The Future of Big Data in Facilities Management: Opportunities and Challenges

Journal:	*Facilities*
Manuscript ID	f-06-2016-0064.R1
Manuscript Type:	Original Article
Keywords:	Big Data, Business Analytics, AEC, Facilities Management, Supply chain, Database

Abstract

Purpose: This paper explores the current condition of the Big Data concept with its related barriers, drivers, opportunities and perceptions in the AEC industry with an emphasis on Facilities Management (FM).

Design/methodology/approach: Following a comprehensive literature review, the Big Data concept was investigated through two scoping workshops with industry experts and academics.

Findings: The value in data analytics and Big Data is perceived by the industry; yet the industry needs guidance and leadership. Also, the industry recognises the imbalance between data capturing and data analytics. Large IT vendors' developing AEC industry focused analytics solutions and better interoperability among different vendors are needed. The general concerns for Big Data analytics mostly apply to the AEC industry as well. Additionally however, the industry suffers from a structural fragmentation for data integration with many small-sized companies operating in its supply chains. This paper also identifies a number of drivers, challenges and way-forwards that calls for future actions for Big Data in FM in the AEC industry.

Originality/value: The nature of data in the business world has dramatically changed over the past 20 years. This phenomenon is often broadly dubbed as "Big Data" with its distinctive characteristics, opportunities and challenges. Some industries have already started to effectively exploit "Big Data" in their business operations. However, despite many perceived benefits, the AEC industry has been slow in discussing and adopting the Big Data concept. Empirical research efforts investigating Big Data for the AEC industry are also scarce. This paper aims at outlining the benefits, challenges and future directions (what to do) for Big Data in the AEC industry with a FM focus.

Keywords: Big Data, Business Analytics, AEC, Facilities Management, Supply Chain, Database

Introduction

Over the past 20 years, data have significantly expanded in a large scale in various dimensions. According to a report from International Data Corporation (IDC), in 2011, the overall created and copied data volume in the world was 1.8ZB (≈ 1021 Bytes), which had increased by nearly nine times within five years (Ganz and Reinsel, 2011). This figure will double at least every other two year in the near future (Chen et al., 2014). Every day, 2.5 quintillion bytes of data are created and 90% of the data today were created within the last two years (IBM, 2012). Zikopoulos et al. (2012) expect data volumes to reach 35 Zettabytes (270 bytes) by 2020. Manyika et al (2011) cite that in 2009, nearly all sectors in the US economy had at least an average of 200 Terabytes of stored data per company with more than 1,000 employees and many sectors exceeded more than one Petabyte in mean stored data per company.

In parallel with this data explosion, the Big Data concept gained momentum in the early 2000s when industry analyst Doug Laney articulated the now-mainstream definition of Big Data as 3V (Volume, Velocity and Variety); and correspondingly, IDC defined it: "Big data technologies describe a new generation of technologies and architectures designed to economically extract value from very large volumes (into petabyte volumes) of a wide variety of data, by enabling high-velocity capture (streaming data), discovery, and/or analysis." (Ganz and Reinsel, 2011). Four other characteristics that are relevant to Big Data: (1) value (the usefulness of data in making decisions), (2) variability (data flows can be highly inconsistent with periodic peaks), (3) complexity (the degree of interconnectedness and interdependence in big data structures spread over data warehouse systems) and (4) veracity (data reliability) (Kaisler et al., 2013; Katal et al., 2013; Zikopoulos et al., 2013). Generally built on Relational Database Management Systems (DBMS), small data sets on the other hand, are in low volumes (usable chunks, in the Gigabyte-Terabyte range), batch velocities, from more limited resources, more structured and accessible to answer a specific question or addresses a specific problem, rather than providing explorative insights (Sagiroglu and Sinanc, 2013). This data continuum, from Small Data to Big Data, can be seen in Figure 1.

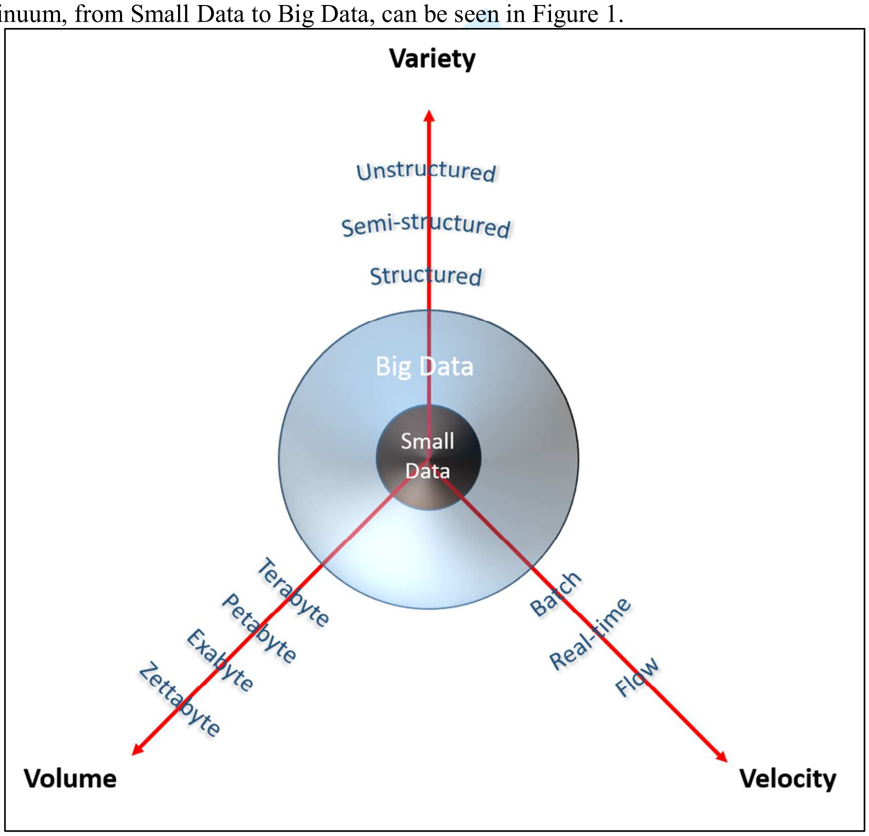

Figure 1. Data continuum

Comment [MOU1]: The location of the figure was changed as per Reviewer 1's 1st comment.

The distinctive characteristics of Big Data (e.g.. the 3Vs and more) presents many technical and social challenges for Business Analytics (BA), which is often referred to as the techniques, technologies, systems, practices, methodologies, and applications that analyse critical business data to help an enterprise better understand its business operations and make timely business decision (Chen et al., 2012). However, its major strengths; discovering new insights and hidden values, also come from a simultaneous analysis of those multi-sourced and unstructured data flows (Chen et al., 2012). For instance, in an article published by Ginsberg et al. (2012), the authors explain the use of Big Data analytics by Google to successfully forecast the spread of an influenza epidemic based on Internet search entries. In summary, discovering explorative insights and complex-interrelations that cannot be seen from smaller data sets, the Big Data information comes from multiple, heterogeneous, autonomous sources with complex and evolving relationships, and keeps growing (Wu et al., 2014).

Business Analytics and Big Data

According to Boyd and Crawford (2012), Big Data analytics is a cultural, technological, and scholarly phenomenon that rests on the interplay of technology, analysis and mythology; the widespread belief that large data sets offer a higher form of intelligence and knowledge that can generate insights that were previously impossible, with the belief of truth, objectivity, and accuracy. To turn the mythology into a reality, a significant value for business and operations, return of investment through Big Data analytics, with an effective deployment of both the technology and analysis should be attained. Otherwise, organisations risk owning huge data chunks without properly utilising those data as Big Data.

Japkowicz and Stefanowski (2016) draws a comparison between the traditional data mining in Small Data and Big Data analysis as illustrated in Table 1. Highlighting that traditional data mining and Big Data are two different concepts, the table essentially summarises the key technical differences between Small Data and Big Data from a BA perspective.

Table 1. Traditional Data Mining (Small Data) vs. Big Data analysis (adopted from Japkowicz and Stefanowski (2016))

	Traditional Data Mining (Small Data sets)	Big Data Analysis
Memory Access	Stored in centralized RAM and can be scanned efficiently several times.	The data may be stored on highly distributed data sources and single scans are common.
Computational Architecture	Serial, centralized processing is sufficient.	Parallel and distributed architectures may be necessary.

