Comparison of Data Warehousing and Big Data Principles from an Economic and Technical Standpoint and Their Applicability to Natural Gas Remote Readout Systems

Vedran Kluk

J. J. Strossmayer University of Osijek
Faculty of Electrical Engineering, Computer Science and Information Technology
vedran.kluk@student.ferit.hr

Dominika Crnjac Milić

J. J. Strossmayer University of Osijek
Faculty of Electrical Engineering, Computer Science and Information Technology
dominika.crnjac@ferit.hr

Zdravko Krpić

J. J. Strossmayer University of Osijek
Faculty of Electrical Engineering, Computer Science and Information Technology
zdravko.krpic@ferit.hr

Abstract – *In natural gas remote reading, a large amount of data is collected, posing a problem of storing and processing such data to the companies involved. Two major technologies have recently appeared, becoming a de facto standard in processing large amounts of data, i.e., data warehousing and big data. Each of these technologies provides different data processing techniques. In this paper, serial data processing and parallel data processing are considered in data warehousing and big data, respectively. The paper analyzes the feasibility of implementing new technologies for processing a large amount of data generated by remote reading of natural gas consumption. The research conducted in this paper was made in collaboration with a local natural gas distribution company. A comparison of potential software vendors has shown that Qlik offers the best software package for the requirements provided by the local natural gas distribution company. Comparison results have also shown that other potential vendors also offer software packages of good quality.*

Keywords – *big data, data warehousing, natural gas consumption, remote readout system*

1. INTRODUCTION

As the amount of data generated increases every day, specialized storage systems for processing large amounts of data were developed.

In the last two years, 90% of the total data in the world has been created. Due to a large increase in the amount of data, business systems have begun to use data warehousing (DW) and big data (BD) solutions. DW solutions are suitable for structured data that companies have been using for decades and they represent numeric values and facts, such as financial results, customer and product characteristics, which are stored in relational databases. On the other hand, BD solutions, along with structured data, also support unstructured data, and they represent digital content, such as images, videos, and text messages, and the business world often considers this kind of data relevant and essential in a decision-making process. Other technologies and approaches do exist, such as data lake (DL), which is conceptually similar to data warehousing, but since it is based on unstructured data, it is not widely used in the business world, but rather by data scientists. Additionally, there are other derivatives of the aforementioned technologies, but since in this paper we deal with a business use case, we place emphasis on the technologies used in the business sector, as in [13].

To the best of our knowledge, there is no study on the applicability of these solutions to data obtained or collected by remote reading of natural gas consumption. The most similar comparison can be found in [1], where the authors describe the potential of using DW and BD solutions in the company. Natural gas reading

for a large number of users generates large amounts of data, which can no longer be analyzed by using classical methods. This paper analyzes the feasibility of introducing DW and BD solutions for analyzing and storing data obtained by reading natural gas consumption. The analysis is done for the purpose of launching a pilot project aimed at introducing new technologies for processing and storing large amounts of data obtained from a local natural gas distribution company.

2. TECHNOLOGIES FOR PROCESSING AND STORING LARGE AMOUNTS OF DATA

Along with the advances in information technology and growing business systems and the emergence of the need for processing large amounts of data, new programming and algorithmic paradigms have appeared, introducing new ways of data processing. Given that large amounts of data cause reduced performance of processing and storage efficiency, it is necessary to reach a compromise between processing details and reasonable performance. The two main principles that successfully solve this compromise and are often used in the business sector are data warehousing and big data. Data warehousing represents modern databases and has been on the market for a long time, while big data is a relatively new way of dealing with large amounts of data. According to [8], what DW and BD have in common are features like data analyses from multiple sources, where the primary goal is to reduce costs for the company using these systems. They also emphasize improved performance when compared to classical databases, which represents part of a DW system, has significantly less storage space and is cheaper than the DW system.

