Big Data in Construction Management Research

Abstract

The paper is a literature review on Big Data in project management of construction projects. The literature shows practical examples of use and potential use of Big Data in construction management research. Big Data has become common as a business term in most businesses. However there is little published management scholarship that tackles the challenges of using such tools, or even that explores the opportunities for new theories and practices that Big Data might bring about. There is a need for broader discussions of Big Data in society and its implications for construction management research.

The construction process can be studied in a number of dimensions. We structure documented and potential use of Big Data related to construction projects in the time perspective of a typical construction project, from concept preparation and brief, through design and construction to use. The other dimension is discussed.

We have identified studies describing Big Data applications and theory. Thematically they fall into three broad categories: 1) New construction equipment; generating, sharing and storing data about use. 2) Data from internal IT systems; such as planning, procurement and BIM can be utilised. Lastly, 3) People generate an increasing flow of information, which can be useful if handled with care. In combination this addresses the life-cycle from concept to decommissioning. We find that the construction phase appear to have received most attention from researches. We also find that several studies are applicable to more than one phase of a construction project. We find a potential for increased use of Big Data methods and applications within construction. While some data and applications have been analysed in isolation previously, there is a potential to combine different types of data.

Keywords: Big Data, Construction management, Construction projects
1. Introduction

Big Data has become common as a business term in most businesses. Olsson et al. (2015) discuss how Big Data can be applied in project evaluation and in project management research. However, there is little published management literature that describes how to tackle the challenges of using such tools, or even explores the opportunities for new theories and practices that Big Data might bring about. There is a need for broader discussions of Big Data in society and its implications for construction management research.

1.1 More quantitative data available

On a worldwide basis, the total amount of digital data created and replicated each year is expected to increase exponentially up to 2020. This is also the case for data that can be applied in project and management. The principles used in Big Data can also be applied on smaller quantitative data sets. This include data stored in company internal IT-systems.

The definition of Big Data is shifting as software tools become more powerful. Big Data was first defined as data sets whose sizes are too large for commonly used software tools to capture, manage and process within a tolerable elapsed time (Manyika et al., 2011, Tien, 2013, Waller and Fawcett, 2013). However, other definitions will probably be needed, as Big Data is becoming a part of commonly used software tools. The uniqueness of Big Data is the volume, velocity and variety, the three V’s (Courtney, 2012, Russom, 2011). The volume refers to the size of data sets, containing a few terabytes to many petabytes. But it is the variety and velocity of the generated data that makes the data sets so big. Variety refers to the variety of sources. In addition, the data are measured and captured in more detail, such a location, time and metadata, giving both structured and unstructured data sets (Russom, 2011, Waller and Fawcett, 2013). The velocity of data refers to the speed at which the data is generated, from being recorded, updated or measured monthly and weekly to more frequent updates such as daily, hourly or continuously (Courtney, 2012). The access to real time or almost real time information makes it possible for a company to be much more agile than its competitors (McAfee and Brynjolfsson, 2012).

Courtney (2012) mentions veracity and value when describing Big Data. Veracity refers to a quality of the data sets, while value is reference to the goal from using the data sets. Veracity is a description of how the measures are reliable, referring to the accuracy and the quality of the generated data. Value refers to turning the Big Data-sets into value for the business. Size does not need to be the only defining part of Big Data, and data can be discussed along more than the mentioned three dimensions. George et al. (2014) points out that there are discussions among practitioners that “big” is no longer the defining parameter, but rather “smart”, including a fine-grained nature of the data (George et al., 2014). The data available to companies are often unstructured Davenport et al. (2012). The sources of digital data can include retail transactions, security cameras, internally registered data in the organisation, time-stamps, GPS-tracking, sensor-data from instrumented machinery and metavalues of documents.
The main reason to carry out data analysis is to derive information from data, knowledge from information, and wisdom from knowledge. And this is the purpose of Big Data. Big Data can give new information and knowledge for decision-making. For instance, Big Data can be used to make more precise predictions, and it follows that better predictions yields better decisions (Jagadish, 2015). McAfee and Brynjolfsson (2012) found that the more companies characterized themselves as data-driven, the better they performed on objective measures of financial and operational results. More and more business activity is digitized, and new sources of information are available, (McAfee and Brynjolfsson, 2012). This also applies to the construction industry.

