CRITICAL ANALYSIS FOR BIG DATA STUDIES IN CONSTRUCTION: SIGNIFICANT GAPS IN KNOWLEDGE

Upeksha Hansini Madanayake and Charles Egbu

School of the Built Environment and Architecture,

London South Bank University, London, UK

Abstract

Purpose – The purpose of this paper is to identify the gaps and potential future research avenues in the big data research specifically in the construction industry.

Design/methodology/approach – The paper adopts systematic literature review (SLR) approach to observe and understand trends and extant patterns/themes in the big data analytics (BDA) research area particularly in construction-specific literature.

Findings – A significant rise in construction big data research is identified with an increasing trend in number of yearly articles. The main themes discussed were big data as a concept, big data analytical methods/techniques, big data opportunities – challenges and big data application. The paper emphasises "the implication of big data in to overall sustainability" as a gap that needs to be addressed. These implications are categorised as social, economic and environmental aspects.

Research limitations/implications – The SLR is carried out for construction technology and management research for the time period of 2007–2017 in Scopus and emerald databases only.

Practical implications – The paper enables practitioners to explore the key themes discussed around big data research as well as the practical applicability of big data techniques. The advances in existing big data research inform practitioners the current social, economic and environmental implications of big data which would ultimately help them to incorporate into their strategies to pursue competitive advantage. Identification of knowledge gaps helps keep the academic research move forward for a continuously evolving body of knowledge. The suggested new research avenues will inform future researchers for potential trending and untouched areas for research.

Social implications – Identification of knowledge gaps helps keep the academic research move forward for continuous improvement while learning. The continuously evolving body of knowledge is an asset to the society in terms of revealing the truth about emerging technologies.

Originality/value – There is currently no comprehensive review that addresses social, economic and environmental implications of big data in construction literature. Through this paper, these gaps are identified and filled in an understandable way. This paper establishes these gaps as key issues to consider for the continuous future improvement of big data research in the context of the construction industry.

Keyword- Construction industry, Built environment, Systematic review, Data analysis, Big data analytics, Knowledge gaps

Paper type- Literature review

Introduction

In essence, “big data” is an artefact generated as a collective intelligence of human individuals shared mainly through the technological environment (Eadie et al., 2013). This environment devices a common platform, where virtually anything and everything can be documented, measured and captured digitally where the digital capture is then transformed into data (Sivarajah et al., 2017). Mayer-Schonberger and Cukier (2012) referred to this process as “datafication”. As the world has now inundated with data, big data analytics (BDA) is increasingly becoming a trending practice and has received substantial attraction from both academics and practitioners regardless of the sector. This massive growth in data generation brings significant opportunities for data scientists to capture meaningful insights and knowledge. Arguably, the accessibility of data and then its management can improve the status quo of many sectors by strengthening existing statistical analytical techniques (Bilal, Oyedele, Akinade et al., 2016; Bilal, Oyedele, Qadir et al., 2016). Notwithstanding the current situation, it is apparent that this trend is going to be improving in the future. In compliance with the concept of “datafication” and the technological advancements, it is conspicuous that the future will majorly rely on the data which is being generated and shared through machines, as machines communicate with each other over data networks- doing so will result less human involvement (van Dijck, 2014).

Having said the state-of-art in big data generally, the construction industry is not an exception to the pervasive digital revolution. Undoubtedly, the industry is dealing with significant amount of data arising from diverse disciplines (i.e. Building Information Modelling (BIM) data) throughout the life cycle of a facility and if they were better harnessed (by discovering the latent patterns, trends and correlations buried inside) it could help deriving useful social, economic and environmental insights that would support datadriven decision-making for competitive advantage. These interests have made construction organisations for the adoption of BDA, with the intention of generating valuable strategic business insights for enhanced decision making and thereby achieve organisational competitive advantage as there is a massive value creation potential from analysis of big data in construction sector (Cook, 2015). While big data appear to be one of the innovative trends for construction industry, it is still developing at a slow-moving pace as it is just starting to see the transformative effects of big data (Cook, 2015). To that end, it is important that the academia informs some of the influential implications of big data in construction for the better understanding of its implementers and/or adopters. This paper therefore investigates the previous big data research in construction literature to find out the gaps that have not being addressed with regards to social, economic and environmental implications as the effort will then could help both industry practitioners as well as academics to carry forward for the continuous improvement of the industry. There is currently no comprehensive survey of the literature, targeting the application of big data in the context of the construction industry tied-up for social, economic and environmental implications. This paper fills the void and presents a wide-ranging interdisciplinary study of fields such as machine learning, data warehousing, data mining, etc., and their applications.

Literature review – big data studies in construction

The built environment and project management research over the past 50 years has focused on emerging research and development topics. "Big Data" is one such concern where research interests include the limitations of utilising big data; challenges associated with its visualisation; the role of improved forecasting in BDA – such as machine learning; access and ownership of data; and the opportunities for experimenting big data techniques for smart city and infrastructure components.

The construction industry generates significant amount of data that can be quickly voluminous from diverse disciplines throughout the life cycle of a facility. BIM is a perfect example for such large data gathering. BIM captures multi-dimensional CAD information systematically to support multidisciplinary collaboration and integration among stakeholders (Eadie et al., 2013).With the emergence of embedded devices and sensors, constructions have even started to produce massive lump of data during the operations and maintenance stage. These are also converted into "big; BIMdata". This situation has lead construction industry to enter the big data era (Bilal, Oyedele, Akinade et al., 2016; Bilal, Oyedele, Qadir et al., 2016).

Optimisation and automation of work processes with more collaboration technologies (digitalisation) were the key concerns during mid-2007. During the last immediate few years, research has begun to discuss more of a lifecycle perspective on costs and benefits of big data based decisions made during design and construction (Levitt and M.ASCE, 2007). Nowadays, considerations have been more focused on many different perspectives of big data as a way of boosting productivity and gain competitive advantages over business rivals (Oyedele, 2016; McGuire et al., 2012; Gandomi and Haider, 2015; Marr, 2017). However, regardless of the approach or the ontological perspective used, many of the articles claim that Big data is a powerful tool that can be considered as a source when "properly managed, processed and analysed, have a powerful potential to generate new knowledge thus proposing innovative and actionable insights for businesses" (Jukić et al., 2015, p. 4). In addition to these extant insights on the current field, it is important to identify the significant gaps and potential future research directions – which are addressed in this study.

Methodology

The methodology is two-fold. The first part is to identify the key themes discussed in the previous big data research, while the second part identifies the significant gaps in existing body of knowledge. First, a systematic literature review (SLR) is carried out to observe and understand past trends and extant patterns/themes in the BDA research area particularly in construction, that lead to identify potential gaps in research. It is important to mention that this review excludes social science research that investigates human behaviour. To achieve the said goal, a comprehensive "keywords" search is been conducted in specialised databases, journals and few conference proceedings. Prospective articles are also used to capture more relevant papers through bibliography. This significantly helped enhance rigour of database as well as to increase the number of papers investigated.

The identified studies are analysed according to their contribution, summarising the key themes discussed, extant knowledge dimensions, thereby identifying limitations, implications and potential further research avenues to fill the knowledge gaps and support the academic community in exploring research themes/patterns.

Thus, in the process of tracing big data studies, a profiling method is employed to analyse peer-reviewed articles published between 2007–2017, extracted from the Scopus and emerald databases. The analysis presented in this paper has identified relevant BD research studies that have contributed both conceptually and empirically to the expansion and accrual of intellectual wealth to the BDA in technology and organisational resource management discipline specifically in construction.

