International Journal of Advances in Mechanical and Civil Engineering, ISSN: 2394-2827 Volume-6, Issue-5, Oct.-2019
http://iraj.in

BUILDING INFORMATION MODELING (BIM) AND BIG DATA ANALYTICS FOR CONSTRUCTION INDUSTRY

[1]GOKHAN DEMIRDOGEN, [2]ZEYNEP ISIK, [3]YUSUF ARAYICI

[1,2]Yıldız Technical University Department of Civil Engineering,
[3]Hasan Kalyoncu University Department of Civil Engineering
E-mail: [1]gokhand@yildiz.edu.tr, [2]zeynep@yildiz.edu.tr, [3]yusuf.arayici@hku.edu.tr

Abstract - The construction industry suffers from inadequate data management of current and past projects. Since, many projects result in problems such as exceeded budgets, time delays, quality problems, conflicts and missing data transfer between stakeholders, Building Information Modeling (BIM) was adopted in the industry to overcome those problems. With the help of BIM, information and analysis results of a project can be collected and used in various decision-making processes. However, a BIM model that consists of semantic and geometric information needs to be integrated with other methods to both manage and achieve necessary information without creating a software bottleneck. Besides, facility managers using BIM in the operational phases of a project need to gather dynamic data that produces too much information. In this context, Big Data Analytics has come into prominence to overcome the data volume, variety, and velocity in building the data resource for BIM model. Thus, in this study it was aimed to make an in depth literature review to reveal the advantages and disadvantages of BIM and Big Data Analytics integration.

Keywords - Building Information Modeling (BIM), BIG Data Analytics.

I. INTRODUCTION

The construction industry currently embraces the data revolution to improve its performance to be more practical, productive, and economical. Also, this revolution is necessitated by complex engineering problems, and analysis of massive data sets is required. For example, according to Reference [1], emerging testing and monitoring systems play an important role in the proliferation of excessive data in civil engineering.

In addition to complex engineering problems, there is tendency to accumulate huge project databases because of multiple data resources that come from various stakeholders. Those data resources can be in various formats, such as "DWG, DOC/XLS/PPT, RM/MPEG and JPEG" that are used to present "project engineering data and project management data" [2]. In that context, Building Information Modeling (BIM) presents comprehensive data storage and management services to construction industry professionals because "BIM is a digital construction process of the project" [3]. The reason for wide applications of BIM in the construction industry is to have data that are related to projects' geometry, spatial relationships, geographic information, quantities and properties of building elements, cost estimates, material inventories, project scheduling, procurement details, and submittal processes [3].
Despite the fact that BIM utilization in the construction industry provides a good opportunity, Reference [4] stated that integration of BIM and as-built documentation is a challenging process because of heterogeneous and unstructured data formats. To overcome those barriers, the property of big data named variety helps users, because variety in big data applications corresponds to "structured, semi-structured, and unstructured" data analysis capabilities [5]. Big data utilization in BIM can be beneficial to discover root causes of poor building performance, to perform real time data queries, to improve the decision-making process, to improve productivity, and to reveal new designs and services in the construction industry, as is the case in every industry [6]. The construction industry has negative productivity records when compared with other industries. Also, reference [6] stated that systematic barriers play an important role in negative productivity. Those systematic barriers can be explained as lack of "talent, IT intensity, data driven mind-set, and data availability". For example, construction companies have limited stored data in the industry (51 petabytes in the USA) and stored data per firm (231 terabytes in the USA). However, the "big data value potential index", which is presented in the figure, shows that the construction industry can benefit moderately from big data.