Comment [MOU2]: The table caption was modified as per Reviewer 1's 2nd comment.

Data Types	The data source is relatively homogeneous. The data is static and, usually, of reasonable size.	The data may come from multiple data sources which may be heterogeneous and complex. The data may be dynamic and evolving.
Data Management	The data format is simple and fits in a relational database or data warehouses.	Data formats are usually diverse and may not fit in a relational database. The data may be greatly interconnected and needs to be integrated from several nodes.
Data Quality	The provenance and pre-processing steps are relatively well documented. Strong correction techniques were applied for correcting data imperfection.	The provenance and pre-processing steps may be unclear and undocumented. There is a large amount of uncertainty and imprecision in the data.
Data Handling	Security and Privacy are not of great concern.	Security and Privacy may matter. Data may need to be shared and the sharing must be done appropriately.
Data Processing	Only batch learning is necessary. Learning can be slow and off-line. The data fits into memory. All the data has some sort of utility.	Data may arrive in a stream and need to be processed continuously. Learning may need to be fast and online. The scalability of algorithms is important. The data may not fit in memory. The useful data may be buried in a mass of useless data.
Result Analysis	Statistical significance results are meaningful. Many visualization tools have been developed Interaction with users is well developed. Answers to specific questions are obtained.	With massive data sets, non-statistically significant results may appear statistically significant. Traditional visualization software may not work well with massive data. Explorative discoveries are common.

Big Data has a potential to generate business value. From a survey of 325 commercial enterprises, Russom (2011) indicates that Big Data analytics can generate value in the form of better targeted influencer marketing, more varied and accurate business insights, recognition of business opportunities, automated decisions for real-time processes, definitions of customer (user) behaviours, customer retention, detection of fraud, quantification of risks, trending for market sentiments, understanding of business (behaviour) changes, better planning and forecasting, resource optimization, identification of root causes of cost and understanding consumer(user) behaviours. Pries and Dunnigan (2015) give similar examples of real-life Big Data benefits on meta knowledge discovery from vast amount of academic publications, curriculum analysis and design in education, maintaining and increasing sales and market share through predictive analysis, increasing security (password control, penetration detection, copyright and patent violation etc.), customer-centric, risk-centric and finance-centric insurance applications, customizing travel experiences, optimization and tracking of logistics assets (freight, vehicles, etc.), and GIS based city-management, crime fighting and resource optimization examples from New York, Chicago, Los Angeles and Baltimore in the US. In parallel with this, a recent research by Gartner (2014) conducted in 2014 indicated that of all the surveyed 302 organisations in the US, 73% of the respondents had invested or planned to invest in Big Data in the next 24 months, up from 64% in 2013. The survey also highlights that the number of organisations stating they had no plans for Big Data investment fell from 31% in 2013 to 24% in 2014.

Comment [MOU3]: The "etc." was corrected as per Reviewer 1's 3rd comment.

However, according to Davenport and Dyche (2013), very few organisations have taken the steps to rigorously quantify the return on investment for their Big Data efforts and that the proof points about Big Data often transcend money represented by cost savings or revenue generation for the long-term. The authors instance a comparative return on investment study in two different environments in 2011 in similar business contexts. The first environment was a high-speed data warehouse appliance employing traditional data warehouse usage (ETL) and data provisioning processes. The second environment was Big Data running on a newer Big Data technology using massively parallel (MPP) hardware. The Big Data project surpasses the traditional project in cumulative cash flow, net present value and internal rate of return by significant margins with a shorter break-even point (Davenport and Dyche, 2013).

While implementing Big Data analytics, those technical issues can be potential barriers: (1) data presentation, (2) data redundancies, (3) data lifecycle management, (4) data security and confidentiality, (5) energy management, (6) interdisciplinary cooperation, (7) inexpert staff, (8) investment and maintenance costs, (9) lack of business leadership, (10) hardships in designing analytic systems, and (11) lack of current database software in analytics (Chen et al., 2014; Sagiroglu and Sinanc, 2013). Also, other more generic concerns include (1) the objectivity and accuracy of Big Data resources, (2) the widespread belief that bigger data sets is always better without giving much concern to methodological issues and data quality, (3) ethical concerns in data use, (4) the highly context dependent nature of Big Data analytics, and (5) that limited access to large volumes of data by small groups in the society can create information divides or concessions (a class of the Big Data rich) (Boyd and Crawford, 2012).

Comment [MOU4]: The location of the reference was changed as per Reviewer 1's 4th comment.

According to McAfee et al. (2012), the main managerial challenges for Big Data are in; (1) creating leadership teams that set clear goals, define what success looks like, and ask the right questions, (2) talent management of the professionals competent in capturing, analysing, inferring from and presenting Big Data sets, (3) management of technology; hardware, which are becoming cheaper (i.e. Big Tables spread on parallel (cheaper) computers), and software that are becoming open-source (i.e. the Java based Hadoop framework which incorporates data map reducing), but also generally are out of the skill sets of IT departments, (4) effective decision making through a flexible organizational structure, and (5) sustaining a company culture that underpins the data-driven organization.

Big Data and smartness in the AEC industry

With the constant flow of large data sets generated by different organisations, individuals and inanimate objects in many different data formats, the current nature of data in the AEC industry is essentially of Big Data (Jiao et al., 2013). Despite this, Brown et al (2011) identified in a report from McKinsey Global Institute that the ease of capturing Big Data's value, and the degree of its potential for the AEC industry is comparatively low, which highlights the magnitude of the challenge the industry is facing in order to obtain a real value from Big Data analytics for its operations. Spithoven et al. (2009) indicate that the traditionally low-tech AEC industry has lower knowledge absorptive capacity to internalise external developments (e.g. academic research or Big Data analytics) and R&D. On the other hand, instancing two possible implementation areas in Business Information Modeling (BIM) based maintenance forecasting for Facilities Management (FM) and project performance analysis, Hardin and McCool (2015) state that the interest in Big Data has been rapidly growing in the AEC industry. Although still immature, leveraging those vast amounts of unstructured data is an asset that can greatly benefit the industry in terms of better resource management and forecasting, risk mitigation and efficiency improvements in their resource use (McMalcolm, 2015; Rijmenam, 2015).

Although Big Data focused research in the AEC industry is currently very limited, some empirical works highlighting the value of data mining over smaller data sets to discover knowledge patterns exist in the literature. For Facilities Management, data mining was utilised to understand facilities' use patterns (Peitgen et al., 1992), maintenance cost patterns and preventive maintenance planning (Liew and Rosenblatt, 2003) and for energy efficiency purposes (Reffat et al., 2006). However, one should note that although Big Data employs some data mining techniques, it goes beyond the analytics and technology employed for smaller data sets as explained in Figure 1 and Table 1. Data management in the AEC industry also presents its own challenges; (1) during the life-cycle of an asset, different organisations collect and own data fragmentally, (2) alongside conventional data media (i.e. paper files, old project drawings on papers), existence of many digital data types and interoperability issues, (3) lack of standardisation in data management, (4) blurred contractual configurations as to the management of data, (5) skill-shortages for effective data utilisation, (6) low penetration of and buy-in for innovation, (7) Small and Medium sized organisations (SMEs) dominating the supply chain that are risk averse and generally lacking the necessary funds for data management investments, (8) hardships faced during the collection of correct as-is data from live construction conditions, and (9) issues associated with demonstrating the business case and return-on-investment for data management to the industry (Soibelman and Kim, 2002; Bakis et al., 2007; Ergen et al., 2007; Chen and Kamara, 2011; Jiao et al., 2013).

Comment [MOU5]: For Reviewer 2 comments to expand on the literature review toward Data Mining and data management

The emergence of connected sensor networks and ubiquitous computing, often explained under the term the Internet of Things (IoT), has recently added new dimensions to the applications of Big Data. Chui et al. (2010) define the IoT as: sensors and actuators embedded in physical objects – from roadways to pacemakers – are linked through wired and wireless networks, often using the same Internet Protocol (IP) that connects the Internet. The emergence of the IoT through data generating inter and intra connected sensor networks in everyday objects coupled with Big Data analytics gives way to the collection of and inference from an immense amount data to create “smart” cities, the new urban environment, one that’s designed for performance through Information and Communication Technologies (ICTs) and other forms of physical capital (Chen et al., 2014).

