

Challenges of Data Analytics for Innovative V's and Security in Bigdata

M. Angelin Rosy^{1*}, R. Deepa², M. Felix Xavier Muthu³

^{1,2}Department of MCA, Er.Perumal Manimekalai College of Engineering, Anna University, Hosur, India

³Department of Mechanical Engineering, St.Xavier's Catholic College of Engineering, Anna University, Nagercoil, India

Corresponding Author: angel_rosym@yahoo.co.in, Tel- 9944579754

Available online at: www.ijcseonline.org

Abstract— Data sets that may be analyzed computationally behavior, interactions are too large. Big data considering the security point of view is safeguarding the user's privacy. A framework for big data security analysis and the semantic technology. The emerging security analysis challenge in big data environment is getting more attention of security. Framework that efficiently processes large volumes of data. The big data analytics is merged with big data security which results in big data security analytics. In this paper, I explain the big data challenges and privacy of outsourced enhancing security be handled by analytics techniques. Big data main problem is privacy of security challenges.

Keywords—Bigdata, Applicationsofbigdata, Privacy, Security, Challenges.

I. INTRODUCTION

We are living in the era of big data. A short history of big data 90% of the available data has been created in the last 2 years and the term big data has been around 2005 [1] the large volume of data both structured and unstructured. The digitalization of our day-to-day activities has resulted in a huge volume of data [2] the processing of such data is made possible by using multiple techniques. Recent years the amount of data generated by Government, Smart cities, E-Commerce, Business, Industries increasing day to day life [3]. This data, called big data, is used by many organizations to extract valuable information either to take marketing decisions track specific behaviour's [4] the traditional security solutions are not capable of handling the security of big data mobility. The big data is a collection of large set of interrelated data set thus it needs its own security from unauthorized access [5]. The main goal of machine learning is to create systems that learn knowledge from data and use that knowledge [6]. Getting data into a big data structure across data sources, getting useful information out of the big data, volume skill availability, and solution cost etc. fig (1.1).



Fig. 1

II. LITERATURE REVIEW

Data are becoming the new raw material of business. “Craig Mundie”, Senior Advisor to the CEO at Microsoft.

A technological perspective, which details the technical aspects of big data. It proceeds by identifying the sources of these data, the utility of these data. Communication technology has led growth of data are becoming the new technical abilities of business analyze and protected. The real revolution is not about the data, it is about the stunning progress in the methods of extracting the data. Output can be highly customized with a bit of error. It becoming commoditized huge data production. Increases data will be easy to come.

III. BIG DATA

Big data refers to data sets that are too large or complex for traditional data. Processing application software to adequately deal with. Data with many cases offer greater statistical power, while data with higher complexity may lead to a higher false discovery rate. Big data was originally associated with three key concepts: Volume, Variety, and Velocity [1] current usage of the term big data tends to refer to the use of predictive analytics, user behavior analytics [2].

IV. APPLICATIONS OF BIG DATA

Big data has increased the demand of information management specialists so much so that Software AG,

ORACLE Corporation, IBM, Microsoft, SAP, EMC, HP and Dell have spent more than \$15 billion on software firms specializing in data management and analytics. Developed economies increasingly use data intensive technologies.

A. Government:

Civil Registration and vital statistical collects all certificates status from birth to death. CRVS is a source of big data for governments.

B. International development:

Advancements in big data analysis offer cost-effective opportunities to improve decision- making in critical development areas such as Health care, Employment, Economic Productivity, Crime, Security and natural disaster and Resource Management.

C. Health Care:

Big data analytics has helped health care improve by providing personalized medicine and prescriptive analytics, Clinical risk intervention and predictive analytics, waste and care variability reduction, automated external and internal reporting of patient data, standardized medical terms and patient registries and fragmented point solutions. With the added adoption of M health, E health and wearable technologies the volume of data will continue to increase. This includes electronic health record data, imaging data, patient generated data. Sensor data and other forms of difficult to process data.

D. Education:

A McKinsey Global Institute study found a shortage of 1.5 million highly trained data professionals and managers and a number of Universities including University of Tennessee and UC Berkeley, have created masters programs to meet this demand. Private boot camps have also developed programs to meet that demand, including free programs like General assembly [2] in the specific field of marketing. One of the problems is that marketing has several sub domains (e.g. advertising, promotions, product development, branding) that all use different types of data.