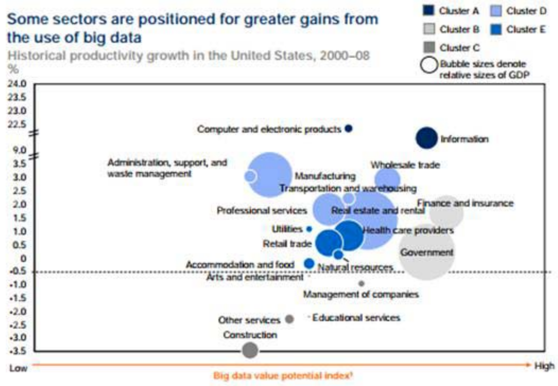

Figure 1. "Some sectors are positioned for greater gains from the use of big data" [6]

As is mentioned by reference [6], a systematic barrier that is related to data availability can be overcome with BIM's features. According to reference [4], the concept of BIM provides a data management platform. With the help of that platform, the participants can retrieve and share necessary information among them. Therefore, the aim of this paper is to clarify the BIM and big data analytics relationship. Also, advantages and disadvantages of the interrelationship will be discussed in the context of the study.

II. BUILDING INFORMATION MODELING (BIM)

According to reference [7]'s study, BIM contains numerous definitions as a result of the users' requirements. Some of these expressions emphasize "virtually constructing a building, planning and maintaining environment, and discovering new knowledge for buildings via what-if analysis." As can be understood from BIM definitions, BIM has promising features for the construction industry in terms of designing and management. Furthermore, BIM provides a common virtual environment in which different stakeholders can identify and review their tasks in the project. Thus BIM allows stakeholders to interact each other.

On the other hand, BIM provides positive outcomes on project success. Some of them were summarized by reference [8] as "cost reduction and control, time reduction and control, communication improvement, coordination improvement, quality increase and control, negative risk reduction, scope clarification, and organization improvement."

In addition to positive outcomes of the project, the BIM model needs information specific to the construction project. According to reference [9], BIM technology consists of four data layers. These layers help data classification to build a construction project in a virtual environment. The layers and their variables and components are given in Figure 2. The figure also shows information in the model that can be extracted to use in another analysis, in another project, or in management activities. Data properties identify "characteristics of BIM objects." Relation is used to define "logical and physical relation between properties". Standards have been developed to enable interoperability between users. Utilization is used to show and improve BIM utilization in different areas.
BIM maturity is also important to achieve necessary information via BIM models. Reference [10] divided the maturity of a BIM into five categories. The names of the categories are ad-hoc, defined, managed, integrated, and optimized maturity levels. According to reference [11], the maturity levels of BIM comprise four categories. In level zero, CAD formats are used, and it is impossible to manage construction projects with CAD files. In level one, there are 2D and 3D models, and they can be used to manage data with stand-alone software. In level two, there are only 3D models, and their data are separately managed with software that is designed to provide a service in the stakeholders' specialty areas. In level three, the BIM model is open to stakeholders to design and retrieve important information on open servers.

Also, BIM utilizations are varied according to various needed information and management activities in construction projects. Some of the examples are 4D BIM (time), 5D BIM (cost), 6D BIM (project life cycle information), an energy model, a waste management tool, and others. Those efforts help construction professionals to create and refine new information with BIM models.

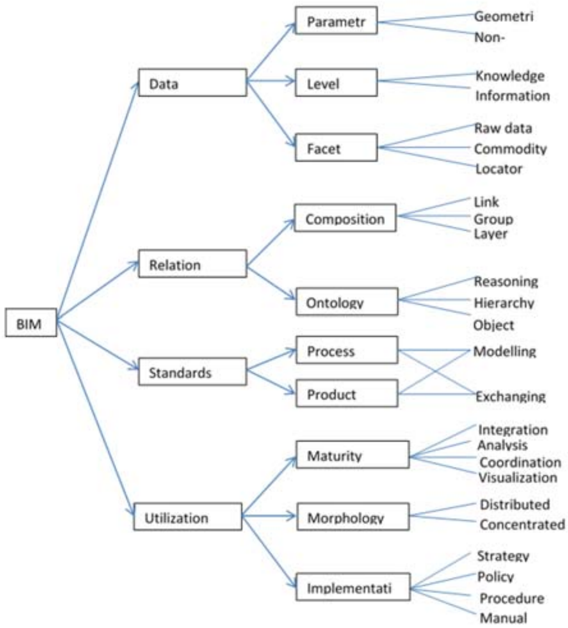

Figure 2. "Variables of BIM framework" [9]

