

# Today's Challenges and Future Scope of Big Data.

Saloni Rajput

Invertis University, Bareilly

## Abstract-

In recent years, Big data has rapidly developed into a hot topic that attracts extensive attention from academia, industry, and governments around the world. In this position paper, we first briefly introduce the concept of big data, including its definition, characteristics. Next, we discuss about some present challenges (Insufficient understanding and acceptance of Big Data, data growth issue, privacy, security issue, Confusion while Big Data tool selection), as well as possible solutions to address these challenges. At the last we discussed about the what other data scientists predict future scope of big data in near future.

## Keywords

(Big data, Data growth, Acceptance of big data, Tool selection).

## 1. INTRODUCTION

The size and number of available data sets has grown rapidly as data is collected by devices such as mobile devices, cheap and numerous information-sensing Internet of things devices, aerial (remote sensing), software logs, cameras, microphones, radio-frequency identification (RFID) readers and wireless sensor networks. Big data has seen normalcy in most businesses today, but that doesn't mean that the journey is always smooth.

## 2. BIG DATA

In 2005 Roger Mougals from O'Reilly Media coined the term Big Data for the first time, only a year after they created the term Web 2.0. It refers to a large set of data that is almost impossible to manage and process using traditional business intelligence tools. Big Data is a collection of small data. In general, it refers to the collection of large and complex datasets which are difficult to process using traditional database management tools or data processing applications. These are available in structured, semi-structured, and

unstructured format in petabytes and beyond. Formally, it is defined from 3Vs to 4Vs. 3Vs refers to volume, velocity, and variety. Volume refers to the huge amount of data that are being generated everyday whereas velocity is the rate of growth and how fast the data are gathered for being analysis. Variety provides information about the types of data such as structured, unstructured, semistructured etc. The fourth V refers to veracity that includes availability and accountability. The prime objective of big data analysis is to process data of high volume, velocity, variety and veracity using various traditional and computational intelligent techniques.

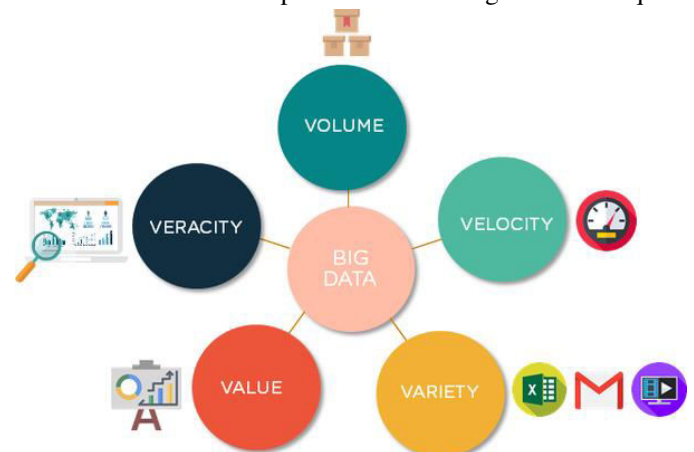


Fig.1- characteristics of bigdata

A 2018 definition states "Big data is where parallel computing tools are needed to handle data", and notes, "This represents a distinct and clearly defined change in the computer science used, via parallel programming theories, and losses of some of the guarantees and capabilities made by Codd's relational model."

## 3. THE PURPOSE OF THIS WORK

This paper explores the challenges and issues in the field of big data and also discusses solutions and future scope of big data.

## 4. CHALLENGES IN BIG DATA

In digital world we face so many challenges in every digital platforms. Opportunities always follow some challenges we need to know various computational complexities, security

threats and computational techniques of big data to analyze big data problems. The analysis of big data presents challenges in sampling, and thus previously allowing for only observations and sampling. Therefore, big data often includes data with sizes that exceed the capacity of traditional software to process within an acceptable time and value. The data in the big data is processed to help the organizations to take decisions on various important issues[1]. All the challenges listed in the above will be discussed in the subsequent section.

### A. Insufficient understanding and acceptance of Big Data

Oftentimes, companies fail to know even the basics what big data actually is, what its benefits are, what infrastructure is needed, etc. Without a clear understanding, a big data adoption project risks to be doomed to failure. Companies may waste lots of time and resources on things they don't even know how to use. And if employees don't understand big data's value and/or don't want to change the existing processes for the sake of its adoption, they can resist it and impede the company's progress.

### B. Data growth issue

One of the foremost pressing challenges of massive Data is storing these huge sets of knowledge properly. The quantity of knowledge being stored in data centers and databases of companies is increasing rapidly. As these data sets grow exponentially with time, it gets challenging to handle. Most of the info is unstructured and comes from documents, videos, audio, text files, and other sources.

