International Journal of Computer Trends and Technology
ISSN: 2231-2803 / https://doi.org/10.14445/22312803/IJCTT-V68I5P105

Volume 68 Issue 5, 14-17, May 2020

Review Article

Big Data - Benefits and its Growth

S.Subatra Devi

Assistant Professor, Department of Computer Applications, Hindustan College of Arts & Science Chennai, Tamil Nadu, India.

Received Date: 31 March 2020
Revised Date: 15 May 2020
Accepted Date: 16 May 2020

Abstract - *Due to the arrival of new technologies, devices and communications, the amount of data generated from the technology is growing rapidly to a huge amount every year. This gives rise to the era of Big Data. The field of big data plays an important role in most fields. Big data is a term for enormous data sets, more varied and complex structures. Now a day's Big data is most preferably used in enterprises, organizations, companies, businesses etc. In this paper, the benefits of big data processing and the growth of big data are analyzed.*

Keywords - *Big Data, Internet of Things, cloud computing.*

I. INTRODUCTION

A huge repository of terabytes of data is generated each day, hour by hour, from modern information systems and digital technologies such as the Internet of Things and cloud computing, and this gives rise to Zettabytes of data. Analysis of these massive data requires a lot of effort at several levels to extract knowledge for decision making. Hence, big data analysis is a current area of research and development.

Big Data is a term that represents a large collection of data that is huge in volume and keeps on growing exponentially time by time. It encompasses the volume of information, the velocity or speed at which it is created and collected and the variety or scope of the data. It treats ways to analyze, systematically extract information from, or deal with the data sets that are too large or complex with traditional data-processing application software. The prime objective of big data analysis is to process data of high volume, velocity, variety, variability and veracity using various traditional and computational intelligent techniques [1].

It is expected that the growth of big data is estimated to reach 25 billion by 2015 [2]. From the perspective of information and communication technology, big data is a robust motivation to the next generation of information technology industries [3], which are broadly built on the third platform, mainly referring to big data, cloud computing, the internet of things, and social business.

More research was carried out by various researchers on big data and its trends [4, 5, 6]. Different analytics methods and tools are applied to big data in various domains [8].

The rest of the paper is organized as follows. Section II deals with the history of big data, Section III, illustrates the importance, and Section IV specifies the working of big data. Section V gives the benefits of big data processing; Section VI illustrates the growth of big data, and Section VII concludes the paper and presents the future work.

II. HISTORY OF BIG DATA

The term "Big Data" was first introduced to the computing world by Roger Magoulas from O'Reilly media in 2005 [9], which was used to define a great amount of data that is so large, fast or complex. Big data was defined for the traditional data management techniques that cannot be managed and processed due to the complexity and size of the data. It is difficult or impossible to process using traditional methods. A study on the evolution of big data on Research and Scientific topics shows that the term "Big Data" was present in research starting with the 1970s but has been comprised in publications in 2008 [7].

The method of accessing and storing a large amount of information for analytics has been around for a long period of time. But the concept of big data gained momentum in the early 2000s when industry analyst Doug Laney [10] articulated the now-mainstream definition of big data as the three V's. Two more V's, namely Variability and Veracity, are additionally considered for big data [12, 13, 14].

A. Volume

Organizations collect data from a variety of sources, including business transactions, smart devices, industrial equipment, videos, social media and more. In the past, storing it would have been a problem – but cheaper storage on platforms like data lakes and Hadoop have reduced the burden.

B. Velocity

With the growth in the Internet of Things, data streams into business at an exceptional speed must be handled in a timely manner. RFID tags, sensors and smart meters are in need to deal with these data in real-time.

C. Variety

Data comes in all types of formats – from structured, numeric data in traditional databases to unstructured text documents, emails, videos, audios, stock ticker data and financial transactions.

D. Variability

In addition to the increasing velocities and varieties of data, data flows are unpredictable – changing often and varying greatly. It's challenging, but businesses need to know when something is trending in social media and how to manage daily, seasonal, and peak time data loads.

E. Veracity

Veracity refers to the quality of data. Because data comes from so many different sources, it's difficult to link, match, clean and transform data across systems. Businesses need to connect and correlate relationships, hierarchies and multiple data linkages. Otherwise, the data can quickly spiral out of control.

III. IMPORTANCE OF BIG DATA

In 2013, IBM noted that the everyday world creates 2.5 quintillion bytes of data. 90% of the data in the world has been created in the last two years [23]. Data is growing at an exponential rate, and the experts of data analytics technology do not have enough knowledge to analyze the enormous amount of data. Big data presents three main aspects to any organization for interest, first is lacking arrangement, second producing new opportunities, and third, technologies used for big data have low cost.

The importance of big data doesn't specify how much data is present but what is being done with the data. Taking data from any source and analyzing it to enable the below-mentioned facts.

- cost reductions
- time reductions
- new product development
- optimized offerings
- smart decision making.