2.1. DATA WAREHOUSING PRINCIPLE

According to [1], data warehouse is a large database that collects data from a variety of sources, with the aim of processing and visualizing data, providing a user with the ability to extract a hidden value from the interdependence of processed data. DW systems differ from other data processing and storage systems because of their own unique version of truth, performance, simplicity and data consistency and integrity.

The structure of the DW system consists of the following several layers: the data source, the integration layer, and the data warehouse, providing reports as a result of the aforementioned, as can be seen in Figure 1, showing the structure of the DW system. Reports are the ultimate result of collected data processing, which can be in the form of tables, graphs, texts, or other forms, depending on which a software tool is used to visualize the data and the purpose of introducing the DW system.

Data from various sources are copied and transferred to the integration layer, where the process of compiling data into the same format occurs. Then there follows a procedure of loading data into DW, after which a general report is created as the final result of collected and processed data.
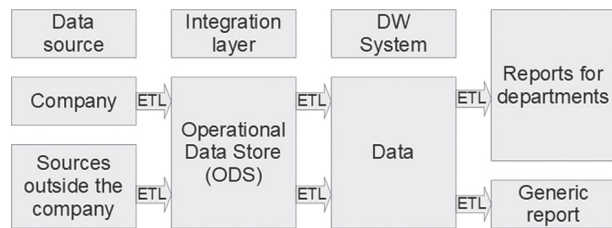

Fig. 1. Data warehousing structure

In addition, there is a specific data management process (i.e., Extraction, Transformation and Loading, or ETL) between the data source, the integration layer, DW and the reporting system. ETL is divided into three phases: decompression, formatting and loading.

Unpacking data from the data source into the integration layer is performed in order to maintain the integrity of each data source. Once the data are in the integration layer, it needs to be in a common format. After formatting, data are uploaded to the DW system to get the final result. Reports are available and can be used for different purposes. Just to illustrate, a company can use reports as guidance for company management when making decisions vital to its business. When the data are uploaded to the DW system, the reports are generated and compiled using data collected from various sources. The reports are properly structured if they are simple and provide specific and concrete information on all relevant issues the DW system end user wants. It is easier to solve the problem on a smaller data set than on a larger one, and this is precisely how the DW generated reports work.

There are several major DW based technology vendors on the market, such as Oracle, Teradata, Amazon Web Services, clouders, etc.

According to [2], DW system vendors and their solutions are presented in Table 1.

Table. 1. Data warehouse vendors and platforms

Data warehouse vendors	Data warehouse platforms
Oracle	Present for decades in the database market. Platform: Oracle Exdata Machine
Teradata	Present for 30 years in the market. Platform: Teradata EDW (enterprise data warehouse) that provides robustness, hybrid storage capability.
Amazon Web Services (AWS)	A front-runner in the cloud data storage Platform: Amazon Redshift
Cloudera	A company known for its Enterprise Data Hub (EDH) platform specialized in data warehousing.

2.2. BIG DATA PRINCIPLE

According to [3], big data is a term that describes large amounts of data (structured and unstructured) collected daily from a variety of sources of interest to the user.

According to [4], the purpose of applying the BD principle in a business is to spot potential problems in the company's daily operations, to provide insight into potential business problems and to anticipate future events accurately.

The BD principle is used when the data cannot be processed in a classic way and/or by using a single computer to provide appropriate results, because of the fact that the computer receives and processes data serially. That is why parallel data processing is required: computers receive smaller amounts of data, process them, and send them to a computer that will compile final results in form of a report. This report gives the user an insight into collected and processed information. The BD principle implies several open source programs working together to process large amounts of data coming from different sources in a better and faster way.

There are different platforms that allow system operation according to the BD principle, e.g. Apache Spark, Apache Storm, Ceph, Disco and others.

