# A Study on the Role of Big Data in Enterprises

Gaurav Sitaula

Bowling Green State University

November 12, 2017

## Introduction

This era belongs to information and technology. Each day millions of data are generated. For a considerable length of time, the overwhelming sum of data created was considered pointless. Nevertheless, data has dependably been a vital piece for each venture regardless of its size. According to the SAS Institute, "Big data is a term that describes the huge volume of data – both structured and unstructured – that inundates a business on a day-to-day basis". The storage for the data in this present technological world is not a big challenge since we can store millions of data in a tiny device. As the significance of using data for any enterprises became apparent, so did the expansion of information storehouses inside a venture [2]. The method of using data is extremely important nowadays for any firm. If it is leveraged properly for analyzing the business insights, it can deliver immense business values leading to informed decisions and strategic advantages.

The current trend in research is on how big data is helping enterprises to gain business insights. The research increasingly suggests the growing importance of big data in managerial, strategic, and operational goal setting processes and the necessity of unified data supervision across the organizations. The review examines the cloud adoptions, big data reduction, big data collaboration, and solving big data challenges in today's intelligent enterprises and concludes that the big data has a significant impact on today's intelligent enterprises. The appropriate use of the analysis of big data improves the effective value creation of the firms which results in higher customer retention rate, lower operating cost, and increased revenue.

**The idea of Big data**:

Big data represents an enormous extent of raw data that is collected by almost every other business or corporate organizations. The accessibility of the Internet has dramatically altered the way enterprises operate. Every business is now becoming a “digital” business which results in data explosion. Multiple research shows that the volume of data in the movement of businesses is accelerating approximately 200 % per year. Furthermore, it is predicted that within next five years, the business data will grow by 800%, out of which, 80% will be raw unstructured data [3]. This huge information is a combination of transactional data and interactive data. The evolving technologies will take care of handling volumes of transaction data. On the other hand, the interactive data which adds variety and velocity features to the ever-growing data reservoir will be causing significant challenges to enterprises [2]. The IBM research performed on a group of 1500 president of different organizations shows that the frustration of CEOs has been increasing due to the lack of opportunity for converting the available data into the established operation plans implementation. Several other research also suggests that in next decade, the United States alone will face a shortage of 190,000 skilled data scientists, managers, and analysts those who can find trends from the enterprise big data. Because of the explosion of big data in the next decade and the rising demand for the skilled and trained workforce, almost all of the industries will be influenced by machine learning and data analytics. Kościelniak and Puto (2015) proposed the integrated data management be highly prevalent within three major business areas: a) customer service, b) product development, and c) marketing departments. These areas constitute the significant portion of data generation of any typical firm. The biggest challenge for the next-generation analyst would be working with unstructured data which is stored across enterprise data-warehouses. For specialized roles of data analysts in data retrieval, exploratory analysis expertise, or predictive model building,

the complexity for each role increases and each person has to collaborate with each other to add value in the business insights [6]. Sukumar and Ferrell (2013) suggest collaboration lifecycle in three magnitudes: (1) comprehending area-specific knowledge, (2) understanding the data model, and (3) understanding the data. Primarily, the collaborative parties should have proper knowledge of data domain before generating any data model. Otherwise, the data model will not drive us to the business insights. After generating the proper data model, we can now use current data to predict the future behavior.

**Big data drives cloud computation for businesses**:

The technology of computational, networking and storage resources to lessen the functional and implementing cost of using extensive automated systems is defined as cloud computing [4]. Liu (2013) analyzes the requirement of cloud computing in the enterprises that are motivated by the invention of big data. The innovation of many technological startups is contributing to the excessive growth of cloud because they are more likely to rent it instead of purchasing whole IT infrastructure due to comparative benefits. In concordance with previous research, Rehman et al. (2016) also emphasize the adoption of cloud computing systems across all business functions to increase efficiency and productivity. Recently, several cloud computing service startups have started offering big data processing facilities for businesses because of the increasing demand of cloud computing systems. The challenges, on the other hand, due to the increasing requirements of storage resources, computational capabilities, and networking ability are yet to be explored and analyzed.

The history of using cloud by enterprises is not very old. Previously, enterprises stored their data in traditional database systems which ran on commodity hardware. There is a gap between the traditional database and the present cloud-based database. Kaur and Sharma (2015)

suggested in their paper that if we wish the cloud to support all sorts of present and upcoming applications then the gap should be bridged.

The large enterprises seek opportunities to implement cloud computations although they have more issues such as securities. The cloud cost, complications for computations, technology transformation and firm's data security against hackers are the complications against the cloud adaptation [1]. The big data applications in every enterprise are the main motivation for the cloud adoption.

**Big data provides business opportunities**:

Big data is now more than an advertising term. Organizations across the different industries are evaluating different methods to make informed business decisions by exploring raw data [2]. Nowadays, businesses collect huge amounts of data from multiple sources to discover secret patterns to optimize the business process. Mohanty, Jagadeesh, and Srivatsa (2015) mention that big data has brought many success stories for companies by enabling innovative services, inventing new database solutions by improving current solutions for data mining, predictive analytics, prescriptive analytics and machine learning techniques. This result produces many new roles such as data scientists, data engineers, data analysts, business analysts, big data architects, and quantitative analysts.