The reason behind the selection of Scopus lies at its comprehensiveness for academic journal articles consisting of nearly 22,000 titles from over 5,000 publishers, of which 20,000 are peer-reviewed journals in the scientific, technical, medical and social sciences (including arts and humanities) (Elsevier, 2017). Besides, it is frequently updating, unbiased and a reliable source that covers the research spectrum with more than 50 per cent of key research publishers (Cheng et al., 2012). Emerald data base is selected for the consistency of data to be derived from AEC sectors.

Research scope

Although there is a growing demand, big data research for construction is still in the nascent stage (Qadir et al., 2016). Therefore, comparatively a paucity of research studies is expected. It is appearing that BD and BDA as a research discipline are still evolving and not yet established. Thus, a comprehensible understanding of the phenomenon, its definitions, classifications and applications are yet to be fully established.

Although there is a notable body of research that provides understanding on different aspects of BD and BDA area, there seems to be a lack of comprehensive and methodical approaches to understand the phenomenon of BD. Further, there is a lack of studies that empirically reports about the practical applications of big data techniques-more precisely the types of BDA methods used in particular construction case studies (projects or organisations). This study not only aims to grasp these gaps, but also to capture the conclusions offered by each review article to analyse, synthesise and present a state-of-the art structured analysis of the normative literature on big data and BDA to support future research avenues.

Data collection process

The SLR presented in this study aims to evaluate the existing research published on BD and BDA in relation to the construction sector. An established profiling approach is used to investigate and analyse different themes discussed. Following search strings were used to identify the relevant articles in the sections of "title, abstract, keywords" of articles belonging to Scopus and Emerald databases. The search is further filtered to subject areas related to construction sector journals only. The main keywords used in the search were "big data" AND "construction". Relevant literature was majorly found in the journal for automation in construction and Journal of Information Technology in Construction. Since those journals are relevant data sources, even though the papers do not use the phrase "big data", it was able to find relevant articles through the "suggestions" facility. Searches were also made with "construction engineering projects" as search keywords, but with limited additional search results. The latter search resulted with more technical data. Few exclusion keywords used to avoid the search being biased to pure engineering/technical papers.

Using the aforementioned search strings, the search was limited to published "journal articles" filtered for the two subject areas in Scopus: engineering and energy, for the time period of 2007–2017. For emerald, the subject areas filtered to Information and communications technology, Information and knowledge management, Management science/operations research and Information systems. In the search results, there were several papers that are not available or do not specifically deal with construction industry. Such papers were excluded by the excluding keywords. Some papers were found within the areas of ICT and software development and data/information management, but are not directly related to the construction industry. Papers only dealt with construction industry/projects were selected among them. The search initially offered with 113 and 93 journal articles to review from Scopus and Emerald respectively. The selected material was subsequently examined in more detail to identify the relevance of articles. Only papers that give a clear contribution or are of clear relevance to big data in the construction industry were selected. However, after cleansing out of the total number of

221 articles, 115 articles were discarded and finally 106 articles were remained and taken forward for further interrogation. The selected 106 articles were then profiled according to following sub-categories:

(1) type of publication (e.g. research or technical paper, literature review, viewpoint);

(2) type of research method(s) employed (e.g. case study, mixed method, analytical);

(3) yearly publications from 2007 until 2017; and

(4) main theme of the article (definition/challenges-opportunities/techniques, etc.).

For the fourth sub-category above, this paper focuses on the themes discussed in selected papers generally and thereby extracts the gaps that have not being addressed. However, the authors suggest a further investigation incorporating few different dimensions such as time, cost quality. An example for time dimension can be; application of big data techniques in stages of a typical construction project. The themes were further sub-categorised as definition/concept, challenges-opportunities and application of big data techniques into construction. The selected papers were carefully observed to identify the key themes discussed and counted into one sub-category among the three (as mentioned above). To all the above sub and sub-sub categories, frequency of citations was calculated. This profiling helps developing and understanding the state-of-the-art of big data research as well as the significant gaps both in theory and practice. If an article speculates more than one topic, it was classified into the category perceived as predominant.

Identification of knowledge gaps is the second part of the analysis, which was again carried out systematically. The future directions, suggestions, limitations, uncovered areas were carefully read and scrutinised in each selected paper and a list of gaps was developed. The list (with short running titles) initially resulted with 26 key gap areas. These were further brought down into seven using simple manual factor analysis. The final list of gaps includes; sustainability, mass-scale data use, contribution to supply chain, skills/knowledge dimensions, qualitative/quantitative data difference, other innovative technology synergies with big data, interoperability issues and whole life cycle perspective; which are detailed described in the subsequent paragraphs.

Literature review analysis

Data analysis

Majority of the papers reviewed were research papers (40 per cent), technical papers (22 per cent) and case studies (20 per cent) predominantly based on case/study research methods (35 per cent). A significant rise in construction big data research is also identified with an increasing trend in number of yearly articles. Figure 1 shows how research interests in big data has risen almost exponentially from 2007 to 2017 implying that BDA is an extremely emerging topic which reports to be in its pinnacle at this time.

According to Figure 2, the main theme discussed in the literature was big data opportunities and challenges/drivers and barriers for adoption (42 per cent) that vary from project domain to firm level as well as to industry and national level. The opportunities as identified by SLR in this study inter alia: value creation, generation of business intelligence for informed business decisions, visualisation of patterns trends-correlations, process optimisation, enhancing the flexibility of supply chain and resource allocation, assets/resources management, productivity growth, competitive advantage, etc.

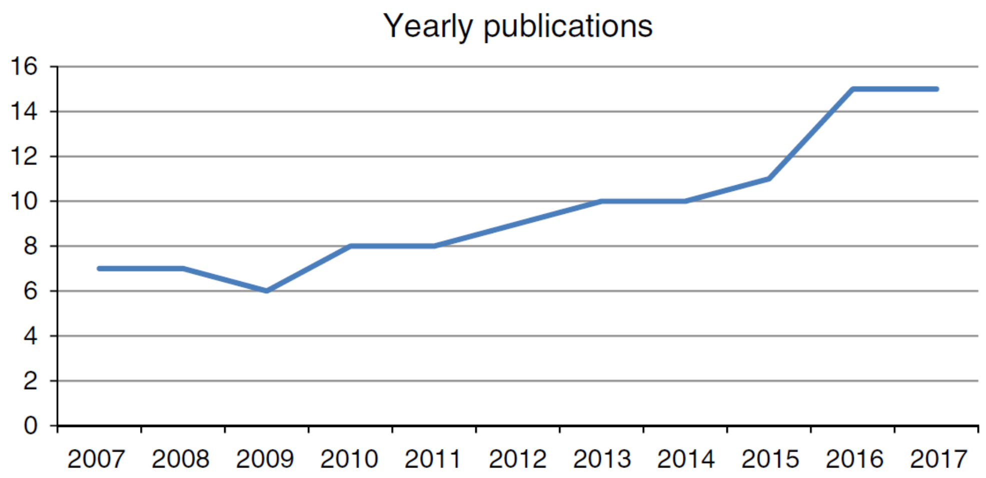

Figure 1- Number of articles

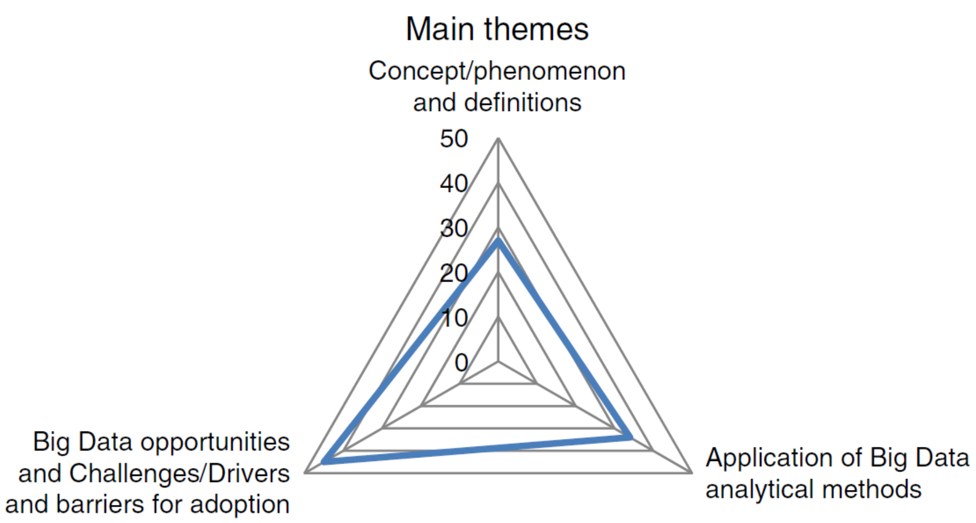

Figure 2- Main themes

The challenges include, data privacy data security, infrastructure and maintenance costs, lack of skills and training, hardware and software complexities for data integration and data synchronising.

