Eighth International Conference on Construction in the 21st Century (CITC-8)
"Changing the Field: Recent Developments for the Future of Engineering & Construction"
May 27-30, 2015, Thessaloniki, Greece.

Big Data in the Construction Project Management

Ogbonna A.Chibueze ph.d
Department of Computer Science, Babcock University, Nigeria.
ogbonnaac@babcock.edu.ng

Komolafe Oyindolapo
Department of Computer Science, Babcock University, Nigeria.
oyindolapokomolafe@yahoo.com

Abstract

The construction and engineering industry generates vast amount of data defined by dimension and attributes. Construction projects are associated with the collection, processing, and exchange of large amounts of data among the stakeholders. The larger the size of the project, the larger the amount of data associated with the project. To generate information from the data, proper predictive analytics and data visualization tools are required – a practice that is not common. The application of big data in construction helps in decision making, resource efficiency improvements, risk management and reporting, financial and cost management and ideas generation. Big data can be applied to every aspect of construction including project scheduling, costing and management etc. The resources required for project scheduling are derived from data. Guidelines for incorporating big data into the construction industry is also analyzed. Therefore, Big data in the construction industry is a business opportunity that needs to be explored

Keywords: Big data, Construction, Data Visualization, project management

1. Introduction

Electronic information has become a major and indispensable part of business in the world. We are surrounded by a mass of data because every second electronic devices and systems are creating, receiving, processing and transmitting information. Data is not generated from a vacuum, it definitely has a source. Data can be of various forms such as sounds, videos, files, barcodes, pictures etc. data can be divided into categories irrespective of its source: structured data, unstructured and semi structure.

- Structured data: this refers to information that's highly organized. The computer can easily search and organize it based on many criteria with the barcode information as an example
- Unstructured data: the information does not have a pre-defined model or do not fit perfectly well in a spreadsheet as exampled by video and audio files
- Semi structured data: it sits between structured data and unstructured data. Example is email.

Data can be of different flavors of qualitative data which is normally descriptive and often subjective and quantitative which is numeric and can either be continuous or discrete. projects by their nature deal with

massive amounts of data based on which projects are managed and subsequently measured for performance. Hence the focus of this paper.

2. Background

In recent years, there has been an unprecedented increase in the quantity and variety of data generated worldwide. The world's information is predicted to reach 40ZB by 2020 (Digital Universe, 2012). We produce around 2,200 petabytes of data every day. This includes 2 million of searches produced by google each minute, 4000 hours of video uploaded to YouTube every hour and 144 billion sent around the world every day (Garside & Fox, 2013). This increase in data, often referred to as a "data tsunami", is driven by the proliferation of social media along with an increase in mobile and networked devices, finance and online retail as well as advances in the physical and life sciences sectors (Duque & O'Driscoll, 2014).

Big Data is a phenomenon defined by the rapid acceleration in the expanding volume of high velocity, complex, and diverse types of data. Big data can be defined by its dimension and attributes.

The dimensions of data include

Volume: refers to the vast amounts of data generated every second. Scale is certainly a part of what makes Big Data big. The internet-mobile revolution, bringing with it a torrent of social media updates, sensor data from devices and an explosion of e-commerce, means that every industry is swamped with data- which can be incredibly valuable.
Velocity: refers to the speed at which new data is generated and the speed at which data moves around. It is the increasing speed at which this data is created, and the increasing speed at which the data can be processed, stored and analyzed by relational databases
Variety: Big Data data is not just numbers, dates, and strings. It also includes geospatial data, 3D data, audio and video, and unstructured text, including log files and social media.

The attributes of data

Visualization: Once data has been processed, one needs to present the data in a manner that's readable and accessible. Hence satisfying the attribute of visualization
Validity: data is generated every second in large volumes, and to acquire meaningful information from the data within the period when the data is useful is referred to as validity.

