African Journal of Computing and Information Systems (AJCIS). Vol. 1 Issue 1 September 2017
*Available online at www.academicinsights.org email: ajcis@academicinsights.org /*
editor@academicinsights.org

# Big Data Analytics; Data-drived Decision Making for Enhanced Enterprise Performance

***Benjamin Kiprono***
*Department of Information Technology, Kaimosi Friends University College*
*Email: bkiprono@kafuko.ac.ke*

**Abstract**

The demand for data driven decision making has consistently grown over the last couple of years. Driven by the desire to compete effectively in a business environment that is ever dynamic; where predicting the behavior of customers and their demand patterns is critical to an enterprise's performance; the dependence on data to provide this information to firms has inherently grown. A key driver for this has been the growth in the volumes of data gathered by firms as well as an increasing acquisition of computer based information systems within organizations. Issues still arise on the slow uptake of these descriptive analytic tools as only a few organizations globally have fully implemented data analytics systems and tools. This paper therefore presents a case for big data analytics for firms. We first analyze the benefits and trends in big data and data analytics and later present a case on why big data analytics is the technology for the future.

**Keywords**: Data analytics; big data; predictive analytics; data mining; knowledge discovery.

## Introduction

The Big Data phenomenon creates tremendous opportunities for society to foster data-driven innovation that can enable faster and better decision-making, building a competitive advantage for our knowledge society [1]. When big data is distilled and analyzed in combination with traditional enterprise data, enterprises can develop a more thorough and insightful understanding of their business, which can lead to enhanced productivity, a stronger competitive position and greater innovation – all of which can have a significant impact on the bottom line [2]. To derive real

business value from this data, an organization needs the right tools to capture and organize a wide variety of data types from different sources, and to be able to easily analyze it within the context of all your enterprise data. This is made difficult largely because of the unstructured form in which this data exists in and by the fact that not all this data is relevant to the firm. The first task therefore for any firm is to make sure that whatever data the firm stores is relevant to the enterprise. Various tools exist in the market that helps a company acquire and organize their data and also dig into this vast data to find meanings and relationships that are relevant to the firm. A good example is Oracle big data appliance and oracle big data connectors' tools that provide a complete and integrated solution to address the full spectrum of enterprise big data requirements [2].

**Defining big data**

According to Big Data can be defined as the data that is so big that it cannot be analyzed or mined for information using the traditional data processing application tools, traditional database and software technologies owing to its big size.

There is no properly agreed definition for big data. IBM [3] opines that this is based on two main reasons:

- The name itself is inaccurate; data does not have to be voluminous to be classified as big data.
- There is a large perception/misconception of what it is, based largely on vendor hype, and sensationalist news stories.

IBM therefore defines big data as "Extracting insight from an immense volume, variety, and velocity of data, in context, beyond what was previously possible." This is based on that too much data is being captured daily, yet this data is meaningless and useless unless insights can be derived from it. Because data has become so voluminous, complex, and accelerated in nature, traditional computing methods no longer suffice.

Companies capture trillions of bytes of information about their customers, suppliers, and operations, and millions of networked sensors are being embedded in the physical world in devices such as mobile phones and automobiles, sensing, creating, and communicating data [4]. Big Data is defined by [1] as huge data sets that are orders of magnitude larger (volume); more diverse, including structured, semi-structured and unstructured data (variety) and arriving faster (velocity) than any organization has had to deal with before. [5] defines big data as data that cannot be processed using traditional data processing tools and processes; while [4] defines "big data" as large pools of data that can be captured, communicated, aggregated, stored, and analyzed; [6] suggests that big data is a collection of data sets that are so large and complex that software systems are hardly able to process them.

The datasets are larger; more varied in structure and in format, and are generated at a faster rate than ever before. The Big Data phenomenon creates tremendous opportunities for the society to foster data-driven innovation that can enable faster and better decision-making, building a competitive advantage for our knowledge society. Big data reflects how an organization identifies, analyzes, and uses new sources of information to make decisions about information that until recently was too difficult, expensive, or time consuming to access.

