TOWARDS AN INTEGRATED FRAMEWORK OF BIG DATA CAPABILITIES IN THE CONSTRUCTION INDUSTRY: A SYSTEMATIC LITERATURE REVIEW

Bernard Tuffour Atuahene[1], Sittimont Kanjanabootra and Thayaparan Gajendran

School of Architecture and Built Environment, The University of Newcastle, Callaghan 2308 NSW, Australia

Generation of data in the construction industry is increasing exponentially, becoming bigger and varied due to the use of sophisticated devices. This is necessitated by the call for automating the construction process. Getting valuable insights from data has taken the centre-stage of digitalization in the business world, through big data applications. Big data describes the application of advanced technologies to mine insights from datasets to solve problems. Studies have recommended the benefits of big data in the construction industry, though scanty, but cognizance is not given to what it takes for the industry to realize to the fullest, the benefits of big data. This study identified big data capabilities across different disciplines required in the construction industry through a systematic literature review. Four major big data capabilities thus organizational strategy, data, technology and people were identified from literature through qualitative synthesis. The research was limited to peer reviewed articles indexed in Web of Science and Scopus. The findings together with antecedents of innovation from the Schumpeter's theory of innovation led to conceptualization of a process framework for big data capabilities in the construction industry. The framework consists of antecedents to innovation, big data capabilities and outcomes. The study addressed the gap of big data capabilities deficit in the construction industry. Practically, the framework can serve as a guide for construction organizations interested in adopting big data in their operations.

Keywords: big data, capabilities, construction, innovation, systematic literature review

INTRODUCTION

Data on construction projects are compiled as progress reports to brief stakeholders on project performance (Omran, 2016). Subsequently, the report informs decisions about the direction of, or changes on the project. Mostly, minute datasets are captured in the report due to reporting templates and inability of stakeholders to process data. Unused datasets becomes idle, occupy space and add-up to waste, though, insight from these datasets can be mined for making well-informed decisions. Generation of data is on the rise due to the use of sophisticated IT devices like sensors, surveillance devices and drones (KPMG International, 2016; Han and Golparvar-Fard, 2017). Currently, data is essential in decision-making and improving processes. Those benefits have

[1] BernardTuffour.Atuahene@uon.edu.au

Atuahene, B T, Kanjanabootra, S and Gajendran, T (2018) Towards an Integrated Framework of Big Data Capabilities in the Construction Industry: A Systematic Literature Review *In:* Gorse, C and Neilson, C J (Eds) *Proceeding of the 34th Annual ARCOM Conference,* 3-5 September 2018, Belfast, UK, Association of Researchers in Construction Management, 547-556.

made data an inevitable resource in firms, necessitating the introduction of Big Data (BD) in the field of data management. In simple terms, BD is the application of sophisticated information technologies to gain valuable insights from large datasets (*volume*), generated at a faster rate (*velocity*) and in different format (*variety*) (Manyika *et al.*, 2011).

BD is intended to enhance business improvement, provided firms are competent and capable to host and manage the BD technologies. Nonetheless, the scanty literature have revealed benefits of BD in construction, cognizance is not given to what it takes for the industry to realize to the fullest, the benefits of BD analytics. This study presents a systematic literature review (SLR) on BD capabilities across different disciplines required in the construction industry.

LITERATURE REVIEW

Schumpeter Theory of Innovation and Big Data

Joseph Schumpeter is one of the progenitors of innovation and economic growth theories. Schumpeter theory of Innovation (Schumpeter, 1934, 2013), business cycle (Schumpeter, 1939) and entrepreneurship (Schumpeter, 1939) are based in economic, political and historical settings. The business cycle theory suggests firms experience fluctuations such as depression and growth due to both internal and external drivers. Nevertheless, firms' ability to make changes to internal drivers through the theory of innovations can lead to economic growth due to the influence of a leader with entrepreneurial qualities. These theories focus on the life, economic growth and production function of business entities.