Application of big data analytical methods (31 per cent) was secondly on the discourse while big data as a concept/phenomenon (21 per cent) is the least discussed theme.

Data synthesis

Big data as a concept. The data concept of being “big” is difficult to define as what appears to be big today may appear to be smaller in the near future (MIT, 2013). Further, no evidence proves that massive data sets are always complex or small data sets are always simple (Sivarajah et al., 2017). Therefore, the complexity of the data is not confined to data size, since data sets will increase in the future. Supporting this argument,McGuire et al. (2012, p. 559) refer big data to data sets whose size is beyond the ability of typical database software tools to capture, store, manage and analyse. A similar definition was provided by Amir Gandomi and Haider (2015) with an explanation of three ‘v’s (volume, variety and velocity). Bilal, Oyedele, Akinade et al. (2016); Bilal, Oyedele, Qadir et al. (2016) also bring out the 3’V’s concept as clear evident in construction data. Because, construction data is large, heterogeneous and dynamic (Aouad et al., 1999), they become voluminous due to large volumes of data gathered at a fast moving speed. These data as suggested by Bilal, Oyedele, Akinade et al. (2016); Bilal, Oyedele, Qadir et al. (2016) include: design data, schedules, enterprise resource planning (ERP) systems, financial data, etc. The latter authors have studied the diversification of these data in such a deep level that they can be categorised into single format (e.g. DWG, DXF, DGN, RVT, ifcXML, XML, ifcOWL, OWL, DOC/XLS/PPT, RM/MPG and JPEG). This dynamic nature of construction data has allowed systems to stream through sensors, RFIDs and BMS (building management sensors). This rich intellectual performance of sensors has been successfully tested by Akhavian and Behzadan (2015) with the use of smartphone sensors and machine learning classifiers to improve the quality and reliability of project decision-making and control. Thus, Bilal, Oyedele, Akinade et al. (2016); Bilal, Oyedele, Qadir et al. (2016) purport that utilisation of these voluminous data in an optimum manner would indisputably be the next frontier of innovation in the construction industry. The authors further distinguish the concept of big data engineering (BDE) and BDA as two concepts that needs careful consideration.

Big data analytical methods/techniques. Amir Gandomi and Haider (2015) describe several popular BDA techniques specifically can be used in construction sector. These techniques include text analytics (data mining) like information extraction (IE), entity recognition (ER) and relation extraction (RE), text summarisation, question answering (QA), sentiment analysis – data mining (Fan et al., 2015), content-based analytics, structure-based analytics (Chen et al., 2016 for cloud-based system framework for structured BIM data); audio analytics (ref) and visual-image/video analytics (Han and Golparvar-Fard, 2017); social media analytics (Tang et al., 2017), GIS analytics (Buffat et al., 2017) and predictive analytics (Li, 2017; Fan et al., 2017).

Bilal, Oyedele, Akinade et al. (2016); Bilal, Oyedele, Qadir et al. (2016) suggest possible BDE and BDA methods that can be successfully applied into construction practices, seeing BDE as an infrastructure provider to support BDA. The authors further introduce big data processing techniques (Map Reduce (MR), Directed Acyclic Graphs (DAG)) and storage techniques (distributed file systems, NoSQL databases) for BDE. Map Reduce process has been widely used in construction as a mean of laser scanning/point scanning for high definition surveying (HDS) where important information can be extracted from LiDAR point clouds such as the location, orientation and size of objects and possible damage. Aljumaily et al. (2016) propose a big data approach-mapreduce process to automatically identify and extract buildings from a digital surface model created from aerial laser scanning data. Han and Golparvar-Fard (2017) use a similar point cloud system to analyse the building performance. On the other hand, for BDA techniques such as statistics: data mining, Machine learning techniques, regression, classification and clustering are reported. The construction industry has employed some of these statistical methods in a variety of application areas, such as identifying causes of construction delays (Chau et al., 2003) learning from post-project reviews (PPRs) (Carrillo et al., 2011), decision support for construction litigation (Jordan and Mitchell, 2015; Mahfouz, 2009), detecting structural damages of buildings (Jiang and Mahadevan, 2008), identifying actions of workers and heavy machinery (Gong et al., 2011; Huang and Beck, 2013).

Chau et al. (2003) in their study of identifying critical factors for construction delays has employed data mining techniques to capture ML algorithms to produce knowledge discovery dataset (KDD). Another study conducted by Buchheit et al. (2000) also showcased a KDD process for a project related to construction of infrastructure. KDD process is reinvented by Soibelman and Kim (2002) to illustrate it is applicability to construction industry in identifying construction disputes such as; delays, cost overrunand quality failures. While Carrillo et al. (2011) employ data mining techniques to use past projects as learning material for risk free future projects, Liao and Perng (2008) attempted to employ association rule mining technique to investigate the prevention extent of workplace as well as occupational health safety threats. A similar study has been conducted by Cheng et al. (2012) using data mining to investigate the occupational injuries in construction sites. Interestingly, the system was competent enough to reveal the most impactful causes as; falls and collapses.

Data warehousing (DW) is another technique employed by many construction-related studies. Chau et al. (2003) and Kimball and Ross (2011) used DW to evaluate construction productivity data by an SQL multi-layer analysis. SQL has been reported for its wide usage specifically in construction for is facilitation for querying partial BIM models query languages such as express query language (EQL) and Building Information Modelling Query Language

(BIMQL) (Bilal, Oyedele, Akinade et al., 2016; Bilal, Oyedele, Qadir et al., 2016). Such studies include Kimball and Ross (2011) and Koonce and Judd (2001).

Machine learning is a type of predictive statistics that is widely applied in construction data predictions. Arditi and Pulket (2005) a closer version of artificial intelligence (AI) allowing a programme to learn from data about specific task automatically and predict the possible future outcomes. Machine learning rule-based learning is an industry-wide application where many researchers found beneficial with a variety of applications such as artificial neural networks methods, case-based reasoning techniques and hybrid methodologies (Ahn and Kim, 2009; Arditi and Pulket, 2005, 2010; Chau, 2005, 2006, 2007; Chen and Hsu, 2007; Cheng et al., 2009; Choi et al., 2014; Du et al., 2010; Pulket and Arditi, 2009; Sanyal et al., 2014). A similar study carried out by Sacks et al. (2018) introduced a time and cost saving automated method of checking building designs for code compliance using machine learning technique. This method is considered to be highly beneficial at the phase of pre-processing and preparing BIM models for checking. Predictive analytics has been widely used in construction activity predictions through simulations. Li (2017) in his investigation of stadium construction used a constrained parametric index analysis model of the progress analysis to monitor, predict and control the resource (i.e manpower, scheduling) progress.