1.2 Use of new quantitative data in construction management

Technology has always been a part of construction management, both in research and practice, but there are new technological demands. This means that new data can support the trend towards an intelligent built environment, covering the whole lifecycle of facilities. Big data has a potential to generate new insights into the costs, designs and processes of construction management. The aim is to develop tools and approaches for intelligent, efficient and sustainable construction management. Such strategies need to be sufficiently flexible to meet requirements resulting from changes in user-demands, technology and other framework conditions, while at the same time increase efficiency. It is a potential to integrate areas such as Computer Aided Facility Management (CAFM), Building Information Modelling (BIM), and Integrated Building Control and Monitoring Systems (BCMS). Such knowledge can later be utilized in decision making support, innovation of technical systems and in the education and training of project managers.

Monitoring activity across a large, complex construction site is particularly difficult because there are so many moving parts, and because the jobs being performed change frequently. In contrast to production industry, most construction sites are also temporary by nature, often challenging the investment in production infrastructure. Several reports document (including Egan, 1998; Lo et al., 2006; Durdyev & Mbachu, 2011) that construction lags behind other industries such as manufacturing in terms of productivity, and blamed the situation on problems with planning, coordination, and communication.

Modern construction equipment also generates data through usage. Producers of equipment such as trains have for some time utilised equipment life-cycle management data which are generated in large scales through the period of production, operation, maintenance. There is broad recognition of value of data and information obtained through analysing it. This type of data is also possible to generate from construction equipment The exponential growth in this type of data means that new measures are needed for data management, analysis and accessibility.

MIT Technology Review (2015) reports on the use of drones to monitor construction progress. Once per day, drones automatically patrol the work site, collecting video footage. The footage is then converted into a three-dimensional picture of the site, which is fed into software that compares it to computerized architectural plans as well as the construction work plan showing
when each element should be finished. The software can show managers how the project is progressing, and can automatically highlight parts that may be falling behind schedule.

A discussion has emerged about use of Big Data and performance measurement for micro management and continuous monitoring of employees Using drones to monitor activity continually can be controversial. Such controversies have occurred related to monitoring of employees in other sectors. As reported by the New York Times (2015), the company Amazon, monitor employees in warehouses using sophisticated electronic systems to ensure they are packing enough boxes every hour. In a similar way, the Amazon also uses a self-reinforcing set of management, data and psychological tools to spur its tens of thousands of white-collar employees to do more and more.

1.3 Purpose and research questions

There is a need for broader discussions of Big Data in society and its implications for management research in general, including construction and project management. George et al., (2014) points out that even though “big data” has become commonplace as a business term, there is little published management scholarship that tackles the challenges of using such tools, or that explores the opportunities for new theories and practices that Big Data might bring about. This paper will give a literature review on different use and potential use of Big Data in construction management research.

The purpose of this paper is to investigate different use and potential use of Big data in construction management research. The construction process can be studied in a number of dimensions. We use structure documented and potential use of Big Data related to construction project along a time axis of a typical construction project, from concept preparation and brief, through design and construction to use. We have discussed different alternatives for another dimension to present the results. Alternative dimensions include represents different stakeholders involved in the project, type of data sources or different types of use of the data.

Our research questions addressed in this paper are:

- Which applications of big data in construction project management have been published in recent years, based on the defined literature search criteria?
- Which time phases of a construction project are recent big data research relevant to?

