

# Role of Big Data in Business and Information Technology

Deepali D. Patil

**Abstract—** This paper aims to evaluate the significance of the impact of big data in various fields and determine its future in those areas. A comprehensive review of literature on influence and trends of big data variables with respect to different fields built hypothetical foundation of the paper. The big data is a tremendous amount of structured and unstructured data that provides major insights into a particular field which further supports the decision-making systems and builds the foundation of company's competitiveness. Data contributes in the processing, evaluation, understanding and decision-making steps. Big data influences every area that constitutes of pools of data. This proposed paper is beneficial to accumulate the knowledge regarding the hidden patterns and insights of big data variables and assess the value of it in major areas such as business and information technology.

**Index Terms—** Big data, Business, Information Technology Hadoop.

## I. INTRODUCTION

The raw format of information and knowledge is represented by data. Data collection from distinct procedures supports in effective decision-making process and contributes to the understanding of the results bearer's requirements of the procedures. Data warehousing provides the essential support for storage, collection and the management of data [1]. Big data aims at extracting valuable information from a data that could be considered useless. This is conducted by the utilizing technologies such as MongoDB, Hadoop, Cassandra, Hive etc. which are key technologies in the big data field. Big data deals with the unstructured, vast and dynamic data and extricates meaningful information that helps in the advancement and growth of the company [2]. Big data has the ability to handle large amount of information, scrutinize it and deliver impressive conclusions that has the power to change the thinking towards businesses and the other fields. Data has transformed into a raw material and it has the powerful probabilistic ways that allows it to deal with dirty or messy data without any difficulties [3]. Big data is ought to be implemented in an automated and navigable data collected from various sources, instead of the systematic and constant approaches. It can be used to derive open access resources for various databases [4]. Big data is and impressive transformation of how the information is processed by the society. The utilization of dealing with tremendous amount of data demands three changes in the approach of data which include working with large amount of data, accepting the messiness and unstructured format and using previous

data for collecting information that helps in anticipating the outcomes of certain queries [5]. Big data contributes highly in the competitiveness and advancement of the companies. It is used to make strategic decisions and appear unique in the market as companies can innovate and attain valuable information from the roots [6]. The big data research is driven by major companies that include Facebook, twitter etc. The data collected from enormous end consumers is used to perform customer, work analysis, fraud determination and systems that facilitate recommendations [7].

## II. IMPACT ON BUSINESS

### A. Big Data and Business

Big data and big data analytics is utilized to give the description of applications that are significantly complex, large and incorporate sets of data and techniques concerned with analytics. These applications need an effective data management and storage software. Big data allows the businesses to grow by supporting in the decision-making process. It also helps in modelling of the predictions, visualization and mining processes of data sets [8]. Big data helps the businesses to utilize large amount of information and analyze it to gain insights of the critical aspects such as sales. These aspects of businesses are analyzed by concentrating on the information gained from social media and customer locations. This information further provides foundation to the cost reduction, product improvement and customer retention processes [9]. Big data tools amplify the critical operations of the business by giving insights that hold great value to the company. It fetches cost saving to the business activities [10]. Many business companies are willing to invest in big data technologies to improve the efficiency of the returns delivered by predictive and analytical investments in the typical data [11]. Companies like Facebook, Twitter, Apple, Google are analyzing the tremendous amount of data generated by consumers to improve the consumer experience through the exploitation of big data tools such as Hadoop [12]. In business the usage of big data must be realistic. The big data tools must focus on attaining the expected outcomes in certain fields while pushing the potentials in other fields. For instance, the unstructured feedback data must be analyzed with a simultaneous attempt to predict the consumer actions with predictive algorithms [13]. The BI tool of Microsoft has the features like cloud access, visualization, business modelling and collaboration with Office 365. This business big data tool is beneficial in predictive analytics [14]. Big data helps the businesses to build the security and privacy services for customers to keep the personal updated data secured from malicious activities [15].

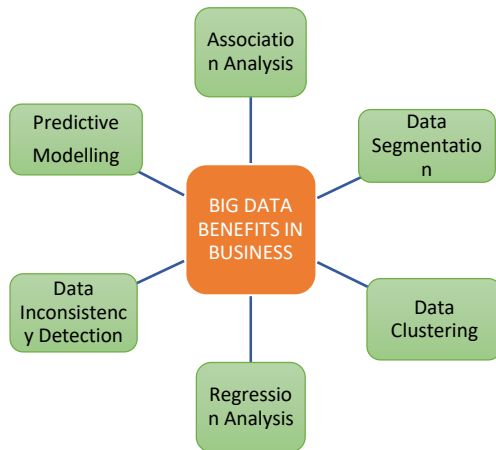


Fig. 1. Big Data Benefits in Business

### B. Big Data Business Objectives

Hadoop clusters deliver storage for terabyte amount of data. The companies tend to adopt the big data approach to achieve certain objectives that foster growth [16]. These objectives are mentioned below:

#### 1) Cost Reduction:

GroupM is a subsidiary for conglomerate of media purchase activity. It utilizes big data tools to track the number of viewers. The decentralized approach for tracking services would result in \$1 million cost each site. The use of big data tools will allow the company to spend only one third of the cost needed for decentralized approach.

