Big Data: A Tool for Development in Developing Nations

Suzzie Adobea Apenteng

Bachelor of Engineering in Computer Engineering with the All Nations University College in Koforidua Ghana

Abstract- Big Data in recent years has become a major topic in the field of Information and Communication Technology. Big Data, which in simple words stands for massive/huge data, unmistakably means big opportunities when utilized in the right ways. Many countries in the developed regions have already begun intensive research on how best they can use Big Data as leverage in growing various sectors of their economy. Whiles most research about big data have been focused on the volume aspect of it, this paper focuses on the value aspect to the developing nations. Countries in the developing regions should not sit by and allow countries in the developed regions to take advantage of Big Data especially in this global economy. Big Data has come to stay hence the need for developing nations to tap into this buzz to bring about massive developments in these regions.

Index Terms- Big Data, Developing Nation, and Volume of Big Data and Value of Big Data

I. INTRODUCTION

Big Data has become the word or term of the day. One should know that Big Data does not necessarily means good data. Big Data could be good and bad at the same time. The worthiness of big data in decision and policy making is very critical hence the need for whoever using big data to process or analyze to extract the needed or critical information which can be employed to make a highly informed decision in various sectors of the economy including energy, healthcare, and education.

II. WHAT IS BIG DATA?

In recent times, there has been a buzz about Big Data. Many researchers and research organizations have come up with several definitions, ideology and theories concerning big data. Software developing companies have also begun looking into developing tools to analyze and extract relevant information from Big Data. In the developed continents like the North America and the Europe's Big Data is being utilized everywhere ---- from the healthcare to policy making and many more. The [1]McKinsey Global Institute (MGI) defines Big Data as large pools of data that can be captured, communicated, aggregated, stored, and analyzed. [2]ISACA also defines Big Data as a trend in technology that is leading the way to a new approach in understanding the world and making business decisions. Another definition of Big Data is that by [3]Edd Dumbill, which is data, that exceeds the processing capacity of conventional database systems. [18]Wikipedia also defines Big Data as the term for a collection of data sets so large and complex that it becomes difficult to process using on-hand database management tools or traditional data processing applications however for the purpose of this study, Big Data will be defined in the most simplistic way as the enormous or massive data available to utilize for national development.

III. ASPECTS OF BIG DATA

Cecere[4] talks on the fact that data keeps on piling up in growing volume, velocity and variety on corporate doorsteps and believes that big data offers new opportunities for the corporation to listen, test and learn, and respond faster.

Figure 1: Aspects of Big Data Source: SAS[17]

The seizure of this new opportunity will require a new form of leadership because it can ignite new business models. However, the ignition of new business models cannot just happen by big data itself but through initiatives that must be aligned to business objectives. NESSI[8] elaborates on the quantum of available data. NESSI[8] believes the amounts of data exploded in the past years were because of new social behaviors, societal transformations as well as the vast spread of software systems. Furthermore, NESSI[8] also comments on how big data have become an important driver for innovation and growth, thus very important to foster productivity growth in Europe since it will not affect only software-intensive industries but also public services like the health, administration and education sectors.

The table below summarizes the main aspect and challenges connected to handling different types of large data sets, and explains how Big Data technologies can help solve them according to NESSI[8].

Table 1: Category of the different types of data. Source: NESSI[8]

Aspect	Characteristics	Challenges and Technology Response
Volume	The most visible aspect of Big Data, referring to the fact that the amount of generated data has increased tremendously the past years. However, this is the less challenging aspect in practice.	The natural expansion of internet has created an increase in the global data production. A response to this situation has been the virtualization of storage in data centres, amplified by a significant decrease of the cost of ownership through the generalization of the cloud-based solutions. The noSQL database approach is a response to store and query huge volumes of data heavily distributed
Velocity	This aspect captures the growing data production rates. More and more data are produced and must be collected in shorter time frames.	The daily addition of millions of connected devices (smartphones) will increase not only volume but also velocity. Real-time data processing platforms are now considered by global companies as a requirement to get a competitive edge
Variety	With the multiplication of data sources comes the explosion of data formats, ranging from structured information to free text.	The necessity to collect and analyze non-structured or semi-structured data goes against the traditional relational data model and query