2. Method

We review previous research and structure the papers based on the time perspective of a typical project. Several approaches were used to identify relevant literature. Searches in the Norwegian library database (Bibsys) were conducted, covering both books and academic journals. Searches were made using several search engines on the Internet, such as Emerald, Science Direct, Wiley Online Library, and Google Scholar. During the database searches, both titles and keywords were examined. In the searches in Google Scholar the entire texts are searches, consequently providing more search results of both relevant and irrelevant papers. Exclusion keywords were
provided in the search, to exclude papers from non-relevant research fields. The collected material was subsequently examined in more detail. The main keywords used in the search were «big data», «construction project/industry», and (construction/project) management. Relevant literature was also found in the journal for automation in construction, as this is relevant technology and data source even though the papers do not use the phrase “big data”. Searches were also made with capital projects and engineering projects as search keywords, but with limited additional search results. Exclusion keywords used were psychology, ecology, biology, constructionism, biology, health and “network construction”. The overview of search keywords, databases and results is given in Table 1.

In the search results, there are several papers that are not available or do not deal with construction industry. Most of these were excluded by the excluding keywords. However, several papers are within the areas of software development, data management, construction of data centre facilities, big data project etc. but are not directly related to the construction industry. The papers where excluded by the search keywords, based on headings, abstracts, or a quick search for construction, management, and big data within the main text. Exclusion criteria were used to exclude studies that are not relevant to answer the research questions. Only papers that give a clear contribution or are of clear relevance to Big Data in construction projects or management are of interest. The papers are limited to 2014 and 2015, and Table 1 show that most research contribution has been published these last couple of years.

Table 1: Overview of search keywords, databases and results.

<table>
<thead>
<tr>
<th>Keyword</th>
<th>Search engine</th>
<th>database</th>
<th>No of papers</th>
</tr>
</thead>
<tbody>
<tr>
<td>«big data» AND construction AND management</td>
<td>Oria</td>
<td>121* (95)***</td>
<td></td>
</tr>
<tr>
<td>«big data» AND «construction project OR industry» AND management</td>
<td>Google Scholar</td>
<td>(591)** (462)***</td>
<td></td>
</tr>
<tr>
<td>«big data» AND «construction project OR industry» AND management -psychology -ecology -constructionism</td>
<td>Google Scholar</td>
<td>(418)** (319)***</td>
<td></td>
</tr>
<tr>
<td>«big data» AND construction AND «project management» -health, -ecology, -constructionism, -“network construction”, -psychology</td>
<td>Google Scholar</td>
<td>838 (581)** (553)*</td>
<td></td>
</tr>
<tr>
<td>«big data» AND AND «construction project OR industry»AND «project management» -psychology -ecology -constructionism -biology</td>
<td>Google Scholar</td>
<td>168 (142)** (103)***</td>
<td></td>
</tr>
</tbody>
</table>


The selected literature is categorised in the time perspective of a typical construction project, from concept preparation and brief, through design and construction to use.

The literature that was found to be relevant to our research questions where structured into the time perspective of a typical construction project inspired by the phases in the RIBA model, from concept preparation and brief, through design and construction to use. The RIBA model
has served as an inspiration to a recently presented Norwegian phase model for construction (bygg.no, 2015). The Norwegian model uses a time frame similar to the RIBA model, but highlights the perspectives of owner, user, supplier and society. The Norwegian model has the following phases: Strategic definition; Concept development; Concept design; Detailed design; Construction; Handover; Use and facility management and finally, Disposal. The selected literature is categorised based on this life cycle perspective.

In addition to the time perspective, we found need for adding a second dimension, and investigated a number of alternatives. We then discuss what division would potentially be suitable for the other dimension. Olsson and Bull-Berg (2015) propose a distinction based on possible data, sorted from low to high degree of structure: raw sensor data, registrations in IT system, data in IT-systems, structured data and finally summarised reports.

3. Literature

We have studied Big Data related to construction projects, and categorised them in a time perspective. The phases are based on a typical construction project inspired by the phases in the RIBA model, and the corresponding Norwegian version; from strategic definition and concept development, through design and construction to use and disposal. The selected literature is categorised based on this dimension, as shown in Table 2.

A significant portion of Big Data is geospatial data, generated from sources such as mobile devices and RFID sensors. Geospatial big data gives both opportunities and challenges, as discussed by Lee & Kang (2015). Related to the construction industry, location aware data can give useful information into urban planning, by providing information in the early project phases on how people use public spaces and infrastructure.