IV. HOW BIG DATA WORKS

Big data gives new insights that open up new opportunities and business models. Big data involves three key actions:

A. Integrate

Big data brings together data from many disparate sources and applications. Traditional data integration mechanisms, namely ETL (extract, transform and load) generally not up to the task. It requires new strategies and technologies to analyze big data sets at terabyte or even petabyte scale. During integration, the data should be fetched, processed, formatted and available in a form that the business analysts can get started with it.

B. Manage

Big data requires storage. The storage solution can be in the cloud, on-premises, or both. The data can be stored in any form based on the user and brings the desired processing requirements and necessary process engines to those data sets based on an on-demand basis. Many people choose their storage solution according to where the data is residing currently. The cloud is gradually gaining popularity because it supports the users current compute requirements and enables them to spin up resources as needed.

C. Analyze

The user's investment in big data pays off when he analyzes and acts on the data. The user should get new clarity with a visual analysis of their varied data sets. Explore the data further to make new discoveries. The user can share the findings with others, build data models with machine learning and artificial intelligence and put the data to work.

V. BENEFITS OF BIG DATA PROCESSING

Big data is characterized by the large volume of different types of data, such as social, web, transactions etc., that are built 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.

The ability to process Big data brings in multiple benefits, such as

- Businesses can utilize outside intelligence while taking decisions: Access to social data from search engines and sites like Facebook Twitter are enabling organizations to fine-tune their business strategies.
- Improved customer service: Traditional customer feedback systems are getting replaced by new systems designed with Big data technologies. In the new systems, Big Data and natural language processing technologies are used to read and evaluate customer responses.
- Early identification of risk to the product/services
- Better operational efficiency

Big data technologies are used for creating a staging area or landing zone for new data before identifying what data should be moved to the data warehouse. In addition, such integration of Big data technologies and data warehouses

helps an organization to offload infrequently accessed data. Fig. 1 shows the benefits of Big data. Increase in usage of Big data in terms of efficiency, making better business decisions, improved customer service and financial savings.

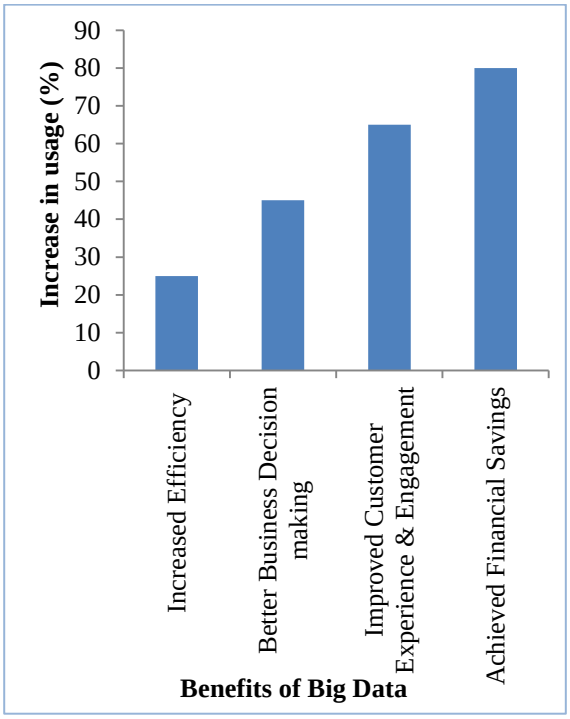

Fig. 1 Benefits of Big Data

VI. GROWTH OF BIG DATA

Worldwide, people are generating many numbers of data each day. Since apps and programs are becoming more complex, there is an increasing need for advanced big data processing, even on a smartphone level. Machine-generated data is produced by a computer without human input. There are lots of ways for businesses to use this data to generate more profits. In recent years, marketers have begun developing data analytics tools that help them understand the market better. Advanced data analytics show that machine-generated data will grow to encompass more than 40% of internet data in 2020. Internet of Things is growing in volume, with more and more apps and data connected to cloud computers.

Analytics illustrates that data grows at an exponential rate. The average day of an internet user is unimaginable without social media. Millions of people are creating a huge mass of online data simply by sharing or commenting on news, posts and articles. Fig. 2 shows the growth of Big data in different years. Big data will grow to a high level in 2020 than in the previous years. In the following years, the usage of the internet grows, and it makes the big data grow to a higher level year by year. Nearly 90 zettabytes will need analysis by 2022, and this may grow more year by year.