#### 2) Development of New Data-Based Offerings:

Big data technologies are key in developing new products and services based on the data available. Google uses big data for search refinement and advertisement placements. It had developed data-based products like Gmail, Google Apps etc.

#### 3) Time Reduction:

Macy's has been able to reduce the time needed for price optimization of around 73 million products from 27 hours to just 1 hour. It helped Macy's in frequent price fluctuations that are favorable in dynamic marketplace.

#### 4) Internal Business Decisions Support:

Big data supports internal business decision-making systems by providing solutions to queries related to product pricing, inventory volume, customer satisfaction etc. Banks use customer transactions, call center recordings to analyze the opportunities and challenges in customer satisfaction.

### C. Business Data Types for Decision Making Systems

If the companies have the ability to analyze data from any source and of any amount and type, then it would result in favorable outcomes that might provide valuable insights regarding the business trends and patterns. This data includes the customer and product data which further incorporate the business customer data and selling-purchasing data respectively [17].

## III. IMPACT ON INFORMATION TECHNOLOGY

Author describes that a new era of data research and exploitation is represented by Big data [18]. The demand for IT products and services is growing fast as it is generating

unique offerings to use goods and services through cloud computing, virtual operations [19]. Author explains that the organizations that use big data are different from other traditional environments of data analytics in three ways [20]:

- The concentration is directed thoroughly towards flows of data
- Data scientists and developers are considered to be more reliable than data analysts
- Analytics is implemented more in business, production and operation fields.

In information technology, big data profoundly increases the power of computation and uses the precision of the algorithms to acquire the connection, analysis and compare data sets that have large amount of data [21]. At the same time, big data methodologies like data mining have given rise to serious privacy security issues where an individual's personal information may be accessed in an unauthorized way [22].

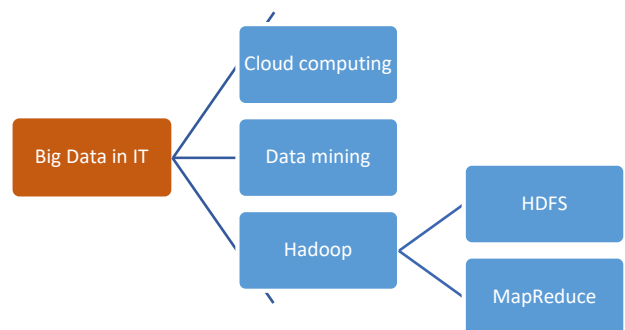


Fig. 2. Big Data Benefits in Information Technology

### A. Cloud Computing

Cloud computing enables expedient and on demand authorized access to the shared computing services that are effortlessly deliverable and do not require much management efforts [23]. Cloud computing allows the companies to invest money and time in the complex computational groups of data and the high-level storage systems. This has been possible due to the reasonable processors and machines based on virtual technologies [24]. Cloud computing has several disadvantages which include the high costs and timespan needed in the uploading and downloading activities of the data to the cloud [25]. The responsibility of the security and privacy of the data shared with cloud is considered as major challenge in the triumph of cloud's data management systems [26]. Data mining is the analysis of large data sets to detect connections that are hidden or unexpected and to recapitulate the data in lucid and beneficial manner [27]. Data mining is utilized for applying the tools and techniques of analytics to methods and procedures of big data sets [28]. The data mining process focuses on retrieving comprehensible knowledge discovery that underpins decision-makers [29]. The applications of data mining include market analysis, market segmentation and management, risk examination and management, detection of frauds and uncommon trends, text and web mining.

## B. Hadoop

Hadoop is a programming language that has a foundation of Java framework and reinforces the handling of huge data sets in a computing arena that is distributed. Hadoop incorporates two major components such as MapReduce and HDFS which process the data parallelly that has distribution over several servers [30]. Hadoop Distributed File System (HDFS) is a storage system that has a fault-tolerance feature. It has capacity to store tremendous amount of data and ensures the availability of the information in severe cases of fault occurrence [31]. MapReduce facilitates the failure handling and parallel processing of patterns for simplification of applications. It performs data distribution and balance loading between several processors [32].