Variability: Variability refers to data whose meaning is constantly changing for different reasons.
Veracity: refers to the messiness or trustworthiness of the data. A term that is similar to the concept of relaibility

Versatility: this refers to how the data can be used and underlying capabilities of the data (Buchholtz & et al, 2014; TechAmerica Foundation, 2012)

3. Big data in construction

The construction industry with its numerous projects generates vast amounts of data and complex information. Effectively managing the complex and enormous data to insure its accuracy and availability in an apt method is crucial to the successful and effective completion of any project. Inaccuracies or holes in information can readily lead to project delays, uneconomical decisions, or even the complete failure of a project. It is for these reasons, therefore, that the efficient collection of onsite construction data and the timely communication of that data is such a critical concern to all members of the construction industry (Pmbook, 2002; Cox & et al, 2002; Shrestha & Jeong, 2013). Even small construction projects have

gigabytes of data for design, estimation, and construction. Large projects are associated with enormous amount of data and in other to produce valuable information from the "big data", proper predictive analytics and data visualization tools are required. Big data analytics allows discovery of knowledge through data (Wactlar, 2012).

|The opportunity for big data in construction is limitless and represents the ability to get business insight from information within the business, as well as from outside sources to compare against similar-sized contractors, Contractors might want comparisons of project profitability, days outstanding of accounts receivable, cash flow and other financial metrics to similar-sized contractors (Kuehner-Hebert, 2014). In the construction industry, the more you know and the more effectively you can process the data you have, the bigger advantage you will have in the marketplace. Technology in the construction industry offers designers, builders, suppliers and manufacturers the ability to make accurate predictions, precisely control business processes, and find opportunities for business that otherwise might have remained hidden (Mineer, 2014).

Construction data can be in digital and manual format and they can occur in various forms as:

Various software programs for planning, designing, scheduling, estimating, bidding, etc;
2D paper drawings to 3D, 4D, and 5D models;
Mechanization and digitization of tools and equipment;
Use of tablets in construction site;
Chronological files of project correspondence and memorandums, including RFIs;
Change order and submittal requests;
Quality control and assurance records;
Construction field activity and progress logs;
Resource and inventory logs, including tracking of labours, equipment, and materials; and
legal contracts and regulatory documents (Shrestha & Jeong, 2013; Cox & et al, 2002; Pmbook, 2002)

4. Impact of big data in construction

In the field of construction projects, big data effects can be felt in the areas of decision making, resource efficiency, risk management reporting, financial and cost management.

Decision making: Big data is tremendously valuable to companies looking to solve their immediate decision-making challenges. Delivering the right data to the right decision maker in a way they can understand - no matter where that data lives or what form it takes - is crucial, and companies need to make this easier in order to get maximum value from big data.

Resource efficiency improvements: big data helps to make changes on-screen to see how they'll affect the entire structure, reducing errors on the construction site and it helps to provide detailed pre-construction analysis that helps prevent problems such as delays, cost overruns and material waste. It also helps to assess and predict energy usage patterns, power requirements, and benefits from using green technology or high-efficiency equipment. It enables the generation of detailed 3D maps of building sites that aid pre-construction planning and allow for clear visualization of the placement and effect of structures (Mineer, 2014).

Risk management and Reporting: Big data enable market events across the globe to be captured in real time via sources such as news, research, graphs, news, social media, visuals etc. this type of information can be useful in risk mitigation decisions. Data from supporting documents, contracts, daily activities etc. can also be used, thereby enabling better risk management

Financial and Cost Management: the use of Big data technology helps to analyze data for pattern more quickly at a much lower cost. Advancement in technology which includes processing power, data warehousing enables speedy organization of data. This reduces the time taken to search since the data is indexed. The emergence of techniques and tools enables the organization to analyze large store of data easily and quickly, using fewer resources and reducing cost.

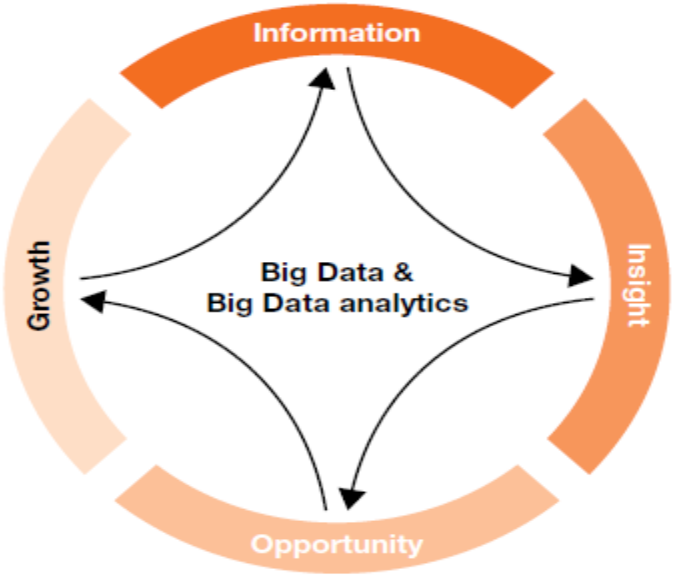

Figure 1: Value Chain of Big data and Big data Analytics. Source PwC (2014)

Ideas Generation: Big data serves as a platform for opportunity. It gives room to conceive and develop ideas. Once the idea is determined to be valuable, it can be put into production. Consumers' data can be harvested from blogs, social media, mobile devices etc. to gain intelligence on customer's changing perceptions, needs, expectation and behavior to new ideas and products. The information received can be channeled into predictive models (PwC, 2014).