In addition to the aforementioned, the platform that is most represented and considered relevant for describing the system operating mode according to the BD principle is called Hadoop, and the programming models that make up the specified platform are HDFS (Hadoop Distributed File System) and MapReduce. According to [5], Hadoop is an open source platform for data storage and application running on interconnected computers. Hadoop supports structured, semi-structured and unstructured data. Also, due to its characteristics, Hadoop has the ability to process a virtually unlimited number of processing tasks and jobs simultaneously. Figure 2 shows the principle of the Hadoop platform.
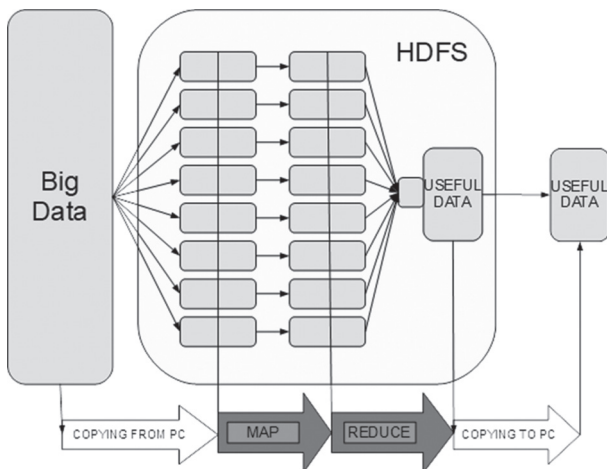

Fig. 2. Working within the Hadoop environment

According to [6], MapReduce consists of two parts: Map and Reduce.

The Map represents an action of gathering data from one format and converting it into data of a different format so that each data format gets a key depending on which data it represents.

After that, the Reduce action distributes data depending on their keys and sends certain information to the user. MapReduce can perform two functions; manage computer resources and data processing. Hadoop has several versions, e.g. 1.0, which includes HDFS and MapReduce, while version 2.2 adds a YARN (Yet Another Resource Negotiator).

In Hadoop Version 2.2, YARN assumed one of the MapReduce functions, i.e., computer resource management in order to further accelerate parallel data processing.

BD is applied in large corporations made up of several smaller, interconnected companies. According to [6], some of the companies that have successfully implemented the BD principle are: Walmart (supermarket management), Rolls-Royce (application in luxury cars production), and Facebook (user observation).

According to [7], there are several major BD solution providers, such as Tableu, IBM, SAP, Qlik, Microsoft, SAS and others.

Table 2 shows vendors using the BD principle and their solutions.

Table. 2. Big data vendors and their software packages

Big data vendors	Big data solutions
Tableu	Started as a college project, offers visualization of all data types from various sources, from Hadoop to Excel files
IBM	Cloud server for storing and processing large amounts of data, Program support name: **DB2, Informix, and InfoSphere**.
SAP	Primary tool: **HANA** (relational database that can run an analysis on 80 terabytes of data).
Google	Main product: **BigQuery**, based on cloud processing of large amounts of data.

2.3. DATA LAKE (DATA SWAMP)

A data lake is a collection of raw data founded on the big data principle. It is usually a single warehouse of all data in a single company which includes reports, visual data analyses, data obtained by machine learning and the like, as in [12]. On the other hand, a data lake can contain data stored as relational databases, as well as half-structured data (e.g. XML, JSON, etc.) and unstructured data (e.g. email, PDF, etc.). A data lake is similar to data warehousing, except that the data contained within is not structured, the purpose of the data within is not yet known, and the data is easily accessible and

quickly updated, [13]. Hence, a data lake is more frequently used by data scientists than by business professionals. Considering the structure of the data in the use case presented in this paper, the data lake technology is not an option. There are other names for the technology, such as a data swamp, as in [13], since the data stored in the data lake is not structured, but rather just "thrown" in the pool of data.

2.4. COMPARISON OF DATA WAREHOUSING AND BIG DATA PRINCIPLES

According to [8], every user managing large amounts of data should have a system that can collect and process data from various sources of interest.