## IV. CONTRIBUTION AND NEW SIGHT

This paper contributes to the identification the emerging trends and interconnects the information, technologies, methods and impacts [33]. This study contributes to the new insights of the accurate composition of big data analytics potential required for transformations that deliver experiential foundation that will encourage the detailed exploration of analysis of big data [34]. In developed countries, big data give insights of the social and economic indicators and enables the anticipation of the trends in socio-economic features. It also helps to derive the potential characteristics of income, price, education and population sector. On the other hand, in developing countries, big data can retrieve instant and timely records and evaluate the untraceable shocks of disasters. It also underpins the poverty, farm inputs and food security influences [35]. In order to reveal the potential of big data, businesses are required to prepare a solid way of utilizing the data for reporting and analytics purposes. By the year 2020, the spending on cloud-based BDA technologies will multiply 4.5 times more than that of the big data and analytics problem-solving tools used inside of the company [36]. Big data has the ability to give rise to the smart cities that incorporate real-time management based on the datasets that are constantly changing, tremendous and interlinked [37].

## V. CONCLUSION

The volume of data generated by machines and automated systems is dominated by the enormous amount of data produced by human beings. This has fueled the need for utilizing the astonishing tool like big data that supports the decision-making systems in complexities. Regardless of the fact that big data is in its initial phase, it has upgraded and renovated the major areas of market by providing it new dimensions to be productive and gain competitive advantages to survive in the market. From this research, we can conclude that the companies have evolved amazingly that they have compiled tremendous data in this period of huge transformation. Big data technologies have certain solutions that solve the complexities of this massive ambiguous data. Based on the impact and hype of big data in major fields, we can infer that big data has a trend of development and has critical challenges that companies have to face in order to expand and grow significantly. Hadoop and data mining technologies are used to arrive at

conclusions that influence the decision-making systems of the company. Big data is a double-edged sword that has the power to give success yet destroy certain things as it has some serious setbacks. These setbacks include data privacy, information leakage and fraudulent activities. Big data has the ability to give huge transformation to economies by delivering changes in various fields. The market has lack of manpower to cope up with the analytical and data science skills. The big data hype is creating massive demand of big data skilled employees who can contribute to the industry's growth. This study signifies the importance of big data in strategic decision making and unique service deliveries. Big data with analytics proves to be a great combination for businesses to get effective predictive analysis results. This research paper enlightens the challenges, benefits and types of data pooled in every field. Big data has the potential of delivering success to companies and driving it towards the innovation. However, it has many drawbacks that demand the companies to change entire approach and implement solid strategic decisions. Sometimes this leads to implementation of incorrect strategic moves that prove to be detrimental to the development of the company. Hence, from the research in this paper, we can conclude that despite of several setbacks, big data has ability to maximize the profitability and retrieve valuable information from junk data that further supports decision-making processes.

## REFERENCES

- [1] Krishnan, K. (2013). *Data Warehousing in the Age of Big Data*. San Francisco, UNITED STATES: Elsevier Science & Technology.
- [2] Zakir, J., Seymour, T., & Berg, K. (2015). BIG DATA ANALYTICS. *Issues in Information Systems*, 16(2).
- [3] John Walker, S. (2014). Big data: A revolution that will transform how we live, work, and think. In: Taylor & Francis.
- [4] Swan, M. (2013). The quantified self: Fundamental disruption in big data science and biological discovery. *Big Data*, 1(2), 85-99.
- [5] Cukier, K., & Mayer-Schoenberger, V. (2013). The rise of big data: How it's changing the way we think about the world. *Foreign Aff.*, 92, 28.
- [6] Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Byers, A. H. (2011). Big data: The next frontier for innovation, competition, and productivity.
- [7] Yang, X., Lu, R., Liang, H., & Tang, X. (2015). Big data research.
- [8] Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business Intelligence and Analytics: From Big Data to Big Impact. *MIS quarterly*, 36(4), 1165-1188.
- [9] Logica BANICA, A. H. (2015). BIG DATA IN BUSINESS ENVIRONMENT. *Buletin științific: Universitatea din Pitești. Seria Științe Economice*, 14(1), 79-86.
- [10] Alfouzan, H. I. (2015). Big Data In Business. *International Journal of Scientific & Engineering Research*, 6(5), 1351.
- [11] Gopalkrishnan, V., Steier, D., Lewis, H., & Guszczka, J. (2012). *Big data, big business: bridging the gap*. Paper presented at the Proceedings of the 1st International Workshop on Big Data, Streams and Heterogeneous Source Mining: Algorithms, Systems, Programming Models and Applications.
- [12] Fan, W., & Bifet, A. (2013). Mining big data: current status, and forecast to the future. *ACM SIGKDD Explorations Newsletter*, 14(2), 1-5.
- [13] Mithas, S., Lee, M. R., Earley, S., Murugesan, S., & Djavanshir, R. (2013). Leveraging Big Data and Business Analytics [Guest editors' introduction]. *IT professional*, 15(6), 18-20.
- [14] Data, B., & Intelligence, B. (2014). TECHPractices.
- [15] Karjoth, G., Schunter, M., & Waidner, M. (2002). *Platform for enterprise privacy practices: Privacy-enabled management of customer data*. Paper presented at the International Workshop on Privacy Enhancing Technologies.
- [16] Davenport, T. H., & Dyché, J. (2013). Big data in big companies.
- [17] Alam, J. R., Sajid, A., Talib, R., & Niaz, M. (2014). A review on the role of big data in business.