Regression models have also been in the use for many years in construction research. The use of regression model often comes with a machine learning technique. Many of the studies have used regression models to predict construction tender prices, cost estimates and material price fluctuations (Cheng et al., 2009; Fallis, 2013; Lau et al., 2010; Narbaev and DeMarco, 2014). Using a similar technique Shrestha et al. (2017) used dynamic items basket (DIB) method, which is based on regression modelling for large amount highway bid data to establish a framework for improved calculation process of Highway Construction Cost Index (HCCI).

Big data opportunities and challenges. There is a considerable body of research on the opportunities and challenges offered by BDA (Bilal, Oyedele, Akinade et al., 2016; Bilal, Oyedele, Qadir et al., 2016; Wamba et al., 2016; Wang et al., 2016; Devlin, 2016).

The actual challenge of tackling with big data as suggested by Mishra et al. (2017, p. 28) “was to deal with diversified data types (variety), timely response requirements (velocity) and uncertainties in the data (veracity)”. Mishra and Sharma (2015) advocate handling both semi-structured and unstructured data is challenging especially when they are not received in a timely manner. The reason as latter authors mention is mainly be due to the lack of insufficient sources needed to gather, store and analyse big data but within a particular time frame. The authors further mention the reliability of data is also a big issue where additionally cleansing methods required to be applied in order to mitigate the uncertainty which may consume much more time and resources. Manyika et al. (2011) state identifying the exact area of applications as one of the biggest challenges in big data. Besides, Data security, privacy and protection, quality of construction industry data sets, cost implications for big data in construction industry, internet connectivity for big data applications, exploiting big data to its full potential are also discussed in the existing literature (Tene and Polonetsky, 2013; Bilal, Oyedele, Akinade et al., 2016; Bilal, Oyedele, Qadir et al., 2016).

The greatest opportunities as mentioned by many articles are, Resource and waste optimisation (Bilal et al., 2015; Bilal, Oyedele, Akinade et al., 2016; Bilal, Oyedele, Qadir et al., 2016; Oyedele, 2016; Oyedele et al., 2013; Lu et al., 2015, 2016), Generative designs and clash detection and resolution (Nima, 2014; Wang and Leite, 2013), performance prediction (Abaza et al., 2004; Kobayashi et al., 2010), visual analytics (Goodwin and Dykes, 2012; Löfström and Palm, 2008), social networking services/analytics (Demirkesen and Ozorhon, 2017; Wolf et al., 2009), personalised services (Liu et al., 2012; Singh et al., 2010), facility management (Isikdag et al., 2013; Liu et al., 2012; Rueppel and Stuebbe, 2008; Taneja et al., 2012), Energy management and analytics (Hong et al., 2012; Linda et al., 2012; Sanyal and New, 2013), big data integration with BIM (Volk et al., 2014), big data integration with IOT and cloud computing (Elghamrawy and Boukamp, 2010), big data for augmented reality (AR) (Williams et al., 2015), use of big data for smart buildings and smart city/urban infrastructure (Khan and Hornbæk, 2011; Liu et al., 2014). Further, Mishra et al. (2017) and Janssen et al. (2017) purport that exploitation of BDA can lead to gain competitive advantage at any level.

Big data application. There are number of studies address the applicability of big data techniques into construction specifically for performance and process optimisation (Eriksson et al., 2017; Becerik-Gerber et al., 2012; Bilal et al., 2015; Bilal, Oyedele, Akinade et al., 2016; Bilal, Oyedele, Qadir et al., 2016; Demchenko et al., 2014; Koseleva and Ropaite, 2017; Lu et al., 2015; Zhang et al., 2015; Qadir et al., 2016; Hao et al., 2015; Alaka et al., 2015; Rathorea et al., 2016; An, 2014).

Motawa (2017) deployed BDA into BIM system to capture buildings operation knowledge, particularly for building maintenance and refurbishment. The proposed big data technique in this study was cloud-based spoken dialogue system and case-based reasoning BIM system. Thus, the study tried to answer problems specific to building maintenance.

In a similar study conducted to investigate the leading Hong Kong construction firms' efficiency by Chiang et al. (2013) used variation in weights for estimation and enhances the adequacy for individual contractor's efficiency scores. A study conducted by Mansouri and Akhavian (2018) discussed the importance of early engagement of stakeholders as well as early incorporation of big data principles into construction projects, to use big data in its full potential. The authors explain how it improves the forecasting/planning process and manage both explicit and tacit knowledge provisions of the project.

Discussion

Significant knowledge gaps- social, economic and environmental implications of big data in construction. Although there are plenty of research in the application of big data in construction. There is a lack in focus on the implications of BDA on to the social, economic and environmental dimensions of sustainability. Subsequent paragraphs critically review some of the literature and discuss these implications.

Social implications

Implication of big data on the society as cited by many of the relevant studies are tied-up to people's quality of living and their behaviours/attitude for the post-occupancy of buildings. Few see big data as a convergence of people, place and technology as it helps to improve understanding on common data environments (Cook, 2015) as well as to control and/or maximise the social interactions through social media (Tang et al., 2017). In opposed to this positive impact, the negative impact on human relationships caused by humans' constant connection with digital data which ends up as an addiction is an under-researched area in construction. As a finding of Tang et al.' s (2017) study, it has been presented the mostly tweeted words by construction workers, but the study does not address how this finding benefit to improve the health and well-being of them. Thus, it is advisable to undertake further research on how social media BDA can be used to prevent threats such as safety, injury or mental illness caused by work-related stress. On adifferent dimension, Big data techniques have proven to monitor and analyse indoor quality of space such as air quality, noise, light, etc. Such empowers occupiers on useful information on health and well-being attributes of spaces and it is sign of improved sustainability to the society. In light of same direction, Zhu and Ge (2014) investigate the social impact of green buildings focusing on user satisfaction of green buildings using a big data post-occupancy evaluation. However, many of these studies related to green building performance evaluation lacks occupiers' subjective as well as objective opinion evaluation. Further, big data with its asset management have massive potential for positive social implications as it allows for meaningful business decisions based upon the life cycle of building not limited to the capital cost alone (Cook, 2015). However, the studies lack the robustness of lifecycle studies as to how these data can be reused for future benefit. The ways to leverage any drop of data, is to use them over and over again, before it becomes stale and that indeed saves massive amount of time as well. Akhavian and Behzadan (2015) profess through mobile sensor data quality and reliability of project decision-making could be improved, Despite the advantages big data offers, some of the social implications like disruption to social interactions, quality assurance, need for mutually agreed standards/guidelines, intellectual property, privacy and security issues are still not entirely addressed. Finally, need for skills, knowledge and training is another important social implication emerged from the review. Currently there is a huge demand for data management capabilities among professionals, help embed use of new technology across the built environment sector and generate new avenues for value creation from the vast amounts of data available in construction industry as well as all other sectors that is linked into it.