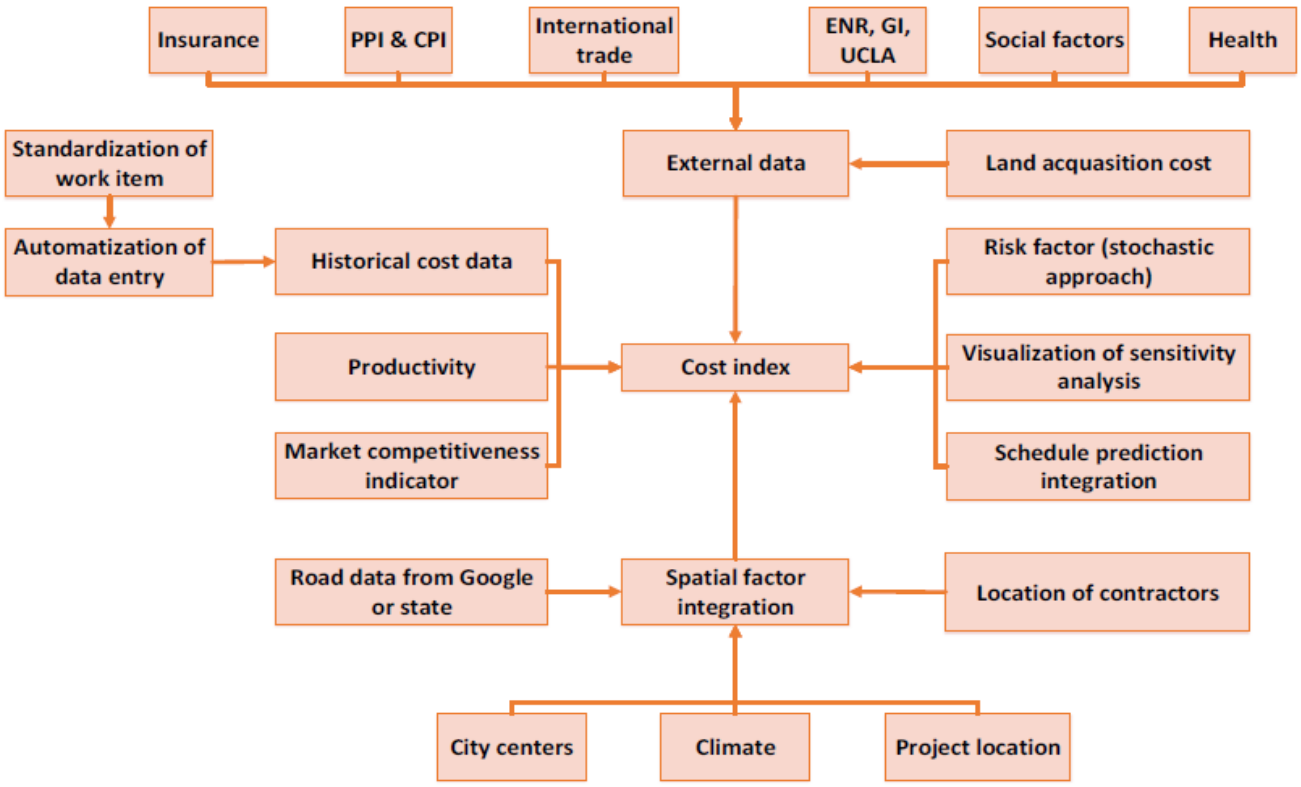

Figure 2: Construction Cost Index, an Application Example of Big data Source Shrestha & Jeong (2013)

5. Big data and project schedule in mega projects

Construction mega projects consist of vast amount of data and managing the data can be strenuous if the proper tools are not available. Scheduling tools make use of updated schedule data to perform schedule network analysis. The big data can be used in Earned Value Management (EVM). EVM is a method of measuring project progress by comparing the actual schedule and cost performance against the planned performance.

$$EVM = f(PV, EV, AC)$$ ---------- Equation 1

Where: PV = Planned Value (budgeted portion of the approved cost estimate)

EV = Earned Value which indicates the value of work actually performed during a particular time period

AC = Actual Cost which indicates the actual money that has been spent for work that has been completed.