According to [1] and [3], four key characteristics define big data, these are;

- **Volume**: Machine-generated data is produced in much larger quantities than non-traditional data.
- **Velocity**: this data is generated in real time, with demands for usable information to be served up immediately.
- **Variety**: Traditional data management processes can't cope with the heterogeneity of Big Data or shadow or dark data, such as access traces and Web search histories.
- **Value**: The economic value of different data varies significantly. Typically there is good information hidden amongst a larger body of non-traditional data; the challenge is identifying what is valuable and then transforming and extracting that data for analysis.

These vast amounts of data are generated by interconnected devices – from PCs and smart phones to sensors such as Radio frequency identification (RFID) readers and traffic cams. The growth of ubiquitous computing has also had profound effect on the gathering of data. Household devices such as fridges, microwaves and even digital television sets can gather data from their environment; analyze it and effectively provide relevant information to their users. This data is also heterogeneous and comes in many formats, including text, document, image, video, and more. Big data reflects not just how an organization identifies, analyzes and uses the data managed within its own walls, but also data that was previously considered inaccessible, including data from new sources of information that may lie outside the control of an organization, to make business decisions [7].

To make the most of big data, enterprises must evolve their IT infrastructures to handle these new high-volume, high-velocity, high-variety sources of data and integrate them with the pre-existing enterprise data to be analyzed [2]. If this is not done, then enterprises may lose the overall benefit that would have been generated from this data.

**Challenges of big data**

The big question for every firm today is how to turn big data into insights and insights into business advantage. A survey carried out by KPMG in the united states [8] found out that implementing the right solutions to accurately analyze and interpret data; identifying the right parameters and reacting in a timely fashion to insights as they arise were some of the main challenges that respondents faced when it comes to turning vast amounts of data into valuable, practical and actionable business intelligence using analytics. [9] identifies five key challenges to big data, these are: Uncertainty of the data management landscape; The big data talent gap; Getting data into the big data platform; Synchronization across the data sources and Getting useful information out of the big data platform.

There is unquestionable evidence that big data presents a myriad of challenges to organizations and to the users of this data. If these issues are not handled well, the results would generally compromise the quality of the information that is generated by these systems. Solutions to this should start right from the beginning. At this point, mechanisms must be put in place to ensure the overall consistency of data to ensure the overall integrity of the information generated.

**Big data analytics**

Big data analytics is the process of collecting, organizing and analyzing big data to discover patterns, relationships between the data and other useful information that can support an organization's decision making process. Big data analytics helps organizations better understand the information contained within the data and will also help identify the data that is most important to the business and future business decisions. Analysts working with big data basically want the knowledge that comes from analyzing the data.

Like conventional analytics and business intelligence solutions, big data mining and analytics tools helps uncover hidden patterns, unknown correlations, and other useful business information hidden within an organization's data. However, big data tools can analyze high-volume, high-velocity, and high-variety information assets far much better and more efficiently than conventional tools and relational databases that struggle to capture, manage, and process big data within a tolerable elapsed time and at an acceptable total cost of ownership (www.qubole.com).

**Benefits of big data analytics**

It is believed that Google and LinkedIn were among the first organizations to experiment on the big data by developing proof of concept and small scale projects to learn if their analytical models could be improved with new data sources. The results of these experiments were largely positive. Today, big data analytics has moved behold experimentation into an approach major organizations use to gain insight into their data. Tom davenport[1] the IIA director of research identifies a number of benefits to firms based on data collection from over 50 firms in the US. These include:

- Faster and better decision making: the major aim of analytics is to improve decision making. Analytics enables both large and small organizations to make faster and better decisions based on big data.

[1] Tom Davenport is an analytics industry pioneer and serves as the IIA Director of Research and faculty leader. He is a Visiting Professor at Harvard Business School, a distinguished professor at Babson College, a research fellow of the MIT Center for Digital Business, and a senior adviser to Deloitte Analytics.

- Development of new products and services: Using insights from data analytics, organizations for long have been able to develop new products and services based on analysis of customer data/information. This includes analysis of web data from web-based systems, customer relationship data, product information databases and market trend analysis.
- Reduction on costs: Big data tools like Hadoop allow businesses to store massive volumes of data at a much cheaper price tag than a traditional database. Companies utilizing big data tools for this benefit typically use Hadoop clusters to augment their current data warehouse, storing long-term data in Hadoop rather than expanding the data warehouse. Data is then moved from Hadoop to the traditional database for production and analysis as needed. Versatile big data tools can also function as multiple tools at once, saving organizations on the cost of needing to purchase more tools for the same tasks[2].