III. BIG DATA ANALYTICS

Big data analytics have an influence on all industries. That impact is so great that it must be used "from one-person companies to Fortune 500 enterprises" [12]. According to a Forbes essay that was written by reference [12], the companies produce data via payment methods, sold products or services, social media, websites, and others. Thus, the companies can comprehend market trends via data storage. Also, the construction industry cannot be separated from the data revolution. Big data analytics is basically used to analyze data and to draw conclusions with the help of these analytics [13].
Big data is also used for massive data sets, because there is a limitation to analyzing massive amounts of data with traditional methods. Therefore big data analytics are seen as a helper tool for users to

understand and analyze complex data sets and different data types [5]. Complex data sets and different data types can be understood from current technology resources (e-mail, social media, sensors, mobile phones, data storage, and others). According to reference [5], humans had produced nearly 5 exabytes of data prior to 2003. However, now humans are producing 2.5 exabytes of data every day with varied data types. To understand hidden values from these data sets or to reach more informed decisions in projects, the construction industry must be integrated with big data analytics to be more realistic and optimized.

The characteristics of big data analytics are defined with the 3V's concept. These concepts are namely volume, variety, and velocity. Volume is defined by the size of data sets. Also, there is discussion about size of data sets that qualifies data sets as big data problems. According to survey results, one terabyte is enough to describe a data set as big data. Variety defines the types of data. Types of data can be in the form of structured, semi-structured, or unstructured. Tabular data is considered to be structured data. Video, images, and other visual data are examples of unstructured data. Also, sensors that are used in the buildings are examples of unstructured data. There is no standard for semi-structured data, but it includes tags to analyze data. The last characteristic of big data analytics is velocity. Velocity defines the speed of data generation from data resources. After practitioners accept big data analytics, some new concepts are added to the 3V concepts. They are namely veracity, variability, and value. While veracity emphasizes reliability of data, variability focuses on varied data generating rates. Value refers to the amount of data that can be used in business as a result of big data analytics [14].

A. Components of BIG Data Analytics

Reference [15] have merged all processes of big data analytics for the health care industry. According to the article, big data architecture was separated into five steps, and all steps included sub-processes. The architecture is given in Figure 3. The figure has been revised according to the construction industry. In the data layer, gathered data from different resources are classified as structured, semi-structured, and unstructured data. In the data aggregation layer, there are three important processes. In the first step, data is read from different resources. In the second step, "transformation engines", gathered data, which can be in different formats (structured, semi-structured, and unstructured data), is converted to a "standard data format". Afterwards, processed data is transferred to databases that are used in big data. Therefore, advanced analysis can be performed. In the "analytics layer", recorded data on databases is processed with parallel computing to make an analysis. In the "information exploration layer", results of analyses are obtained by users to make informed decisions. In the last step, the -"data governance layer", "data management, data life-cycle management, and data privacy and security" are handled.

The Apache Hadoop platform is commonly used in the literature. Apache Hadoop architecture allows parallel processing to analyze big data sets. Users can benefit from different servers or computer clusters to realize parallel processing. To realize parallel calculation or processing on different servers or computer clusters, the Hadoop platform uses the Hadoop Distributed File System (HDFS). HDFS stores data, but it divides the data into different nodes [16]. On the different nodes, a MapReduce procedure is used to get necessary information from data sets. MapReduce consists of two processes, namely Map and Reduce. In the Map process, data that are on different nodes are converted to key and value pairs. As a result of the Map process, outputs are combined with reduce task [17]. Obtained outputs are stored again on the HDFS. However, HDFS and MapReduce (MR) are not the only examples of big data components. Tachyon can be used instead of HDFS, and Directed Acyclic Graphs can be used instead of MapReduce. The counterparts can be investigated according to the users' problems.

