THE USE AND THE FUTURE OF BIG DATA ANALYTICS IN SUPPLY CHAIN MANAGEMENT

Wojciech Zdrenka*

* Transport and Logistics Department, Poznan University of Economics, Poznan, al. Niepodległości 10, 61-875 Poznan, Poland, Email: wojciech.zdrenka@ue.poznan.pl

Abstract: Global computerization and informatization of enterprises, Internet popularization and fast-growing number of mobile devices has caused the rapid growth of data generated by the society. There never was so much data in the whole humans history. Forecasts shows that in 5 years the growth rate of data being generated will increase by several times. From one side, easy-access to company's economic environment and customers' data enables better decision taking, but from the other side huge amount of data leads to the "information noise" which may be a cause of incorrect conclusions and finally wrong decisions. Due to this phenomenon, companies have faced completely new challenge – development of company's competitive edge through the analysis of huge amount of unstructured and changing data bases – so-called "Big data". Analysis that were difficult or even not possible to conduct couple years ago, today are supporting companies on every day basis thanks to Big data analysis.

Paper type: Research Paper

1. INTRODUCTION

The field of Big data analytics still isn't standardized by any organization so many definitions of Big data can be found in the literature.

The often-quoted definition is the one that was defined by McKinsey Global Institute – "Big data refers to datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyze" (McKinsey Global Institute, 2011).

Big data has not only technological aspect, the economic dimension of Big data emphasize by Reinsel and Gantz, describing Big data as "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" (Reinsel & Gantz, 2011).

Complex definition was presented by Gartner company, which defines Big data as a "high-volume, high-velocity and/or high-variety information assets that demand cost-effective, innovative forms of information processing that enable enhanced insight, decision making, and process automation" (Beyer & Laney 2012).

Regardless of the definition, Big data characteristics relates to the three main elements (Berman 2013): high volume, high velocity, high variety.

High volume is one of the main characteristics of Big data. It can be defined as the physical capacity which is used to store the data (usually measured in terabytes and petabytes), or the number of records, transactions, tables or files (Kubina, Varmus & Kubinova 2015). Since companies are gathering more and more data, the marginalism can be observed – every single bite stored, bring less and less marginal utility.

High velocity is a speed at which Big data are being generated. It can be described as the time between the data has been registered till the moment of the decision has been made. The importance of the speed of data processing will be increasing. There already are some fields where decision must be made in almost the same moment when the data is registered (Kubina, Varmus, Kubinova 2015).

High variety of information is also a key factor of Big data. Traditional data sets have given structure – the structure was defined before the data has been collected. There is more and more of unstructured and semi-unstructured content in recent years. This type of data has different form and formats. (Kubina, Varmus & Kubinova 2015). It includes not only files but also all the data being generated online by websites, applications and social networks including text, and users' behaviours – mouse clicks, changing or closing of browser tabs or even time spend on different webpages.

Over the course of time, to the model of 3Vs was added a forth one – Value. High value is understood as the significance of knowledge hidden in Big data (Abaker & Hashem, 2015).

2. THE USAGE OF BIG DATA

2.1. Big data in management

The amount of data is growing continuously. The amount of created and replicated information till 2011 surpassed 1.8 zettabytes (1.8 trillion gigabytes) – growth by 9 times in just five years. According to the forecasts till 2020 the volume of data will grow up to 40 zettabytes (IDC, 2011).

One of the main cause of so the rapid increase of the data volume is the increasing number of sources that are generating the data. This sources can be divided into 4 main categories (Kelly, 2014):
- Social media – Currently there is over 700 million of Facebook users, more than 250 million users of Twitter and over 156 million of blogs. Every new post or tweet increases the volume of data.
- Machine generated – Machine data are information generated automatically from a hardware or software such as computers, medical devices, or other machines, without human intervention.
- Transactions – Every day billions of transactions are executed online. The transactions are done not only by humans but also by the machines (e.g. automatic transactions on stock exchange). Information about each of these transactions is registered and stored by all the interested parties. This generates even more data.
- Sensing – Sensing devices like wages, thermometers or moisture meters, not only measure physical quantities now. More and more devices register and stores the measures in digital form to provide the data to IT systems. The more digital sensing devices will be used, the more data will be generated.
- Internet of Things – Internet of Things is a set of physical devices like smartphones, tablets, smartwatches, buildings, vehicles and other digital devices embedded with sensors and software, connected via the Internet. Because the number of these devices is constantly increasing, the number of the data generated increases as well.