A change in the production function, either marginally or wholly can improve a firm's economic performance (Schumpeter, 1934). It may include altering the quantity of materials, new equipment, change in supply of commodities and adopting different managerial strategies. These changes are termed as innovation. Impliedly, innovation is adjusting operational routines with the purpose of improving efficiency. External and internal drivers can motivate organizations to change its production function, but Schumpeter theory of innovation considered internal factors. Because innovations are not a response to the external demands but the need for change in firms (Schumpeter, 1934). In perspective, these drivers are within the control of the firm and affects its production function. However, the authors believe external drivers also influence process change in firms. Example is Revit and AutoCAD from Information System (IS) discipline, which are external to the construction industry but impacted the design of construction projects. Also, the effectiveness of BD in other sectors can motivate the construction industry to revisit the poor handling of data. External factors are included to the theory based on the aforementioned reasons. Schumpeter theory of Innovation is used because BD application can improve construction processes and economic growth of firms through the use of data. Furthermore, the theory becomes instrumental in understanding the antecedent for BD applications in the construction industry.

Generally, the use of sophisticated devices and the inability of traditional data analytic technologies to process voluminous and variety of datasets are the main drivers for BD applications (Oracle Corporation, 2014). The construction industry is using sophisticated devices to generate voluminous datasets like videos and images, which traditional technologies are incapable of processing. The internal drivers for BD application in the construction industry include the demand for changes in data

management, use of sophisticated devices, innovations and the need for data-driven management of the construction process. The benefits of these changes are pushed down to customers (Schumpeter, 1934, 1939). Process improvement experienced in healthcare (Wang and Hajli, 2017) and business processes (Wamba *et al.*, 2017). As well as pressures from project environment as witnessed from BIM (Dainty *et al.*, 2017) can motivate the construction industry to explore BD application.

Dynamic Capabilities (DC) and Big Data

The datasets and BD technologies constitute BD application. Drivers for BD application were identified through the Schumpeter theory of innovation, albeit, it is essential to identify the capabilities required for its implementation. Eisenhardt and Martin (2000, p.1107) defined DC as

> …the firm's processes that use resources - specifically the processes to integrate, reconfigure, gain and release resources - to match and even create market change…

The definition emphasized the aptness and ability of firms' to respond or make changes in the business environment. Gradually, the construction industry is becoming data driven, requiring firms to adopt processes capable of addressing data mismanagement. Capabilities are developed to enable firms to improve its processes to be competitive in the market (Wamba *et al.*, 2017). Processes, position and path are the essential elements in defining capabilities needed in firms (Teece *et al.,* 1997). Processes involve the integration of activities, learning and experimenting to know the best and quicker way of reconfiguring and executing tasks to meet the dynamism in the market. Position comprises of the processes and assets of the firm including technological, financial, institutional, firms' boundaries and the market. Path involves the past, present and future opportunities available to a firm by assessing the firm's competence and capabilities. Eisenhardt and Martin (2000, p.1106) made three observation of DC. Firstly, DC is adopting the "*best practices"* common amongst similar firms. Secondly, DC involves the manipulation of firms' resources to achieve new value. Thirdly, DC is developed in response to market dynamism. The firm's resources in that context are reorganized and reconfigured to respond to the changing environment (Gajendran *et al.*, 2014). For example, the business world is witnessing technological advancement like BD analytics which requires the manipulation of firm's resources or outsourcing to achieve data-driven management.

The combination of the essential elements of DC thus processes, position and path indicate that sources of data, BD infrastructure, BD analysis and the value from data are the basic capabilities derived from the definitions of BD. Discussing further the essential roles and relevance of the aforementioned components point to the capabilities required to exploit them. Amit and Schoemaker (1993) described capabilities as the ability of a firm to achieve its intended goals through judicious use of resources. Developing capabilities in firms tend to be inspired by demand from clients, paradigm shifts in the industry, external innovators and the firms desire to affect change by addressing challenges. Scholars from other disciplines have explored capabilities needed to run an effective and efficient BD (Mneney and Belle 2016) and BD analytics (Gunasekaran *et al.*, 2017). Different variables were considered in the various scholarly pieces albeit the authors believe all the variables can be subsumed into organizational strategy (Organization), BD infrastructure (Technology), Data and People (BD analyst). Through organizational capabilities, the BD analysts process the datasets using the big data technologies to produce knowledge. Meanwhile, these knowledge become irrelevant until management

experience is applied to understand, decode and interpret the knowledge in improving business value (Bradlow *et al.*, 2017). The decoded knowledge assist management in decision making, improving the competitive advantage and processes of firms (Wamba *et al.*, 2017). However, the relevance of these constructs in construction is still indeterminate and to get a better understanding we have adopted a SLR methodology to develop a theoretical framework.

