Achieving Competitive Advantage Through Big Data.
Strategic Implications

[1]*Vincent Charles* and [2]*Tatiana Gherman*

[1]CENTRUM Católica Graduate Business School, PUCP, Lima, Peru,
[2]Universidad del Pacífico, Lima, Peru

Abstract: Big data is, without doubt, a hot topic nowadays, moreover because the development of new technology makes it possible to analyze all available ever-growing data which easily amasses terabytes of information. However, big data poses both a big opportunity and a big challenge at the same time, as the power does not truly rely on processing it anymore, but rather on how it can be transformed into meaningful knowledge that can be used to make intelligent decisions. The big data age requires a big data mindset. Hence, the implications of big data are, not to say the least, revolutionary and in this article, we discuss the meaning of big data and argue over its strategic implications for the firm. We believe that in order for big data to become an asset and serve business effectively, trying to define the dimensions of big data is of utmost importance. In consequence, the present article presents a perspective which combines the conventional and more technical dimensions of big data with the more challenging dimensions of context, connectedness and complexity. It is argued that by understanding this expanded perspective, decision makers can much better appreciate the value added which big data has to offer and how it can become a competitive advantage for the firm.

Key words: Biodata · Opportunity · Dimentions · Decision

INTRODUCTION

One should try doing a short exercise: just Google "Big Data is…" and realize the suggestions given by the latter. Some of them include: *Big Data is the new oil*, *Big Data is the future*, *Big Data is a big deal*, *Big Data is big business* and so on. There is no doubt that *big data* is one of the new buzz words nowadays, as recently, the IT industry and not only, has begun to extensively use the term big data in the context of the increasing data-rich landscape of today. Moreover, it is very interesting to note that the data scientist position has been appraised as *the sexiest job of the 21st century* by the Harvard Business Review [1]. However, when it comes to big data, there is still much research to be done.

But what does big data actually mean? Usually, big data, also known as evidence-based decision-making [2], refers to the ability to collect and analyze vast amounts of data generated worldwide. [3] stated the following: "Big data are high volume, high velocity and/or high variety information assets that require new forms of processing to enable enhanced decision making, insight discovery and process optimization." And according to [4], big data can be described as extremely large data sets which are difficult to manage with conventional information technologies, especially when trying to analyze or visualize the data. The reality is that over the past few decades the amount of data available captured by corporations, but also by the social media and Internet (e.g. Facebook, Twitter, Google, e-bay, Amazon, census data, etc.), has increased exponentially.

The size of these sets has been continually growing over the years and is becoming more difficult to analyze. IBM stated that 2.5 quintillion bytes of data are created every day globally and 90% of the data in the world has been created in the last two years alone [5]. By way of example only, Facebook processes 10 Terabytes of data every day, Twitter processes 7 Terabytes and Google processes more than 24 Petabytes; there were 2 billion internet users in 2011 and it is expected that annual internet traffic will increase by 667 Exabytes by the end of 2013.

Corresponding Author: Vincent Charles, CENTRUM Católica Graduate Business School, PUCP, Lima, Peru.

All of these known companies and many more have access to vast amounts of data, which are generated daily by their clients. Facebook, for example, has over one billion active users worldwide [6] posting what they like and what they do every day and Amazon has 282 million visitors a month to their web page [7], leaving information about what products they like and how often they buy it. So, knowing how to analyze these data and discovering all the information hidden inside becomes a very important goal for the companies; knowledge becomes power which companies can leverage to gain a competitive advantage. But, of course, the question is: how can they do it?

Ever since the massification of internet access and the rise of e-commerce, companies have gained increasing amounts of details about their sales, manufacturing, logistics, all-sort-of-costs, but most importantly about their clients (which goes from their personal information to their shopping preferences and driving behavior), just to mention a few. However, in many cases, these data are not properly managed and the opportunity to harness information as a competitive catalyst to create better products, to make better marketing campaigns, or to gain a 360º view of the customers is lost. In this context, "the ability to harness the ever-expanding amounts of data is completely transforming our ability to understand the world and everything within it" [8]. Big data analysis improves the decision making process to make better, more frequent, timelier and more granular fact-based decisions. Big data is important as it allows to understand not only the value of collecting the data, but also the value of keeping it because it can be reused for a numerous of secondary purposes.

