



4th World Conference on Business, Economics and Management, WCBEM
Use of big data for competitive advantage of company

Milan Kubina^a, Michal Varmus^{a*}, Irena Kubinova^b

^a Faculty of Management Science and Informatics, University of Zilina, Univerzita 8215/1, 010 26 Zilina, Slovakia

^b Faculty of Management Science and Informatics, University Science Park, Univerzita 8215/1, 010 26 Zilina, Slovakia

Abstract

Information is present all around us in different kind of forms. Huge amounts of information are produced every day in big variety of forms. Information plays crucial role for people and for companies as well. Their survival in the highly competitive environment depends upon possessing the right information in the right time. Companies must have information concerning their customers, products, environment and themselves also. The aim of the paper is to show the way to reach competitive advantage by using big data and innovation approaches based on the analysis of literature and conducted research. During the research were used these methods: content analysis, documents study, comparative analysis, process analysis, statistic analysis, empirical research and more. One of the main results is model of information system for work with big data, that is base for successful correct using of information to reach competitive advantage in the market.

© 2015 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Peer-review under responsibility of Academic World Research and Education Center

Keywords: Big Data, management, decision-making, information system, efficiency, technologies, company.

1. Introduction

Information and data are very important and in future will be even more important component of success of all organizations operating on market. One of the main challenges of the business is the need to get the right information as fast as possible to the right person in the right time and in the easiest way possible. There is a huge amount of data produced everywhere. Just Google receives 2 million search queries every minute and Facebook users post around 700 thousand pieces of content in the same amount of time.

* Michal Varmus. Tel.: +421-41-513-4471

E-mail address: michal.varmus@fri.uniza.sk

There are some analytical tools which help us to orientate in this huge amount of data, to make decisions easier and to speed up the processes. But traditional tools can practically work just with structuralized data and cannot work with unstructuralized data. But there is also need to cope with huge amount of unstructuralized data, mainly with the data in the form of text. That means that we have to use and work on more advanced analytical methods to prepare this data, which will help us to better understand not just our own business, customers or suppliers, but also relationships between subjects on market and so on. One of the well-known tools is Hadoop. This tool is used for the work with Big Data.

Big Data is an important term of present. This term changed the meaning of data, what can be mined of data and the utilization of data for company's growth. Social networks are playing the huge role in this. Nowadays people can be online everywhere. They can and do share information about the products, about their desires, they also complain about products etc. This information is really important for companies and have big added value. But Facebook and social networks are not everything. There is a lot more, modern person uses lot of communication channels, he communicates through emails, SMS, Twitter, he writes comments and all these activities leave his indelible trace. And this trace is the trace which has potential for business. It can help to better know the customer through advanced analytical methods of unstructuralized data. The aim of this project is to provide glance into Big Data sector, provide information about Big Data as a whole, about technologies used for processing them, potential legal and ethical risks and managing the Big Data. There are also shown some use cases for better understanding of the topic.

2. Concept of Big Data.

Most companies have been storing and using huge amounts of information. With technical progress, which is visible nowadays especially in the field of information technology, requirements and emphasis on storing, analysing and processing information has increased significantly. According to IBM, we produce 2.5 quintillion (2,5 x 10¹⁸) bytes of data every day and therefore 90% of data in the world today was formed during the last two years. It originates from many sources - from messages, social media posts, invoices, sensors, digital pictures, videos and many others. They all represent the concept of Big Data (IBM, 2013).

Before defining the term Big Data, I would like to define and explain the terms Big Data, which are closely related to this topic and are mentioned very often in this article. To find the exact and homogeneous definition of the term Big Data is practically impossible. The concept is not formalized or unified. For the better understanding I will provide few well-known definitions of Big Data.

One traditional and quite popular is definition created by McKinsey Global Institute (McKinsey Global Institute, 2011):

“Big Data refers to data sets whose size is beyond the ability of typical database software tools to capture, store, manage and analyse.”

This definition is pointing on the problem, which causes Big Data for companies. The amount and type of this data is so huge and varied that it's not possible to work with it in traditional ways and with traditional tools as the companies were used to.