RESEARCH METHODOLOGY

SLR is conducted to form the basis for new research, identify research gaps and provide summary evidence of existing scholarly works in a discipline (Santos and Da Silva, 2013). SLR is used in this study to identify BD capabilities in literature, which can be further explored in the construction industry. The research question (RQ) for the study is, *what are the capabilities required for the adoption and implementation of Big Data?* The key terms in the RQ became the search terms as shown in Figure 1. Scopus (362) and Web of Science (228) databases were searched. The time period span from the creation of the databases to 3rd January 2018. The downloaded citations and articles were managed and processed using Endnote X8. In all 590 articles were downloaded from the two databases as shown in Figure 1. The PRISMA flowchart (Figure 1) illustrates the processes involved in a systematic literature review to screen and select eligible articles for final extraction and analysis (Moher *et al.*, 2009).

Database	Searched terms	Results
Scopus	BD and Capabilities and Implementation	281
	BD and Competence and Implementation	12
	BD and Capabilities and adoption	64
	BD and Competence and adoption	5
WoS	BD and Capabilities and Implementation	178
	BD and Competence and Implementation	7
	BD and Capabilities and adoption	40
	BD and Competence and adoption	3
	Total	590

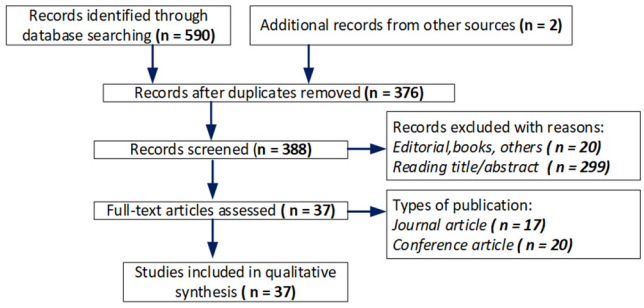

Figure 1 PRISMA flowchart and databases searched

Duplicates and articles not written in English were removed from the Endnote as well as editorials, books, book section. The included articles for qualitative synthesis were peer-reviewed conference and journal publication because high level research findings are disseminated through those medium. Articles on general Information systems and general reviews were excluded after reading the abstracts. Finally, thirty-seven (37) articles were further synthesized according to the classification - organization, data, technology and people - as indicated in earlier sections.

RESULTS AND DISCUSSION

Organizational strategy, data, technology and people were identified under the DC section of the study. The four capabilities were further decomposed after reviewing the thirty-seven articles (Table 1). Some of the selected 37 articles have not been included due to page restriction. The capabilities are discussed below:

Organizational strategy: Firms are required to respond to internal and external changes. The capacity of the firm to run the BD system is ingrained in its functions. Changes can happen in firms provided management shows its commitment (Table 1). For instance, top management commitment significantly influences the acceptance of BD predictive analytics in firms (Gunasekaran *et al.*, 2017). Apparently, it is easy to demand change but implementing change can fail abysmally due to individuals'

reluctance. Nevertheless the business agility and the culture of the firm can facilitate the rate of adoption or acceptance. Investing in the BD system cannot be ignored since it involves monetary commitment from firms. The combined efforts of investment, commitment, culture and business agility as well as external factors inform firms to sanction research and development goals, which subsequently leads to innovation in a firm or industry.

Table 1 Big Data capabilities

Capabilities	Decomposed capabilities	References
Data	Real time data devices (sensors)	Swan (2013); Intrieri *et al.,* (2017); Lee and Tso (2017)
	Availability data	Gupta and George (2016)
	Structured and unstructured	Leyens *et al.,* (2017)
	Data quality	Adrian *et al.,* (2017); Kim and Park (2017)
	Data standardization	Kim and Park (2017)
People	Managerial (Business) skills	Wamba *et al.,* (2017); Gupta and George (2016)
	Technical skills	Kim and Park (2017); Wamba *et al.,* (2017); Gupta and George (2016)
Technology	Real time data processor	Liu *et al.,* (2014)
	Big data infrastructure (Hadoop, fault detection etc.)	Mneney and Van Belle (2016); Intrieri *et al.,* (2017); Moyne and Iskandar (2017); Wamba *et al.,* (2017); Bisson *et al.,* (2016);
	Scalability	Vargas-Perez and Saeed (2017); Ghiţ *et al.,* (2013)
	Query capabilities	Borkar *et al.,* (2016); Istephan and Siadat (2016); Papakonstantinou (2016)
	Data mining techniques	Boehm *et al.,* (2016); Kopczynski *et al.,* (2017)
	Analytics capability (analytical, traceability, predictive	Kim and Park (2017); Wang and Hajli (2017)
	Data privacy and security	Adrian *et al.,* (2017)
	Visual analytics	Jayasingh *et al.,* (2016)
Organization strategy	Management support	Hallman *et al.,* (2014); Chen *et al.,* (2015); Coleman *et al.,* (2016); Mneney and Van Belle (2016); Kim and Park (2017)
	Financial resource	Adrian *et al.,* (2017); Kim and Park (2017)
	Data-driven culture	Gupta and George (2016); Kim and Park (2017)