It is vital to understand the very meaning of the word big data, as it is essentially a misnomer, comprising two very important but different aspects. On one hand, it denotes huge data sets which come in volumes and varieties never seen before and which cannot be analyzed or processed by means of the existing conventional tools, requiring, therefore, the development of new technology. On the other hand and most importantly, when talking about big data, it is not the actual volume of data that represents the main issue in discussion, but rather the phenomenon that we are trying to record and understand and the hidden patterns and complexity of the data itself. In consequence, the definition of big data encompasses any data set that is too complex in the sense that is lacking a pattern and which is too large to be managed in traditional relational databases. The difficulty and the real challenge concerns capture, storage, analysis, handling, viewing and sharing of data, inter alia, as we are dealing mainly with unstructured data sets, in other words, with a complete chaos. And most importantly, 80% of world´s data is unstructured [5].

Each industry is now capable of using the information in ways that was not able to before because of technology limitations and can benefit from using big data for an improved decision-making. But even with available technology, big data poses a big opportunity and also a big challenge, as the power does not fundamentally rely on processing it, but rather on how one extracts value from it. The potential to leverage big data is unlimited and whatever the debate over the definition, there is no doubt that the implications of big data are revolutionary and will lead to new opportunities in the coming years. In the big data age, there is only to win for people with the right big mindset.

The Dimensions of Big Data: By reviewing the existent literature, it was found that big data spans four dimensions:

Volume: The vast amounts of data generated every second, which includes both simple and complex analytics and which pose the challenge of not just storing it, but also analyzing it [9].

Velocity: The speed at which new data is generated and moves around as compared to the time window of translating it into intelligent decisions [9]. Consider the importance of this particular dimension when taking into account that "for time-sensitive processes such as catching fraud, big data must be used as it streams into your enterprise in order to maximize its value" [10].

Variety: The increasingly different types of data, structured and unstructured, from a large number of diverse data sources (e.g. from financial data to social media feeds, from sensor data to click streams and log files) [9]. Big data refers mainly to unstructured datasets, which are not owned by a company and which can be generated anonymously or not, both by users via web, but also by sensors, cameras, monitoring solutions, equipment, etc. in the most various formats and standards.

Veracity: Statistics show that 1 in 3 business leaders don't trust the information they use to make decisions [5]. It is quite important to note that all too often the data can be a complete mess, some data being intrinsically inaccurate and other data becoming inaccurate through processing errors [11].

The Competitive Advantage of Big Data: Companies in almost every industry are focused nowadays on exploiting data for competitive advantage, on how to extract value from it. The volume and various aspects of data have so far outstripped the capacity of manual analysis and conventional databases, in which case new platforms and big data tools are being developed. Moreover, the development of new algorithms makes it possible to connect datasets in such a way that broader and deeper analyses are possible when compared to past situations.

One of the most notable advantages of big data in general is related to the fact that all observations can be used in the analysis. In the *small* data age, we might have used an expensive sampling technique, but now we do not have to sample anymore. In consequence, one does not need to have pre-established hypotheses. By having recorded all the data points there are about a phenomenon, one can run regression analyses and change the hypotheses accordingly.

"Large-scale data gathering and analytics are quickly becoming a new frontier of competitive differentiation" [12]. In other words, the analysis of large data sets will establish the path for future competition among organizations, fostering productivity, growth and innovation. What each company chooses to do with the data it generates can make the difference between progress and stagnation. However, not too many business leaders truly understand the potential that big data poses over their companies. The following lines pretend to draw attention upon the most important benefits resulting from the use of big data, in an attempt to highlight its true potential and, therefore, competitive advantage.

If history has taught us something is that in general, executives take action and try to develop novel strategies for their organizations only when those organizations pass through a crisis or a threat, but not before. The same rule seems to apply to the use of big data. Even though the users of big data have increased over the past few years, most big data strategies have erupted at times of sorrow. Hence, leaders and decision makers must understand that big data allows taking action before the organization undergoes a crisis.