Systems and principles processing large amounts of data in real time, and collecting data and providing processing results within a reasonable time can be compared by using the following parameters: sustainability and implementation complexity, availability, modularity, storage speed and data processing, and business benefits.

System sustainability is seen in terms of human resources. Investment in people capable of using company resources effectively saves time and money that can be invested in other projects improving the overall business operation. When the sustainability of the desired system is defined, availability, another important parameter of system evaluation, has to be considered.

According to [9], the system is available when users can work without interruptions. A system availability benchmark is customer own satisfaction and their subjective impression on whether something is available or not. The most common sources of interruption (a decrease in availability) are: hardware, software, and the human factor. Service availability interruptions are reduced by improving the quality of system components, both hardware and software. If user-operated systems are not fully automated but require occasional human intervention, there are possible problems caused by the very same person using the system, not hardware or software support. DW system availability is often questioned due to data duplication problems in databases. Data are not synchronized in all databases, which increases data processing time. In order to address this problem, it is necessary to create an on-line copy of the data, which is also used by an off-line copy.

An on-line data copy is used in the ETL process and other database maintenance activities to achieve the impression of DW system availability, along with data processing speed.

An off-line data copy is a standard database created when the data is collected. By introducing an online data copy, the required synchronization is achieved and the data duplication problem is eliminated.

On the one hand, DW systems are seen as a solution to the problem of processing and storing large amounts of data. On the other hand, Hadoop is viewed as a framework designed to apply the BD principle in data processing and is considered a potential solution to the problem of processing and storing large amounts of data. It consists of two parts: HDFS (in charge of storing and serving data), and MapReduce (in charge of data processing). The HDFS system is based on computer distribution processing the data in parallel, allowing the entire system to work if a failure occurs on one of the computers processing data. In this case, computers processing the data in parallel take over tasks from a computer that might be malfunctioning.

2.5. ADVANTAGES AND DISADVANTAGES OF DATA WAREHOUSING AND BIG DATA PRINCIPLES

The difference between data warehousing and big data principles is that DW uses serial data processing, while the BD principle requires parallel data processing and, in addition, BD supports unstructured data. There are also hybrid systems that allow DW to use parallel data processing.

Table 3 shows the advantages of data warehousing and big data.

Table. 3. Advantages of data warehousing and big data

Advantages of data warehousing	Advantages of big data
• Analysis of multiple data sources in one place • Efficient analysis of large amounts of data • Proper decision making and greater profit potential for the company • Improved business intelligence • Improved performance compared to classical databases • Timely access to data • Consistency and data standardization • Business cost reduction	BD includes all aforementioned advantages of DW, but also the following: • Less time needed to process data • New product development • Understanding the market and its conditions • Improvement of company's business operations through the right interpretation of the collected information • Partner trust by using new data analysis methods

There are three common disadvantages of introducing DW and BD technologies, i.e., a long process of introducing these into daily operations, fear of the unknown, and the need for educating people who will use this technology. The most significant disadvantages of DW are serial data processing and misestimation of the time necessary to carry out the ETL process.

Table 4 shows the disadvantages of DW and BD.

On the other hand, if BD is used, there is a problem of visualizing large amounts of data collected from different sources. Also, it is necessary to determine the quality of the collected data in terms of their ability to improve company's business performance.

Table. 4. Disadvantages of data warehousing and big data

Disadvantages of data warehousing	Disadvantages of big data
• High maintenance costs • Serial data processing • Misestimation of the time required to complete the ETL procedure • A long process of introducing the DW system into daily operations of the company	• Visualization of large amounts of data • Fear of change (the unknown) • Potential security attacks due to confidential business information

3. CASE STUDY: PILOT PROJECT APPLYING DATA WAREHOUSING AND BIG DATA PRINCIPLES TO THE REMOTE NATURAL GAS METERING SYSTEM

New technologies are desirable in business operations of many companies. A local natural gas distribution company recognized this and presented a pilot project aimed at introducing remote natural gas reading (Project 1). As part of Project 1, a sub-project referring to the introduction of new technologies for data storage and processing of the collected remote readings was also presented, which is defined as large data sets (Project 2). For additional information, please refer to [10]. According to [10], if remote reading of natural gas consumption is not profitable, storage and data processing is pointless. The study was carried out as part of this paper and the applicability and viability of Project 1 were tested because the results are necessary for future projects aimed at introducing technologies for processing and storing large data from other sources.