Therefore from the above function, the following can be deduced from the data: Schedule Variance (SV), Schedule Performance Index (SPI), SPI Trend Analysis, Cost Variance and Cost Performance Index (CompTIA Project, 2009).

6. Using big data

The use of Big datais often considered a challenge to many project stakeholders, hence the need for some guidelines. In order to successfully use big data the following needs to be understood;

Recognizing that Big data is not a problem but a business opportunity ;

Plan on leveraging Big Data by defining business statement that describe how Big data will be used to create new capabilities and improve existing ones;

Training current employees and employing employees with knowledge on Big Data;
Education on both the value and how-to of making Big Data driven, fact based business decisions
Preparing the environment by acquiring the necessary infrastructures;
Collection of data through internal sources, Crowdsourcing and Third parties;
Learn how to combine third-party data with internal data assets for decision making;
Processing and analyzing the data received;
Visualizing the information processed in a more presentable manner;
making decision based on the analysis and visualization and;
Applying the decision made. (Shrestha & Jeong, 2013; PwC, 2014).

7. Conclusion

Construction and Engineering industry generates voluminous and complex data. The data can be analyzed for prediction and decision making. Big data analytics and visualization tools can be applied to construction to generate information from the data. Without proper tool to analyze the data, the data is useless. The larger the data, the more accurate the predictions.

REFERENCES

Buchholtz S, Bukowski M. and Śniegocki A. (2014): Big and open data in Europe. A growth engine or a missed opportunity? The Warsaw Institute for Economic Studies (WISE Institute)

CompTIA Project (2009): Monitoring and Controlling Projects Constraints. CompTIA Project Certification (2009 objectives).

Cooper S. (2014): Are asset owners and the construction industry really ready for 'Big Data'? Accessed from http://www.conjectblog.com/are-asset-owners-and-the-construction-industry-really-ready-for-big-data/ Accessed on January 13, 2015

Cox, S., Perdomo, J., & Thabet, W. (2002). Construction field data inspection using pocket PC technology. International Council for Research and Innovation in Building and Construction CIB w78 conference Aarhus School of Architecture. Conference Proceedings – distributing knowledge in building. Retrieved from http://itc.scix.net/data/works/att/w78-2002-69.content.pdf Accessed on January 13, 2015

Digital Universe Study (on behalf of EMC Corporation) (2012) Big Data, Bigger Digital Shadows, and Biggest Growth in the Far East. In: http://idcdocserv.com/1414

Duque Barrachina and O'Driscoll (2014): A big data methodology for categorising technical support requests using Hadoop and Mahout. Journal of Big Data 2014 1:1.

Garside W., and Cox B (2013): Exploring the world of data, Big data Storage for Dummies. A wiley brand. 4-5

Kuehner-Hebert K (2014): Using Big Data to Predict a Contractor's Next Move. Construction's Crystal Ball retrieved from http://www.constructormagazine.com/constructions-crystal-ball/#.VLOy3Nso80M Accessed on January 13, 2015

Mineer D. (2014): How Technology and Big Data Has Improved the Construction Industry. Retrieved from http://blog.constructionmonitor.com/2014/11/17/technology-and-big-data-has-improved-the-construction-industry/

Pmbook (2002): Organization and Use of Project Information. Retrieved from http://www.ce.cmu.edu/~cth/pmbook/14_Organization_and_Use_of_Project_Information.html Accessed on January 13, 2015

PwC (2014): "fs viewpoint, where have you been all my life? How the financial services industry can unlock the value in Big Data". Retrieved from http://www.pwc.com/fsi

Shrestha J. & Jeong D(2013) Big Data, Predictive Analytics, and Data Visualization in The Construction Engineering, Data Analytics Laboratory for Project and Infrastructure Management (DALPIM), IOWA State university retrieved from http://works.bepress.com/joseph_shrestha/2 Accessed on January 13, 2015

TechAmerica Foundation (2012): Demystifying Big Data: A Practical Guide to Transforming the Business of Government. Prepared by TechAmerica Foundation's Federal Big Data Commission

Wactlar H. (2012): Big Data R&D Initiative. CISE Directorate National Science Foundation. Retrieved from http://www.nist.gov/itl/ssd/is/upload/NIST-BD-Platforms-05-Big-Data-Wactlar-slides.pdf Accessed on January 13, 2015