Business generates huge amounts of data, which up until few years ago it required major investment to be archived. And even when archiving did not represent a problem, still there would be the issue of only being able to analyze a small percentage of the data, not fully taking advantage of its potential. Today´s advancement in technology has solved this difficulty, as companies are able, for example, to use the collected data and conduct controlled experiments to make better management decisions; moreover, the sophisticated analytics can considerably improve the decision-making process. Hence, the only thing left is for businesses to decide what they want to find out. Data monetization is the promise of big data and it represents an advantage for both small and large companies.

As previously mentioned, Harvard Business Review [1] has rated data scientists as the sexiest job of the 21st century and this can only be due to one reason: the scarcity of data scientists. As [1] clearly stated,

"If capitalizing on big data depends on hiring scarce data scientists, then the challenge for managers is to learn how to identify that talent, attract it to an enterprise and make it productive. None of those tasks is as straightforward as it is with other, established organizational roles. Start with the fact that there are no university programs offering degrees in data science. There is also little consensus on where the role fits in an organization, how data scientists can add the most value and how their performance should be measured."

According to a McKinsey Global Institute report on big data [13], the United States alone will need over 1.5 million more data-savvy managers and analysts to take advantage of big data. And this is a big thing already.

However, in order for big data to turn into a competitive advantage of a firm, one needs to take into consideration the wider context in which the data were created, as too many times companies gather data without a thorough understanding of it. Or even worse, too many analyses have created stories including certain data and discarding other data, which led to consider the conclusions unreliable. One can take as an example the quite recent financial crisis, when "it became clear that banks and rating agencies had been relying on models which, although they required a vast amount of information to be fed in, failed to reflect the financial risk in the real world" [14]. Hence, a significant judgment is necessary to decide what is right and what is wrong in this new era of data analytics and one must remember at all times that data itself isn´t the solution, but rather the means to obtain the solution [2].
Also, as [11] stated,

"It's natural to expect a return from our investment in data and analytics, but there's a sneaky side effect. People will "find" what they think you want. Saying upfront that

you expect a 10% uptick in revenue can cause people to find a short-term 10% growth that's not there for the long-term, to be so busy looking for the 10% that they'll miss a potential 100% gain, or miss negative correlations entirely."

Hence, it is important to note the ethical challenge that big data raises. As the internet is being invaded by more and more technological applications, the only thing one has to do is to fill in some numbers and the program gives you the results you are looking for. Hence, one does not even have to know about basic statistical analysis; the complex algorithms do all the work. But in this context, big data raises a big problem:

The ability of lots of people to run regressions (or plug in values into web applets) also raises the real possibility that a lot of shoddy crunching is going to be done. A little knowledge can be a dangerous thing. Just because you can get a statistical program to run a regression and produce a prediction, doesn´t mean that the result will be valid. The key question is: valid compared to what? Even crude statistical analysis is likely to do at least as well as traditional experts in making predictions. But the weekend Super Cruncher isn´t likely to do as well as people who have solid social science or statistical training. In the world of democratized data mining, decision makers all the more will want to seek out Super Crunching audits to make sure that their statistical model hasn´t somehow gone terribly wrong [15].

Maybe now it could be the right time to consider that the big data age might require the creation of a new World Trade Organization (WTO) which would oversee the flow of digital goods and services.

The challenge of big data is to capture information about market, customers, products and services and to analyze it with the objective of gaining a competitive advantage. For now, companies can keep performing their normal activities without having to worry too much about the phenomenon of big data, but if they want to know what the market thinks about their products and where their active customers are, then the approach should be changed. Big data should be seen as an opportunity, a catalyst for a marketing strategy based on data accuracy once considered impossible. Only by taking into account that about 90% of the data available in the world today has been created in the last two to three years, the picture of the near future becomes more blurred. However, one thing is clear: in the future, big data will allow companies to obtain answers to questions now considered impossible.
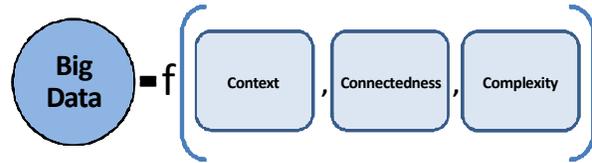

Fig. 1: The challenging dimensions of Big Data

The Big Data Model Expanded: The previously mentioned ideas helped us derive the following logical model (see Fig. 1) for big data, which goes beyond its four more conventional and technical dimensions of volume, velocity, variety and veracity. As such, big data becomes a function of complexity, connectedness and context, being these, the actual challenging dimensions brought out by big data and the ones that would assist the decision maker in converting big data into a competitive advantage.