Continuously increasing amount of data requires not only the development of storage but also the development of processing and analyzing capacities. Right processing and analyzing is the key factor that enables to take the right decision and to build the knowledge, based on the collected data.

One of the examples of modern researches based on Big data analysis is the research conducted by the Global Pulse organization. Global pulse has analyzed discussion themes in Twitter data to investigate what indicators can help understand people's perceptions and concerns around food, fuel, finance and housing in the US and Indonesia. The analysis based on the publicly available data from Twitter for July 2010 through October 2011 in Javanese/Bahasa Indonesia and English. The goal of the research was to determine which factors might be present in social

media data that could shed light on how populations cope with global crises, such as commodity price volatility or the continuing global economic crisis (Crimson Hexagon, 2011).
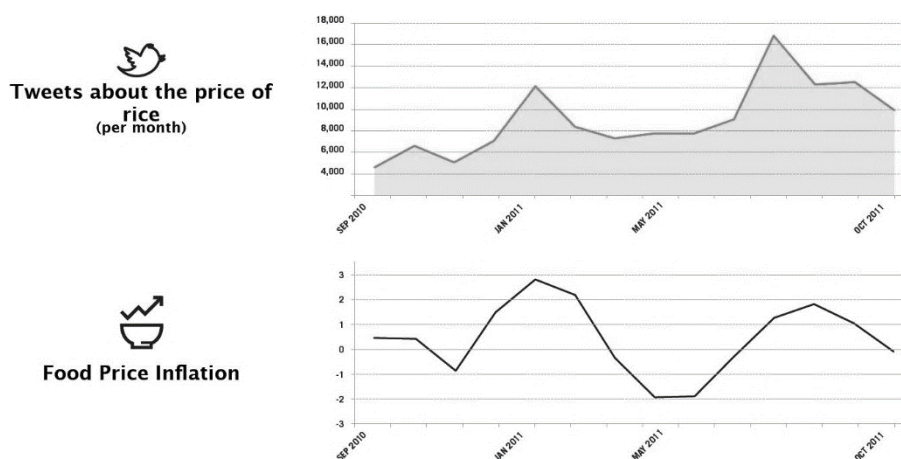
Fig. 1. Volume of tweets per month about the price of rice from October 2010 to October 2011 in Bahasa Indonesia / Javanese and monthly inflation rate for the food basket in Indonesia from October 2010 to October 2011; (Crimson Hexagon, 2011)

Big data analyzing of social media may help in better understanding of Internet users' behaviour and discover social and economics laws. According to the report of Crimson Hexagon Big data analysis of social media can improve three main fields (Crimson Hexagon, 2011):

- Early warning – Information from social media allows to detect the anomalies, trends and events earlier. The quicker it will be detected, the better the response to emerging crises can be.
- Real–Time Awareness: More accurate and up-to-date picture of what a population needs and wants can lead to better, more effective planning and implementation of development programmes.
- Real-Time Feedback: Sooner understanding of changing needs of society allows for rapid, adaptive course correction in development programmes and current policies — making impact evaluation and response more agile.

Big data has also significant value in enterprise management. Thanks to development of methods of Big data analyzing companies are able to take decisions based on the knowledge that wasn't available some time ago. Big data analytics can transform key organizational business processes, such as (Schmarzo, 2013):

- Procurement – Big data analytics can improve the methods of identification and evaluation of suppliers.