With the effective management of resources through intelligent management, visionaries hope that cities will drive a higher quality of life for citizens, drive down waste, and improve socio-economic conditions (Stimmel, 2015). Hollands (2008) identifies five main characteristics of a “smart city” of this nature from the literature: (1) widespread embedding of ICT into the urban fabric, (2) business-led urban development and a neoliberal approach to governance, (3) a focus on social and human dimensions of the city from a creative city perspective, (4) the adoption of smarter communities agenda with programmes aimed at social learning, education and social capital, (5) and a focus on social and environmental sustainability. Many city governments now employ real-time analytics in their data centres to manage the aspects of how a city functions and is regulated in dynamic traffic management (adjusting traffic lights and speed limits), efficient use of the police force, environmental monitoring (air pollution, water levels, seismic activity), and general city information visualized on the city map (e.g. energy monitoring, weather, air pollution, public transport delays, public bike availability, river level, electricity demand, the stock market, Twitter trends in the city, traffic camera feeds, the happiness level etc.) (Kim et al., 2012; Kitchin, 2014).

Another closely related use of those technologies is in “smart grids”, in which electrical grids for power delivery are managed through sensors, readers and data analytics for optimized energy load balancing and forecasting, improved efficiency and sustainability in energy management (Amin and Wollenberg, 2005). Also, in the maintenance of industrial facilities and power plants, where the reliance on sophisticated electro-mechanic equipment is high, Big Data analytics can yield significant savings through predictive maintenance (Song et al., 2013; Chappel, 2014).

From the AEC industry’s perspective, Big Data analytics, with other emerging technologies such as the IoT, can translate into (1) reduction in asset maintenance costs, (2) better tailored or customized services to individuals and groups, (3) more accurate forecasting, (4) risk mitigation, (5) resource levelling and optimization, (6) performance evaluation, (7) informed investment decisions and on a greater scale, “smarter” cities with predictive solutions (Reffat et al., 2006). Alongside “smart cities” and “smart grids”, it can be inferred that Big Data will soon cascade down to the realisation of “smart” design offices, construction sites and facilities. Despite the potential of Big Data, there is little research conducted to understand the Big Data phenomenon in the AEC industry. The rest of this paper will therefore look into the barriers, drivers and potential applications of such data in the AEC industry, particularly for Facilities Management.

Comment [MOU6]: The location of the reference was changed as per Reviewer 1’s 5th comment.

Methodological approach

In order to explore the potential drivers and challenges of Big Data within the AEC industry, two industry workshops were held, bringing together a total of 200 participants in partnership with key industry players and professional bodies. Table 2 shows further details about these workshop events.

Table 2. Workshop Details

Workshop	Date	Venue	Number of attendees
Workshop 1 Big Data Challenge and Opportunities in the AEC industry	February 2015	In partnership with Asite Solutions and CIOB - Google Campus, London	130
Workshop 2 Big Data in Facilities Management	June 2015	In partnership with CIOB and RICS - Media City Manchester	70

Workshop 1 – Big Data challenges and opportunities in the AEC industry

This workshop was held in collaboration with Asite Solutions, which is an IT solution provider within the AEC industry. The workshop brought together an average of 130 people, 70% of which were from the AEC industry, while 20% were from academia and 10% from the IT sector. The workshop aimed at assessing the level of understanding of Big Data, identifying the drivers and challenges perceived by the attendees. Amongst the participants, 73% considered themselves neither novice nor expert in Big Data, while 18% considered themselves as experts and 12% as novices. The day-long workshop invited a number of key speakers to share their understanding of Big Data from the technological, and industrial perspectives and to share some of the applications used in the AEC industry. The participants were then asked to engage in a clicker exercise in order to give their views of the Big Data drivers and challenges. This workshop identified Facilities Management as a discipline with the greatest potential for the integration that can benefit from the utilisation of Big Data in the AEC industry. The outcome of this workshop provoked the organisation of a follow on industry event, hence, workshop 2.

Workshop 2 - Big Data in Facilities Management

This workshop was held in collaboration of a number of industry partners, and sponsored by the CIOB and RICS. An average of 20% of the attendees were from academia while 70% of the attendees were from the AEC industry (with the majority of FM expertise) and 10% from other industries. This paper shares the outcome of the focus group discussions and analysis of the participants feedback identify around the opportunities and challenge of Big Data in FM.

Results and analysis

Workshop 1 – Big Data challenges and opportunities in the AEC industry

Following a number of presentations around Big Data, its theories and concepts, the workshop participants were asked to select their preferred choice by using the clickers in relation to the questions below. These results were further discussed with the support of a panel of experts.

Challenges facing Big Data

The participants were asked to select one of the factors that they perceived to be the most challenging for utilising Big Data in the AEC industry. The results showed that 55% of the participants found that obtaining, structuring, managing the data is one of the most challenging factors. 24% thought the need for changes in the culture and attitude towards sharing project data, while 12% voted for the need to understand the business value (See Figure 2).

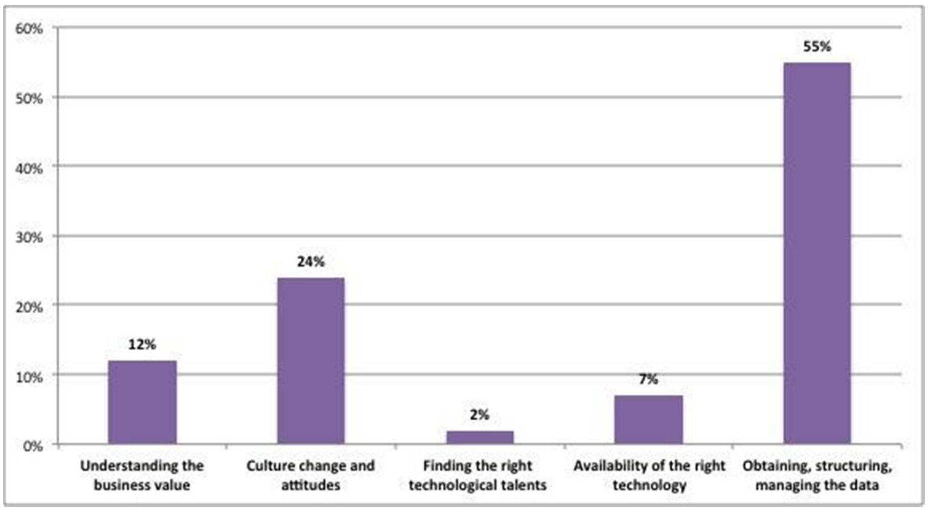

Figure 2. Challenges facing Big Data in the AEC industry

It can therefore be argued that the majority of these challenges are somewhat interlinked, whereby changing the way people work, and appreciating the business value of the processes they engage in and how these processes can be a vehicle to provide the organisation with the relevant business intelligence, this may then help with the way data is handled, exchanged and stored.

Technological challenges

The participants were asked to select what they perceive to be the most technologically challenging for Big Data. The responses showed that 67% of the participants felt that making use of structured (which are in silos) and unstructured data

Comment [MOU7]: The typo was corrected as per Reviewer 1's 7th comment.

(in different formats) are the most technologically challenging factors, followed by capturing the right data and real-time delivery to the right people at 19%, with 14% voting for having data privacy and security to access and deploy data (see Figure 3).