Context: Defined as the ability to understand and to successfully respond to the new challenges posed by the massification of Internet access, as big data should not start with the creation of new technology, but rather with tackling existing business problems. Moreover, understanding the context is essential when dealing with big data, because raw data could mean anything without a thorough understanding of the context that explains it and turns it into a story about the firm´s performance, as previously stated. The context allows big data to convert into knowledge, i.e. meaningful information that can be used to make intelligent decisions.

Connectedness: Defined as the ability to understand data in its wider context and within its ethical implications. In an increasingly connected world, a company which organizes its internal data even with the best business intelligence software, might not achieve its true potential if its data remains isolated from the rest of the data universe, including news, social media and all other relevant sources. Information managers must fundamentally rethink their approach to data, engaging and building effective relationships among all the dimensions of information management.

"The business's demand for access to the vast resources of big data gives information managers an opportunity to alter the way the enterprise uses information. IT leaders must educate their business counterparts on the challenges while ensuring some degree of control and coordination so that the big-data opportunity doesn't become big-data chaos, which may raise compliance risks, increase costs and create yet more silos" [16].

Complexity: Defined as having the skills to survive and thrive in the face of complex data, which will become a core competence of enterprises that are increasingly using new forms of information in their attempt to look for patterns that support their business decisions. It refers to acquiring the abilities to identify the key data and differentiate the information that truly has an influence on the organization. Also the organization faces the challenge of understanding how its decisions impact on the stakeholders affected by the use of the information, both positively and negatively. And probably, one of the important abilities related to the complexity of the use of data refers to being able to learn from mistakes, understand the interdependency of actions, the range of implications and the ethical basis on which business decisions are being made when it comes to big data.

CONCLUSION

Big data refers to large data sets from various sources whose size makes it challenging to capture, manage and process using conventional information technologies [17]. There has always been technological progress, but what is new is the increasingly rapid and continual transformation of data into meaningful knowledge and therefore, into concrete applications that would in turn, create a competitive advantage. "This phenomenon is known as technoscience and firms play an important role in it. By mastering the methods and tools of scientific rationalization, they have been able to make use of it in their economic strategies" [18].

Big data is creating new opportunities and is therefore, an element of economic power, as it promises to revolutionize the way firms will leverage data-driven strategies to innovate, compete and capture value from deep and up-to-real-time information. Yet, many firms are still struggling to realize the full value of the data they have at their disposal [19, 20].

The main goal of big data analysis should not be merely to understand the numbers by running regression analyses, but to understand the behavior behind the numbers; and this is the true challenge posed by big data in order to affect change and gain competitive advantage [21]. The present paper contributes, therefore, to raise awareness over the potential of big data by means of drawing attention upon its more challenging dimensions of context, connectedness and complexity. By understanding this expanded perspective, decision makers can much better appreciate the value added which big data has to offer and how it can become a competitive advantage for the firm.