Businesses utilize big data primarily for marketing analytics, customers' segmentation, and customization [4]. Rehman et al. (2016) analyze different sources of big data that an enterprise can utilize to improve the total business performance. Enterprises obtain data from online websites in the form of customers' feedback and product reviews. Cluster analysis, also a part of big data, is used to segment the customers to optimize business process models that help enterprises to target specific market sectors where they can achieve maximum profit. Furthermore, big data obtained

from supply chain systems can be used to improve resource management and lead time of products and services. Similarly, if an organization analyzes the big data obtained from human resource management systems properly, it would help the organization to devise effective and competitive compensation plan to attract potential employees and to decrease employee turnover rate. Overall, businesses integrate and analyze big data from diverse internal sources to improve their performance.

The large volume of enterprise data (Big data) also plays a vital role in the managerial and strategic process. Kościelniak and Puto(2015) indicate the stages of the informed decision-making process by analyzing the responses given by the Polish business managers in the authors' experiment. By deeply analyzing the four stages of decision-making processes, the authors concluded that those organizations which use data to provide innovative solutions in decision-making processes can address customers' expectations and help stay ahead in the market competition.

**Intelligent enterprise uses big data effectively**:

Modern enterprises use a huge amount of data from internal and external sources. Business operations such as production, supply chain, sales and services, and human resource management are examples of the internal source of data. Whereas the customer-generated data, market segmentation, surveys, customer reviews and ratings, and transactional histories are handled by outbound data sources [4]. Those sums of data, commonly called big data, should be processed in an intelligent way that it leads to a better result with lower investment. For example, an enterprise may extensively use the artificial intelligence and machine learning to analyze and predict the results based on previous knowledge and history. Internet of Things (IoT), big data analytics, and

deep learning is largely used by intelligent enterprises to handle large amounts of data and to draw meaningful conclusions from the data.

The intelligent enterprises make efficient use of machines to reduce the analysis cost without compromising the overall result or performance of the system. The article by Rehman et al. (2016) discusses the "Big Data Reduction" practices for enterprises with a focus on reducing cloud computing cost as well as effective value creation. They emphasize multiple elements of data reduction techniques: a) "dimensionality reduction" which essentially concentrates on the techniques of obtaining valuable information out of millions of data b) "Compression/Decompression" that focuses the methods for reducing data movement within the network and less storage requirement for large datasets c) "Redundancy Elimination" which explains the elimination of duplicate data. In big data domain, knowledge discovery, supervised learning, unsupervised learning, and heuristics discovery have been largely researched and used by intelligent enterprises. Intelligent enterprises are in pace with the advancement in technology, so most of the data processing tasks are automated with a single built script. The analysis and visualization time and cost are also negligible in comparison to old-fashioned data manipulation techniques. So, all the enterprises that hold a large volume of data should use modern tools and techniques for the advanced but understandable result.

**Conclusion:**

In this technological era, the data is an inevitable part of any organization. Every business organization is adopting big data analytics for not only managing their scarce resources but also staying ahead in the market competition. Firms are now transforming their data storage technology from hardware-based systems to cloud servers. This transformation might generate issues such as securities for their business data, but on the other hand, it is cost-efficient. The enterprises are now

becoming intelligent enterprises because they use artificial intelligence and machine learning techniques for handling day to day activities. So, most of the research is focusing on how big data is helping enterprises to gain business insights. Majority of them suggest the importance of big data in managerial and strategic processes and the necessity of integrated data management in businesses. With this analysis, we can conclude that the big data has a significant impact on today's intelligent enterprises. The appropriate use and analysis of big data are improving the effective value creation of the firms. At the same time, the challenges due to the increasing requirements of storage resources, computational capabilities, and networking ability are still needed to be explored and analyzed.

**References:**


[1] Liu, Huan. 2013. Big data drives cloud adoption in enterprise. *IEEE Internet Computing* 17 (4): 68-71.

[2] Mohanty, Soumendra, Madhu Jagadeesh, Harsha Srivatsa, Ohio Library and Information Network, and Inc Books24x7. 2013;2015;. *Big data imperatives: Enterprise big data warehouse, BI implementations and analytics*. 1st ed. Berkeley, California: Apress.

[3] Kościelniak, Helena, and Agnieszka Puto. 2015. BIG DATA in decision making processes of enterprises. *Procedia Computer Science* 65: 1052-8.

[4] Rehman, MHU, V. Chang, A. Batool, and YW Teh. 2016. Big data reduction framework for value creation in sustainable enterprises. *International Journal of Information Management* 36 (6): 917-28.

[5] Srinivasan, Venkat, 1956, and Ohio Library and Information Network. 2016. *The intelligent enterprise in the era of big data*. 1st ed. Hoboken, New Jersey: John Wiley & Sons.

[6] Sukumar, Sreenivas R., and Regina K. Ferrell. 2013. 'big data' collaboration: Exploring, recording and sharing enterprise knowledge. Information Services & use 33 (3-4): 257.

[7] Kaur, Pankaj Deep, and Gitanjali Sharma. 2015. Scalable database management in cloud computing. Procedia Computer Science 70 : 658-67.