DE GRUYTER OPEN

Management Systems in Production Engineering

2017, Volume 25, Issue 4, pp 251-254

Date of submission of the article to the Editor: 04/2017
Date of acceptance of the article by the Editor: 08/2017

DOI 10.1515/mspe-2017-0036

BIG DATA IN THE INDUSTRY - OVERVIEW OF SELECTED ISSUES

Sylwia GIEREJ
Bialystok University of Technology

Abstract:
This article reviews selected issues related to the use of Big Data in the industry. The aim is to define the potential scope and forms of using large data sets in manufacturing companies. By systematically reviewing scientific and professional literature, selected issues related to the use of mass data analytics in production were analyzed. A definition of Big Data was presented, detailing its main attributes. The importance of mass data processing technology in the development of Industry 4.0 concept has been highlighted. Subsequently, attention was paid to issues such as production process optimization, decision making and mass production individualisation, and indicated the potential for large volumes of data. As a result, conclusions were drawn regarding the potential of using Big Data in the industry.

Key words: *Big Data, Industrial Internet of Things, Industry 4.0*

INTRODUCTION

The current level of economic development is influenced by advanced information technologies. Due to the development of software and mobile platforms, the conditions in which companies operate today have definitely changed. Achievments in the Internet of Things area have made it possible to collect data that was difficult or even impossible to obtain a few years ago . The amount of data currently collected is huge and is constantly growing at a very fast pace. This is the reason for the growing interest in the Big Data concept, which is to manage and analyze high-speed mass flow data and diversity. It has been recognized as one of the key technologies in smart management strategy and future global development. It should be borne in mind, however, that with such large datasets it is essential to analyze and evaluate which resources are actually useful. Maintaining data on your own or external servers entails costs. That is why it is so important for companies to focus not only on data collection but mainly on their effective analysis and elimination of overdue and worthless data. Due to the flow rate and the variety of data produced, distributed and used over a few seconds to a few hours, as well as the structural and unstructured form of data, it appears that existing data analysis methods do not cope with the proper management of data. That is why Big Data should be understood not only as a mass data collection, but also as human resources, organizations, and the right technologies to get valuable information. In Big Data analysis, it is crucial to extend the perspectives and thought horizons, allowing you to change your field of vision, take a closer look at information resources, and then skillfully select those that are relevant and further analyzed [5].

Available statistics and reports show that the global demand for enterprise experts specializing in large datasets is growing rapidly. According to McKinsey & Company data in 2018 in the United States, there will be lacking 14-19 thousand Big Data analysts, and 1.5 million managers who can practice this knowledge effectively in order to make effective decisions. This demonstrates that businesses recognize the potential of acquiring, processing and analyzing mass data collections and utilizing information obtained to optimize processes and plan their strategies [7]. The following article discusses the Big Data concept, taking into account the opportunities and risks of its use and its role in the fourth industrial revolution. Next, selected options for exploiting the potential of large data sets in industry were presented.

OVERVIEW OF THE BIG DATA CONCEPT

By analyzing available literature, you can come across several definitions explaining the term Big Data. So far, none of them has been officially accepted and fully accepted. The first publications on the subject of large volumes of data come from the late 1990s. The authors of one of the first definitions are M. Cox and D. Ellsworth, according to which Big Data is a large dataset that needs to be expanded to extract information values [3].

With the passage of time and technological advances, the definitions associated with Big Data evolved, gaining more attributes. Technologies that allow the processing of large volumes of data of a completely different nature (quantitatively and qualitatively), such as Hadoop or HBase, make it much easier to implement exploration techniques and support other data science efforts. Consequently, Big Data should be understood not only as data sets but also as a set of techniques for processing and analyzing them. Below are the main attributes of Big Data [11]:

- volume – this feature refers to the quantity and means the high dynamics of data growth requiring advanced technologies for efficient processing,

- velocity – data is streaming delivered to the organization in real time, which requires sufficient computing power to provide fast response time and isolate only relevant information,
- variety – refers to the heterogeneity of the data, the lack of structuring, occurrence in different formats (numerical data, text, image, sound) generated by various sources (internal data of the organization, external data from online sources: social networking sites, purchased databases, etc.),
- variability – represents a change in the intensity of data over time and indicates the subordination of flows to certain periodic cycles and trends, eg. the increase in traditional and electronic sales during Christmas, increased interest in hotel services during the holiday season, increased traffic in social media during the election parliamentary,
- complexity – is directly related to diversity and is characterized by a variety ways of systemating data in which are structured data (such as phone numbers, credit cards), mixed data (eg. XML files) and unstructured data (eg. documents, video, photos),
- value – refers to the ability to capture unique information from large data sets that significantly influence organizational efficiency by supporting decision-making and can be an additional revenue stream through the monetization of your information assets.