		languages. This reality has been a strong incentive to create new kinds of data stores able to support flexible data models
Value	This highly subjective aspect refers to the fact that until recently, large volumes of data where recorded (often for archiving or regulatory purposes) but not exploited.	Big Data technologies are now seen as enablers to create or capture value from otherwise not fully exploited data. In essence, the challenge is to find a way to transform raw data into information that has value, either internally, or for making a business out of it.

In Zikopoulos et al[7] book on Understanding Big Data, we come to realize that IBM uses V[3] (Velocity, Variety and Volume) excluding Value as described by NESSI[8]

Irrespective of what both Zikopoulos et al[7] and NESSI[8] say about the aspect of Big Data, one can certainly understand why Big Data is trending in the world of IT. SAS[17] also makes it clear that the real issue here is not about how you can acquire large amounts of data but rather what you do with the data that counts. This tends to agree with NESSI[8] on the inclusion of Value as an aspect of Big Data.

IV. WHO IS A DEVELOPING NATION?

A developing country, also called a less-developed country, according to Wikipedia is a nation with a lower living standard, underdeveloped industrial base, and low Human Development Index (HDI) relative to other countries. A key word from this definition is the Human Development Index, which according to Wikipedia again is a composite statistic of life expectancy, education, and income indices used to rank countries into four tiers of human development.

From the figure below, the majority of the African countries have a low HDI and few like Ghana and South Africa are on a medium HDI. One can say that if any country from this category begins to see the various opportunities and benefits of Big Data and decide to tap into this recent IT Buzz, the country in the long run, in its process of using Big Data for the development of its economy will eventually take the country's HDI a step forward.