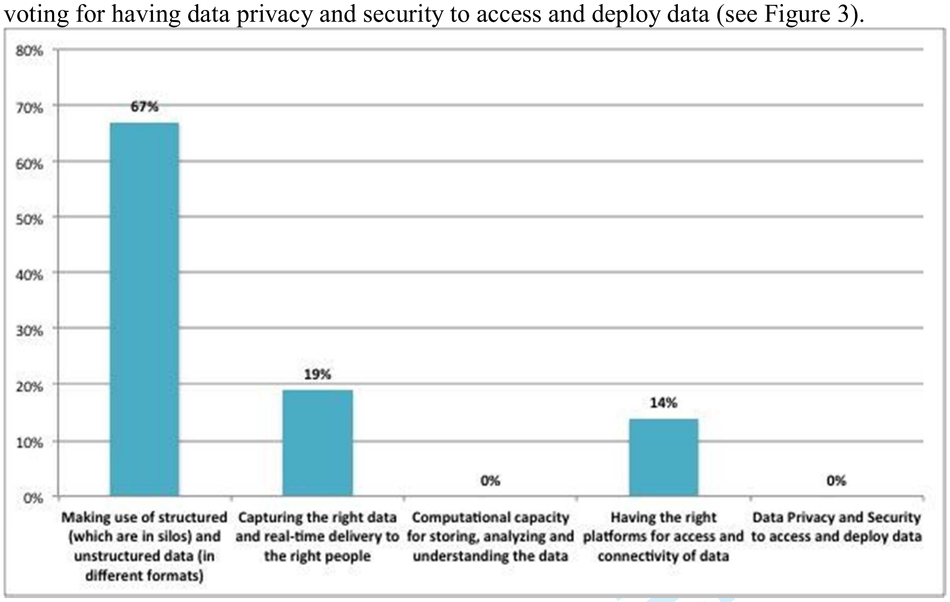

Figure 3. Technological challenges for Big Data in the AEC industry

Interestingly, factors relating to computational capacity for storing, analysing and understanding the data, and data privacy and security to access and deploy data did not gather any score. The panel discussions around these issues indicated that the participants felt that the technology is available and it is out there, however, the main challenge is for organisations to utilise these technologies to serve their organisational goals.

Operational challenges facing Big Data

Comment [MOU8]: The terminology was corrected as per Reviewer 1's 6[6h] comment.

When asked about the biggest operational challanges that face the utilisation of Big data, 42% of the participants felt that finding the right talent that is capable of working with new the technology and interpreting the AEC processes is the most significant barrier. This was followed by the cost of the technology (26%) and meeting the high cost of recruiting technologists (16%) (see Figure 4).

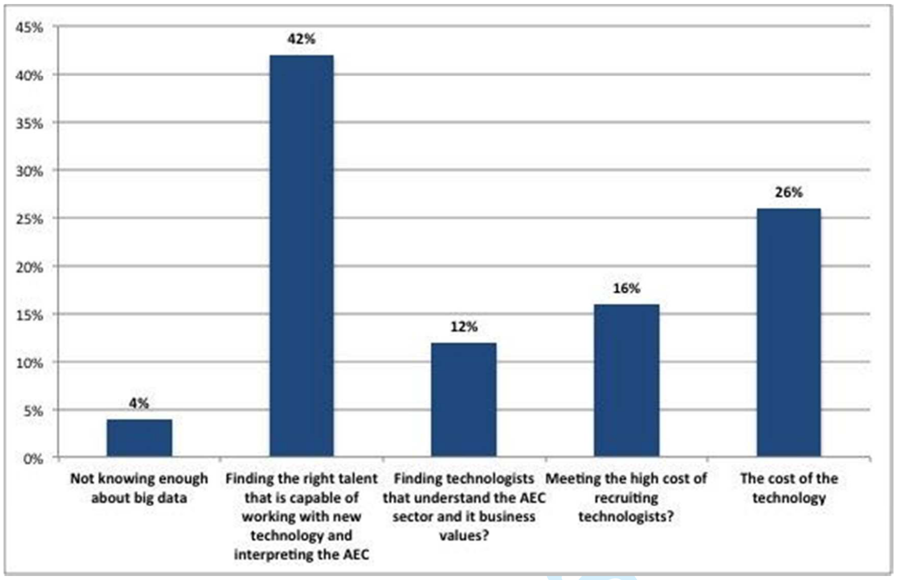

Figure 4. Operational challenges facing Big Data in the AEC industry

Comment [MOU9]: The terminology was corrected as per Reviewer 1's 6[6h] comment.

It therefore seems that organizations lack the technological talents that are capable of understanding the balance between the AEC processes and aspects of the technology that can enhance these processes, and that without a good understanding of the business value that Big Data can bring, organisations may be missing out on the long term investment in Big Data.

Data challenges

The participants were asked to identify the most challenging issue about the data. The response percentages were fairly close, whereby 37% and 34% of the participants felt the structuring of the data and the uniformity and consistency of the data are the most significant challenges before Big Data; while a relatively close percentage to these (29%) thinks the existence of paper driven documents is a major challenge (see Figure 5).

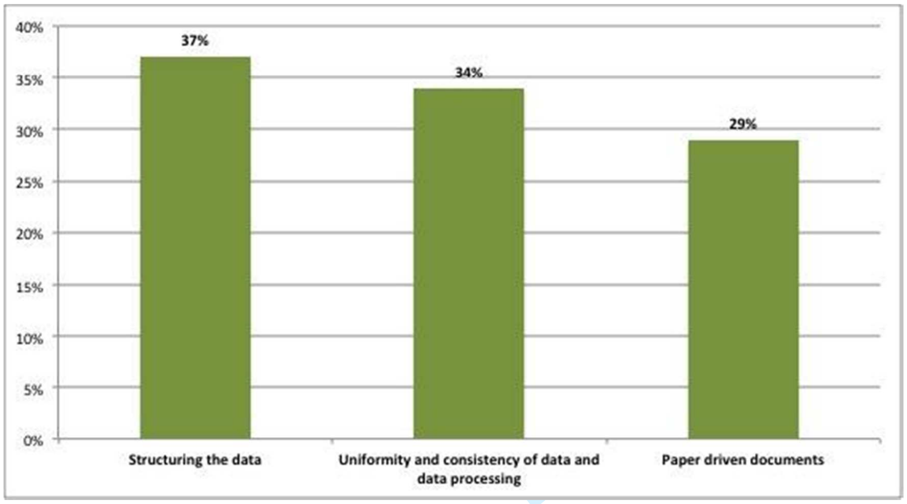

Figure 5. Data challenges for Big Data in the AEC industry

During the panel discussions of these results, strong arguments were put forward in relation to the fragmented nature of the construction process, with the existence of the project information in different forms and formats.

AEC readiness for Big Data

The participants were asked to assess the AEC industry's readiness for Big Data. Only 33% of the participants felt that the industry is ready to utilise Big Data, while 67% felt that the industry is not yet ready (see Figure 6).

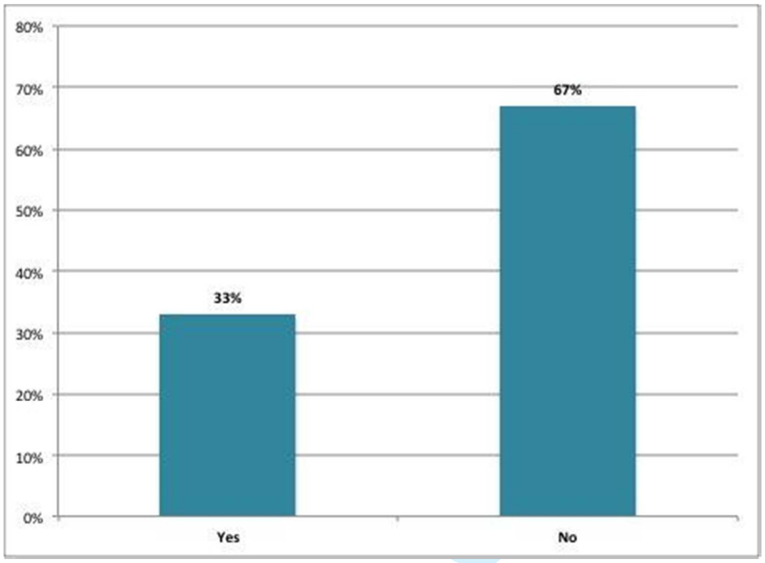

Figure 6. Readiness of the AEC industry for Big Data

Potential for integrating Big Data

The participants were asked in which of the project aspects do they see the most potential in integrating Big Data in the AEC industry. The results show that the use of Big Data has the biggest potential use for the Facilities and Operations Management aspect of the project life-cycle, followed by the sustainability and reduction of energy consumption aspect, with little discrepancy between this aspect and the feasibility, design, procurement and construction aspect (see Figure 7).