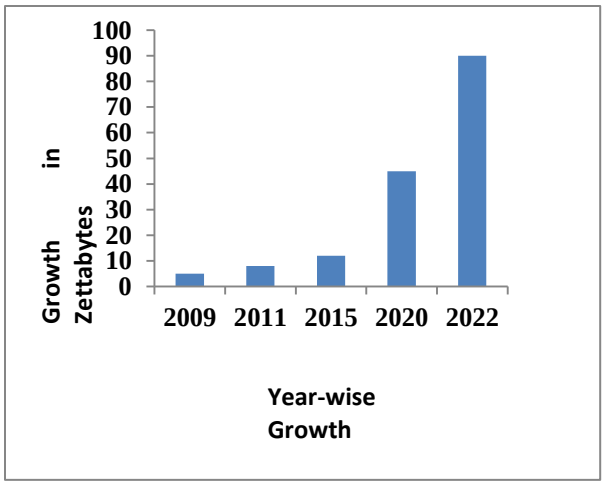

Fig. 2 Growth of Big Data

VII. CONCLUSION

In recent years, big data has played a needful role in developing technology. In this paper, the importance of big data and the working of big data are discussed. The main aim of this paper is to view the benefits of big data in various fields and its growth in different years. Big data is a powerful tool that makes things easier in various fields. Big data used in many applications is a current area of research and development and is an evolving field where much of the research is being done.

REFERENCES

[1] M. K.Kakhani, S. Kakhani and S. R. Biradar, Research issues in big data analytics, International Journal of Application or Innovation in Engineering & Management, 2(8) (2015)228-232.

[2] C. Lynch, Big data: How do your data grow?, Nature, 455(2008) 28-29.

[3] X. Jin, B. W.Wah, X. Cheng and Y. Wang, Significance and challenges or big data research, Big Data Research, 2(2)(2015)59-64.

[4] C. L. Philip, Q. Chen and C. Y. Zhang, Data-intensive applications, challenges, techniques and technologies: A survey on big data, Information Sciences, 275 (2014) 314-347.

[5] K. Kambatla, G. Kollias, V. Kumar and A. Gram, Trends in big data analytics, Journal of Parallel and Distributed Computing, 74(7)(2014) 2561-2573.

[6] S. Del. Rio, V. Lopez, J. M. Bentez and F. Herrera, On the use of map-reduce for imbalanced big data using random forest, Information Sciences, 285(2014)112-137.

[7] H. Moed, The Evolution of Big Data as a Research and Scientific Topic: Overview of the Literature, (2012). Research Trends.

[8] Nada Elgendy and Ahmed Elragal, Big Data Analytics: A Literature Review Paper, Springer International Publishing, (2014) 214-227.

[9] http://strata.oreilly.com/2010/01/roger-magoulas-on-big-data.html

[10] Doug Laney, 3D Data Management: Controlling Data Volume, Velocity, and Variety, Gartner, 949(2001).

[11] C.L. Philip Chen, and Chun-Yang Zhang,Data-intensive Applications, Challenges, Techniques and Technologies: A survey on Big Data, Information Science, 275(2014) 314-347.

[12] Yuri Demchenko, Cees de Laat, and Peter Membrey, Defining Architecture Components of the Big Data cosystem, International

Conference on Collaboration Technologies and Systems (CTS), pages (2014)104 - 112.

[13] The 7 pillars of Big Data, A White Paper of Landmark Solutions, (2016) from https://www.landmark.solutions/Portals/0/LMSDocs/Whitepapers/The 7 pillars of Big data Whitepaper.pdf.

[14] Eileen McNulty, Understanding Big Data (2016). fhttp://dataconomy.com/seven-vs-big- data.

[15] Thomas H. Davenport, Jill Dyche, Big Data in Big Companies, International Institute for Analytics,(2013).

[16] J. Manyika, M. Chui, B. Brown, J. Bughin, R. Dobbs, C. Roxburgh, and A. Byers, Big data: The next frontier for innovation, competition, and productivity, Mickensy Global Institute,(2011) 1–137.

[17] A. Katal, M. Wazid, R. H. Goudar, Big data: Issues, challenges, tools and good practices, in Proc. of IEEE Int. Conf. on Contemporary Computing, (2013)404–409.

[18] Dong, X.L. Srivastava, D. Data Engineering (ICDE), Big data integration, IEEE International Conference, 29(2013)1245–1248.

[19] Tekiner F. and Keane J.A., Systems, Man and Cybernetics (SMC), Big Data Framework, IEEE International Conference, (2013) 1494–1499.

[20] Chun-Wei Tsai, Chin-Feng Lai, Han-Chieh Chao & Athanasios V. Vasilakos Big data analytics: a survey, Journal of Big Data, 2(21), (2015).

[21] Uthayasankar, Sivarajah, Muhammad Mustafa, Kamal, Zahir Irani, Vishanth Weerakkody, Critical analysis of Big Data challenges and Analytical methods, Journal of Business Research,70(2017)263-286.

[22] Xiaolong Jin, Benjamin W. Wah, Xueqi Cheng and Yuanzhuo Wang, Significance and Challenges of Big Data Research, Big Data Research, 2(2) (2015)59-64.

[23] IBM, Bringing Big Data to the Enterprise: What is Big Data?, (2016).