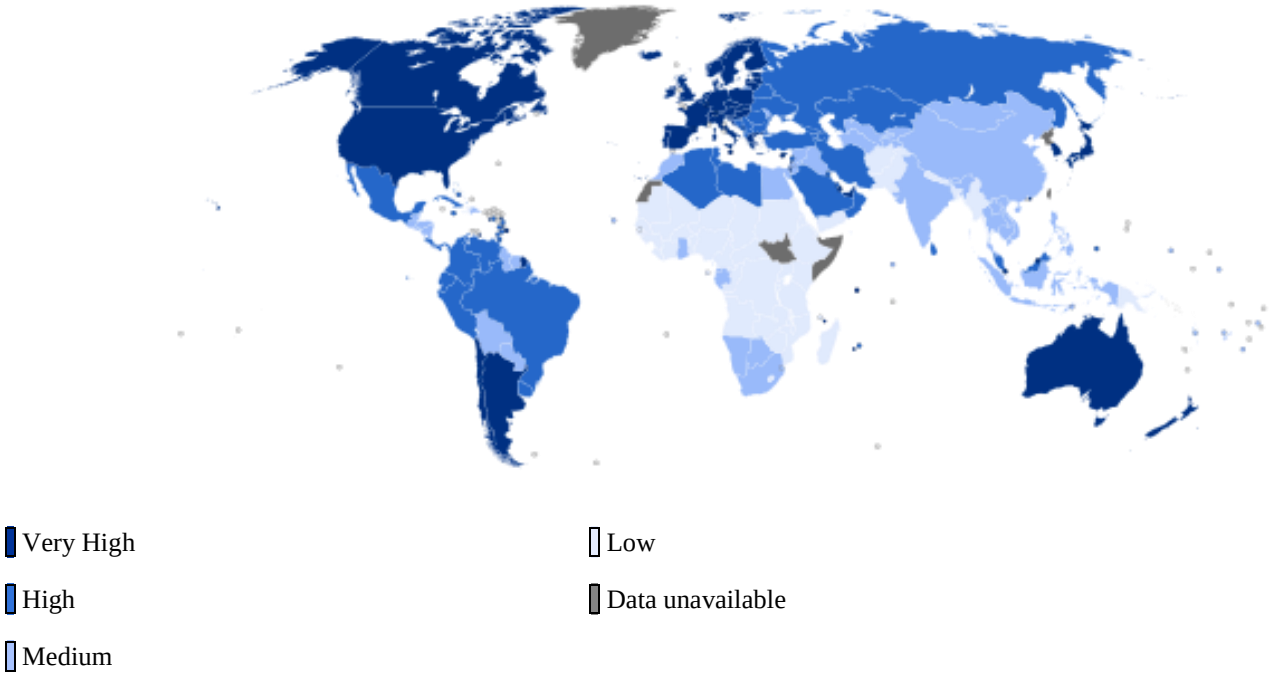

Figure 2: World map by quartiles of Human Development Index in 2013.
Source: Wikipedia[19]

V. WHY BIG DATA FOR DEVELOPMENT IN A DEVELOPING NATION?

A. Education

Big Data has the potential to revolutionize not just research, but also education[9]. Lets imagine being in a world where we have unrestricted access to a very huge database with zillions of data concerning student's academic performances over the years. What can we do with such massive data? We can quite a lot with it by picking on the value aspect of the data which could be used to design one of the most effective approaches to education, which could start from reading, math, to an advanced level such as the tertiary etc. These available data extracted will not just benefit the student's but also help the ministry or department in-charge of education to know how to direct its resources and train the educationists in its country.

B. Health Sector

The health of every nation's citizen is critical to the growth of the economy. Healthy citizens are likely to bring about growth in production. In the US economy, health care is a large and important segment according to MGI. The health sector of the US economy, like many other economies faces tremendous challenges. Every health sector has multiple and varied stakeholders ranging from the pharmaceutical to the patients. Although the many sectors under healthcare seem to be intertwined, they all have different interests and business incentives. For example, the interest of the National Health Insurance Scheme (NHIS) of Ghana is definitely different from that of the Ghana Medical Association although both have a common goal that is to provide better and quality health services to the patient. In the era of big data, an economy must take an advantage of the massive information in growing its health sector in infrastructure wise, research on diseases and medical instrumentation wise as well.

C. Security

The security of a country is very critical. In a country full of insecurities, investors find it very hard to invest in that country and even the indigenes also find it the same causing many indigenes to migrate to other countries. I believe that if the security forces and the government use big data to extract the many digitized messages from the very huge datasets, it will help the government and its security council members come up with better strategies and plans to keep its citizen safe and the nation a harbor for investments.

D. Public Sector

According to MGI[1], the public sector can be put into the category of being a large part of the global economy facing tremendous pressure to improve its productivity. When the government / the public sector realize that they have access to large pools of data of which they can take advantage of its powerful ways to improve performance and transparency, they will be awe-stricken.

E. Manufacturing

Big data can be used in this arena to examine the multiple points in the value chain be it from the beginning of acquiring and bringing the products to the market to after-sales. Big data can be used for research and development (RD) that helps to know the market trend, the people's taste for specific goods and many more.

F. Transportation & Energy Sector

Big Data can also be used to revamp the transportation industry in most developing countries. The energy sector of an economy is trivial to the survival of the economy in respect to industrialization. If an economy decides to take advantage of Big Data to grow this sector of the economy, then the country is headed for a very strong and robust industries and investments.

G. Telecommunication

Another area that could benefit from big data is the telecommunication sector. With an increase in the penetration rate of mobile phones in many developing countries such as Ghana who has a 63% penetration rate per 100 habitants, big data should not be ignored[5]. Many telecommunication industries in the developing countries can use big data to determine the trend of how its consumer uses his or her mobile phone. A consumer could use it for communicating with friends and families only while others use it for mobile money transfer and other stuffs altogether.