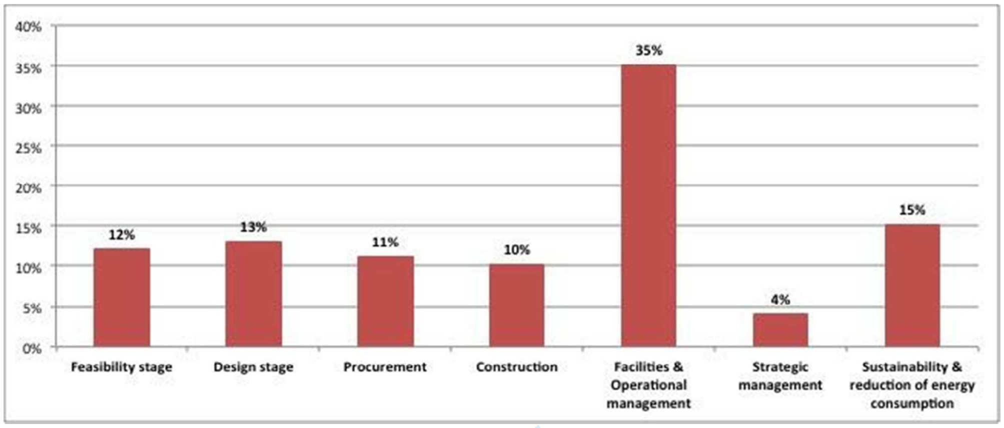

Figure 7. Potential for integrating Big Data in the AEC industry

These results were key to exploring the potential of Big Data in Facilities Management, which is the centre of this paper.

Workshop 2 - Big Data in Facilities Management

Given the participants responses to the earlier questions, and their identification of Big Data as a tool which can bring in value to the Facilities Management (FM) discipline, this section shares the outcome of the 2nd workshop on Big Data and FM. Following a number of keynote presentations by discipline experts from the industry and academia during this workshop, the participants were asked to provide written feedback on what they perceived to be the;

1. Drivers for the integration of Big Data and FM
2. Challenges for the integration of Big Data and FM
3. Steps the FM sector should take for a better use of Big Data

The rest of this section discusses theses findings based on the three research questions above.

Drivers for the Integration of Big Data and FM

Figure 8 below show a summary of the participants' responses, which were grouped in three main categories;

- *Category I - Drivers relating using Big Data in direct relation to the nature of the FM discipline:* The participants came up with a list of factors that they perceived as the drivers for using Big Data for FM. These evolved around improving the decision making process, for example, the effective use of heating and cooling systems within spaces, as well as providing more sustainable design solutions that can help reduce energy consumptions. Big Data is also seen as a mean for helping improve the decision making process, not only at the design stage, where working on large scale designs for smart cities, but also with making decisions on managing the facilities in terms of predicting maintenance dates or identifying faults and problems within the facilities, and the consequences of these decisions on nearby facilities. Giving organisations a competitive advantage is also another important driver for using Big Data.

- *Category III - Drivers relating to the use of Big Data based on external factors:* Some of the external factors mentioned by the participants that could drive organisations to integrate Big Data evolve around understanding the role of Big Data in FM through live examples or shared case studies and practices. The participants also felt that there are no current policies that force organisations to adopt certain practices in order to develop an infrastructure for Big Data within organisations. Also, it was identified that there is a need for more research to be carried out to show how the benefits of Big Data can be utilised for FM.

- *Category III - Drivers relating to the technological factors*: The participants' responses evolved around organisations being equipped with systems and platforms that have the capacity and capability to capture large volumes of data, with a level of intelligence that can react to this data input in real time, as well as developing the right IT skills within organisations that will deal with the data.

These results align with the nature of the FM discipline where, for example, maintenance data or energy level monitoring data can be actively and proactively captured by different processes and systems, with intelligent simulations embedded within them that can assist the end user with managing the facilities in a more efficient and effective way.

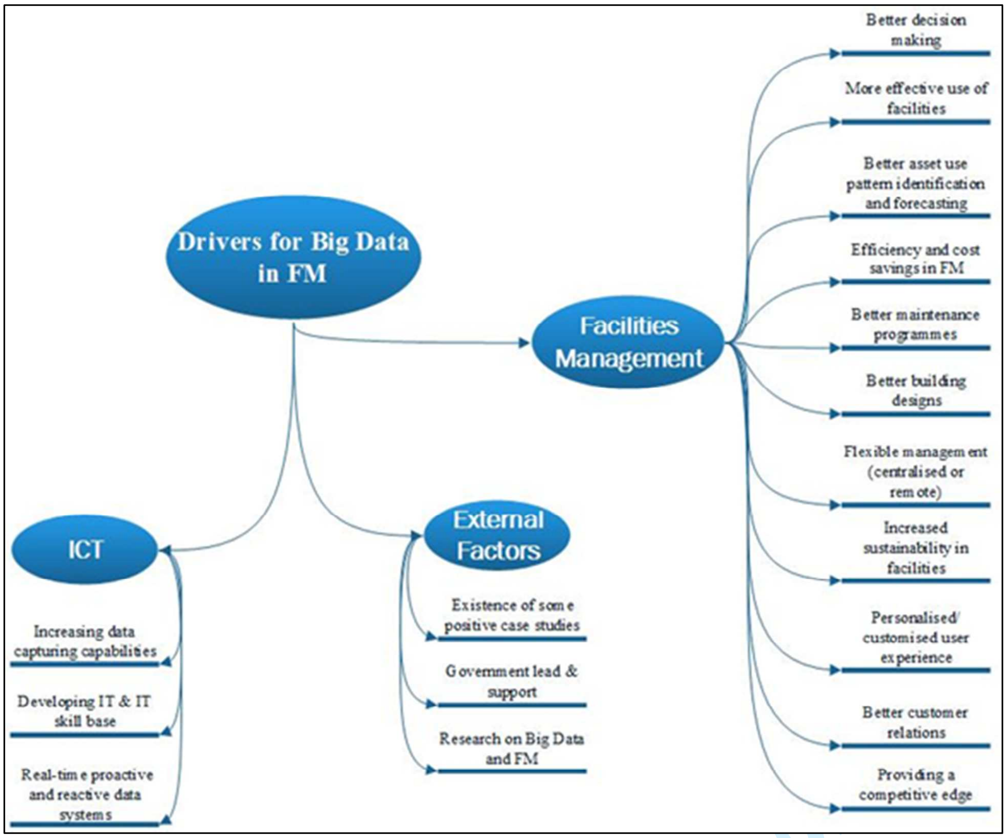

Figure 8. Drivers for the integration of Big Data and FM

It is apparent from the participants' responses that there are a number of drivers that recognise the relevance of Big Data to the FM discipline, to provide more sustainable and easily managed facilities, but with the support of the relevant skills, platforms and policies.

Challenges for the Integration of Big Data and FM

When asked about the challenges facing the utilisation of Big Data in FM, the participants' responses were boxed in three main categories (see Figure 9);

- *Category I - Challenges facing the FM discipline:* A number of challenges were identified under this category, which evolved around producing a meaningful value for FM out of extensive data that are generated from different processes out of pre and post construction, and without doubt is closely linked to the lack of knowledge and skills in differentiating between the value of these data, and how it could be utilised for FM. One of the interesting points that was highlighted is the inconsistency of data generated (i.e. the handover data from construction to maintenance/ operations), which makes its reliability questionable under different circumstances, and that the field of Big Data in FM is still in its infant stages and needs to grow in maturity. It was discussed that the FM discipline itself tends to be known for its short term thinking by dealing with the problems there and then and after the construction process had been completed, at the same time (and to its nature) that FM discipline tends to lack collaboration and inter-linkage of data with other processes and disciplines within the same project or at a larger scale in order to deliver "smartly" managed facilities.

- *Category II - Challenges based on external factors:* Under this category the participants identified some of the external factors that relate to ethical issues, one of which being, who owns which data and how much of it can be used by FM managers without compromising data protection acts. For example, the government's data protection act and, ethical and privacy issues around the use of personal data were underlined. That said, one of the main challenges mentioned has to do with the possible bias in the generated data which could be due to wrong inputs or environmental challenges. The participants also highlighted the lack of good examples that can guide the discipline to implement Big Data in FM.