- Product Development – The real-time feedback provides information how the products are being used. This information can speed up the development process and improve the new products launches.
- Manufacturing – The data registered automatically by the sensors (e.g. machine statuses, process indicators, etc.) can improve the manufacturing process by the identification of the root causes of quality problems.
- Distribution – The information about external factors such as weather, holidays, and economic conditions, can be used for better inventory levels evaluation and supply chain activities optimization.
- Marketing – By the analysis of the big sets of data about customers' behaviour, engagement and sales figures, Big data analysis can optimize the marketing promotions and campaigns.
- Pricing and Yield Management – The customers' behaviour analysis combined with the information about the inventory and with the information about external factors, can improve the pricing and yield management process – especially in the area of "perishable" goods.
- Merchandising – Based on current buying patterns, inventory levels and product interest insights gleaned from social media data, Big data analysis can improve the merchandise markdown optimization.
- Sales – The analysis of huge amount of data can influence the way the sales resources are assigned, optimize the product mix and improve commissions modelling.
- Store Operations – The prediction buying patterns compared with local demographic, weather, and events data, may lead to inventory levels optimization.
- Human Resources – The Big data analysis can enable the identification of the characteristics and behaviours of the most successful and effective employees.

The usage of Big Data analytics in above mentioned areas may lead to the improvement of companies' effectiveness. The analysis of combined information about customers' online behaviour, consumers' preferences, effectiveness of marketing campaigns, macroeconomics factors or even the information about the weather may lead not only to changes in companies' operations but to a change of the whole companies' current strategy.

An example how Big data analysis affects companies' strategies may be the Netflix case-study. Netflix engaged over 800 IT developers to build an advanced system of algorithms to monitor customers' behaviour. Based on the results of the analysis Netflix has changed the profile from DVD rental shop to the company offering video content online. Currently Netflix uses Big data analytics for price setting and evaluation of web tv series popularity (Xu, Frankwick & Ramirez, 2016).

2.2. Supply chain management with Big Data

Big data analytics has a great impact in Supply Chain Management (SCM) especially (Wamba, Akter, Edwards, Chopin & Gnanzou, 2015). The cause of increasing attention of Big data analytics in SCM is its complexity and the influence that SCM have on the overall performance of companies. The SCM area is facing the challenges which may lead to inefficiency and wastage, e.g. delayed shipments, rising fuel costs, inconsistent suppliers or ever-rising customer expectations. Companies expects that Big data analysis will bring profits through the process visibility and effectiveness improvement, global supply chains integration and demand management improvement. In the area of strategic planning and management of supply chains, Big data analysis is especially important. It supports sourcing decisions, supply chain configuration, and design and development of products or services (Wang, Gunasekaran, Ngai & Papadopulos, 2016).

Wang, Gunasekaran, Ngai and Papadopulos, (2016) specify in the studies a concept of Supply Chain Analysis (SCA) which is an application of Big Data Analytics (BDA) in Supply Chain Management (SCM). As a result of their studies, the researches presents a maturity framework of SCA. The framework defines 5 different levels of SCM analysis based on different supply chain goals, including:

- Functional SCA – The functional level of supply chain analysis helps keeping the costs on low level by synchronization of activities and processes and integration between supply chain partners.
- Process-based SCA –The Process-based SCA level compose tools and techniques used to solve problems with integration of internal supply chain processes within an organization. The purpose of this SCA level is to help companies to achieve operational effectiveness in supply chain processes.
- Collaborative SCA – The goal of the collaborative SCA level is to improve the communication integration of the key activities between different partners in the supply chain.
- Agile SCA – The Agile SCA level are techniques used to improve the ability to respond to changes. The purpose of real time monitoring is to speed up in addressing changing customer demands, and short lead times related to the transformation of supply chains, if needed.
- Sustainable SCA – the purpose of the sustainable level of supply chain analysis is to provide the appropriate information that can be used for effective and efficient decision-making on sustainability issues.
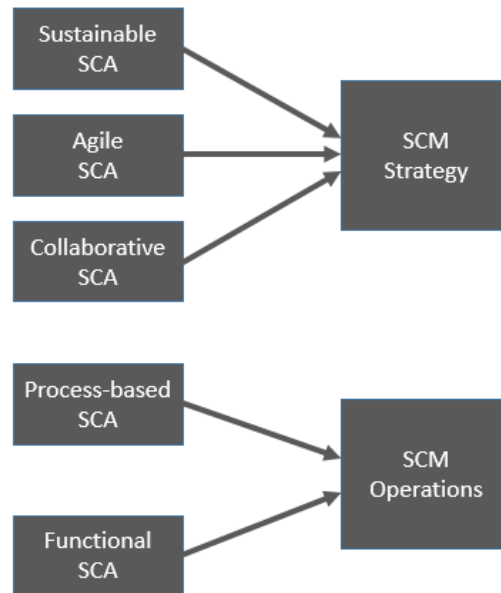
Fig. 1. SCA Maturity framework (Wang, Gunasekaran, Ngai & Papadopulos, 2016)

Due to the fact that Big data is generated by big companies, Big data analytics in Supply Chain are used mainly by logistics operators and e-retailers currently.