In the literature you can already meet the concepts of Big Data 1.0 and Big Data 2.0. This is reflected in the analogy to the acquisition of Internet technology by the business sphere. At the early stages of Web 1.0, businesses used Web technologies primarily to network, enable e-commerce, and improve the efficiency of their operations. Referring to the Big Data 1.0 concept, it can be stated that the current purpose of the business is to build the right technology base to process large volumes of data to support current business. The economics of the Web 2.0 era have begun to make great use of the possibilities of interaction on the Web. Companies have recognized this potential and are currently using it extensively to improve their offerings and more accurate adjust to consumers needs. Similarly, after mastering the technology of flexible data mass processing, organizations in the Big Data 2.0 era will take full advantage of the information potential they will be able to achieve [9].

THE ROLE OF BIG DATA IN THE FOURTH INDUSTRIAL REVOLUTION

The term Big Data usually coincides with the concept of Industry 4.0 and the Industrial Internet of Things, which are often mistakenly understood as identical concepts. To avoid discrepancies in the rest of this article, it need to clarify the difference between them. Industry 4.0 is a broader and it generalizes term for the concept of "industrial revolution" in connection with modern mutual use of automation, data processing and exchange, and techniques. The differences between Industrial Internet Things and Industry 4.0 are not big. However, it can formulate the opinion that the road to Industry 4.0 leads through the Industrial Internet Things. It runs on a smaller scale (sensors) and connects many applications, which become the infrastructure base for larger (automation) applications involved with Industry 4.0 [12].

Concept of Industry 4.0 first appeared in 2011 and referred to the initiative of the German government known as the Future Project. It assumed that the company's production system would be built from an IT platform and intelligent, autonomous digitally controlled machines. However, due to the specificity of production systems, dependent of branch, Industry 4.0 has not been generalized. It should be considered individually, depending on the objectives of the enterprise. Previous industrial revolutions were linked to advances in the mechanics of production, electricity, information technology and the reduction of physical labor. The current revolution in the industry is based on the idea of creating intelligent digital factories. Production systems are not only characterized by autonomy but also have the potential for self-configuration, self-monitoring and self -improvement [14].

The foundation of the planned industrial change is the use of cyber-physical systems (CPS) that generate and use mass data sets. They are the next generation of intelligent manufacturing systems consisting of machines and devices, information systems, monitoring systems, PLCs, controllers and sensors. The CPS system model is defined as "5C", and the essence of its operation is the sharing of data and information used to acquire knowledge by machines and devices. The model shows the next phases of CPS implementation (Fig. 1). As a result of this process, the machine will use its historical data as well as data from other devices to diagnose and predict faults and to improve the process itself. It can therefore be stated that CPS systems are one of the data sources classified as Big Data collections.

The Industrial Internet of Things is an essential element for the emergence of innovative changes, which are the result of the development of Industry 4.0. It is thanks to the use of solutions enabling remote control of production machines, means of transport and access to IT systems from practically every place in the world, there is a technological revolution in various sectors of the industry. With the ever-increasing use of sensors, it is possible to obtain industrial data and transfer them over the network to high-performance computing. This significantly streamlines process monitoring and optimization. This is another source of mass data and is the basis for the creation of a digital factory. The key issue is the protection against cybercrime, as data is primarily stored in the cloud [13].

Another source of Big Data is data resulting from the design of virtual reality. Thanks to the development of innovative devices that allow digital mapping of real objects, it is possible to collect information about them and their convenient processing. An example might be a product offered by Microsoft. HoloLens is the world's first holographic wireless computer. Allows you to freely control the holographic objects - move them, change their shape, set in space. All equipment is housed in special glasses, no wires or additional equipment are needed. HoloLens gives an almost tangible way to get acquainted with the technology used in Volvo cars. Thanks to HoloLens, customers will be able to make a realistic configuration of their chosen car model [8].

Based on the information presented, it can be concluded that Big Data is not only one of the core elements of Industry 4.0 but the essence of this industrial revolution. Any technological solution is a tool for generating, collecting and processing data, but the information resulting from their analysis is crucial and invaluable to the business.