Caron (2015) focus on the importance of an early engagement of stakeholders to improve the forecasting/planning process and manage the project. Engagement of stakeholders from the stage of the project life cycle, in which they may be involved in or be impacted by the project, can significantly increase the available amount of both explicit and tacit knowledge. Managing stakeholders is critical in large construction projects, and in infrastructure major projects this is especially difficult due to the diversity of stakeholders and interests. A study by Bakht et al. (2014) evaluate stakeholders’ impact on infrastructure megaprojects through analysis of big data captured from online social media.

During the project phases, from early preparation to handover and the liability period, a project produces a lot of documents. Kähkönen et al. (2015) study content and characteristics of current practice of Electronic data/document management systems (EDMS) in building construction projects. This includes project cost, number of files and file accessing actions (open, view, download). They looked at 15 building development projects, which used the same data system. Collection of new data and analytical approaches has potential to develop new insights in project management maturity, as examined by Williams et al. (2014). In particular two big data analytical techniques is highlighted to have potential to develop understanding of maturity in
organisations. These are social network analysis and text analysis. The study by Whyte (2015) analysed change management practices in three separate organisations, including a large construction project, who all deliver complex projects, rely on digital technologies to manage large data-sets, and use configuration management to establish and maintain integrity. The organisations use for instance configuration management in the concept, product identification and definition stages, and as control at the “as-built” stage before handover. The authors conclude that the unstructured, uncontrolled nature of big data presents challenges to complex projects that deliver assets. Martinez-Rojas et al. (2015) demonstrate that suitable data handling facilitates improves the decision-making process and helps to carry out successful project management. The main information and communication technologies were analyses, and proposals that exist in the literature focused on the management of information and knowledge from a general point of view in the field of project management were reviewed.

Barista (2014) presents some early successful applications of data-driven design and planning applications, and how Building Teams can benefit from this. For instance, designers can capture and analyse data from key building performance metrics, such as energy use intensity, to optimize early prototypes. Based on feedback data from building occupants, firms can evaluate design concepts against the real world, and help the building team to understand how people interact with spaces. Redmond et al. (2015) studied how social network analysis and energy usage analyses can be a source to create integrated models for green building design. The main objective was to highlight green building technologies, while at the same time engaging end-users and harnessing their collective knowledge in building design. Several papers also treat opportunities and challenges in combining BIM and Big Data (e.g. Chen et al., 2015).

Several papers cover the construction phase. Akhavian et al. (2015) investigate the prospect of using built-in smartphone sensors as data collection and transmission nodes in order to detect detailed construction equipment activities. The method demonstrated a perfect success in recognizing the engine off, idle, and busy states of construction equipment. The work by Teizer (2015) outlines early results for vision-based sensing technology for tracking of temporary assets on infrastructure construction sites. Research and practical industry applications demonstrate promising work towards automated visual recording and progressing of temporary construction resources. Guo et al. (2015) show that big data can be used for behaviour observation in China metro construction. The suggested framework was verified in an example to be able to analyse semantic information contained in images effectively, extract worker's unsafe behaviour knowledge automatically.

Automation is field in construction that uses new data sources and technologies. M2M installed on construction machines could be used to recommend overhauls to end users at the optimum timing according to Vanzulli et al, 2014. To track the progress of earthwork processes at underground construction sites, Bügler et al. (2014) suggests a novel method that combines two technologies based on computer vision, photogrammetry and video analysis. Combining these data sources allows exact measurement of the productivity of the machinery and determining site-specific performance factors. The construction quality of the material roller-compacted concrete, used in construction of storehouse surfaces, is affected by factors such as the roller.
compaction, concrete temperature and construction climate. Liu et al. (2015) proposes a real-time construction quality monitoring method, to provide the construction operations on site with timely collection and comprehensive analysis of construction data from the construction process. Wang et al. (2015) presents a method for automatic object recognition and rapid 3D surface modelling, including point cloud data collected from a construction jobsite. Yang et al. (2015) reviews state-of-the-art vision-based construction performance monitoring methods. According to Skibniewski et al. (2015), the construction and operation of infrastructure systems have opportunities for improvement through research on robotics and automation. Automated equipment could cut waste, improve job safety and the overall quality of construction projects. Performance monitoring can be made more effective with tools that better characterize the extent to which construction plan are being followed and the extent to which workers and equipment are being fully utilized.