When it comes to natural gas consumption, the following data were used in the analysis and collected by means of remote reading: the current state of the device that collects data, temperature, daily natural gas consumption and unauthorized tampering detection. Other data collected are not of major importance to project cost-effectiveness analysis. One such data is a GSM summary.

In addition, the process of remote consumption natural gas metering takes place such that all these data are collected on an hourly basis and communication between data collecting devices (data loggers, DL) and software via the GPRS/SMS protocol takes place every day at the exact same time and this communication is open for a certain limited period of time (a couple of minutes). Communication is then interrupted to save DL battery. After that, a new data collection cycle starts in order to collect information so that the data could be sent through the communication channel at the same time the next day.

Fig. 3. Remote readout of natural gas consumption

From the description of the data collection process, it can be seen that it results in a large amount of data, especially taking into account a large number of consumers. The local distribution company is responsible for 10,633 consumers (gas meters); if each DL processor reads at least 5 different data sets, there are 53,135 different data sets stored in just a few minutes and they need to be processed in the shortest time possible.

Currently, the number of consumers connected to the local natural gas distributing company remote natural gas consumption metering system is not large and data processing is done in a classic way. However, for future needs, if most of the distribution network users decided to connect to the natural gas consumption remote metering system, data flows would drastically increase, which will have to be addressed by using special methods and technologies. Another example can also be found in Croatia, where the biggest electricity provider HEP Group uses the big data principle to collect, store, structure and process the data collected by means of remote reading electrical energy consumption. The data collected have a structure similar to the data collected by remote reading of natural gas. The same company also plans to introduce the big data principle for another similar purpose, i.e., to model the behavior and habits of electrical vehicle charging station users, [15].

By incorporating BD and DW for the use case presented in this paper, processing the data collected by remote reading of natural gas proves its feasibility by the fact that these technologies are a natural option for processing the data obtained by various remote reading systems. E.g., sources such as [14] denote the complex nature of the analysis of data obtained by remote readout systems.

3.1. DATA STORAGE AND PROCESSING SOFTWARE REQUIREMENTS

Each project has certain goals and requirements that must be met to make the project successful. Important requirements of the future pilot project are as follows: product prices, delivery time and the way data are stored.

Product cost is one of the most important factors (an optimal price is formed on the invested-gained principle), the delivery time defines the importance of meeting the deadlines for mutual satisfaction of both customers and suppliers, while the way data are stored can be divided into two groups: locally (on the premises) and in the cloud (Cloud). Local vendors have an advantage when it comes to storing collected data locally.

Other requirements for a future pilot project include modularity, investment durability, an SLA (Service Level Agreement) contract, data storage and data processing speed, and database snapshot.

Modularity represents an upgrading option without changing the overall system (changing certain parameters, dynamically refreshing and generating reports,

and adding new functionalities). Investment durability refers to the availability of the installed equipment for the purpose of project execution and must be designed to last for a minimum of 5 years, for both application and circuitry.

It is also necessary to define a Service Level Agreement (SLA) contract regulating system maintenance service availability, customer support, the ability to educate your own team for proper system use and certification, system downtime, addressing complaints and the deficiency submission deadline.

3.2. POTENTIAL VENDORS FOR THE FUTURE PILOT PROJECT

Based on prerequisites imposed on software solutions for processing and storing large amounts of data, the local gas distributing company considered the following three providers: Microsoft, Qlik and SAS (Statistical Analysis System). Each of these potential bidders has conducted storage and data processing projects, and each project has its own specific characteristics. Table 5 shows a comparison of potential bidders.