Data: Processes in firms generate data. Overtly, identifying the what, how and where data can be found is a required BD capability. The first step in data capabilities -what - relates to the forms of data generated and gathered to solve an essential matter. Wenger and Sinha (2013) indicated that the data collected should conform to the firm's strategy. Tacitly, the data should be consistent with the expected outcome and possibly have a standard format, whilst there might be the possibility of having unstandardized data. For example, it is possible to have email communications not following a standardized format, unlike issuance of receipts on transactions. The second issue in data capabilities (*how*) deals with the manner of generating data. It is during this stage that data generating and capturing technologies (DGCT) takes the centre-stage. DGCT includes sensors, smart devices and unmanned aerial vehicles (Table 1). There is not much difference between *what* and *how*. It is important to note that these require firms' commitment to invest in DGCT. The last step critically looks at where data is generated in the firms' processes. For instance, Han and Golparvar-Fard (2017) used unmanned aerial vehicles to capture images and videos of

activities during construction. Succinctly, the business strategy towards the adaption of BD focuses on the firms' processes which needs serious improvement. An example in the construction industry can be on the construction supply chain, because of the fragmentation of the construction industry which is believed to have a negative effects on the performance of the industry. Though, all the steps in these capabilities are equally relevant but the *where* to get the data, perhaps dictates the other steps.

Technology: the centrality of the BD *buzz* is using IT infrastructures, which have the capacity to handle volumes of data unlike the traditional ones. Different infrastructures are made available by developers of BD analytics, such as Apache and Hadoop Spark (Table 1). The adapted platform should commiserate and compatible with the data gathered. The technology capability comprises of the hardware and software to manage the datasets (Kim *et al*., 2011). Wamba *et al.,* (2017) added that the BD technology have to be compatible to the firms' operations. The ability to apply machine learning techniques to the datasets should be an essential feature of the technology. Examples include using both supervised and unsupervised machine learnings for the purposes of clustering and classifying both structured and unstructured data. The presentation of the knowledge generated through the data analysis process like visual analytics is also vital in having a robust technology (Jayasingh *et al.*, 2016).

People: The implementation of the BD is meaningless if there is no BD expert in the firm. The competencies of the analyst should include technical knowledge, technology management skills, business and relational knowledge (Table 1). This capability might be outsourced, however firms can have an in-house personnel with expertise in managing the BD technologies. The ability of the BD experts to understand the construction processes together with the BD knowledge is an added advantage.

The Integrated Big Data Capabilities Framework

Figure 2 illustrates a holistic framework comprising of drivers of innovation (Schumpeter), BD capabilities (DC), experience and expected outcome on BD analytics in the construction industry. BD applications are highly spoken of in the data management literature and among managers.