V. BIG DATA PRIVACY

Big data privacy also called information privacy that deals with the ability organizations. Data in a computer system can be shared with third parties. Privacy big analytics is still challenging due to either the issues of flexibility along with effectiveness [3]. Privacy risks multiply as large quantities of personal data are collected over longer periods of time draws attention to the relative weakness of data protections in the corporate and public sectors and provides practical for protecting privacy when collecting and managing Government data over extended periods of time [3, 4]. Fig (5.1)

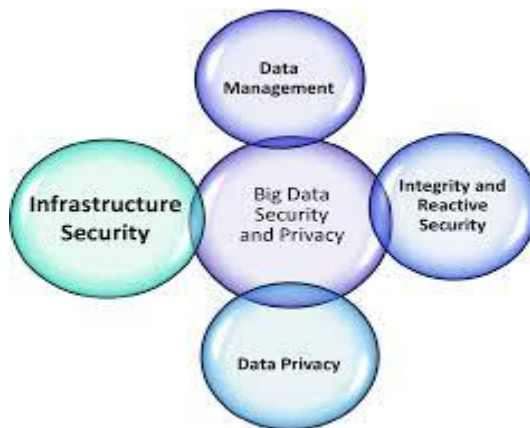


Fig. 2

VI. EIGHT TYPE OF DATA PROTECTION

1. Lawfully processed.
2. Limited process.
3. Secure.
4. Not transferred to other countries.
5. Not excessive data.
6. Accurate.
7. Not kept for longer than data.
8. Processed in line.

VII. BIG DATA SECURITY

Big data security is the processing of data and analytics process challenges are negatively affect. Big data security is the collective term for all the measures and tools used to both the data and analytics processes. One of the most common security tools is encryption a relatively simple tool that can go a long way. Big data security is a constant concern environment for better security [5].

VIII. CHALLENGES OF BIG DATA

The handling of big data is very complex. Some challenges faced during its come to managing and analyzing large volumes of data it comes with its own set of issues. [6, 7] fig (8.1)

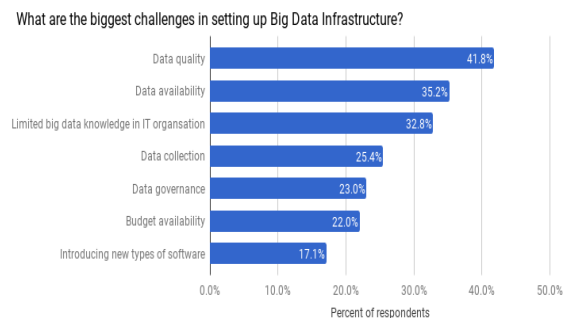


Fig. 3

Data Quality:

Inconsistent and Incomplete.

Discovery:

Finding patterns is difficult.

Storage:

Data is massive.

Analytics:

Analyzing data is difficult.

Security:

Authorization and Authentication.

Lack of talent:

Finding test.

IX. CONCLUSION

In this sense, when a data set is so large that it cannot be easily analyzed by the user, data summarization is a particular interest to the data distribution on the different dimensions. It has been proven to be a useful and effective technique supporting data analysis of large amounts of data. The availability of big data, low cost commodity hardware and new information management. Big data analytics has affected the field of computational was created. Big data will continue to improve the quality of everyday life even though there will always be challenges to overcome.

REFERENCES

- [1]. Robinson, Department of Computer Science.
<https://www.robinson.gsu.edu/org.com>.
- [2]. Priyank Jain and Manasi Gyanchandani.
<https://doi.org/10.1186/S40537-016-0059.st>.
- [3]. Data Flair Team.
<https://data-flair.training/blogs/Bigdata>.
- [4]. Berkman Klein Center.
<https://Cyber.harvard.edu/2018/03/bigdataprivacy>.
- [5]. Science Soft Research Center.
<https://www.scnsoft.com/blog/security>.
- [6]. Sisense Data Center.
<https://www.flydata.com/the6-challengingbigdata>.
- [7]. Dataversity.
<https://www.dataversity.net/bigdatachallenges>.