B. Benefits of BIG Data Analytics

According to reference [18]'s report, big data analytics can be used in "customer relations, business intelligence, and many analytic applications". Reference [18] performed a survey in order to identify benefits of big data analytics in the report. According to the results, "social-influencer marketing, customer-base segmentation, recognition of sales and market opportunities, customer behaviors, and consumer behavior from clickstreams" were identified as benefits of big data analytics in costumer relations. In the business intelligence part, "accurate business insights, an understanding of business change, better planning and forecasting, and the identification of root causes of cost" were identified as important benefits of big data. In analytic applications, "fraud detections, risk quantification, market trends, and real-time decisions" were identified as benefits of big data.

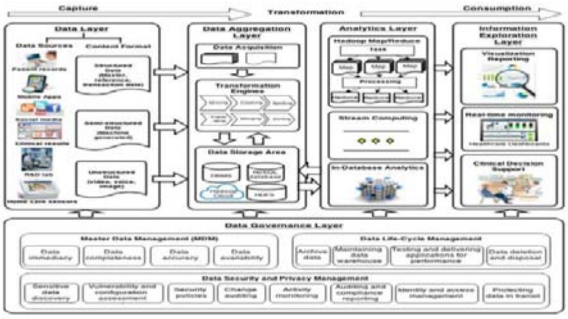

Figure 3. BIG Data Analytics architecture in construction projects (adapted from [15])

In [19], benefits of big data were summarized as "increasing operational efficiency, informing strategic direction, better customer service, identifying and developing new products and services, enhanced customer experience, identifying new markets, faster go to market and complying with regulations". In addition to them, reference [5] stated that "becoming more competitive and extracting meaningful value" could be provided with big data analytics.

When reference [15] were studying big data utilization in the health care industry, they also mentioned benefits of big data analytics. The authors divided the benefits of big data into five parts. They are namely "IT infrastructure benefits, operational benefits, managerial benefits, strategic benefits, and organizational benefits". Those parts were given with their sub-dimensions. According to the sub-dimensions, big data analytics have impacts on cost decrease in operations, time decrease, increase in quality, efficient resource management, increase in productivity, informed decision making, increase in performance, cost leadership, product differentiation, and an overview of working style and organizational learning.

As can be understood from the above benefits, big data analytics has found wide application areas in different industries, and it presents important features for use in the construction industry. As is known, projects are generally managed with conventional methods in the construction industry. In addition to conventional methods, the fragmented structure of construction projects and early-stage utilization of BIM result in data loss. This data loss was summarized by reference [20] in Figure 4. According to their study, while the bold line represents utilization of non-collaborative BIM efforts, the thin layer represents a fully collaborative environment with advanced BIM utilization. The bold line also shows data loss between project phases. To overcome data loss, a fully collaborative BIM environment must be used. Also, if a BIM environment is combined with big data analytics, the hidden values from BIM and all data can be mined, managed, classified, and stored for use in different projects.

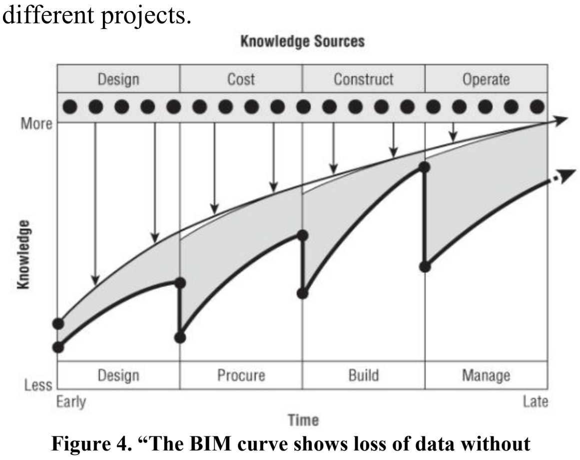

Figure 4. "The BIM curve shows loss of data without interoperability at project milestones" (Bernstein, [adapted from [20]])

IV. LITERATURE REVIEW FOR THE STUDIES THAT COMBINE BIM AND BIG DATA ANALYTICS

In the literature, reference [21] conducted a comprehensive literature review to clarify big data importance in the construction industry. According to their study, big data has relationships with many disciplines. They summarized them in Figure 5. The authors thereby investigated the studies that were performed with these traditional methods to show applicability of big data analytics in the construction industry. In other words, in the literature, the authors have benefited from some methods that have relationships with big data.