Economic implications

Cook (2015) in his paper based upon the 2015 RICS/SPR Cutting Edge conference purports that there is major change impacting from the rise of digital technology to creating value from big data. One of the main areas of impact as he suggested would be changes to job roles and business structures, requiring continuous learning and greater flexibility and adaptability for economic stability of the construction industry as its greater potential to many add value and generate more income. Again, Big data with its potential for future prediction models thrives longer term view on total life-cycle costs rather than short term financial implications, (Cook, 2015). However, the studies that address the economic value big data do not exactly evaluate and state as to how big data could be exploited to thrive competitiveness. The effectiveness of big data cannot be measured just by accumulating large volumes of data; it is more of the use cases or industrial problems that dictate the usefulness of these technologies (Bilal, Oyedele, Akinade et al., 2016; Bilal, Oyedele, Qadir et al., 2016). Thus, an evidently proven case study analysis would help practitioners to see how big data could maximise competitive edge to stay on top of the market. Progress prediction model produced by Li (2017) and Shrestha et al. (2017) is good examples of how big data improve estimation and bidding process. Again, how this index can be used to improve bidding success is an undiscussed area. Further, Cook (2015) emphasises, as a result of big data hype, businesses are starting to be keener on grouping and sharing economies with short term needs such as entrepreneurships. On what complex and dynamic ways big data facilitate these sharing economies is still unexplored. BDA supported by mobile technology with its geospatial capability adds a valuable dimension (Akhavian and Behzadan, 2015) that is already of great interest for real estate and building developers (Buffat et al., 2017; Shrestha et al., 2017). By understanding data and big analytical techniques, these businesses can better understand the current and future client demand and better target their new potential clients – which is a massive implication for sectoral economy as a whole. Nevertheless, client behaviour analysis is very common in sectors like retail but not much in construction. The cost implications of big data, considered amongst the low-profit-margin businesses (construction industry) is another under-researched area. Chau et al. (2003) mention that adoption of big data incorporates costly endeavours such as data centre purchases and software licensing, such costly add-ons to projects are more likely to be opposed difficult to be defended on low-profit margins. Hence it would be more worthwhile conducting more “quantified” research that presents the business case on the extent of financial return for big data investment which would help strategic decision makers for their investment decisions. More studies on cost-benefit analysis of using big data technologies in construction projects are required to this end. Since BDA is often involved with large data sets, it would be beneficial if these percentages of revenue making could be examined in larger scale (predominantly considering the entire construction industry). The current body of knowledge predominantly limits to micro levels studies like project (Han and Golparvar-Fard, 2017; Zhang et al., 2015; An, 2014) level.

It was noticeable that many of the studies used "Building Information Modelling-BIM" (Han and Golparvar-Fard, 2017; Motawa, 2017) and "Internet of things-IOT" (Akhavian and Behzadan, 2015) interchangeability to explain the application of big data as there is a closer relationship between the technologies in terms data sharing. However, considering the dynamic and competitive environment of today's world, it is imperative to embrace innovative technologies and their synergies (Demirkesen and Ozorhon, 2017). Thus, yet, no research has been undertaken exploring the synergies between them.

Environmental implications

Numerous studies have conducted addressing the issues with energy demand and presented new approaches for building stock modelling (Buffat et al. 2017; Fan et al., 2015, 2017; Moreno et al., 2016; Mathew et al., 2015; Yu et al., 2016; Sanyal et al., 2014) using geographical/spatial data sets such as building footprints and digital elevation models and building automation systems. The model helps controlling the building heat demand for various climate conditions and improves energy efficiency and conservation – which shows a positive environmental impact of BDA. This improves understanding of the impact of climate change in future years and allows designers ascertain heating and cooling loads in different parts of the world which ultimately saves massive amount of energy. Further, an accurate estimation of future climate conditions supported by an energy simulation has the advantage of decreasing the extensive supply and demand for energy without neglecting the extremes and variations of the possible climate changes (Nik, 2016; Nik et al., 2016). However, in most of the energy studies reviewed, selecting a suitable measure is an important step- typically a multi-criteria decision-making procedure (Nik, 2016), affected by several factors related to economy, availability, etc., which have not been considered in most of the work. It is also apparent that more research required to be produced as solutions to disasters which is a major effect of climate change. Disaster resilience and coordination in humanitarian operations and supply chains, is a shortage in existing body of literature. Another study conducted by Chen and Lu (2017) looked in to the use of big data sets to minimise demolition waste in Hong Kong which had numerous environmental implications including mitigating adverse impacts – i.e. land deterioration, resource depletion and various forms of pollution such as noise, dust, air and discharge of toxic waste. It is worthwhile noting that, many of the suggested waste reduction tools are yet to be validated with use cases. Although there are many research studies claiming how BDA encourage efficient process efficiency and optimisation (Motawa, 2017; Li, 2017) there is a lack of connection between how it actually makes agile and adaptable to dynamic business environments. There is a gap identified in the areas of how BDA contributes supply chain design by focusing on main characteristics of supply chain including agility, adaptability, alignment and integration.

Summary of findings – gaps in existing body of knowledge for future research

Social

- How social media bid data analytics can be used to prevent threats such as safety, injury or mental illness caused by work-related stress.
- The negative impact on human relationships caused by humans' constant connection with digital data which ends up as an addiction.
- How to leverage any drop of data in life cycle studies by reusing data.
- Green building post-occupancy evaluation should consider occupiers' both subjective and objective opinions.
- Issues related to Quality assurance, intellectual property, privacy and security.
- Need for mutually agreed standards/guidelines.
- Current and future needs for skills, knowledge and training for big data.

Economic

- Use cases for the exploitation of big data in its full potential for competitive advantage (at every level i.e. project, organisation, sector).
- How bidding progress prediction models can be used to improve bidding success rate.
- In what complex and dynamic ways big data facilitate sharing economies within organisations.
- Client behaviour analysis to predict market conditions.
- Business case on the extent of financial return for big data investment which would help strategic decision makers for their investment decisions.
- Cost-benefit analysis of using big data technologies in construction projects
- Synergies between BDA, BIM and IOT.

Environmental

- Predictive analytics for Disaster risk reduction and climate change.
- Multi-criteria decision-making procedure for energy studies.
- Validation of waste reduction tools with use-cases.
- Supply chain agility and adaptability.

Conclusion

The main purpose of this paper is to identify the gaps and potential future research avenues in the big data research in the context of construction industry. The authors observe a greater demand as well as prospects for increased use of big data methods and applications within construction and highlight that as the need for this research. Gaps in knowledge in current research efforts are identified through an SLR. After in-depth discussion on gaps-in-knowledge, insights in different research areas are discussed in line with the social, environmental and economic implications to provide a comprehensive big picture toward which big data- related areas need further attention for the competitive advantage of the construction industry. The implications reflect both positive and negative impacts while the issues around big data, i.e. data privacy is still an on-going research topic that needs further investigation. Some of the gaps identified inter alia: impact of big data on health and wellbeing of construction workers, need for mutually agreed standards/guidelines, contribution to supply chain agility, impact of big data in disaster risk reduction, skill/knowledge dimensions. It also interesting that many authors have purported that there is a potential to combine different types of innovative technologies with big data techniques to maximise the effectiveness potential, although they have not precisely outlined the "how" part. It is crucial that the real power of big data is properly discoursed (i.e. how the predictability helps generating valuable insights and thereby informed decisions) through academic channels. To that end, identification of contemporary gaps aids continuous improvement.

It is worthwhile mentioning that the SLR presented in this paper limits to peer-reviewed journal articles only. In fact, there has been increasing number of indexed conferences in big data, smart systems and digital information and communication technology-related areas. However, this has been identified as a limitation of this paper and therefore the conclusions may have impacts on the generalisability as well as the representation of the sample of papers reviewed. The paper acknowledges this as further Research Avenue for future investigation.

References

Abaza, K.A., Ashur, S.A. and Al-Khatib, I.A. (2004), "Integrated pavement management system with a markovian prediction model", Journal of Transportation Engineering, Vol. 130 No. 1, pp. 24-33.

Ahn, H. and Kim, K. (2009), "Bankruptcy prediction modeling with hybrid case-based reasoning and genetic algorithms approach", Applied Soft Computing, Vol. 9 No. 2, pp. 599-607.

Akhavian, R. and Behzadan, A.H. (2015), "Construction equipment activity recognition for simulation input modeling using mobile sensors and machine learning classifiers", Advanced Engineering Informatics, Vol. 29 No. 4, pp. 867-877.

Alaka, H., Oyedele, L., Bilal, M., Akinade, O., Owolabi, H. and Ajayi, S. (2015), "Bankruptcy prediction of construction businesses: towards a big data analytics approach", 2015 IEEE

First International Conference on Big Data Computing Service and Applications, San Francisco, CA, 30 March-3 April, pp. 347-352.