- *Category III - Technological challenges:* Similar challenges to the drivers for adopting Big Data can be depicted under this category, which revolve around the security and scarcity of the data to be instantly available via technological solutions, and channelling the data from the different stages of a project using relevant technological solutions (for example sensors) so that the data can be linked and connected to the FM stages, as well as the compatibility of the data generated from other processes.

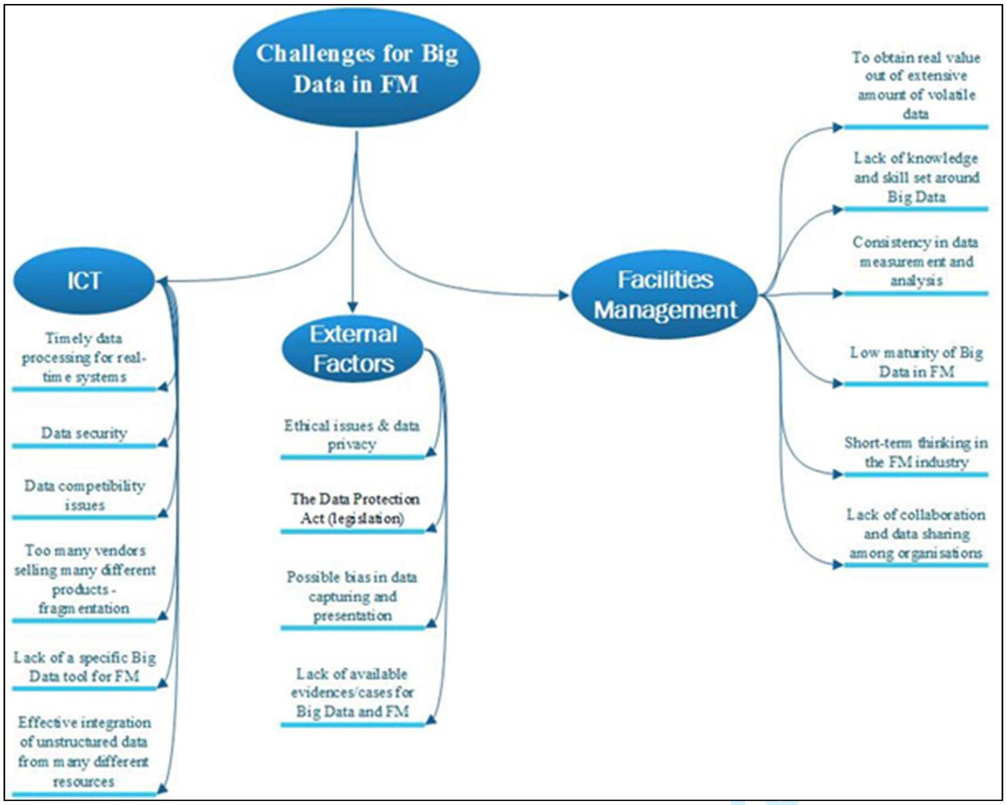

Figure 9. Challenges for the integration of Big Data and FM

Therefore, the main challenges facing Big Data for FM seem to focus mostly on the data itself. While the inter-linking between the data generated from different processes seems to pose one of the biggest challenges, the need for suitable platforms for data processing, data protection, data accuracy and data security also seems to add to these challenges.

Steps the FM discipline should take for a better use of Big Data

The participants' responses were also categorised as shown below (see Figure 10);

- *Category I - Steps within the FM discipline:* The participants highlighted the need for recruiting individuals who are equipped with the right discipline-special skills as well as the relevant technological skills to help with the integration of the technology and the processes that feed into FM. There is also a need for a change in the culture of working and increasing the awareness of the application of Big

Data internally within the organisation and to draw lessons from other industries such as the retail industry, where the Big Data concept has started to play an important role in predicting patterns in customer behaviour and customising customer experience. There also needs to be a well defined strategy that allows for clear data inputs, and value of outcomes and benefits.

- *Category II - Steps reliant on external factors:* Besides some of the factors that were previously mentioned, such the assurance around data protection issues and the presence of good practice cases, there is a desire for the formation of a Big Data industry group for FM in order to set the agenda for the discipline, in a similar way that this happened for BIM. A regulatory body will also encourage utilising some standardised ways for setting up Big Data sets for FM for a better integration of the data through the project phases and processes.

- *Category III - Technological steps:* One of the valuable points made under this point is the need for developing FM specific data assessment (analysis) models. These models could be supported with distributed computer system architectures (grids) and parallel processing of Big Data sets from different organisations in the industry. Those connected systems can be used, for instance, for identifying the level of occupancy within a facility and the associated energy consumption that is relevant to such occupancy and how certain intelligent adjustments could be made accordingly.

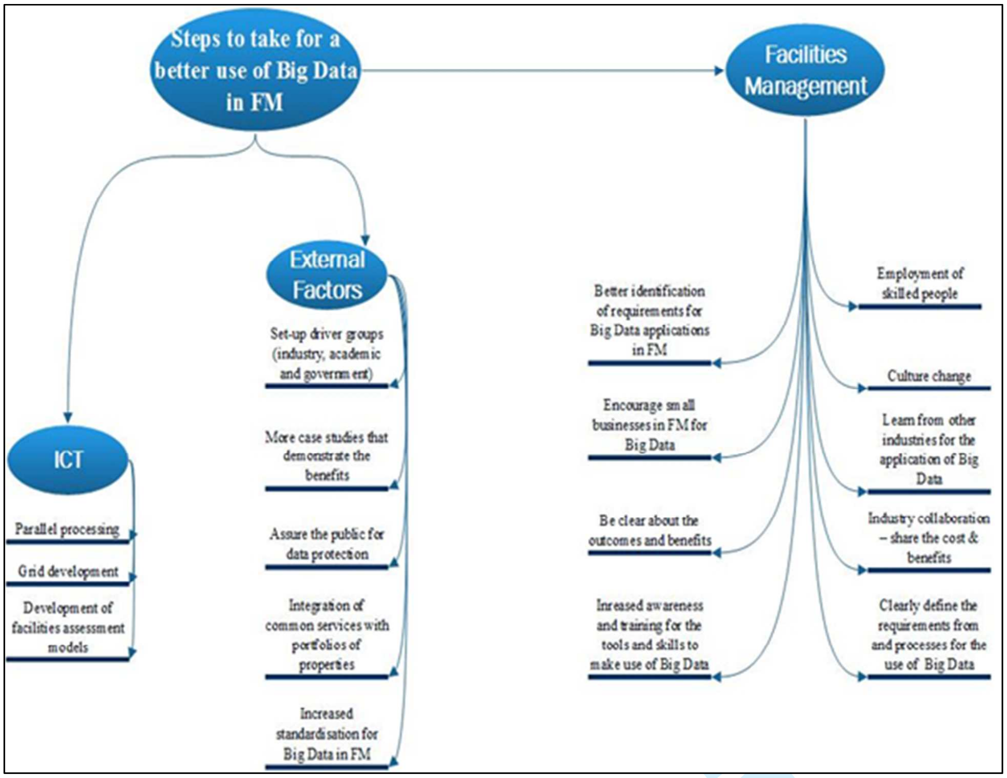

Fig 10. Steps the FM discipline should take for a better use of Big Data

The results therefore call for a genuine need for cultural changes in the way the FM discipline operates in terms of data sharing and handling, with a need for a regulatory body or a focus group that can drive this change and bring about new models of data analysis and processing.

Discussions around Big Data in FM

Following the evaluation of the displayed questions, the attendees engaged in open ended discussions that evolved around these challenges, drivers and opportunities, reaching some important conclusions as shown below;

- The majority of the attendees think that the Big Data concept is something new to the FM sector.
- Information and data should be treated as assets.
- The role of big players in disseminating Big Data in FM was highlighted.
- The attendees' expectation for the timeframe in which FM organisations can make use of Big Data vary greatly from within 12 months to more than 3 years.

