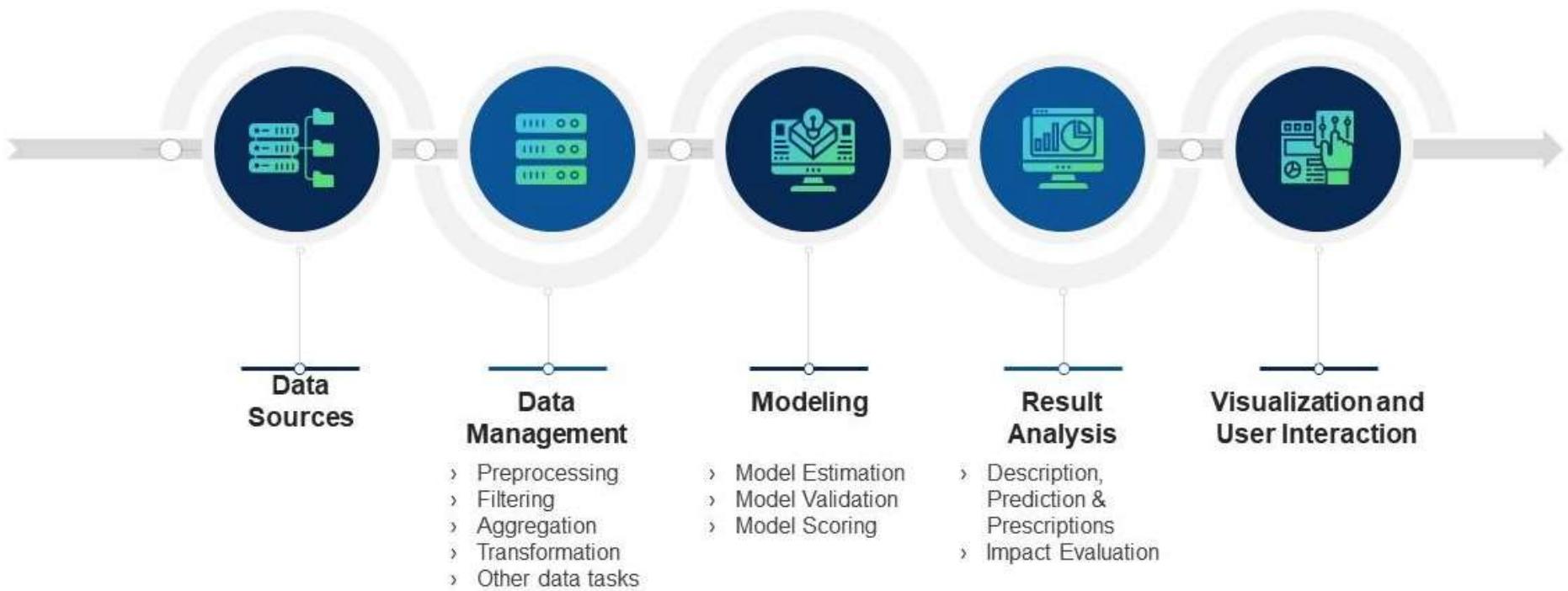
# Big Data Technologies

---



# Big Data Workflow

---



# Forms/ Type of Big Data



## Unstructured



Analog data



GPS tracking  
information



Audio/ video  
streams

01

## Structured



Databases



Data  
warehouses



Enterprise  
systems

02

## Semi-Structured



XML



E- Mail



EDI

03

Data that does not reside in fixed locations generally refers to free-form text, which is ubiquitous.

Data that resides in fixed fields within a record or file.

Between the two forms where “tags” or “structure” are associated or embedded within unstructured data.

# Data Analytics Process

---



# Impact of Big Data

---

## Healthcare



It allows us to find new cures and better understand and predict disease patterns. This leads to saving more lives.

## Science



It creates new possibilities and ways to conduct research which would otherwise be impossible, helping us to make new discoveries.

## Security



Police forces use big data tools to predict criminal activities, conduct investigations and ultimately to catch criminals faster.

## Business



It helps us to improve and optimize the ways we do business by making data-driven decisions.

# Top big data trends in 2023



## Edge computing

Explosive growth in data generated from cloud systems, sensors, smart devices and video streaming is driving adoption of edge computing. Data processing is done on the periphery of the network as close to the originating source as possible.



## Cloud and hybrid cloud computing

Cloud computing enables organizations to process nearly limitless amounts of data. Hybrid cloud approaches are being developed to enable companies in regulated industries to take advantage of cloud's economic and technical advantages.



## Data lakes

These large repositories store structured and unstructured data in its native format. Data scientists often extract just what's needed for a project, eliminating costly ETL processes required of centralized data warehouses.



## Machine learning and AI technologies

Machine learning and other AI technologies are revolutionizing big data analytics. AI's ability to ingest and analyze massive amounts of structured and unstructured data is being used by companies to optimize and improve business operations.

# BIG DATA CHALLENGES





# Questions