Aljumaily, H., Laefer, D.F. and Cuadra, D. (2016), “Big-data approach for three-dimensional building extraction from aerial laser scanning”, Journal of Computing in Civil Engineering, Vol. 30 No. 3, pp. 1-10.

An, Q. (2014), “The effective classification process analysis of the big data in construction project”, Applied Mechanics and Materials, Vol. 44 No. 50, pp. 1749-1751.

Aouad, G., Kagioglou, M., Cooper, R., Hinks, J. and Sexton, M. (1999), “Technology management of IT in construction: a driver or an enabler?”, Logistics Information Management, Vol. 12 No. 2, pp. 130-137.

Arditi, D. and Pulket, T. (2005), “Predicting the outcome of construction litigation using boosted decision trees”, Journal of Computing in Civil Engineering, Vol. 19 No. 4, pp. 387-393.

Arditi, D. and Pulket, T. (2010), “Predicting the outcome of construction litigation using an integrated artificial intelligence model”, Journal of Computing in Civil Engineering, Vol. 24 No. 1, pp. 73-80.

Becerik-Gerber, B., Jazizadeh, F., Li, N. and Calis, G. (2012), “Application areas and data requirements for BIM-enabled facilities management”, Journal of Construction Engineering and Management, Vol. 138 No. 3, pp. 431-442.

Bilal, M., Oyedele, L.O., Qadir, J., Munir, K., Akinade, O.O., Ajayi, S.O., Alaka, H.A. and Owolabi, H.A. (2015), “Analysis of critical features and evaluation of BIM software: towards a plug-in for construction waste minimization using big data”, International Journal of Sustainable Building Technology and Urban Development, Vol. 6 No. 4, pp. 211-228.

Bilal, M., Oyedele, L.O., Akinade, O.O., Ajayi, S.O., Alaka, H.A., Owolabi, H.A., Qadir, J., Pasha, M. and Bello, S.A. (2016), “Big data architecture for construction waste analytics (CWA): a conceptual framework”, Journal of Building Engineering, Vol. 6 No. 2, pp. 44-156.

Bilal,M., Oyedele, L.O., Qadir, J., Munir,K.,Ajayi, S.O.,Akinade,O.O., Owolabi, H.A.,Alaka,H.A. and Pasha, M. (2016), “Big data in the construction industry: a review of present status, opportunities, and future trends”, Advanced Engineering Informatics, Vol. 30 No. 3, pp. 500-521.

Buchheit, R., Garrett, J., Lee, S. and Brahme, R. (2000), “A knowledge discovery framework for civil infrastructure: a case study of the intelligent workplace”, Engineering with Computers, Vol. 16 No. 4, pp. 264-274.

Buffat, R., Froemelt, A., Heeren, N., Raubal, M. and Hellweg, S. (2017), "Big data GIS analysis for novel approaches in building stock modelling", Applied Energy, Vol. 208 No. 2017, pp. 277-290.

Carrillo, P., Harding, J. and Choudhary, A. (2011), "Knowledge discovery from post-project reviews", Construction Management and Economics, Vol. 29 No. 7, pp. 713-723.

Chau, K.W. (2005), "Predicting construction litigation outcome using particle swarm optimization", in Ali, M. and Esposito, F. (Eds), Innovations in Applied Artificial Intelligence, Lecture Notes in Computer Science, IEA/AIE, Springer, Heidelberg, Berlin, Vol. 3533.

Chau, K.W. (2006), "Prediction of construction litigation outcome – a case-based reasoning approach", in Ali, M. and Esposito, F. (Eds), Innovations in Applied Artificial Intelligence, Lecture Notes in Computer Science, IEA/AIE, Springer, Heidelberg, Berlin and Annecy, pp. 548-553.

Chau, K.W. (2007), "Application of a PSO-based neural network in analysis of outcomes of construction claims", Automation in Construction, Vol. 16 No. 5, pp. 642-646.

Chau, K.W., Cao, Y., Anson, M. and Zhang, J. (2003), "Application of data warehouse and decision support system in construction management", Automation in Construction, Vol. 12 No. 2, pp. 213-224.

Chen, H.M., Chang, K.C. and Lin, T.H. (2016), "A cloud-based system framework for performing online viewing, storage, and analysis on big data of massive BIMs", Automation in Construction, Vol. 71 No. 2016, pp. 34-48.

Chen, J.H. and Hsu, S.C. (2007), "Hybrid ANN-CBR model for disputed change orders in construction projects", Automation in Construction, Vol. 17 No. 1, pp. 56-64.

Chen, X. and Lu, W. (2017), "Identifying factors influencing demolition waste generation in Hong Kong", Journal of Cleaner Production, Vol. 141 No. 2017, pp. 799-811.

Cheng, C.W., Leu, S.S., Cheng, Y.M., Wu, T.C. and Lin, C.C. (2012), "Applying data mining techniques to explore factors contributing to occupational injuries in Taiwan's construction industry", Accident Analysis & Prevention, Vol. 48 No. 2012, pp. 214-222.

Cheng, M.Y., Tsai, H.C. and Chiu, Y.H. (2009), "Fuzzy case-based reasoning for coping with construction disputes", Expert Systems with Applications, Vol. 36 No. 2, pp. 4106-4113.

Chiang, At.H., Li, J., Choi, T.N.Y. and Man, K.F. (2013), "Evaluating construction contractors' efficiency in Hong Kong using data envelopment analysis assurance region model", Journal of Facilities Management, Vol. 11 No. 1, pp. 52-68.

Choi, S., Kim, D.Y., Han, S.H. and Kwak, Y.H. (2014), "Conceptual cost-prediction model for public road planning via rough set theory and case-based reasoning", Journal of Construction Engineering and Management, Vol. 140 No. 1, pp. 401-426.

Cook, D. (2015), "RICS futures: turning disruption from technology to opportunity", Journal of Property Investment and Finance, Vol. 33 No. 5, pp. 456-464.

Demchenko, Y., De Laat, C. and Membrey, P. (2014), "Defining architecture components of the big data ecosystem", 2014 International Conference on Collaboration Technologies and Systems, Minneapolis, MN, 19-23 May, pp. 104-112.

Demirkesen, S. and Ozorhon, B. (2017), "Impact of integration management on construction project management performance", International Journal of Project Management, Vol. 35 No. 8, pp. 1639-1654.

Devlin, B. (2016), "Cultivating success in big data analytics", Cutter, IT Journal, Vol. 29 No. 6, pp. 8-16.

Du, Y., Wen, W., Cao, F. and Ji, M. (2010), "A case-based reasoning approach for land use change prediction", Expert Systems with Applications, Vol. 37 No. 8, pp. 5745-5750.

Eadie, R., Browne, M., Odeyinka, H., McKeown, C. and McNiff, S. (2013), "BIM implementation throughout the UK construction project lifecycle: an analysis", Automation in Construction, Vol. 36, pp. 145-151.

Elghamrawy, T. and Boukamp, F. (2010), "Managing construction information using RFID-based semantic contexts", Automation in Construction, Vol. 19 No. 8, pp. 1056-1066.

Elsevier (2017), Scopus Content Overview, Elsevier, Amsterdam, available at: www.elsevier.com/ solutions/scopus/content (accessed 15 January 2018).

Eriksson, C., Cheng, I., Pitman, K., Dixon, T., Van De Wetering, J. and Sexton, M. (2017), "Smart cities, big data and the built environment: what's required?", available at: www.rics.org/Global/RICSSmart- Cities-Big-Data-REPORT-2017.pdf (accessed 12 February 2018).

Fallis, A. (2013), "Applying regression analysis to predict and classify construction cycle time", Journal of Chemical Information and Modeling, Vol. 53 No. 9, pp. 1689-1699.