According to reference [21], the construction industry can benefit from big data analytics to solve problems related to "resource and waste optimization, generative designs, clash detection and resolutions, performance predictions, visual analytics, social networking services, personalized services, facility management, energy management and analytics, Internet of Things, and smart buildings". The study also clarified opportunities for integration of BIM and big data analytics. Briefly, integration of BIM and big data analytics can be used for project data management, construction process management, waste management and optimization, management of FM activities, design optimization, and energy management.

Another literature review related to BIM and big data analytics was performed by reference [22]. The understanding from the study is that it is important for a cloud-based dynamic BIM to utilize big data analytics and its data mining techniques [22]-[23], because it is believed that online processing of dynamic data flow can be solved with cloud-based approaches. Also the study emphasized that life cycle energy management can be provided with big data and BIM collaboration.

Reference [24] performed a study that combined BIM and Building Management Systems (BMS) to improve the design and operation phases in terms of energy performance. The study emphasized that much information is gathered with the help of BMS, but the collected information is not used very well, because integration of BIM and BMS needs to be used with the analyzed data. Also, the authors emphasized this situation as a "big data syndrome". However, they did not benefit from big data analytics; they benefited from automated process solution. It shows that integration of big data analytics and BIM is inevitable. In addition to the study of reference [24], reference [25] made a literature review to identify the studies that are related to environmental sustainability analysis of construction projects with GreenBIM applications. The study stated that much information is generated and stored through the life cycle of the project to enable building performance to be more environmentally sound. However, the current studies that have taken place in the literature have not given place to big data analytics. In other words, integration of big data analytics and BIM with the aim of GreenBIM will be future topics for the construction industry.

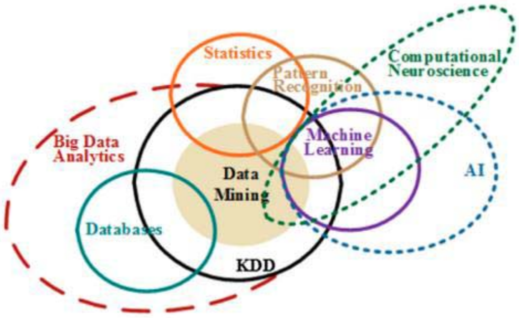

Figure 5. Multidisciplinary nature of Big Data analytics [21]

Another study, performed by reference [26], stated that calculation of life cycle assessment of building energy could be solved with BIM and big data analytics. In the study, the solution was given as a framework, but the integration of big data and BIM is ambiguous and unclear because the authors did not identify what they stored and processed in the databases by big data analytics.

Additionally, facility management specific to safety management was provided with integration of BIM and big data analytics [27]. The authors made that integration because integration of wireless sensors into BIM is difficult. Therefore, the authors benefited from big data analytics to make an analysis, which was gathered from sensors, for safety management. Obtained results were visualized on BIM models so that facility managers could monitor and manage with the help of those technologies to prevent hazards in the buildings.

Finally, reference [28] performed a study that integrated BIM and big data analytics. In the study, BIM models were uploaded to cloud systems, and necessary information could be queried and obtained with utilization of the big data application. Also, dynamic data that were obtained from sensors could be analyzed with big data to manage the facilities in the study.

As a result of literature review, advantages and disadvantages of BIM and Big Data Analytics are presented in Table 1.