John Gantz has described Big Data in his article as following (Reinsel - Gantz, 2011):

“Big Data technologies describe a new generation of technologies and architectures, designed to economically extract value from very large volumes of a wide variety of data, by enabling high velocity capture, discovery and/or analysis.”

This definition more describes Big Data as a global term for huge and complex sets of technologies, which are aimed to manage and analyze mostly unstructured data which are important for the operation and development of a company. However, this quotation doesn't describe one important aspect of the whole concept, which was described by Douglas Beyer (Laney - Beyer, 2012):

“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.”

That taking into account, Big Data represents data sets which cannot be stored or processed with regular technologies and tools, what means they are so hard to process that it is not possible to keep them stored on one place but it is necessary to distribute them on a larger set of memory devices.

2.1. Basic characteristics of Big Data

Despite the fact that the definitions of Big Data are not consistent, they all touch in the point of the three key elements (3 V's of Big Data), which distinguish them from the regular data. Firstly, it is the volume of the data itself, then it is the velocity what is the frequency of generating new data and lastly variety as the format of the data. More characteristics can be added because process of getting to know the Big Data is just at the beginning.

Volume, or in other words the amount of data, is the primary attribute of Big Data. It can be described as the physical capacity, which is being used itself (In basic definitions, this capacity is measured in terabytes to petabytes), or it can be quantified with the number of records, transactions, tables or files.

The more the companies are focusing on considering the data as an asset, the more does the unwillingness to getting rid of the data and reducing its size grow. This issue is also connected with a loss of the relative value of each additional byte and vice-versa, with a growth of the relative value of each additional unit of data storage needed for the preservation of the data.

Velocity can be described as the speed at which Big Data are being generated. However, it can be understood as the time it takes to transfer the data from the moment of retrieval till the moment of being analyzed and the decision is being made itself as well.

The speed can be considered as a characteristic which is still growing its importance, but it can be assumed that it will become the most significant out of the other ones. This can be supported with the fact that there are already some fields which need to process the data in real time and make the decisions based on these data almost immediately.

The key factor of Big Data is the **variety** of information. Traditional data that have given structure – are generally referred to as structured data. Typical example of this data is data which are being saved into databases storage structures. But during recent years, there is more and more of unstructured content and semi structured. Unstructured content can have many forms and it is hard to define them in detail. This category includes most of the modern formats including image and audio data, reports from social networks, blogs and information about geographical location, records of web clicks, data available on internet and many others.

The objective of Big Data is to mine all this data and synchronize, unify, integrate them into the form suitable for further business processing.

3. Use of big data for competitive advantage of company

So why are Big Data or Big Data technology so important for companies? First of all they help companies to understand and to extract the meaning from the whole range of information which is in the world. Companies and organizations gathered and stored data, which were part of each transaction. This information was primarily used for tracking or forecasting the future. Today this data are exploding. It is possible to gather information about every customer, who visits your website. So marketers can gather information about every customer's conversation about their product or brand.

These sources are treasure for companies, because they can provide a glance into customer's minds. However this requires implementation of new processes, technologies and mechanism called Big Data. The data are here and they hide valuable information, it is just needed to find the way, how to mine it.

It is suitable to use comparison, which is in recent years very actual. And that is, that if we take into consideration troubles with pollution of the Earth, where data are accumulated as well as garbage. It is up to us, if we leave the garbage lying and gathering in the ground and let them pollute the planet henceforth, or we take them and recycle them. The same trouble is with Big Data, information in companies accumulates same as garbage and lie in storage, nobody is using them, and they just "pollute" the pc. We have to use this data for our benefit. We are surrounded by it everywhere where we look. If we that "garbage" take and sort into needed form, we will be able to use it and increase the value of it in future.

Big Data solutions are ideal for analyzing raw structured data, but also for analyzing semi structured and unstructured data from all the sources described in previous chapter. Big Data solutions are also ideal when all, or majority of data need to be analyzed or when sampling of data is not as effective as a larger set of data.