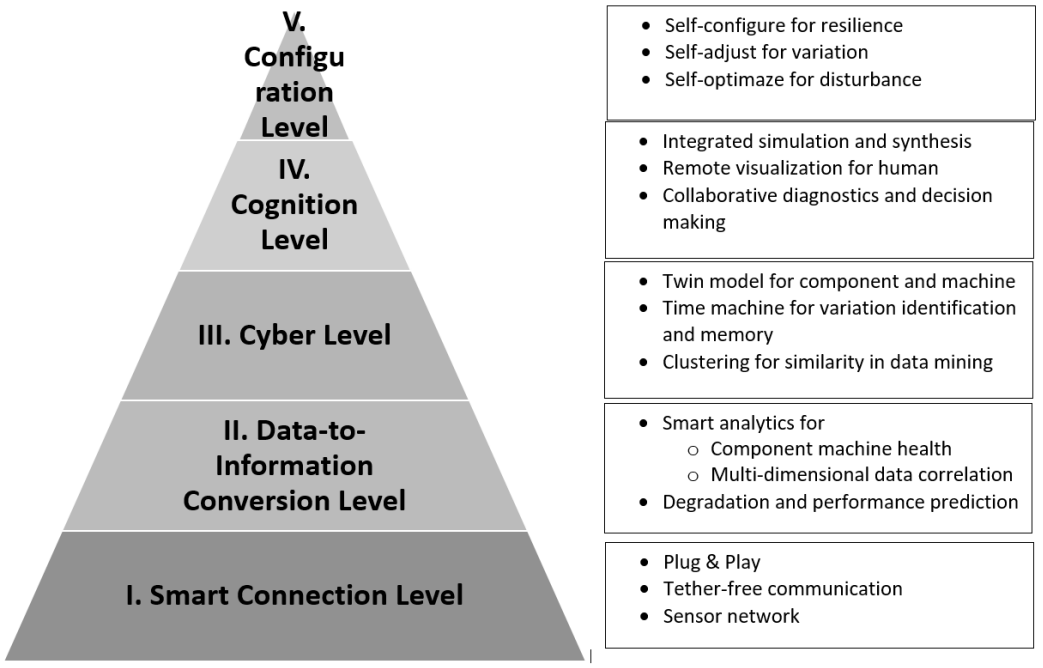

Fig. 1 5C Architecture for Designing Cyber-Physical Systems in Manufacturing
Source: [6].

POSSIBILITIES OF USING BIG DATA ANALYSIS IN INDUSTRY

The potential of analyzing large volumes of data is seen as a breakthrough in virtually all areas of life. Its use is seen as a chance to improve and facilitate many processes. Fields that currently hope for the development of Big Data technology are industry, finance and medicine. Below are presented selected potential areas for using Big Data in the manufacturing sector.

The first possible use of mass data analysis results is to improve the efficiency of processes. With the implementation of sensors, it is possible to obtain and manage material traffic data, which significantly reduces labor, storage and logistics costs. It is also important to minimize unnecessary work in the workflow of the value chain so that the process can be restructured and optimized. Prior to the digital revolution productivity gains were achieved by replacing human labor with devices and computers. At present, machines, products and even materials are intelligent thanks to the use of sensors. This means that they are able to transmit data that is analyzed in real time. This information is used to develop more detailed forecasts, which certainly contributes to more effective management [4, 10].

The use of Big Data analytics is very important in the context of decision-making. Erik Brynjolfson and his collaborators made a research that have definitely shown that data-driven decision making positively influences their position. As a result, research has developed a measure that charts companies based on the extent to which they use data analysis in decision-making processes. It turns out that the greater the use of data, the more efficient the enterprise. It should be mentioned that interference factors are also included in the study. The study found that one standard deviation above the Data-Driven Decision (DDD) average is associated with a 4-6% increase in productivity. In addition, DDD contributes to a higher return on assets, equity, more efficient use of assets, and increased market value [1]. An example may be the use of the presented concept in the fuel industry. Exploration and planning of oil and gas fields is accomplished through the implementation of a sensor network in the earth's crust. This is to precisely locate and structure the fields of natural resources. This results in lower infrastructure and transportation costs [2].

Another concept, whose development is undoubtedly dependent on Big Data analysts, is mass customization. In other words, mass customization of production will be possible thanks to the appropriate application of the data. Until recently, this idea was difficult to achieve, especially in large factories. Producing individual articles with specific characteristics was often simply unprofitable. This was associated with every design, which was more costly and time consuming with technically complicated products. The current strategies of many companies are aimed at the most precise alignment with customer expectations and delivering the product exactly as it is needed. This is a competitive strategy provided that the organization has the resources to respond immediately to customer requests. Such high flexibility of production is possible thanks to the use of elements of Industry 4.0. The construction of a system consisting of adequate IT resources and appropriate infrastructure will enable the acquisition of data directly from the customer. Ordering information through the appropriate data processing technology will be the basis for the automatic design of the product. Then the data generated on the basis of the project will create the corresponding production order to be delivered. The key to such a system is the ability to perform various tasks and the speed of performing such tasks. If one of these factors is not met, the idea of mass production individualisation will most likely be unfeasible. To achieve this, effective data processing technology is essential [15].