- The attendees think that gathering accurate and comprehensive data is a challenge.
- The attendees strongly agree that Big Data offers significant opportunities to enhance FM processes.
- The attendees underlined that Big Data would require considerable IT investments and IT expertise.
- It was discussed that the scope of this investment should have covered the technology, business process and human resources aspects to Big Data.
- The target should be to obtain a real value from Big Data.
- Despite the considerable IT investment, the attendees questioned and were somewhat suspicious of the reliability of the outcomes from Big Data (What if Big Data conclusions are wrong or irrelevant?).
- The need for a tool or system for the measurement of the return of investment on Big Data in FM was underlined.
- The statement Big Data will only be valuable when all parties and stakeholders are open and trusting enough to share data was strongly agreed.
- The attendees found an imbalance between data collection and data analysis. Large amount of data can be collected in various ways; however, the analysis of this data for value is lagging behind.
- Large organisations in the FM sector can lead the way for Big Data and take their supply chains with them.
- Data sharing among organisations is a problem. The current contractor-subcontractor relationship structure can be a barrier for Big Data.
- Some ethical concerns and data privacy issues were discussed.
- The benefits of the synergy between Big Data and FM should be better illustrated to the industry with case studies, industry groups and best practices.
- The positive relationship between FM and visualisation technologies was affirmed.

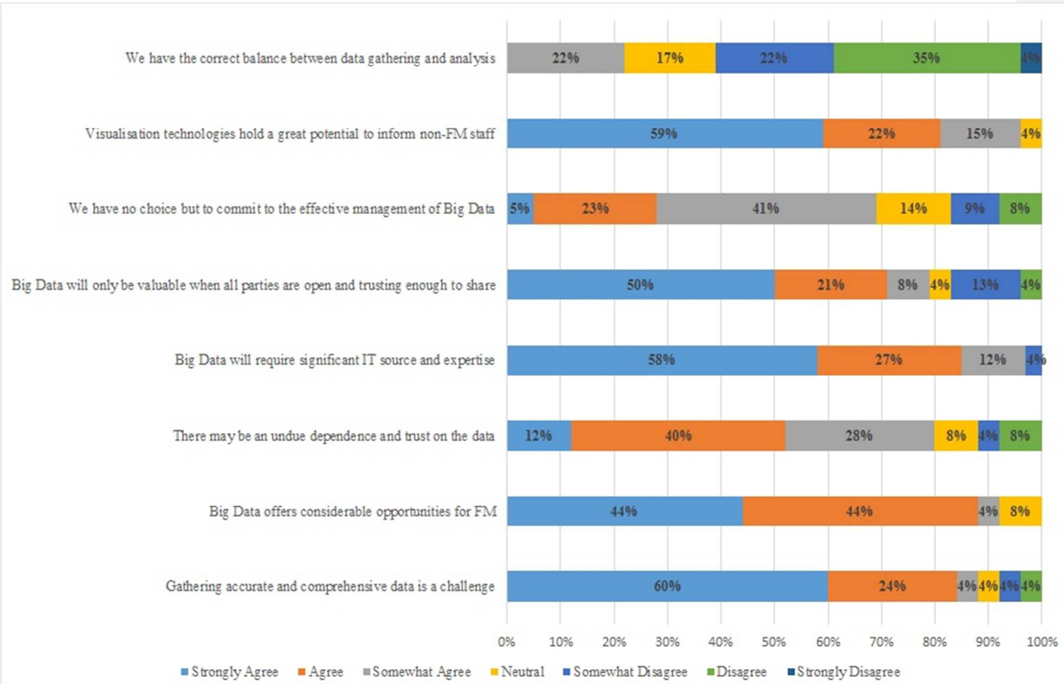

Fig 11. Evaluation of Big Data in FM

Comment [MOU10]: The location of the figure was changed as per Reviewer 1's 8th comment.

The participants were also asked to evaluate some Big Data related statements based on a Likert scale, ranging from "Strongly Agree" to "Strongly Disagree". The participants strongly agreed that gathering accurate and comprehensive data within the FM sector is a significant challenge. On the other hand, according to the participants, alongside visualisation technologies, Big Data offers many opportunities for the FM sector. The participants also thought that Big Data will only be valuable when all parties are open and trusting enough to share data; even so, it will still require significant IT source and expertise. The findings also highlighted the current imbalance between data gathering and data analysis in the FM sector. The complete survey findings can be seen in Figure 11.

Discussion and conclusion

This paper takes an overall look at the data mining and Big Data concepts with an emphasis on FM in the AEC industry, which was found missing in the current literature. Firstly, the identified BA opportunities for the industry to empirically record their benefits or challenges can be applied in future research efforts. Secondly, further discussions on how to overcome those identified barriers, some of which are more generic and some of which are more AEC specific, are needed from an AEC industry perspective. Thirdly, this paper focuses more on the FM side of BA. Thus, more in-depth analysis on the use of data mining and Big Data for the others phases in the asset life-cycle, such as design and construction, can be addressed in future research.

Comment [MOU11]: To address Reviewer 1's 9th comment.

The AEC industry essentially operates in a data rich business context and those business data are becoming bigger. Therefore, with developing technologies and analytic techniques, data mining and Big Data opportunities hold great potentials to add value to the industry's asset-life cycle (end products/"smart" assets) and business processes. In many other industries, those BA concepts have already become a part of their core business efforts with recorded benefits. By shifting from the current intuitive decision making culture to Data Driven Decision (DDD) making, in order to better exploit the opportunities, those BA concepts should be truly perceived as business competencies by the AEC industry. However, from the findings in this paper, it can be argued that the value in data analytics and Big Data is perceived by the industry; yet the industry needs guidance and leadership for an established DDD culture. Another important finding is the recognition of the imbalance between data capturing and data analytics, which indicates that the industry needs to make a better use of their data in hand. Also, there is a degree of scepticism as to whether the necessary investment for Big Data, both technology-wise and training-wise, can be justified through any significant return on investment. For those concepts to be better utilised in the industry's day-to-day business operations, end products and decision makings and to overcome those doubts, they should be supported by benchmarking efforts from other industries, best practice examples, knowledge transfer partnerships between academia and the industry, community of practice organisations as the concepts' champions, awareness increasing events and publications, and governmental guidance and impetus. These efforts are particularly relevant to disseminating Big Data, as an emerging concept, in FM. Some challenges captured for Big Data are also very much in-line with the generic issues identified from the literature on data management in the AEC sector, such as fragmented data structures, skill shortages and the lack of the business case for the necessary Big Data investment. However, the study also revealed some more Big Data specific challenges, such as the need for parallel computational processing, and ethical and legislative issues.

Comment [MOU12]: For Reviewer 2 comments to compare and contrast the findings against the existing literature.

As for the ICT side, large vendors' developing AEC industry focused analytics solutions and better interoperability among different vendors will contribute to the BA culture in the industry. Also, those vendors can direct their attention more to the AEC industry with publications, workshops and training events to raise awareness and clarify confusions.

The general concerns for Big Data analytics (i.e. data privacy, data protection, data bias etc.) mostly apply to the AEC industry as well. Additionally however, the industry suffers from a structural fragmentation with many small-sized companies operating in its supply chains. The data generally come from those smaller fragments. This fragmentation and the industry characteristics can lead to a significant barrier before an extensive penetration of BA. In this regard, as underlined in the findings, the leadership of large organisations and clients in the industry can be of vital importance in unifying those data resources and promoting a data sharing culture among their supply chains. Expecting small-medium sized companies to embark on extensive BA efforts in short-term does not seem viable. However, some sort of a mutually beneficial data-analytics alliance between large organisations and smaller organisations in the industry, in which the analytics from the larger organisations on the data provided by the smaller organisations, can be formed.

In parallel with the data fragmentation issue, BIM can present a collaborative, unified data repository for data sharing in the AEC industry. As the use of BIM has been extending through the asset life-cycle, from design to demolition, the integration of data analytics with BIM models can yield valuable business insights and help to overcome the fragmentation problem. Also, BIM models can present a data visualisation platform for data mining and Big Data analytics. Nevertheless, it should be underlined that this integration will bring along further complexities in terms of ICT integration and compatibility, business analytics and training needs for the industry's professionals. Although the BIM concept presents many data manipulation opportunities (e.g. simulation, ERP integration, automated safety and regulatory code checks etc.), those opportunities also considerably increase the amount of data the industry needs to cope with.