Advantages	References
Time reduction for analysis	[28]
Storage capacity	[26], [28]
Query on model	[28]
Data mining based on BIM model	[28]
High reliability in analysis	[21]
Automated design	[21]
Increase in BIM utilization in facility management	[21]
Increase BIM models' utilization in operation and maintenance stage	[21]
Design improvement	[22]
Cost decrease in modeling as-built BIM model	[22]
Disadvantages	References
Investment for IT systems	[22]
Security	[29]
Inconsistency in data collection	[29]
Transferability	[29]
Fragmented structure of construction industry	[30]
Personal and organizational qualification	[30]
Requirement for analyst	[31]
Components of BIG Data Analytics (Such as iterative process in MR)	[32]

Table 1. Advantages and disadvantages of integration of BIM and BIG Data Analytics

V. CONCLUSION

The study focused on clarifying the interrelationship between BIM and big data analytics. Therefore, both methods and their relationships were investigated from the literature. The literature review showed that there are few studies that combine BIM and big data analytics. That has also arisen from the infancy of big data analytics in the literature. According to the studies, big data analytics and BIM look promising in terms of both increasing BIM utilization and solving chronic issues in construction projects.

Also, advantages of integration of BIM and big data analytics were discussed in the study. The literature review showed that big data analytics helps to store and mine BIM data and dynamic data that are used for facility management. Also, this integration will increase the effectiveness of BIM in the facility management process, because real-time decision-making is possible with big data analytics. Also, the studies showed that point cloud technology is beneficial in creating BIM models.

Disadvantages of integration depend on unsolved problems of BIM and big data analytics separately. For example, if MR is used in an iterative process, it needs to read and write data again and again, causing data to increase on the hardware. Also, while level one and level two BIM utilizations are very new for the construction industry, dissemination of big data that requires advanced BIM and programming skills seems unrealistic in the construction industry. Additionally, while construction companies have difficulty in finding engineers who have BIM skills, engineers need to have BIM and analytical ability to process results with the big data revolution. Also, because BIM investment is found to be costly by construction companies, it is hard to believe that construction companies will invest in big data analytics.

Further study will be made on energy management with BIM and big data analytics to increase energy efficiency in facilities.