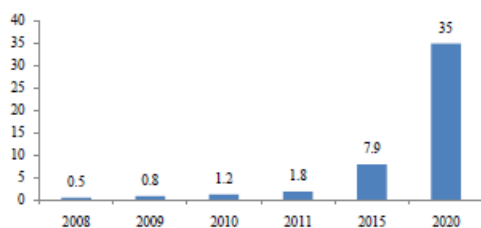


Fig. 2. The forecast of global data growth

Companies choose modern techniques to handle large data sets, like compression, tiering, and deduplication. Compression is employed for reducing the number of bits within the data, thus reducing its overall size. Deduplication is the process of removing duplicate and unwanted data from a knowledge set. Data tiering allows companies to store data in several storage tiers. It ensures that the info is residing within the most appropriate space for storing. Data tiers are often public cloud, private cloud, and flash storage, counting on the info size and importance. Companies also are choosing Big Data tools, like Hadoop, NoSQL, and other technologies.

### C. Privacy, Security issue

Big data comes with security issues—security and privacy issues are key concerns when it comes to big data. Privacy is an important issue in applying Big data technologies for analytics. Bad players can abuse big data—if data falls into the wrong hands, big data can be used for phishing, scams, and to spread disinformation. They have many security issues like no encryption support for the data files, weak authentication between client and server, data at rest unencrypted which can cause privacy threats. Data security is usually put on the back burner, which is not a wise move at all as unprotected data can fast become a serious problem. Stolen records can cost millions

for a company. Overall, companies should surely overcome the big data privacy challenges and security challenges that are like a hurdle to them. It is a reality that if there is an intention to collect, store, and use big data, then investment in adequate security needs to be done[3].

### D. Confusion while Big Data tool selection

These questions bother companies and sometimes they are unable to find the answer. They end up making poor decisions and selecting an inappropriate technology. As a result, time, efforts, money and work hours are wasted. It can be easy to get lost in the variety of big data technologies now available on the market.

## 5. SOLUTIONS

Big Data provides a new method to traditional data analysis, which has a variety of technologies, including Hadoop and MapReduce, cloud computing, grid computing and so on[4]. In competitive business environment, every organization must utilize the information and manage and process the data in hand to generate a good business activities and better management of the resources. Big data is not just about volume and from various sources it is about its other characteristics such as size, speed of data, structure and quality and new-generation analytic technologies that help organizations get more value from their information assets[7]. The major challenge in this case is to pay more attention for designing storage systems and to elevate efficient data analysis tool that provide guarantees on the output when the data comes from different sources. Big data is considered as an important tool to generate fundamental input to decision-making and competitive advantage. Every leader and manager in any organization want all of their action and decisions is based on accurate and precise information. Therefore, big data analytics is the best solution that distills terabytes of low-value data, transforming them into a single bit of high-value data[2].

## 6. FUTURE SCOPE

Recently it was announced that, Indian Prime Minister's office is using Big Data analytics to understand Indian citizen's sentiments and ideas through crowd sourcing platform [www.mygov.in](http://www.mygov.in) and social media to get a picture of common people's thought and opinion on government actions[5]. Google is launching the Google Cloud Platform, which provides developers to develop a range of products from simple websites to complex applications. It enables users to launch virtual machines, store huge amount of data online, and plenty of other things. Basically, it will be an one stop platform for cloud based applications, online gaming, mobile applications, etc. All these required huge amount of data processing where Big Data plays an immense role in data processing.

The majority of big data experts agree that the amount of generated data will be growing exponentially in the future. In this data age 2025 report for seagate, IDC forecaste the global datasphere will reach 175 zettabytes by 2025. To help you understand how big it is.

## 7. CONCLUSION

As we know in today's digital world, new challenges keep coming for companies every day and data scientists make every effort on their solutions. The first task to our companies should be that they can tell their employees what is big data in actual. Big data has become that thing which will go on increasing in future, meanwhile some people will do it for wrong work and some good things now developers have to do that in the coming time they have to be ready in advance.

## REFERENCES

- [1] Bhawna Dhupia, M. Usha Rani, "Research Challenges in Big Data solutions in different Applications."
- [2] Syaiful Hisyam Saleh, Raihan Ismail, Zaharuddin Ibrahim, Norhayati Hussin. "Issues, Challenges and Solutions of Big Data in Information Management: An Overview".
- [3] Prarthana T.V, Jebah Jayakumar, Computer Science and Engineering B.N.M Institute of Technology, Bangalore, India "Challenges and Opportunities with Big Data" Vol. 6 Issue 07, July – 2017.

[4] B.V.Hemalatha, Assistant professor, Theni Kammavar Sangam College of Arts and Science, Theni, Tamilnadu, India. R.Vijayalatha, Research scholar, Manonmaniam Sundaranar University, Tirunelveli, Tamilnadu, India, "Big Data Technologies, Challenges and Future Scope". Vol 5, Issue 3, March 2018.

[5] T.Karunakar, M.Nagarjuna, K.Hema Concepts, Applications and Future Scope of Big Data. International Journal of Scientific & Engineering Research Volume 8, Issue 5, May-2017.

[6] D. P. Acharjya Kauser Ahmed P, School of Computing Science and Engineering, VIT University, Vellore, India. "A Survey on Big Data Analytics: Challenges, Open Research Issues and Tools".

[7] <http://dx.doi.org/10.6007/IJARBS/v8-i12/5240>