- [18] IBM, Zikopoulos, P., & Eaton, C. (2011). Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data: McGraw-Hill Osborne Media.
- [19] Bughin, J., Chui, M., & Manyika, J. (2010). Clouds, big data, and smart assets: Ten tech-enabled business trends to watch. *McKinsey quarterly*, 56(1), 75-86.
- [20] Davenport, T. H., Barth, P., & Bean, R. (2012). How big data is different. *MIT Sloan Management Review*, 54(1), 43.
- [21] Boyd, D., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, communication & society*, 15(5), 662-679.
- [22] Xu, L., Jiang, C., Wang, J., Yuan, J., & Ren, Y. (2014). Information security in big data: privacy and data mining. *IEEE Access*, 2, 1149-1176.
- [23] Mell, P., & Grance, T. (2011). The NIST definition of cloud computing.
- [24] Schadt, E. E., Linderman, M. D., Sorenson, J., Lee, L., & Nolan, G. P. (2010). Computational solutions to large-scale data management and analysis. *Nature reviews. Genetics*, 11(9), 647-657. doi:10.1038/nrg2857
- [25] Chen, C. P., & Zhang, C.-Y. (2014). Data-intensive applications, challenges, techniques and technologies: A survey on Big Data. *Information Sciences*, 275, 314-347.
- [26] Agrawal, D., Das, S., & El Abbadi, A. (2011). *Big data and cloud computing: current state and future opportunities*. Paper presented at the Proceedings of the 14th International Conference on Extending Database Technology.
- [27] Hand, D. J. (2007). Principles of data mining. *Drug safety*, 30(7), 621-622.
- [28] Linoff, G. S., & Berry, M. J. (2011). Data mining techniques: for marketing, sales, and customer relationship management: John Wiley & Sons.
- [29] Freitas, A. A. (2001). Understanding the crucial role of attribute interaction in data mining. *Artificial Intelligence Review*, 16(3), 177-199.
- [30] Sreedhar, C., Kavitha, D., & Rani, K. A. Big Data and Hadoop. *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)* Volume, 3.
- [31] Bhosale, H. S., & Gadekar, D. P. (2014). A review paper on Big Data and Hadoop. *International Journal of Scientific and Research Publications*, 4(10), 1-7.
- [32] Yang, X. Y., Liu, Z., & Fu, Y. (2010). *MapReduce as a programming model for association rules algorithm on Hadoop*. Paper presented at the Information Sciences and Interaction Sciences (ICIS), 2010 3rd International Conference on.
- [33] De Mauro, A., Greco, M., & Grimaldi, M. (2015). *What is big data? A consensual definition and a review of key research topics*. Paper presented at the AIP conference proceedings.
- [34] Wang, Y., & Hajli, N. (2017). Exploring the path to big data analytics success in healthcare. *Journal of Business Research*, 70, 287-299. doi:https://doi.org/10.1016/j.jbusres.2016.08.002
- [35] di Bella, E., Leporatti, L., & Maggino, F. (2018). Big Data and Social Indicators: Actual Trends and New Perspectives. *Social Indicators Research*, 135(3), 869-878. doi:10.1007/s11205-016-1495-y
- [36] Chaudhari, P., & Patel, B. (2017). Future of Big Data. *International Research Journal of Engineering and Technology*, 4(1), 595-597.
- [37] Kitchin, R. (2014). The real-time city? Big data and smart urbanism. *GeoJournal*, 79(1), 1-14. doi:10.1007/s10708-013-9516-8

**Deepali Patil** received B.M.S degree from Mumbai University. Presently, she is a student of MS in Technology Management at University of Bridgeport, USA. Her research interest areas include Big Data, Data mining, Data warehousing, Marketing and Management. Previously she has published one research paper in International Journal.