CONCLUSIONS

The digital revolution and the development of web technologies have contributed to the rapid growth of information resources. The result of this phenomenon is the formation of the Big Data concept, which involves the de-

velopment of technologies that enable the processing of large data volumes. It is a fundamental element of the fourth industrial revolution. Utilizing solutions in the field of Industrial Internet Things, while applying the solutions proper for efficient management Big Data can significantly increase the efficiency of manufacturing processes. Decision-making based on data has a huge impact on the performance of your business. More and more real is the realization of ideas that seemed impossible, such as the mass customization of production. Currently, the main challenge is to implement the Big Data 2.0 assumptions, that is to master flexible data processing to take full advantage of their capabilities.

ACKNOWLEDGMENT

The research was conducted within S/WZ/1/2014 project and was financed from Ministry of Science and Higher Education funds.

REFERENCES

[1] E. Brynjolfsson, M. Lorin and H. Heekyung. (2011, April 22). *Strength in Numbers: How Does Data-Driven Decisionmaking Affect Firm Performance?* [Online]. Available: https://ssrn.com/abstract=1819486

[2] *Całkowicie optyczny czujnik sejsmiczny* [Online]. Available: http://www.automatyka.pl/wiadomosci-i-komunikaty/calkowicie-optyczny-czujnik-sejsmiczny-48198-10

[3] M. Cox and D. Ellsworth, "Managing Big Data for Scientific Visualization", *ACM SIGGRAPH*, Vol. 97, May 1997, pp. 5.1-5.17.

[4] K. Kiraga, „Przemysł 4.0: 4. Rewolucja Przemysłowa Według Festo", *Logistyka*, vol. 17, no. 12, pp. 1603-1605, 2016.

[5] H. Lee and I. Sohn, *Fundamentals of Big Data Network Analysis for Research in Industry*. Chichester: John Wiley & Sons Ltd., 2016, pp. 25-38.

[6] J. Lee, B. Bagheri and H. Kao, "A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems", *Manufacturing Letters*, vol. 3, pp. 18-23, Jan 2015.

[7] McKinsey Global Institute Website. (2011, May 1). *Big data: The next frontier for innovation, competition, and productivity* [Online]. Available: http://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/big-data-the-next-frontier-for-innovation

[8] Microsoft Hololens Website. *Transform your business* [Online]. Available: https://www.microsoft.com/en-us/hololens/commercial-overview [Access: 2017.01.05]

[9] F. Provost and T. Fawcett, *Data Science for Business*. Sebastopol: O'Reilly Media, 2013.

[10] A.K. Stasiuk-Piekarska and M.K. Wyrwicka, „Organising – Still An Important Function Of Production Management", *Research In Logistics & Production*, vol. 5, no. 2, pp. 129-142, April 2015.

[11] M. Tabakow, J. Korczak and B. Franczyk, „Big Data – Definitions, Challenges And Information Technologies", *Business Informatics*, vol. 1, no. 31, pp. 138-153, 2014.

[12] A. Varghese and D. Tandur, "Wireless requirements and challenges in Industry 4.0", in *Int. Conf. on Contemporary Computing and Informatics*, Mysore, India, 2014, pp. 634-638.

[13] J. Wieczorkowski and M. Jurczyk-Bunkowska, „Big Data Jako Źródło Innowacji w Zarządzaniu i Inżynierii Produkcji", in *Innowacje w Zarządzaniu i Inżynierii Produkcji*, vol. 1, R. Knosala, Ed. Opole: Oficyna Wydawnicza Polskiego Towarzystwa Zarządzania Produkcją, 2017, pp. 134-144.

[14] P. Wittbrodt and I. Łapuńka, „Przemysł 4.0 – Wyzwanie Dla Współczesnych Przedsiębiorstw Produkcyjnych", in *Innowacje w Zarządzaniu i Inżynierii Produkcji*, vol. 2, R. Knosala, Ed. Opole: Oficyna Wydawnicza Polskiego Towarzystwa Zarządzania Produkcją, 2017, pp. 793-799.

[15] P. Zawadzki and K. Żywicki, „Smart Product Design And Production Control For Effective Mass Customization In The Industry 4.0 Concept", *Management and Production Engineering Review*, vol.7, no. 3, pp. 105-112, Nov. 2016.

Sylwia Gierej, MA
Bialystok University of Technology, Faculty of Management
International China and Central-Eastern Europe Institute of Logistics and Service Science
Tarasiuka Street 2 16-001 Kleosin, POLAND
e-mail: s.gierej@pb.edu.pl