REFERENCES

1. Davenport, T.H. and D.J. Patil, 2012. Data Scientist: The Sexiest Job of the 21st Century. Harvard Business Review. Retrieved July, 10, http://hbr.org/2012/10/ data-scientist-the-sexiest-job-of-the-21st-century.
2. Hammond, K.J., 2013. The Value of Big Data Isn´t the Data. Harvard Business Review. Retrieved July, 13, http://blogs.hbr.org/cs/2013/05/the_value_of_big_data_isnt_the.html.
3. Beyer, M.A. and D. Laney, 2012. The Importance of 'Big Data': A Definition. META Group (now Gartner). Retrieved August, 1, http://www.gartner.com/Display Document?id=2057415 and ref=clientFriendlyUrl.
4. Olhorst, F.J., 2012. Big Data Analytics: Turning Big Data into Big Money. John Wiley and Sons, Inc.
5. IBM, 2012. Simply Good Design. 2012 IBM SOA Architect Summit. Retrieved July, 20, https://www950.ibm.com/events/wwe/grp/grp004.nsf/vLookupPDFs/SOA%20Design%20Principles%20and%20Development%20Q4%202012_Final/$file/SOA%20Design%20Principles%20and%20Development%20Q4%202012_Final.pdf.
6. Facebook Investor Relations, 2013. Facebook Reports First Quarter 2013 Results. Retrieved July, 9, http://investor.fb.com/releasedetail.cfm?ReleaseID=761090.
7. ComScore, 2011. Amazon Sites Visited by 1 in 5 Global Internet Users in June. Retrieved July, 8, http://www.comscore.com/Insights/Press_Releases/2011/8/Amazon_Sites_Visited_by_1_in_5_Global_Internet_Users_in_June.
8. Morris, J., 2013. Is this the ultimate definition of "Big Data"? LibertyCalling.net. Retrieved July, 11, http://libertycalling.net/home/categories/technology/item/1149-is-this-the-ultimate-definition-of-big-data.
9. Laney, D., 2001. 3D Data Management: Controlling Data Volume, Velocity and Variety. Applications Delivery Strategies. META Group (now Gartner). Retrieved August, 1, http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf.
10. IBM, 2013. What is big data? Retrieved July, 20, http://www-01.ibm.com/software/data/bigdata.
11. Redman, T.C. and B. Sweeney, 2013. Seven Questions to Ask Your Data Geeks. Harvard Business Review. Retrieved July, 13, http://blogs.hbr.org/cs/2013/06/seven_questions_to_ask_your_da.html.

12. Bughin, J.J. Livingston and S. Marwaha, 2011. Seizing the potential of 'big data'. McKinsey Quarterly. Retrieved July, 10, http://www.mckinsey.com/insights/business_technology/seizing_the_potential_of_big_data.

13. Manyika, J.M. Chui, B. Brown, J. Bughin, R. Dobbs, C. Roxburgh and A. Hung Byers, 2011. Big data: The next frontier for innovation, competition and productivity. McKinsey Quarterly. Retrieved July, 10, http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation.

14. New rules for big data, 2010. The Economist. Retrieved July, 14, http://www.economist.com/node/15557487.

15. Ayres, I., 2007. Super Crunchers: Why Thinking-by-Numbers Is the New Way to Be Smart. Bantam Dell.

16. Gartner Says Solving 'Big Data' Challenge Involves More Than Just Managing Volumes of Data Gartner, 2011. Gartner Press Release. Retrieved August, 1, http://www.gartner.com/newsroom/id/1731916.

17. The Rise of Industrial Big Data, 2012. Leveraging large time-series data sets to drive innovation, competitiveness and growth-capitalizing on the big data opportunity. GE Intelligent Platforms White Paper. Retrieved August, 3, http://www.ge-ip.com/download/the-rise-of-industrial-big-data/13476.

18. De Woot, P., 2009. Should Prometheus Be Bound? Corporate Social Responsibility. Palgrave MacMillan, pp: 16.

19. Abou-Deif, M.H., M.A. Rashed, M.A.A. Sallam, E.A.H. Mostafa and W.A. Ramadan, 2013. Characterization of Twenty Wheat Varieties by ISSR Markers, Middle-East Journal of Scientific Research, 15(2): 168-175.

20. Kabiru Jinjiri Ringim, 2013. Understanding of Account Holder in Conventional Bank Toward Islamic Banking Products, Middle-East Journal of Scientific Research, 15(2): 176-183.

21. Muhammad Azam, Sallahuddin Hassan and Khairuzzaman, 2013. Corruption, Workers Remittances, Fdi and Economic Growth in Five South and South East Asian Countries: A Panel Data ApproachMiddle-East Journal of Scientific Research, 15(2): 184-190.