Table. 5. Comparison of potential vendors

SAS	Microsoft
Present in financial and telecommunications industries, more than 40 years of experience in providing analytical software solutions. Interesting to large and rich industries dealing with large amounts of data.	Builds the DW/BD platform on existing solutions: • SQL Server • Analytics Services • Reporting Services • Power BI All these services are platforms for themselves and have their own way of working and their own language.
Qlik	
Full-stack platform where all layers up to analytics and visualization are integrated into one on-premise solution. The Qlik platform does not use a classical database for data storage; it uses a proprietary flat-file infrastructure write system optimized for speed and ease of data entry and retrieval.	

Microsoft has a huge infrastructure in terms of servers that can handle large amounts of data in a short period of time. As such, Microsoft is convenient for storing and data processing which is good for companies without their own infrastructure. However, data privacy in cloud mode is questionable. The advantage of cloud mode storage is that, according to their needs, the users can expand storage capacity as well as computing power needed to process increasing amounts of data. At the user's request, the process of expanding data storage capacity is available in less than 60 seconds.

Another potential solution provider is SAS, which has been present in major industries such as the financial and telecommunications industries for over 40 years providing analytical software solutions. Some of SAS applications for storing and processing large amounts of data are SAS Data Loader for Hadoop (used for data upload), SAS Enterprise Guide, SAS Visual Analytics (used for data analysis) and SAS Visual Statistics (for visualization of processed data), forming together a platform capable of processing large amounts of data and visualizing processing results for the end user.

The third potential bidder is Qlik, which offers a full-stack platform where all layers up to an analytical layer and visualization are integrated into a single on-premise solution. Server parallelization aimed at increasing the speed is possible, indicating the application of the BD principle.

Once the desired characteristics of the pilot project are sent to potential bidders, the best offer for the company is selected. One specific issue that the local natural gas distribution company deals with is storage and data processing after remote natural gas consumption reading. Potential bidders must comply with the requirements and decide whether they want to participate in the project.

According to publicly available information in [12] and analysis of the necessary requirements, it can be seen that SAS is by far the most expensive option; however, it offers the most effective data processing tools, which is interesting to major companies. After SAS there follows Microsoft and then Qlik as the most favorable bidder. All potential bidders offer required storage methods (cloud and methods using their own infrastructure). All manufacturers also offer modularity in terms of software support, with Qlik being the most flexible one. All three bidders also offer a database snapshot as well as high speed data processing that can be further regulated depending on the contract type. With all that being said, Qlik is a logical choice for the current needs of a local natural gas distribution company in terms of storage and processing of large amounts of data due to its own flat-file infrastructure, data processing speed and favorable cost.

3.3. FEASIBILITY ANALYSIS OF THE FUTURE PILOT PROJECT

Once the data are collected, it is necessary to store and process them for informational purposes. Since we are currently dealing with a small number of consumers connected to the remote natural gas consumption reading system, data processing can be carried out traditionally by using human resources. However, due to a large amount of data, parallel computer data processing is used, providing results within a reasonable time. With the introduction of the DW system and the BD principle, certain resources are invested, which seems illogical at first glance. The question is why the consumption of natural gas is not measured on a monthly basis. The answer lies in the fact that the local natural gas distribution company owns a 595 km distribution network, which is not easily covered by a small number of people and the data collection requires a certain period of time. Also, measuring once a month is not per-

fect because the distribution company does not have any insight into gas consumption for the entire month, and huge losses can occur in a month, especially in businesses whose share in the overall consumption will most definitely grow in the future. If the problem of remote measuring, storage and data processing is viewed from an aspect that has just been mentioned, it is obvious that ultimately it is very profitable to introduce remote natural gas consumption reading as well as the system for storing and processing natural gas consumption data.