Table 2: Categorisation of literature in the time perspective of a construction project.

<table>
<thead>
<tr>
<th>Strategic definition</th>
<th>Concept dev.</th>
<th>Concept design</th>
<th>Detailed design</th>
<th>Construction</th>
<th>Hand-over</th>
<th>Use</th>
<th>Disposal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lee et al. (2015)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caron (2015), Bakht et al. (2014)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kähkönen et al. (2015)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Redmond et al. (2015)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Williams et al. (2014), Whyte (2015)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Barista (2014)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chen et al. (2015)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skibniewski et al. (2015)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lu et al. (2015)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Lu et al. (2015) study the performance of construction waste management for the construction categories building, civil engineering, demolition, foundation, and maintenance and renovation. The study develops a set of KPIs/WGRs (waste generation rate) using an available big data set...
on construction waste management in Hong Kong. Demolition is found to be the most wasteful works that generate both non-inert and inert construction waste. There have been published several examples of evaluation of buildings in use, using quantitative methods based on Big Data approaches. The study by Hong et al. (2015) reviews the state-of-the-art in the major phases for a building's dynamic energy performance, focusing on the operation and maintenance phase. Ioannidis (2015) presents Big Data and visual analytics techniques for comparing building performance under different scenarios and design. Data that provide useful information is energy consumption, building geometry, and space occupancy. Isikdag (2015) provides a method for facilitating the GIS based fusion of information residing in digital building “Models” and information acquired from the city objects. The virtual BIM sensors in the proposed design pattern will provide geometric and semantic information together with information related to the state of the building elements, and the information can be used to represent the building within a GIS environment, and city monitoring/management.

According to Naderpajouh et al. (2015), effective frameworks to facilitate data-driven decision-making are noticeably lacking in the construction industry. The developed hierarchical definition for health of the construction industry was used to propose a framework to benchmark, interpret and analyse data associated with the status of the health of the industry.

4. Concluding discussion

Our first research question was related to which applications of big data in construction project management that have been published in recent years, based on the defined literature search criteria. We have identified studies describing Big Data applications and theory. Thematically they fall into three broad categories: 1) New construction equipment; generating, sharing and storing data about use. 2) Data from internal IT systems; such as planning, procurement and BIM can be utilised. Lastly, 3) People generate an increasing flow of information, which can be useful if handled with care. In combination this addresses the life-cycle from concept to decommissioning. We find that Big Data have been used in energy management. Data from internal IT systems, such as planning, procurement and BIM can be utilised. People generate an information flow, which can be useful but also treated with care to safeguard personal integrity. HVAC and electricity management systems generate large volumes of data that can be applied for life-cycle management of the equipment, but also for describing the use of the building. None of the capabilities described in this study are entirely novel in inception nor unique to the construction industry, producers of major equipment such as train rolling stock have for some time utilised equipment life-cycle management data, which are generated in large scales through the period of production, operation, maintenance.

Our second research question was which time phases of a construction project that recent big data research is relevant to. The literature is presented based on a project phase perspective. We find that the construction phase appear to have received most attention from researches. We also find that several studies are applicable to more than one phase of a construction project.
We find a potential for increased use of Big Data methods and applications within construction. While some data and applications have been analysed in isolation previously, there is a potential to combine different types of data. Typically, the power of data integration is hard to demonstrate in limited and small pilots, but requires critical mass before providing return. Big Data can be used to make more precise predictions, and it follows that better predictions yields better decisions. While data related to different engineering disciplines, such as energy, have been analysed in isolation previously, there is a potential to combine different types of data. This creates opportunities, but also challenges related to personal privacy. Big data appear to have a potential to generate new insights into the costs, designs and processes in project management. The aim is to develop tools and approaches for intelligent, efficient and sustainable construction project management.

References


