# Big Data

## What is Big Data?

**Big Data** simply refers to a **large amount of data** which is of structured, semi-structured or unstructured nature. The data pool is so voluminous that it becomes difficult for an organization to manage and process it using traditional databases and software techniques. Therefore, big data not only implies the enormous amount of available data but it also refers to the entire process of gathering, storing, and analyzing that data.

Big Data is characterized by **6Vs** as described in the following figure:



## Evolution of Big Data

- In today's world where data is the new oil, big data can hypothetically be assumed as a system that acquires crude oil and makes fuel out of it.

- With the advent of the **Internet of Things (IoT)**, data is now readily available. According to **DOMO's report**, "Over 2.5 quintillion bytes of data are created every single day, and it's only going to grow from there. By 2020, it is estimated that 1.7 MB of data will be created every second for every person on earth."

## Significance of Big Data in India

- With a population of around 1.3 billion, Big Data holds a significant position in the Indian context.
- As per the study conducted by **NASSCOM**, "the Indian analytics industry is predicted to reach $16 billion mark by 2025."

## Applications of Big Data

- **Governance:** Big data is used in governance for the following purposes:
  - **Power Discoms** use data collected from last-mile to cut down their *Aggregate Technical & Commercial losses*.
  - **Security agencies and Police forces** use big data to,
    - Prevent cyber-attacks
    - Enhance security systems
    - Detect card-related fraud cases
    - Predict criminal activities, e.g – *Crime and Criminal Tracking Networks and Systems (CCTNS)*
  - In improving the quality of **education**.
  - **Disaster Management:** In understanding and mitigating the risks of disasters.
- **Economy:** Big data can provide huge benefits to various sectors of the economy like,
  - In **Insurance Sector** to improve customer experience & ensure their right to claim
  - In **Banking Sector** to manage financial data
  - To capture the production, price statistics, & **calculate** the resultant **GDP**
  - **Evade risks & minimize losses** for financial firms
  - **Tax officials** catching hold the tax evaders using **Project Insight**
  - **Deregistration** of lakhs of **shell companies**
  - In preventing **money laundering** & curbing **terrorism financing** by identifying the safe havens
- **Health Care:** Big Data in health care caters the following benefits:
  - Predicting diseases,
  - Prescribing medicines,
  - Optimizing treatment,
  - Using clinical data to improve patient care,
  - In critical Diagnostic tests,
  - Finding new cures (R&D)
- **Agriculture and Food:**
  - Seed Selection
  - Geo-Tagging to keep the track record of agricultural assets in the country
  - Weather Forecasting
  - Irrigation & effective water management
  - Food Processing
  - Identification of Crop Diseases

- **Digital Space:**
  - In the **telecom** sector- connecting the hinterland areas and bringing them to the mainstream,
  - On **Social Media** for targeting platform users,
  - **Artificial Intelligence –** Controlling home appliances,
  - Analyzing & Improving individual performance (at work, sports, or home) using wearable devices.

## Government Initiatives & Interventions

- **NITI Aayog** is currently working on a plan in collaboration with the private players to develop the **'National Data & Analytics Platform'**, which will act as a single source of sectoral data for citizens, policymakers, and researchers.
- **'Big Data Management Policy'** drafted by **CAG** for auditing large chunks of data generated by the public sector in the states and the union territories, is a great start.
- **Ministry of Statistics and Programme Implementation** has proposed to establish a **'National Data Warehouse on Official Statistics'** which will leverage technology and use big data analytical tools to improve the quality of **macro-economic aggregates**.
- Use of **Direct Benefit Transfer in MGNREGA & Aadhaar** for **authentication** & availing welfare scheme benefits helps in the filtering of ghost beneficiaries.
- The Ministry of Agriculture has signed a deal with the ISRO to use satellites for **mapping of agricultural assets**.
- Smart City Mission, Digital India, digital economy proposals like BHIM app, inter alia are important government initiatives that are using Big Data for achieving good governance in the country.

## Challenges

- **Invasion of Privacy:** Big Data Analytics leads to the **twin problems of digitization, which are, Data privacy and Net neutrality**.
- **Data Security:** Several incidents of leakage of Aadhaar data underlined the need for the government to increase the security and safety of virtual data it collects from its citizens.
- **Technical Challenges:** As the name suggests, Big data involves several inherent limitations like,
  - Inefficient infrastructure for data collection & management of such humongous data,
  - Storage & computational problems,
  - Scalability & streaming issues.
- **Challenges related to Governance:** For effective policymaking using Big Data, the government needs to adopt a consistent dynamic approach. Constant evaluation of the generated feedback and flexibly changing the related policy structures will percolate the benefits to the lowest level.

## Way Forward

- **Ethical issues** related to data privacy need to get addressed.
  - Data protection law must incorporate some of the best practices followed in the World.
  - It must **strengthen cybersecurity** in order to safely utilize the large pool of virtually available data.
- **Economically,** the alliance of big data with blockchain technology will potentially change the way we see big data for now.
- **Infrastructural Requirement:** To effectively analyze the large chunks of available data, well-equipped **data centers** are needed. It is essential to segregate the relevant data from the irrelevant data pool.

- In a world where most of the governments & businesses rely more on new-age data sources such as satellite imagery, parking images, and night lights, Big Data governance could be a game-changer for India, where policies have for long been under-implemented.

Big Data is like a double-edged sword. It will be wise to handle it with utmost care and conscience.