Logistics operators use Big data analytics to decrease the costs by increasing the efficiency of operations. One of the examples of using Big data analytics in SCM on operational level is DHL dynamic routing system. DHL uses Big data analysis to recalculate the routes depending not only on addresses, but also based on the current traffic situation. The telematics databases are tapped to change the delivery routes automatically according to the current traffic conditions.

Another example of Big data analysis use is UPS. According to the International Institute for Analytics UPS "tracks data on 16.3 million packages per day for 8.8 million customers, with an average of 39.5 million tracking requests from customers per day. The company stores over 16 petabytes of data UPS registers information about 16,3 million packages per day for 8,8 million customers. Every single day the company receives from customers 39,5 tracking requests. UPS stores over 16 petabytes of data. Much of its recently acquired big data, however, comes from telematics sensors in over 46,000 vehicles. The data on UPS package cars (trucks), for example, includes their speed, direction, braking, and drive train performance. The data is not only used to monitor daily performance, but to drive a major redesign of UPS drivers' route structures. This initiative, called ORION (On-Road Integrated Optimization and Navigation), is arguably the world's largest operations

research project. It also relies heavily on online map data, and will eventually re-configure a driver's pickups and drop-offs in real time. The project has already led to savings in 2011 of more than 8.4 million gallons of fuel by cutting off 85 million miles of daily routes. UPS estimates that saving only one daily mile driven per driver saves the company $30 million, so the overall dollar savings are substantial. The company is also attempting to use data and analytics to optimize the efficiency of its 2000 aircraft flights per day" (Davenport & Dyche, 2013).

E-retailers are seeking better fulfilment strategies using Big data analytics. Anticipatory shipping is one of them. Amazon, one of the greatest e-retailers world-wide, has a patent for the algorithm-based system, which, which may be used to ship products before the customer will place the order. In 2014 Amazon's "method and system for anticipatory package shipping" patent describes a method for shipping a package of one or more items "to the destination geographical area without completely specifying the delivery address at time of shipment", with the final destination defined route. The patent uses a forecasting model which predicts the shipping address base on the data about Amazon's customers' activity, including time on site, duration of views, links clicked and hovered over, shopping cart activity and wish lists. Based on this information, Amazon's algorithm provides „decision support for speculative shipping of items" (Ulanoff, 2014). Thanks to the algorithm, the time between placing the order by the customer and the time when the parcel is being shipped to the customer.

Predictive algorithms based on Big data analytics are used by e-retailers for margin and offer management. Big data analytics supports better understanding of customers' behaviours and preferences. Amazon was one of the first companies which implemented dynamic price setting – constantly price adjustment on millions of SKU's on a daily basis base on the competition, demand, etc. Burberry.com, the online store, collects various data such as e.g. browsing patterns, login counts, or purchase history, to present personalized offer to each client (Ilieva, Yankova & Klisarova, 2015).

2.3. Perspectives for future

The development of Big data technologies is the object of interests of most of the Governments and International Organizations. In Europe, it's European Commission (EC). EC launched Digital Agenda – the document in which EC emphasize how Big data can bring value and opportunities to European markets. According to EC improved analytics and processing of data, especially Big Data, will make it possible to (European Commission, 2015):
- transform Europe's service industries by generating a wide range of innovative information products and services;
- increase the productivity of all sectors of the economy through improved business intelligence;

- address many of the challenges that face our societies more adequately;
- improve research and speed up innovation;
- achieve cost reductions through more personalized services;
- increase efficiency in the public sector.

European Commission concentrates its activities around the research and practice aiming to develop the traditional industries such as manufacturing, logistics and energy with help of the Big data technologies. Recent years EC supported more than 30 thousand of projects related to Big data in Service, Manufacturing and Supply Chain Management.