Fan, C., Xiao, F. and Zhao, Y. (2017), "A short-term building cooling load prediction method using deep learning algorithms", Applied Energy, Vol. 195, pp. 222-233.

Fan, C., Xiao, F., Madsen, H. and Wang, D. (2015), "Temporal knowledge discovery in big BAS data for building energy management", Energy and Buildings, Vol. 109, pp. 75-89.

Gandomi, A. and Haider, M. (2015), “Beyond the hype: big data concepts, methods, and analytics”, International Journal of Information Management, Vol. 35 No. 2, pp. 137-144.

Gong, J., Caldas, C.H. and Gordon, C. (2011), “Learning and classifying actions of construction workers and equipment using bag-of-video-feature-words and Bayesian network models”, Advanced Engineering Informatics, Vol. 25 No. 4, pp. 771-782.

Goodwin, S. and Dykes, J. (2012), “Visualising variations in household energy consumption”, IEEE Conference on Visual Analytics Science and Technology 2012 – Proceedings, pp. 217-218.

Han, K.K. and Golparvar-Fard, M. (2017), “Potential of big visual data and building information modeling for construction performance analytics: an exploratory study”, Automation in Construction, Vol. 73, January, pp. 184-198.

Hao, J., Zhu, J. and Zhong, R. (2015), “The rise of big data on urban studies and planning practices in China: review and open research issues”, Journal of Urban Management, Vol. 4 No. 2, pp. 92-124.

Hong, I., Byun, J. and Park, S. (2012), “Cloud computing-based building energy management system with ZigBee sensor network”, Proceedings – 6th International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing, IMIS 2012, pp. 547-551.

Huang, Y. and Beck, J.L. (2013), “Novel Sparse Bayesian learning for structural health monitoring using incomplete modal data”, ASCE International Workshop on Computing in Civil Engineering, Los Angeles, CA.

Isikdag, U., Zlatanova, S. and Underwood, J. (2013), “A BIM-oriented model for supporting indoor navigation requirements”, Computers, Environment and Urban Systems, Vol. 41, September, pp. 112-123.

Janssen, M., van der Voort, H. and Wahyudi, A. (2017), “Factors influencing big data decision-making quality”, Journal of Business Research, Vol. 70, January, pp. 338-345.

Jiang, X. and Mahadevan, S. (2008), “Bayesian probabilistic inference for nonparametric damage detection of structures”, Journal of Engineering Mechanics, Vol. 134 No. 10, pp. 820-831.

Jordan, M.I. and Mitchell, T.M. (2015), “Machine learning: trends, perspectives, and prospects”, Science, Vol. 349 No. 6245, pp. 255-260.

Jukić, N., Sharma, A., Nestorov, S. and Jukić, B. (2015), “Augmenting data warehouses with big data”, Information Systems Management, Vol. 32 No. 3, pp. 200-209.

Khan, A. and Hornbæk, K. (2011), “Big data from the built environment”, Proceedings of the 2nd International Workshop on Research in the Large – LARGE 2011, pp. 29-34.

Kimball, R. and Ross, M. (2011), The Data Warehouse Toolkit: The Complete Guide to Dimensional Modelling, Nachdr, Wiley, New York, NY.

Kobayashi, K., Do, M. and Han, D. (2010), “Estimation of Markovian transition probabilities for pavement deterioration forecasting”, KSCE Journal of Civil Engineering, Vol. 14 No. 3, pp. 343-351.

Koonce, D.A. and Judd, R.P. (2001), “A visual modelling language for EXPRESS schema”, International Journal of Computer Integrated Manufacturing, Vol. 14 No. 5, pp. 457-472.

Koseleva, N. and Ropaite, G. (2017), “Big data in building energy efficiency: understanding of big data and main challenges”, Procedia Engineering, Vol. 172 No. 2017, pp. 544-549.

Lau, S.C., Lu, M. and Ariaratnam, S.T. (2010), “Applying radial basis function neural networks to estimate next-cycle production rates in tunnelling construction”, Tunnelling and Underground Space Technology, Vol. 25 No. 4, pp. 357-365.

Levitt, R.E., M.ASCE (2007), “CEM Research for the next 50 years: maximizing economic, environmental, and societal value of the built environment”, Journal of Construction Engineering and Management, Vol. 133 No. 9, pp. 619-628.

Li, H. (2017), “Estimation of stadium construction schedule based on big data analysis”, International Journal of Computers and Applications, Vol. 41 No. 4.

Liao, C.W. and Perng, Y.H. (2008), “Data mining for occupational injuries in the Taiwan construction industry”, Safety Science, Vol. 46 No. 7, pp. 1091-1102.

Linda, O., Wijayasekara, D., Manic, M. and Rieger, C. (2012), “Computational intelligence-based anomaly detection for building energy management systems”, Resilient Control Systems (ISRCS), 2012 5th International Symposium, pp. 77-82.

Liu, J., Yao, R., Wang, J. and Li, B. (2012), “Occupants’ behavioural adaptation in workplaces with noncentral heating and cooling systems”, Applied Thermal Engineering, Vol. 35 No. 1, pp. 40-54.

Liu, X., Iftikhar, N. and Xie, X. (2014), “Survey of real-time processing systems for big data”, Proceedings of the 18th International Database Engineering & Applications Symposium on - IDEAS 2014, pp. 356-361.

Löfström, E. and Palm, J. (2008), “Visualising household energy use in the interest of developing sustainable energy systems”, Housing Studies, Vol. 23 No. 6, pp. 935-940.

Lu, W., Chen, X., Ho, D.C.W. and Wang, H. (2016), “Analysis of the construction waste management performance in Hong Kong: the public and private sectors compared using big data”, Journal of Cleaner Production, Vol. 112, pp. 521-531.

Lu, W., Chen, X., Peng, Y. and Shen, L. (2015), “Benchmarking construction waste management performance using big data”, Resources Conservation and Recycling, Vol. 105, Part A, pp. 49-58.

McGuire, T., Chui, M. and Manyika, J. (2012), “Why big data is the new competitive advantage”, Ivey Business Journal, Vol. 76, July/August, pp. 1-4.

Mahfouz, T.S. (2009), Construction Legal Support for Differing Site Conditions (DSC) Through Statistical Modeling and Machine Learning (ML), Graduate Theses and Dissertation No. 10698, Lowa State University.

Mansouri, S. and Akhavian, R. (2018), “The status quo and future potentials of data analytics in aec/fm: a quantitative analysis of academic research and industry outlook”, Proceedings of Construction Research Congress, American Society of Civil Engineers, New Orleans, LA, 2-4 April, pp. 90-100.

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C. and Byres, A.H. (2011), Big Data: The Next Frontier for Innovation, Competition, and Productivity, McKinsey Global Institute, San Francisco, CA, June, p. 156.

Marr, B. (2017), Big Data Analytics- What it is and Why it Matters, SAS, Cary, NC, available at: www.sas.com/en_gb/insights/analytics/big-data-analytics.html

Mathew, P.A., Dunn, L.N., Sohn, M.D., Mercado, A., Custudio, C. and Walter, T. (2015), “Big-data for building energy performance: lessons from assembling a very large national database of building energy use”, Applied Energy, Vol. 140, 15 February, pp. 85-93.

Mayer-Schonberger, V. and Cukier, K. (2012), “Big data: a revolution that will transform how we live, work, and think”, CEUR Workshop Proceedings, Houghton Mifflin Harcourt, Boston, MA.

Mishra, D., Luo, Z., Jiang, S., Papadopoulos, T. and Dubey, R. (2017), “A bibliographic study on big data: concepts, trends and challenges”, Business Process Management Journal, Vol. 23 No. 3, pp. 555-573.

Mishra, R. and Sharma, R. (2015), “Big data: opportunities and challenges”, International Journal of Computer Science and Mobile Computing, Vol. 4 No. 6, pp. 27-35.