Similar to the government's BIM mandate (i.e. BIM Level 2 to be achieved in public projects by 2016), which has helped the BIM technology to disseminate in the AEC industry in the UK, the government can also gradually lead the direction to data mining and Big Data in the AEC industry. This will contribute to the industry's catching-up with other industries in terms of BA. Also, the data protection act may need to be discussed, reviewed and modified in cooperation with industry bodies to allow BA to disseminate further in the AEC industry.

References

Amin, S. M. and Wollenberg, B. F. (2005), "Toward a smart grid: power delivery for the 21st century", *Power and Energy Magazine*, Vol. 3 No. 5, pp. 34-41.

Bakis, N., Aouad, G. and Kagioglou, M. (2007), "Towards distributed product data sharing environments—progress so far and future challenges", *Automation in Construction*, Vol. 16 No. 5, pp. 586-595.

Boyd, D. and Crawford, K. (2012), "Critical questions for big data: provocations for a cultural, technological, and scholarly phenomenon", *Information, Communication & Society*, Vol. 15 No. 5, pp. 662-679.

Brown, B., Chui, M. and Manyika, J. (2011), "Are you ready for the era of 'Big Data'", *McKinsey Quarterly*, available at: http://www.mckinsey.com/business-functions/strategy-and-corporate-finance/our-insights/are-you-ready-for-the-era-of-big-data (last accessed 29 February 2016).

Chappel, J. (2014), "Big data creates predictive maintenance opportunities", *Power Grid International*, available at: http://www.elp.com/articles/powergrid_international/print/volume-19/issue-4/features/big-data-creates-predictive-maintenance-opportunities.html (last accessed 29 February 2016).

Chen, H., Chiang, R. H. and Storey, V. C. (2012), "Business intelligence and analytics: from big data to big impact", *MIS Quarterly*, Vol. 36 No. 4, pp. 1165-1188.

Chen, M., Mao, S. and Liu, Y. (2014), "Big data: A survey", *Mobile Networks and Applications*, Vol. 19 No. 2, pp. 171-209.

Chen, Y. and Kamara, J. M. (2011), “A framework for using mobile computing for information management on construction sites”, *Automation in Construction*, Vol. 20 No. 7, pp. 776-788.

Chui,M., Loffler, M. and Roberts, R. (2010), “The internet of things”, *McKinsey Quarterly*, Vol. 2, pp. 1–9.

Davenport, T. H. and Dyché, J. (2013), “Big Data in big companies”, *International Institute for Analytics*, available at: http://www.sas.com/resources/asset/Big-Data-in-Big-Companies.pdf (last accessed 29 February 2016).

Ergen, E., Akinci, B. and Sacks, R. (2007), “Life-cycle data management of engineered-to-order components using radio frequency identification”, *Advanced Engineering Informatics*, Vol. 21 No. 4, pp. 356-366.

Gantz, J. and Reinsel, D. (2011), Extracting value from chaos*, IDC Iview*, Vol. 1142, pp. 1-12.

Gartner (2014), "Survey Analysis: Big Data Investment Grows but Deployments Remain Scarce in 2014", available at http://www.gartner.com/document/2841519 (last accessed 29 February 2016).

Ginsberg, J., Mohebbi, M. H., Patel, R. S., Brammer, L., Smolinski, M. S. and Brilliant, L. (2009), “Detecting influenza epidemics using search engine query data”, *Nature*, Vol. 457 No. 7232, pp. 1012-1014.

Hardin, B. and McCool, D. (2015), *BIM and Construction Management: Proven Tools, Methods, and Workflows*, John Wiley & Sons, Indianapolis, USA.

Hollands, R. G. (2008), “Will the real smart city please stand up?”, *City*, Vol. 12 No. 3, pp. 303–320.

IBM (2012), *What Is Big Data: Bring Big Data to the Enterprise*, available at: http://www-01.ibm.com/software/data/bigdata/ (accessed 29 February 2016).

Japkowicz, N. and Stefanowski, J. (2016), “A machine learning perspective on Big Data analysis”. In *Big Data Analysis: New Algorithms for a New Society*, Springer International Publishing, pp. 1-31.

Jiao, Y., Wang, Y., Zhang, S., Li, Y., Yang, B. and Yuan, L. (2013), “A cloud approach to unified lifecycle data management in architecture, engineering, construction and facilities management: Integrating BIMs and SN”, *Advanced Engineering Informatics*, Vol. 27 No. 2, pp. 173-188.

Kaisler, S., Armour, F., Espinosa, J. A. and Money, W. (2013), “Big data: Issues and challenges moving forward”, *In IEEE’s 46th International Conference on System Sciences (HICSS),* pp. 995-1004.

Katal, A., Wazid, M. and Goudar, R. H. (2013), “Big data: issues, challenges, tools and good practices”, *In IEEE’s Sixth International Conference on Contemporary Computing (IC3)*, pp. 404-409.

Kim, S. A., Shin, D., Choe, Y., Seibert, T. and Walz, S. P. (2012), “Integrated energy monitoring and visualization system for Smart Green City development: Designing a spatial information integrated energy monitoring model in the context of massive data management on a web based platform”, *Automation in Construction*, Vol. 22, pp. 51-59.

Kitchin, R. (2014), “The real-time city? Big data and smart urbanism”, *GeoJournal*, Vol. 79 No. 1, pp. 1-14.

Liew, M. P., & Rosenblatt J. (2003). Using Data Mining Techniques for Improving Building Life Cycle.

Manyika, J., Chiu, M., Brown, B., Bughin, J., Dobbs, R. and Roxburgh, C. (2011), Big Data: The Next Frontier for Innovation, Competition, and Productivity, *McKinsey Global Institute*, available at: http://www.mckinsey.com/~/media/McKinsey/Business%20Functions/Business%20Technology/Our%20Insights/Big%20data%20The%20next%20frontier%20for%20innovation/MGI_big_data_full_report.ashx (last accessed 29 February 2016).

McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D. J. and Barton, D. (2012), "Big data. The management revolution", *Harvard Business Review*, Vol. 90 No. 10, pp. 61-67.

McMalcolm, J. (2015), "How Big Data is transforming the construction industry", available at http://www.constructionglobal.com/equipmentit/399/How-big-data-is-transforming-the-construction-industry (last accessed 29 February 2016).

Peitgen, H. O., Jurgens, H. and Saupe, D. (1992), Chaos and Fractals: New Frontiers of Science, Springer-Verlag, New York, USA.

Pries, K. H. and Dunnigan, R. (2015), *Big Data Analytics: A Practical Guide for Managers*, CRC Press, USA.

Reffat, R. M., Gero, J. S. and Peng, W. (2006), "Improving the management of building life cycle: A data mining approach", *In CRC Research Conference*, Brisbane.

Rijmenam, M. (2015), "Big Data can help construction companies deliver projects on time", available at: https://datafloq.com/read/big-data-construction-companies-deliver-projects-t/143 (last accessed 29 February 2016).

Russom, P. (2011), *Big Data Analytics, TDWI Best Practices Report, Fourth Quarter*, pp. 1-35, avaialble at: http://www.sas.com/content/dam/SAS/en_us/doc/research2/big-data-analytics-105425.pdf (last accessed 29 February 2016).

Sagiroglu, S. and Sinanc, D. (2013), "Big data: A review", *In IEEE International Conference on Collaboration Technologies and Systems (CTS),* pp. 42-47.

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.

Song, Y., Zhou, G. and Zhu, Y. (2013), "Present status and challenges of big data processing in smart grid", *Power System Technology*, Vol. 37 No. 4, pp. 927-935.

Stimmel, C. L. (2015), *Building Smart Cities: Analytics, ICT, and Design Thinking*, CRC Press, UK.

Spithoven, A., Clarysse, B. and Knockaert, M. (2011), "Building absorptive capacity to organise inbound open innovation in traditional industries", *Technovation*, Vol. 31 No. 1, pp. 10-21.

Wu, X., Zhu, X., Wu, G. Q. and Ding, W. (2014), "Data mining with big data", *Knowledge and Data Engineering, IEEE Transactions*, Vol. 26 No. 1, pp. 97-107.

Zikopoulos, P., DeRoos, D., Parasuraman, K., Deutsch, T., Corrigan, D. and Giles, J. (2013), *Harness the Power of Big Data*, McGraw-Hill, USA.