In other regions similar support is provided. In United States Big data based studies and researches are supported mainly by the U.S. National Science Foundation which in 2015 awarded more than 11 thousand of projects. In Canada the major foundation body is the National Science and Engineering Research Council of Canada (NSERC) who awarded approximately almost 8 thousand projects in last 3 years. The level of support of government in Asia Pacific region differentiates among the countries. Considering two largest population countries India and China example we can conclude that in the region of Asia the support in Big data studies is much lower than in Europe or North America. The National Natural Science Foundation of China (NSFC) and Ministry of Science and technology of People's Republic of China (MOST) – two major foundation bodies, from 2009 to 2015 founded 346 projects. In India, there in the same time 146 projects were founded (Zhong, et al., 2016).

Observing the interests of Governments, Public Organizations and Companies we can conclude that the Big data technologies and the methods of data analysis will be improved in the following years. Because Supply Chain Management plays a very important role in companies, we can assume that the area of logistics will be in main focus of the future of Big data. Further development of Big data technologies will require to handle the current issues and constrains that researches are facing. Zhong, Dai, Qu, Hu & Huang in their studies, points out areas for improvement and predicted future solutions which may be taken (Zhong et al., 2016):

- Data collection methods – Efficient and effective data capturing plays essential role in Big data analytics. Even though there are many innovative tools and techniques to collect data automatically, manual methods are widely uses in SCM. Paper- or Excel-based data collecting systems are prone to be incomplete, inaccurate, and untimely - thus, decisions based on such data are usually unreasonable, unexecutable and unpresentable. The development of data collection techniques will include the popularization of such as technologies as Auto-Id, Internet of Things (IoT) and wearable devices. The main focus of data collection methods in SCM will focus on smart devices like multi-functional ubiquitous devices and smart robotic collectors. These technologies enable the automatic data collection of information that were hard or even impossible to register by humans because of the characteristics of data (very large amount of specific data) or be-

cause the characteristics of environment where the data has to be collected (extremely high/low temperatures or harmful environments with nuclear radiations).

- Data transmissions – Collected data requires to be transferred which includes wired and wireless methods of data transmissions. The speed of transmission relies on structure of the network. Current transmission service of networks – e.g. 4G, don't provide sufficient transmission speed to operate with Big data. The development of high speed networks influences the use of Big data in SCM. According to the Zhong R.Y., the future of Big Data transmission in SMC will mainly include: Smart powerful network, Advanced encryption methods and Synchronization of central and distributed network architecture.

- Data storage – The possibility to store huge amount of data is another factor which influence the Big data usage. As long as the technology will not offer easy and affordable way to store and maintain the data, the Big data appliance in SCM will be limited. Two main dimensions of data storage are considered: Biological storage media and Learnable storage mechanism.

- Processing technologies for Big Data – Currently processing of data (e.g. querying) may be very time consuming. Zhong R.Y. emphasize that the processing of great amount of data may be an issue that resolving requires not only more powerful computers but also new, advanced data processing models and algorithms. Current tools and techniques are not sufficient. The authors anticipate that the future of Big data processing technology will be concentrated on such as aspects as Smart cloud based infrastructure and Intelligent processor.

3. CONCLUSION

The importance of the analysis of huge amount of varied and diverse data in SCM grows quickly. Companies are using Big data analysis not only for one-time decisions but make use of it on daily basis improving operations and reducing costs. Currently managers can take decisions supported by data, in matters in which not so long time ago they have to rely on own gut feeling. Companies can act more quickly and more accurate thanks to the knowledge that Big data analytics gives.

The forecasts say that the data will be growing and the methods and technologies of storing, processing and analyzing of data will develop. The development of the Internet of Things and Machine-to-Machine connections may be one of the leading factors which will have the influence on the amount of data produced – in SCM especially. These factors may lead to the situation where companies that are willing to build great competitive advantage will have to make use of the advantage that Big data analysis brings to the market. The companies which want to suc-

ceed will have to analyze not only internal data generated by the operations, but also the external sources of data like social media and social networking services, machines or everyday life devices. The more data we produce the greater influence on the building of competitive advantage the data analysis will have.