MIT (2013), “The big data conundrum: how to define it?”, MIT Technology Management Review, available at: www.technologyreview.com/s/519851/the-big-data-conundrum-how-to-define-it/ (accessed 3 March 2018).

Moreno, M.V., Dufour, L., Skarmeta, A.F., Jara, A.J., Genoud, D., Ladevie, B. and Bezian, J.-J. (2016), “Big data: the key to energy efficiency in smart buildings”, Soft Computing, Vol. 20 No. 5, pp. 749-1762.

Motawa, I. (2017), “Spoken dialogue BIM systems – an application of big data in construction”, Facilities, Vol. 35 No. 14, pp. 787-800.

Narbaev, T. and De Marco, A. (2014), “An Earned Schedule-based regression model to improve cost estimate at completion”, International Journal of Project Management, Vol. 32 No. 6, pp. 1007-1018.

Nik, V.M. (2016), “Making energy simulation easier for future climate – synthesizing typical and extreme weather data sets out of regional climate models (RCMs)”, Applied Energy, Vol. 177 No. 9, pp. 204-226.

Nik, V.M., Mata, E., Sasic Kalagasidis, A. and Scartezzini, J.-L. (2016), “Effective and robust energy retrofitting measures for future climatic conditions – reduced heating demand of Swedish households”, Energy and Buildings, Vol. 121 No. 2016, pp. 176-187.

Nima, K.O. (2014), “Comparison of machine learning techniques for developing performance prediction models”, Computing in Civil and Building Engineering, pp. 1222-1229.

Oyedele, L.O. (2016), “Big data and sustainability: the next step for circular economy”, Presentation Material, 14th International Annual Conference, World Association for Sustainable Development (WASD), Canary Wharf, London, 20-22 July.

Oyedele, L.O., Regan, M., von Meding, J., Ahmed, A., Ebohon, O.J. and Elnokaly, A. (2013), “Reducing waste to landfill in the UK: identifying impediments and critical solutions”, World Journal of Science, Technology and Sustainable Development, Vol. 10 No. 2, pp. 131-142.

Pulket, T. and Arditi, D. (2009), “Construction litigation prediction system using ant colony optimization”, Construction Management and Economics, Vol. 27 No. 3, pp. 241-251.

Qadir, J., Ahad, N., Mushtaq, E. and Bilal, M. (2016), “SDNs, clouds and big data : mutual opportunities”, Automation in Construction, Vol. 12 No. 3, pp. 1-6.

Rathorea, M.M., Ahmad, A.A., Paul, A.A. and Rho, S. (2016), “Urban planning and building smart cities based on the internet of things using big data analytics”, Computer Networks, Vol. 101 No. 4, pp. 63-80.

Rueppel, U. and Stuebbe, K.M. (2008), “BIM-based indoor-emergency-navigation-system for complex buildings”, Tsinghua Science and Technology, Vol. 13 No. 1, pp. 362-367.

Sacks, R., Bloch, T. and Katz, M. (2018), “Application of machine learning for automated code compliance”, 17th International Conference on Computing in Civil and Building Engineering Proceedings, Tampere, 5-7 June.

Sanyal, J. and New, J. (2013), “Simulation and big data challenges in tuning building energy models”, 2013 Workshop on Modeling and Simulation of Cyber-Physical Energy Systems (MSCPES), pp. 1-6.

Sanyal, J., New, J., Edwards, R.E. and Parker, L. (2014), “Calibrating building energy models using supercomputer trained machine learning agents”, Concurrency Computation Practice and Experience, Vol. 26 No. 13, pp. 2122-2133.

Shrestha, K.J., Jeong, H.D. and Gransberg, D.D. (2017), “Multidimensional Highway Construction Cost Indexes using dynamic item basket”, Journal of Construction Engineering and Management, Vol. 143 No. 8, pp. 1-11.

Singh, H., Muetze, A. and Eames, P.C. (2010), “Factors influencing the uptake of heat pump technology by the UK domestic sector”, Renewable Energy, Vol. 35 No. 4, pp. 873-878.

Sivarajah, U., Kamal, M.M., Irani, Z. and Weerakkody, V. (2017), “Critical analysis of big data challenges and analytical methods”, Journal of Business Research, Vol. 70 No. 1, pp. 263-286.

Soibelman, L. and Kim, H. (2002), “Data preparation process for construction knowledge generation through knowledge discovery in databases”, Journal of Computing in Civil Engineering, Vol. 16 No. 1, pp. 39-48.

Taneja, S., Akcamete, A., Akinci, B., Garrett, J.H., Soibelman, L. and East, E.W. (2012), “Analysis of three indoor localization technologies for supporting operations and maintenance field tasks”, Journal of Computing in Civil Engineering, Vol. 26 No. 6, pp. 708-719.

Tang, L., Zhang, Y., Dai, F., Yoon, Y., Song, Y. and Sharma, R.S. (2017), “Social media data analytics for the US construction industry: preliminary study on Twitter”, Journal of Management in Engineering, Vol. 33 No. 6, pp. 1-15.

Tene, O. and Polonetsky, J. (2013), “Big data for all: privacy and user control in the age of analytics”, Northwestern Journal of Technology and Intellectual Property, Vol. 11 No. 5, pp. 240-273.

Van-Dijck, J. (2014), “Datafication, dataism and dataveillance: big data between scientific paradigm and ideology”, Surveillance and Society, Vol. 12 No. 2, pp. 197-208.

Volk, R., Stengel, J. and Schultmann, F. (2014), “Building Information Modeling (BIM) for existing buildings- literature review and future needs”, Automation in Construction, Vol. 38, pp. 109-127.

Wamba, S.F., Gunasekaran, A., Akter, S., Ren, S.J., Dubey, R. and Childe, S.J. (2016), “Big data analytics and firm performance: effects of dynamic capabilities”, Journal of Business Research, Vol. 70 No. 8, pp. 356-365.

Wang, H., Xu, Z., Fujita, H. and Liu, S. (2016), “Towards felicitous decision making: an overview on challenges and trends of big data”, Information Sciences, Vol. 367 No. 7, pp. 747-765.

Wang, L. and Leite, F. (2013), “Knowledge discovery of spatial conflict resolution philosophies in BIMenabled MEP design coordination using data mining techniques: a proof-of-concept”, Computing in Civil Engineering, pp. 419-426.

Williams, G., Gheisari, M., Chen, P.-J. and Irizarry, J. (2015), “BIM2MAR: an efficient BIM translation to mobile augmented reality applications”, Journal of Management in Engineering, Vol. 31 No. 1, pp. 401-432.

Wolf, T., Schröter, A., Damian, D. and Nguyen, T. (2009), “Predicting build failures using social network analysis on developer communication”, Proceedings – International Conference on Software Engineering, pp. 1-11.

Yu, Z., Haghighat, F. and Fung, B.C.M. (2016), “Advances and challenges in building engineering and data mining applications for energy-efficient communities”, Sustainable Cities and Society, Vol. 25 No. 8, pp. 33-38.

Zhang, Y., Luo, H. and He, Y. (2015), “A system for tender price evaluation of construction project based on big data”, Procedia Engineering, Vol. 123, pp. 606-614.

Zhu, B.-F. and Ge, J. (2014), “Discussion of the evaluation method and value of green building’s POE in the era of large data”, Journal of Harbin Institute of Technology (New Series), Vol. 21 No. 5, pp. 10-14.

Further reading

Hedges, K. (2017), “Closing the gap between big data and the design and construction industries with Building Information Modeling (BIM) schema”, Journal of Architectural Engineering Technology, Vol. 1 No. 1, pp. 11-18.

Corresponding author

Upeksha Hansini Madanayake can be contacted at: madanayu@lsbu.ac.uk