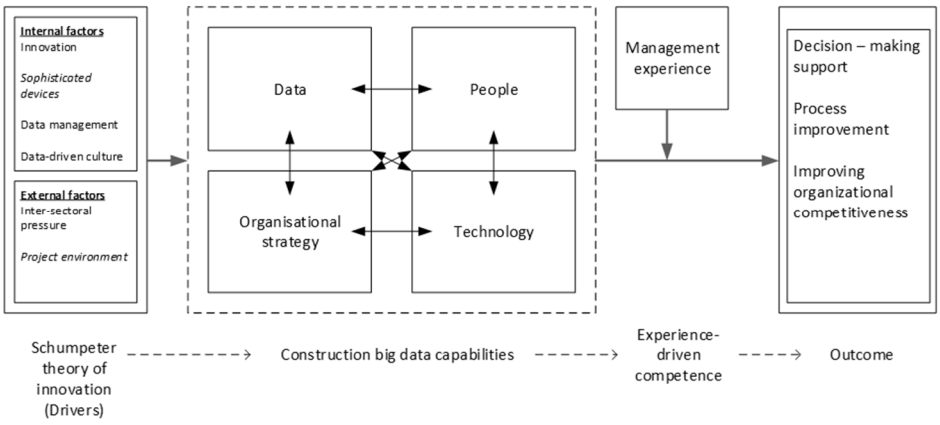

Figure 2 Integrated Big Data capabilities framework

There is a school of thought on whether or not, BD takes the role of managers in making strategic decisions. Bradlow *et al.,* (2017) argued that BD cannot replace the role of managers but a complimentary tool to assist managers. Further, the research

revealed that managers need to rely on theory to identify latent variables in spite of the overwhelming volume of data. This makes it easier for managers to seek for the essential insight needed in a time period rather than being "led astray" by the data.

Figure 2 proposed an integrated framework comprising of Schumpeter theory of innovation and BD capabilities discussed in an earlier section of this paper. Schumpeter theory of innovation argues that innovation in organizations are motivated by two main drivers - internal and external - but the theory focused on only internal drivers of the firm, which leads to changes in the economics of an entity. The assumption of innovation includes the setting-up of new plant and equipment; new entity for a specific purpose; a leader to lead the change (Schumpeter (1939); and new source of raw material supply changing the production function (Schumpeter (1934). Importing ideas and products from one industry to the other - IT devices from IS to construction - is an innovation. The acceptance and success story of innovation encourage competitors to explore its usefulness. Technological advancement in the construction industry like BD contributes to changes in the managerial function of production in an economic entity. Albeit, the construction industry does not necessarily have to set up new entity to make changes as assumed by Schumpeter (1939).

The internal drivers - innovation, sophisticated devices, data management and data-driven management - are motivation for construction firms to use BD (Figure 2). Sophisticated Devices used in the automation of construction generate voluminous and varied real time and secondary datasets which are difficult to process using the traditional data analytic technologies. Moreover, the construction industry do not necessarily mine datasets for improving its processes but store these data as evidence for litigation purposes.

Having data-driven culture can motivate construction firms to use BD in predicting profit margins, overheads and contingency sums for project and alternative ways of reducing the duration and cost of projects. Externally, the impact of BD in other sectors to improve its processes becomes an incentive to use BD in construction.

The identification and strengthening of the four capabilities (Table 1) are expected to make construction firms proficient to use BD. Knowledge is produced after the BD experts process the generated datasets in the BD technologies. Knowledge becomes useless if it is not decoded and interpreted to solve organizational problems. Therefore the success of the BD also depends on management ability to comprehend the knowledge generated from the datasets. This informs the decision to introduce management experience as a critical factor in the theoretical framework above. The application of management experience on the produced knowledge can assist in decision making, improve construction processes and make organizations competitive.

CONCLUSION

Drivers and capabilities are two most important factors needed to adopt and implement BD application in the construction industry. The generation of exponential volume of data continue to edge firms to adopt advance data analytics like BD application to mine valuable insights to the benefits of the firm. The study identified four BD capabilities - data, technology, organizational strategy and people - from the perspective of dynamic capabilities through a SLR. These capabilities together with drivers of innovation (Schumpeter theory of innovation), experience and expected outcome of BD analytics led to the development of an integrated framework for big

data capabilities. This study has highlighted the gap of BD analytics capabilities deficit in the construction industry. Practically, the framework can serve as a guide for construction organizations interested in adopting big data in their operations.

REFERENCES

Adrian, C, Abdullah, R, Atan, R and Jusoh, Y Y (2017) Factors influencing to the implementation success of big data analytics: A systematic literature review. In: *5th International Conference on Research and Innovation in Information Systems, ICRIIS*, 16-17 July, Langkawi Island, Malaysia.

Amit, R and Schoemaker, P J H (1993) Strategic assets and organizational rent. *Strategic Management Journal*, 14(**1**), 33-46.

Bisson, M, Phillips, E and Fatica, M (2016) A CUDA implementation of the pagerank pipeline benchmark. *In*: *2016 IEEE High Performance Extreme Computing Conference, HPEC 2016*, 13-15 September, Waltham, MA, USA.