ACKNOWLEDGEMENTS

REFERENCES

Abaker I. & Hashem T. (2015), The rise of "big data" on cloud computing: Review and open research issues, Information Systems, January 2015, Vol.47, pp. 98–115.

Amazon Technologies, Inc. (2012), Patent US8615473 – Method and system for anticipatory package shipping.

Berman J. (2013), Principles of Big Data: Preparing, Sharing, and Analyzing Complex Information, Elsevier Science.

Beyer M. & Laney D. (2012), The Importance of 'Big Data: https://www.gartner.com/doc/2057415/importance-big-data-definition (access 19 September 2016).

Crimson Hexagon (2011), Monitoring Perceptions of Crisis-Related Stress Using Social Media Data, http://www.unglobalpulse.org/projects/twitter-and-perceptions-crisis-related-stress (access 10 September 2016).

Davenport T. & Dyche J. (2013), Big Data in Big Companies, International Institute for Analytics, May.

European Commission (2015), Making Big Data work for Europe, https://ec.europa.eu/digital-single-market/en/making-big-data-work-europe (access 5 September 2016).

IDC (2011), Extracting Value from Chaos, https://www.emc.com/collateral/analyst-reports/idc-extracting-value-from-chaos-ar.pdf, (acces 2 September 2016).

Ilieva G., Yankova T. & Klisarova S. (2015), Big data based system model of electronic commerce, Trakia Journal of Sciences, Vol. 13, Suppl. 1, pp. 407–413.

Kelly J. (2014), Big Data: Hadoop, Business Analytics and Beyond, http://wikibon.org/wiki/v/Big_Data:_Hadoop,_Business_Analytics_and_Beyond (access 15 September 2016).

Kubina M., Varmus M. & Kubinova I. (2015), Use of big data for competitive advantage of company, Procedia Economics and Finance, 26 (2015), pp. 561–565.

McKinsey Global Institute (2011), Big data: The next frontier for innovation, competition, and productivity: http://www.mckinsey.com/business-functions/business-technology/our-insights/big-data-the-next-frontier-for-innovation (access 10 September 2016).

Reinsel D. & Gantz J. (2011), Extracting value from chaos: http://www.emc.com/collateral/analyst-reports/idc-extracting-value-from-chaos-ar.pdf (access 10 August 2016).

Schmarzo B. (2013), Big Data. Understanding how data powers big businesses, Wiley.

Ulanoff L. (2014), Amazon Knows What You Want Before You Buy It, http://mashable.com/2014/01/21/amazon-anticipatory-shipping-atent/#T44JGMf_0Zqk (access 10 August 2016).

Wamba S., Akter S., Edwards A., Chopin G. & Gnanzou D. (2015), How 'big data' can make big impact: findings from a systematic review and a longitudinal case study. Int. J. Prod. Econ., Volume 165, July 2015, pp. 234–246.

Wang G. Gunasekaran A., Ngai E. & Papadopulos T. (2016), Big data analytics in logistics and supply chain management: Certain investigations for research and applications, Int. J. Production Economics 176/2016, pp. 98–110.

Xu Z., Frankwick G.L. & Ramirez E. (2016), Effects of big data analytics and traditional marketing analytics on new product success: A knowledge fusion perspective, Journal of Business Research, Volume 69, Issue 5, May 2016, pp. 1562–1566.

Zhong R.Y. et al. (2016), Big Data for supply chain management in the service and manufacturing sectors: Challenges, opportunities, and future perspectives. Computers & Industrial Engineering, In press.

BIOGRAPHICAL NOTES

Wojciech Zdrenka is a PhD student of Logistics and Transport Department at Poznań University of Economics and Head of Business Analyst Team of a consulting company. From the very beginning, he is professionally related with business process improvements and business analytics. He specializes in process-based approach to management. He achieved certification in Projects Management – APMG-International (Prince 2 Foundation), SCRUM Alliance (Certified Scrum Master) and Supply Chain Management – APICS (BSCM). The main fields of his researche intrerests is the influence of new technologies on supply chain management. His lately papers refers the cross-boarder e-commerce issues and the influence of Big data on the supply chains.