Table 2: Mobile phone penetration measured by the number of mobile phones per 100 habitants. Source: Google Fusion Tables[5]

Country	Penetration Rate (%)
Botswana	96
Ghana	63
Mauritania	66
Kenya	49
Nigeria	47
Angola	44
Tanzania	40

VI. CONCLUSION

According to the McKinsey Global Institute Report on Big Data, the most developed regions such as Europe have the biggest potential to create value through the use of big data. However in my opinion, the developing regions can tap into the many opportunities available through the use of big fata to create value and in the long run help in the development of our continent.

REFERENCES

[1] McKinsey Global Institute, " Big data: The next frontier for innovation, competition, and productivity (Report Style)." 2011.

[2] ISACA. "Big Data Impacts and Benefits (Whitepaper)." March 2013.

[3] E. Dumbill. "What is Big Data: An Introduction to the Big Data Landscape (Article)." Strata Oreilly, January 11, 2012. http://strata.oreilly.com/2012/01/what-is-big-data.html

[4] L. Cecera, "Big Data Handbook: How to Unleash the Big Data Opportunity (Report Style), " 2013.

[5] UN Global Pulse. "Big Data for Development: Challenges & Opportunities." New York. May 2012

[6] H. G. Miller and P. Mork. “From Data to Decision Value Chain of Big Data.” IEEE Computer Society, 2013. pp. 57 - 59

[7] P.C. Zikopoulos, C. Eaton, D. DeRoos et al. “Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data (Book Style).” McGraw-Hill, 2012.

[8] NESSI. “Big Data: A New World of Opportunities (White Paper).” 2012.

[9] D. Agrawal, P. Bernstein et al. “Challenges and Opportunities with Big Data (Whitepaper).”

[10] Y. Noguchi. “The Search for Analysts to Make Sense of Big Data.” National Public Radio. November 30, 2011. http://www.npr.org/2011/11/30/142893065/the-search-for-analysts-to-make-sense-of-big-data

[11] S. Lohr. “The Age of Big Data (Article).” New York Times, February 11, 2012. http://www.nytimes.com/2012/02/12/sunday-review/big-datas-impact-in-the-world.html

[12] S. Matteson. “Big Data Basic Concepts and Benefits Explained (Article). “ TechRepublic, September 25, 2013. http://www.techrepublic.com/blog/big-data-analytics/big-data-basic-concepts-and-benefits-explained/

[13] B. Gregory. “Finding relevance in big data and using it effectively (Article).” Humanipo, September 20, 2013. http://www.humanipo.com/news/32483/finding-relevance-in-big-data-and-using-it-effectively/

[14] A. Tatevossian. “ ‘Big Data’ for development: What is it, and why you should care (Article).” Devex, July 16, 2013. https://www.devex.com/news/big-data-for-development-what-is-it-and-why-you-should-care-81453

[15] V. Gopalkrishnan et al. “Big Data, Big Business: Bridging the Gap (Deloitte Analytics).” Deloitte & Touche Financial Advisory Services Pte Ltd. 2012.

[16] J. M. Tien. “Big Data: Unleashing Information (Paper).” J Syst Sci Syst Eng., June 2013

[17] SAS Institute Inc. “Big Data Meets Big Data Analytics (Whitepaper).” 2012.

[18] http://en.wikipedia.org/wiki/Big_data retrieved on 7th April, 2014

[19] http://en.wikipedia.org/wiki/Human_Development_Index retrieved on 7th April, 2014

[20] http://en.wikipedia.org/wiki/Developing_country

AUTHORS

First Author – Suzzie Adobea Apenteng holds a Bachelor of Engineering in Computer Engineering with the All Nations University College in Koforidua Ghana. Suzzie is a Bronze Member of the Information Systems Audit and Control Association (ISACA). Suzzie’s research interest lies in Big Data, Gender Related Issues in Information Technology and Technology for Development geared towards the Developing Nations.