Boehm, J, Liu, K and Alis, C (2016) Sideloading - Ingestion Of large point clouds into the apache spark big data engine. *In*: *23rd International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences Congress, ISPRS 2016*, 12-19 July, Prague, Czech Republic.

Borkar, D, Mayuram, R, Sangudi, G and Carey, M (2016) Have your data and query it too: From key-value caching to big data management. *In*: *2016 ACM SIGMOD International Conference on Management of Data, SIGMOD 2016*, June 26-July 1, San Francisco, USA.

Bradlow, E T, Gangwar, M, Kopalle, P and Voleti, S (2017) The role of big data and predictive analytics in retailing. *Journal of Retailing*, 93(**1**), 79-95.

Chen, D. Q, Preston, D S and Swink, M (2015) How the use of big data analytics affects value creation in supply chain management. *Journal of Management Information Systems*, 32(**4**), 4-39.

Coleman, S, Gob, R, Manco, G, Pievatolo, A, Tort-Martorell, X and Reis, M S (2016) How can SMEs benefit from big data? Challenges and a path forward. *Quality and Reliability Engineering International*, 32(**6**), 2151-2164.

Dainty, A, Leiringer, R, Fernie, S and Harty, C (2017) BIM and the small construction firm: A critical perspective. *Building Research and Information*, 45(**6**), 696-709.

Eisenhardt, K M and Martin, J A (2000) Dynamic capabilities: What are they? *Strategic Management Journal*, 21(**10/11**), 1105 - 1121.

Gajendran, T, Brewer, G, Gudergan, S and Sankaran, S (2014) Deconstructing dynamic capabilities: The role of cognitive and organizational routines in the innovation process. *Construction Management and Economics*, 32(**3**), 246-261.

Ghiţ, B, Iosup, A and Epema, D (2013) Towards an optimized big data processing system. In: *13th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing, CCGrid 2013*, 14-16 May, Delft, The Netherlands.

Gunasekaran, A, Papadopoulos, T, Dubey, R, Wamba, S F, Childe, S J, Hazen, B and Akter, S (2017) Big data and predictive analytics for supply chain and organizational performance. *Journal of Business Research*, 70(**sC**), 308-317.

Gupta, M and George, J F (2016) Toward the development of a big data analytics capability. *Information and Management*, 53(**8**), 1049-1064.

Hallman, S, Rakhimov, J, Plaisent, M and Bernard, P (2014) BIG DATA preconditions to productivity. *In*: *2014 IEEE 13th International Conference on Trust, Security and Privacy in Computing and Communications*, 24-26 September, 727-731.

Han, K K and Golparvar-Fard, M (2017) Potential of big visual data and building information modelling for construction performance analytics: An exploratory study. *Automation in Construction*, 73, 184-198.

Intrieri, E, Bardi, F, Fanti, R, Gigli, G, Fidolini, F, Casagli, N, Costanzo, S, Raffo, A, Di Massa, G, Capparelli, G and Versace, P (2017) Big data managing in a landslide early warning system: Experience from a ground-based interferometric radar application. *Natural Hazards and Earth System Sciences*, 17(**10**), 1713-1723.

Istephan, S and Siadat, M R (2016) Extensible query framework for unstructured medical data-A big data approach. *In*: *15th IEEE International Conference on Data Mining Workshop, ICDMW 2015*, 15-17 November, Atlantic City, NJ, USA.

Jayasingh, B B, Patra, M R and Mahesh, D B (2016) Security issues and challenges of big data analytics and visualization. *In*: *2nd International Conference on Contemporary Computing and Informatics, IC3I 2016*, 14-17 December, Amity University, Noida, India.

Kim, G, Shin, B, Kim, K K and Lee, H G (2011) IT Capabilities, process-oriented dynamic Capabilities and Firm Financial Performance. *Journal of the Association for Information Systems*, 12(**7**), 487-517.

Kim, M K and Park, J H (2017) Identifying and prioritizing critical factors for promoting the implementation and usage of big data in healthcare. *Information Development*, 33(**3**), 257-269.

Kopczynski, M, Grzes, T and Stepaniuk, J (2017) Hardware supported rule-based classification on big datasets. *In*: *International Joint Conference on Rough Sets, IJCRS 2017*, 3-7 July, Olsztyn, Poland, 655-668.

KPMG International (2016*) Building Technology Advantage: Global Construction Survey 2016*. KPMG.