**Big data for competitive advantage**

The potential for big data has grown exponentially in the last few years. By analyzing the data and gaining insights from it, business enterprises can improve the business operations significantly, optimize customer engagement, capitalize new sources of revenue and make informed decisions. Insights from big data can also be used to prevent fraud and threats to your businesses and customers. Since big data is high-volume, high-velocity and high-variety information asset, innovative and cost-effective methods must be employed to process information to offer better insight and help in the decision-making process.

**The future of big data**

The capabilities of big data hold great promise in the future. As competition grows, and customer sophistication increases, big data will offer the much needed customer and market intelligence that every firm will require to drive its competitiveness. The capability to predict customer behavior will

[2] **Hadoop** is an open-source framework that allows an organization to store and process big data in a distributed environment across clusters of computers using simple programming models. It is designed to scale up from single servers to thousands of machines, each offering local computation and storage.

enable firms to offer the right products to the customers at the right time, place and in the right quantity. This is a major capability that big data analytics is bound to offer.

The big question is, what does big data hold for the future? Will it continue to grow or is it going to fade as time goes? According to [10], Many market leaders are already using big data and big data analytics in ways that seem futuristic to their lagging competitors. According to them, most questions about the future of big data are often questions about realizing the value of big data as quickly as possible. They thus argue that big data for the future will aim to look at; the monetization of big data; who will use big data; what new business problems big data will solve and how big data will drive better and faster performance management models. [11] opines that the volume of data gathered by firms will grow as data analytics improves; machine learning will increase as data as-a-service will become more popular. He also opines that security and privacy challenges will be enhanced and privacy controls and procedures will be revised to handle the emerging challenges. [12] predicts that the future of big data lies on the internet of things (IOT) this he argues is because the sheer size and speed of the data collected when every device involved in your business process is online, connected and communicating can strain the sturdiest network infrastructure. To solve this, edge technology will continue to gain popularity. This will enable analytics to be performed at the point or close to where the data is captured. With this, cases of dark data will be eliminated instead of storing large files with data that has no business sense. This will also enable real time response to situations especially in the case of CCTV cameras. It is clear therefore that the future of big data seems promising. The continued reliance on information systems and data driven decision making will continually create a market for big data analytics. This will therefore enable continued dependence on data analytics for business decision making support.

**References**

[1] Intel, "Big data policy position paper," 2012.

[2] Oracle, "Oracle: Big Data for the Enterprise," Oracle, Redwood, White Paper 2013.

[3] IMB, *Information Governance principles and practices for a big data landscape*, First edition

ed. New York, Unites States: IBM Red books, 2014.

[4] MCKINSEY, "Big Data: The Next Frontier For Innovation, Competition and Creativity," MCKINSEY Global Institute, New York, 2011.

[5] Paul Zikopoulos, Chris Eaton, Diroos Dirk, Tom Deutsch, and George Lapis, *Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data*, 1st ed., Steve Sit, Ed. New York, United States: MCGRAW-HILL, 2012.

[6] Marc J. Schniederjans, Dara G. Schniederjans, and Christopher M. Starkey, *Business Analytics Principles, Concepts and Applications*, 1st ed., Gill Editorial Services, Ed. New Jersey, Upper Saddle River: Pearson Education, Inc., 2014.

[7] Viewpointe, "dark data, big data, your data: Creating an action plan for information governance," United states, 2013.

[8] KPMG, "Big data top challenges," KPMG LLP, Delaware, 2015.

[9] David Loshin, "Addressing five emerging challenges of big data," www.progress.com, white paper N.D.

[10] www.teradata.com, "The future of Big Data," *www.teradata.com/bigdata*, extracted on 17/11/10 2016.

[11] Kevin Patel, "Taking a Peek into the Future of Big Data," *www.insidebigdata.com*, November 2016.

[12] Bernard Marr, "Will 'Analytics On The Edge' Be The Future Of Big Data?," *Forbes.com*, no. Extracted online on 17/11/2016, August 2016.

[13] IBM, "Business analytics for banking: 3 ways to win," Ottawa, 2010.

[14] Gartner, "A step-by-Step Approach to Successful Business Intelligence," 2011.