4. CONCLUSIONS

The most well-known principles for processing and storing large amounts of data are data warehousing and big data. These principles have been implemented by more and more companies as the amount of data they operate with is on the increase. In this paper, in co-operation with the local natural gas distribution company, we offer a feasibility study conducted through a pilot project for the introduction of the data warehousing and big data technologies, using data collected by remote natural gas consumption reading. As part of this paper, the requirements for the aforementioned pilot project have been outlined, with the main aim of applying the DW and BD principles to remote natural gas consumption reading and measuring systems. Based on the defined requirements, the existing appropriate software solutions are described using publicly available information. The bidders described are Microsoft, Qlik and SAS. Also, by comparing the two aforementioned principles, it can be seen that they can be used together since they basically complement each other. The cost-effectiveness of the project for natural gas consumption remote reading as well as the pilot project for the introduction of a large amount of data processing and storing was also analyzed. Finally, the results show that potential vendors can be adequate for carrying out the future pilot project because they show that they have platforms capable of processing large amounts of data in a fast and efficient manner.

5. REFERENCES

[1] A. Wicks, Data Warehousing - A Quick Summary, https://www.youtube.com/watch?v=zTs5zjSXnvs&t=10, (accessed: 2018)

[2] B. Inmon, Big Data Implementation vs. Data Warehousing, http://www.b-eye-network.com/view/17017 (accessed: 2018)

[3] J. Walker, Top 5 data warehouses on the market today, http://www.monitis.com/blog/top-5-data-warehouses-on-the-market-today/ (accessed: 2018)

[4] R. Buyya, R. N. Calheiros, A. V. Dastjerdi, "Big Data - Principles and Paradigms", 1st Ed., Elsevier, 2016.

[5] M. Rouse, Hadoop Distributed File System (HDFS), http://searchdatamanagement.techtarget.com/definition/Hadoop-Distributed-File-System-HDFS (accessed: 2018)

[6] B. Marr, Big Data: What is Hadoop - An Easy Explanation for Absolutely Anyone, https://www.bernardmarr.com/default.asp?contentID=1080 (accessed: 2018)

[7] B. Marr, "Big Data in Practice – How 45 Successful Companies Used Big Data Analytics to Deliver Extraordinary Results", Wiley, 2016.

[8] A. Patrizzio, Big Data Companies, https://www.datamation.com/big-data/big-data-companies.html (accessed: 2018)

[9] Educba, Difference between Big Data and Data Warehouse, https://www.educba.com/big-data-vs-data-warehouse/ (accessed: 2018)

[10] UN, Big Data for Sustainable Development, http://www.un.org/en/sections/issues-depth/big-data-sustainable-development/index.html (accessed: 2018)

[11] V. Kluk, "A comparison of data warehouse and big data principles and their application to systems of remote natural gas and electrical energy meter reading", J. J. Strossmayer University of Osijek, Faculty of Electrical Engineering, Computer Science and Information Technology, Osijek, Croatia, Master thesis, 2018. (In Croatian)

[12] D. Harris, From Data Lakes to HTAP: 3 Alternatives to OLAP Data Warehouses, Software Advice, https://www.softwareadvice.com/resources/olap-data-warehouse-alternatives/#lakes (accessed: 2019)

[13] E. Hitchcock, BI vs. Big Data and Data Warehouses vs. Data Lakes, Datameer, https://www.datameer.com/blog/video-bi-vs-big-data-data-warehouses-vs-data-lakes/ (accessed: 2019)

[14] M. Zhanyu, X. Jiyang, L. Hailong, S. Qie, S. Zhongwei, Z. Jianhua, G. Jun, "The Role of Data Analysis in the Development of Intelligent Energy Networks", IEEE Network, Vol. 31, No. 5, 2017, pp. 88-95.

[15] Hrvatska elektroprivreda, bigEVdata project description, http://elen.hep.hr/HEP-eMOBILNOST-ciljevi.aspx (accessed: 2019; in Croatian)