REFERENCES

[1] A. H. Alavi, and A. H. Gandomi, “Big data in civil engineering. Automation in Construction”, vol. 79, pp. 1-2, 2017.
[2] Y. Jiao, Y. Wang, S. Zhang, Y. Li, B. Yang, and L. Yuan, “A cloud approach to unified lifecycle data management in architecture, engineering, construction and facilities management: Integrating BIMs and SNS”, Advanced Engineering Informatics, vol. 27(2), pp. 173-188, 2013.
[3] S. Azhar, “Building information modeling (BIM): Trends, benefits, risks, and challenges for the AEC industry”, Leadership and management in engineering, vol. 11(3), 241-252, 2011.
[4] J. Park, and H. Cai, “WBS-based dynamic multi-dimensional BIM database for total construction as-built documentation”, Automation in Construction, vol. 77, pp. 15-23, 2017.
[5] S. Sagiroglu, and D. Sinanc, “Big data: A review”, In Collaboration Technologies and Systems (CTS), 2013 International Conference, pp. 42-47, May 2013.
[6] J. Manyika, M. Chui, B. Brown, J. Bughin, R. Dobbs, C. Roxburgh, and A.H.Byers, “Big data: The next frontier for innovation, competition, and productivity”, McKinsey Global Institute, 2011.
[7] B. Succar, “The five components of BIM performance measurement”, In CIB World Congress, May 2010.
[8] D. Bryde, M. Broquetas, and J. M. Volm, “The project benefits of building information modelling (BIM)”, International journal of project management, vol. 31(7), pp. 971-980, 2013.
[9] Y. Jung, and M. Joo, “Building information modelling (BIM) framework for practical implementation”, Automation in construction, vol. 20(2), pp. 126-133, 2011.
[10] B. Succar, “Building information modelling framework: A research and delivery foundation for industry stakeholders”, Automation in construction, vol. 18(3), pp. 357-375, 2009.
[11] M. Bew, and M. Richards, “BIM Maturity Model”, Construct IT Autumn 2008 Members’ Meeting. Brighton, UK, 2008.
[12] K. Lahiri, “How Is Big Data Changing The Business Landscape?” Web site: https://www.forbes.com/sites/quora/2018/03/12/how-is-big-data-changing-the-business-landscape/#7f3c0c293db2
Access date: 27/12/2018
[13] S. Walker, “Big data: A revolution that will transform how we live, work, and think”, 2014.
[14] A. Gandomi, and M. Haider, “Beyond the hype: Big data concepts, methods, and analytics”, International Journal of Information Management, vol. 35(2), pp. 137-144, 2015.
[15] Y. Wang, L. Kung, and T. A. Byrd, “Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations”, Technological Forecasting and Social Change, pp. 126, 3-13, 2018.
[16] M. Rouse, “Hadoop Distributed File System (HDFS)”.Web site: http://searchdatamanagement.techtarget.com/definition/Hadoop-Distributed-File-System-HDFS
Access date: 27/12/2018
IBM, “What is MapReduce?”
Web site: https://www.ibm.com/analytics/hadoop/mapreduce
Access date: 27/12/2018
[17] P. Russom, “Big data analytics”, TDWI best practices report, fourth quarter, vol. 19(4), pp. 1-34, 2011.
[18] C. P. Chen, and C. Y. Zhang, “Data-intensive applications, challenges, techniques and technologies: A survey on Big Data”, Information Sciences, vol. 275, pp. 314-347, 2014.
[19] P. Read, E. Krygiel, and J. Vandezande, “Mastering Autodesk® Revit® Architecture 2013”, John Wiley & Sons, 2012.
[20] M. Bilal, et al., “Big Data in the construction industry: A review of present status, opportunities, and future trends”, Advanced engineering informatics, vol. 30(3), pp. 500-521, 2016.
[21] K. Gopalakrishnan, A. Agrawal, and A. Choudhary, “Big Data in Building Information Modeling Research: Survey and Exploratory Text Mining”, MOJ Civil Eng, vol. 3(6), 2017.
[22] T.W. Huang, “Application Case of Big Data Analysis on Dynamic BIMs with Parameters for Use Scenarios” [Master’s Thesis], National Taiwan University of Science and Technology, Taiwan, 2016.
[23] A. H. Oti, E. Kurul, F. Cheung, and J. H. M. Tah, “A framework for the utilization of Building Management System data in building information models for building design and operation”, Automation in Construction, vol. 72, pp. 195-210, 2016.
[24] J. K. W. Wong, and J. Zhou, “Enhancing environmental sustainability over building life cycles through green BIM: A review”, Automation in Construction, vol. 57, pp. 156-165, 2015.
[25] Y. Yuan, and Z. Jin, “Life cycle assessment of building energy in big-data era: theory and framework”, In Network and Information Systems for Computers (ICNISC), pp. 601-605, January 2015.
[26] M. Arslan, Z. Riaz, and S. Munawar, “Building Information Modeling (BIM) Enabled Facilities Management Using Hadoop Architecture”, 2017 Proceedings of PICMET '17: Technology Management for Interconnected World, Portland, USA, 2017.
[27] H. M. Chen, K. C. Chang, and T. H. Lin, “A cloud-based system framework for performing online viewing, storage, and analysis on big data of massive BIMs”, Automation in Construction, vol. 71, pp. 34-48, 2016.
[28] Ciklum. (2017). LIMITATIONS OF BIG DATA ANALYTICS. Web Site: https://www.ciklum.com/blog/limitations-of-big-data-analytics/
Access date: 27/12/2018
[29] D. Migilinskas, V. Popov, V. Juocevicius, and L. Ustinovichius, “The benefits, obstacles and problems of practical BIM implementation”, Procedia Engineering, vol. 57, pp. 767-774, 2013.
[30] A. Labrinidis, and H. V. Jagadish, “Challenges and opportunities with big data”, Proceedings of the VLDB Endowment, vol. 5(12), pp. 2032-2033, 2012.
[31] D. Singh, and C. K. Reddy, “A survey on platforms for big data analytics”, Journal of Big Data, vol. 2(1), pp. 8, 2015.

★★★