According to the McKinsey Global Institute, there are 5 ways of how Big Data create value (McKinsey Global Institute, 2011):

- It can create transparency by being more widely available to the new potential.
- It enables companies to set up experiments. For example experiments for process changes, they can create and analyze large amounts of data from these experiments to identify possible performance improvements.
- Big Data can be used to create a more detailed segmentation of customers to customize actions and prepare specific services.
- Analysis of Big Data can support human decision making by pointing to hidden correlations or some hidden risks. An example can be a risk or fraud analysis engines for insurance companies. Low decision making can be even automated to those engines in some cases.
- Data can also enable new business models, products and services or can improve the existing ones. Data about how products and services are used can be used to develop and improve new versions of the product.

By using Big Data and utilizing its benefits can companies gain big competitive advantage and get ahead of their rivals. Big Data offers businesses much bigger growth potential than traditional technologies, even though it is still much less understood. Companies holding further from this concept can allow their competition, which has understood the importance of Big Data faster, to gain a leading position on the market. Organizations shouldn't underestimate the importance of this concept.

4. Conclusion

The term Big Data refers to large amounts of data, mainly unstructured and semi-structured, which comes from a lot of different sources. It is not possible to process this data with traditional approaches and tools. This concept has become very popular recently, when companies are trying to understand the customer behaviour as much as possible and technology progress is galloping forwards. Companies have realized that a lot of value can be found inside huge amounts of unstructured data and with lowered costs on storage place have also barriers for handling Big Data significantly reduced.

Big Data brings several advantages to companies. It allows higher transparency of information inside organizations, enables broader, deeper and more accurate insight, therefore improves decision - making as well. It gives companies the possibility to create more complex and complete image about their customers and therefore offer more accurately tailored products and services.

Acknowledgements

This paper was supported by the Slovak scientific grants VEGA 1/0621/14, VEGA 1/0363/14 and KEGA 035ZU-4/2013.

This paper is supported by the following project: University Science Park of the University of Zilina (ITMS: 26220220184) supported by the Research & Development Operational Program funded by the European Regional Development Fund.

References

- Big Data insights and opportunities*. (2013). Retrieved April 02, 2015, from: <http://www.euolanresearch.com/wp-content/uploads/2014/05/2ndBigData.pdf>
- Dumbill E., *Defining Big Data*. Forbes. (2014). Retrieved February 14, 2015 from <http://www.forbes.com/sites/eddumbill/2014/05/07/definingbig-data/>
- IBM. *What is Big Data* (2014). Retrieved February 14, 2015 from <http://www-01.ibm.com/software/data/bigdata/what-is-bigdata.html>
- Kaisler, S., Aemour, F., Espinosa, J. A., Money, W. (2013). *Big Data: Issues and Challenges Moving Forward*. In Proceedings of the 46th Hawaii International Conference on System Sciences, HICSS '13, pages 995–1004.
- Laney, D. Beyer, M. (2012). *The Importance of Big Data: A Definition*. Retrieved April 01, 2015, from: <https://www.gartner.com/doc/2057415>

- Madden, S. (2012). *From Databases to Big Data*. Retrieved February 17, 2015 from: <http://www.computer.org/csdl/mags/ic/2012/03/mic2012030004.pdf>
- McKinsey Global Institute. (2011). *Big Data: The next frontier for innovation, competition, and productivity*. Retrieved February 14, 2015 from http://bigdatawg.nist.gov/MGI_big_data_full_report.pdf
- Minelli, M. (2013). *Big data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Businesses*. John Wiley and Sons, Hoboken, N.J. 224 s.
- Reinsel, D. Gantz, J. *Extracting value from chaos*. (2011). Retrieved April 02, 2015 from: <http://www.emc.com/collateral/analyst-reports/idc-extracting-value-from-chaos-ar.pdf>
- You Can't Spell "Big Data" Without "V"*. (2014). Retrieved April 02, 2015 from: <http://researchaccess.com/2014/07/big-data-without-v/>
- Zikopoulos, P. C., Eaton, Ch., Deroos, D., Deutsch, T., Lapis, G. (2012). *Understanding Big Data, Analytics for Enterprise Class Hadoop and Streaming Data*. The McGraw-Hill Companies.