Lee, T and Tso, M (2017) A universal sensor data platform modelled for real-time asset condition surveillance and big data analytics for railway systems: Developing a 'Smart Railway' mastermind for the betterment of reliability, availability, maintainability and safety of railway systems and passenger service. *In*: *15th IEEE Sensors Conference, SENSORS 2016*, Orlando, FL, USA

Leyens, L, Reumann, M, Malats, N and Brand, A (2017) Use of big data for drug development and for public and personal health and care. *Genetic Epidemiology*, 41(**1**), 51-60.

Liu, X, Lftikhar, N and Xie, X (2014) Survey of real-time processing systems for big data. *In*: *18th International Database Engineering and Applications Symposium, IDEAS 2014*, 07-09 July, Porto, Portugal.

Lu, W S, Chen, X, Peng, Y and Shen, L Y (2015) Benchmarking construction waste management performance using big data. *Resources Conservation and Recycling*, 105, 49-58.

Manyika, J, Chui, M, Brown, B, Bughin, J, Dobbs, R, Roxburgh, C and Byers, A H (2011) *Big Data: the Next Frontier for Innovation, Competition and Productivity*. McKinsey Global Institute.

Mneney, J and Van Belle, J P (2016) Big Data capabilities and readiness of South African retail organisations. *In*: *6th International Conference on Cloud System and Big Data Engineering (Confluence)*, 14-15 January, Amity University, Uttar Pradesh, Noida, India.

Moher, D, Liberati, A, Tetzlaff, J, Altman, D G and Group, T P (2009) Preferred reporting items for systematic review and meta-analyses: The PRISMA statement. *PLoS Med*, 6(**7**).

Moyne, J and Iskandar, J (2017) Big data analytics for smart manufacturing: Case studies in semiconductor manufacturing. *Processes*, 5(**3**).

Omran, B A (2016) *Application of Data Mining and Big Data Analytics in the Construction Industry*. Ohio, USA: PhD thesis, submitted to the Graduate School of the Ohio State University.

Oracle Corporation (2014) *Information Management and Big Data - a Reference Architecture*. Redwood Shores, CA: Oracle Corporation.

Papakonstantinou, Y (2016) Semi structured models, queries and algebras in the big data era. *In*: *2016 ACM SIGMOD International Conference on Management of Data*, 26 June - 1 July, San Francisco, California, USA.

Santos, R E S and Da Silva, F Q B (2013) Motivation to perform systematic reviews and their impact on software engineering practice. *In*: *2013 ACM / IEEE International Symposium on Empirical Software Engineering and Measurement*, 10-11 October, Baltimore, MD, USA.

Schumpeter, J A [Trans. Redvers Opie] (1934) Theory of economic development: An inquiry into profits, capital, credit, interest and the business cycle. Burlington, MA: Harvard University Press.

Schumpeter, J A (1939) Business cycles: A theoretical, historical and statistical analysis of the capitalist process. New York: McGraw-Hill Book Company.

Schumpeter, J A (2013) *Capitalism, Socialism and Democracy*. Abingdon: Routledge.

Swan, M (2013) The quantified self: Fundamental disruption in big data science and biological discovery. *Big Data*, 1(**2**), 85-99.

Teece, D J, Pisano, G and Shuen, A (1997) Dynamic capabilities and strategic management. *Strategic Management Journal*, 18(**7**), 509-533.

Vargas-Perez, S and Saeed, F (2017) A hybrid MPI-OpenMP strategy to speed up the compression of big next-generation sequencing datasets. *IEEE Transactions on Parallel and Distributed Systems*, 28(**10**), 2760-2769.

Wamba, S F, Gunasekaran, A, Akter, S, Ren, S J-f, Dubey, R and Childe, S J (2017) Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Business Research*, 70(**SC**), 356-365.

Wang, Y and Hajli, N (2017) Exploring the path to big data analytics success in healthcare. *Journal of Business Research*, 70, 287-299.

Wang, Y, Kung, L, Ting, C and Byrd, T A (2015,) Beyond a technical perspective: Understanding big data capabilities in health care. *In*: *48th Hawaii International Conference on System Sciences*, 5-8 January.

Wenger, R and Sinha, V (2013) *The Value of Big Data: How Analytics Differentiates Winners*. Georgia, USA: Bain and Company, Inc.