

Big Data in Self Evaluating Construction Domain Using EOC Indices

J Lokesh^{1*}, S Sakthivel² and T Rajesh³

^{1,2,3}Department of Computer Science, GTEC, Vellore, Tamil Nadu, India

www.ijcseonline.org

Received: Feb/11/2015

Revised: Feb/22/2015

Accepted: Mar/17/2015

Published: Mar/31/2015

Abstract— With the growth of technologies on one side the challenges increased gradually on other sides. When a new technology is introduced there will be a shortage of resource talents in understanding the insight view to take benefits of one from that. There are some very serious challenges the construction industries facing that are motivating new approaches to how we design, operate, and maintain buildings and infrastructure. The new technologies are designed to address challenges in the construction industries especially in both operational and maintenance sectors. Big data is a tool for transformation of manual to automated process uses vast size of data or information that it exceeds the capacity of traditional data management technologies. Inadequate insight knowledge (IK) about generating large datasets is one of the most important constraints in any organizations. This paper focus mainly on construction domain rather than other domains including retail, health care, manufacturing, finance and housing, etc. and their importance of big data technology on construction sites and self evaluating their needs to reduce/mange risks, high returns, intensity, knowledge adequacy using two indices: potential values and ease of capture(EOC).

Keywords— Insight Knowledge; Potential Values;Ease of Capture

I. INTRODUCTION

As the building companies is digitized rapidly. The construction sector has lagging among other domains both in improvements and use of new modern IT innovations. Every one believe the slow improvements in any business sectors or in other domains are cause of cultural and also any improvements must be driven by individuals manually, and mandated by government regulations[5]. The construction companies at most part have failed in adoption of standardized best practices. Such types of sector can't easily pack up and move if their management strategies fail. Big Data is a disruptive force causing many companies to totally rethink their strategies for leveraging data across the enterprise, and innovative companies are powering their analytic and operational processes with real-time data.



As with most paradigm shifts, there are many challenge that are unstoppable, causing acceleration towards changes and adoption. This is true especially where big data and real time are concerned. one of the greatest challenges with managing the project correctly is keep the updated, correct information from several construction sites and distribution of information's to their parties at time[2]. Leveraging the vast amount of information can be a great benefit to the construction industries.

The benefits of construction companies are vast. Fortunately, there are already construction companies that are working on a big data strategy make use of advantages from big data technology [3]. Big data offers opportunities to various industry domains to manage risk and achieve on time .The most important values of big data are reflected in three aspects – information consumption, economic transformation and up gradation.

Meanwhile, two parallel breakthroughs have further helped accelerate the adoption of solutions for handling Big Data: The availability of Cloud based solutions has dramatically lowered the cost of storage, amplified by the use of commodity hardware. Virtual file systems, either open source or vendor specific, helped transition from a managed infrastructure to a service based approach [10].

When dealing with large volumes of data, it is necessary to distribute data and workload over many servers. Additionally, the need to address data integration and management challenges is vital as system becomes more

Corresponding Author: J Lokesh, lokesh_cse@gtec.ac.in
Department of Computer science and Engineering, GTEC, Vellore, Tamil Nadu, India

diverse [3]. Also Making data available for immediate insight requires planning and a strategy to leverage the best possible platform(s) for success. New designs for databases and efficient ways of support massively parallel processing have led to a new generation of products like the so called noSQL databases and the Hadoop map-reduce platform [1].

It is a need of some technology to share ongoing information transparently between the stack holders involved in the construction or operation sites. Big data has some characteristics that business and data scientists need to be concerned with.

a) Volume

Big data implies enormous volumes of data. Such data is generated by machines, networks and human interaction on systems like social media the volume of data to be analyzed is massive.

b) Variety

Variety refers to the many sources and types of data both structured and unstructured. We used to store data from sources like spreadsheets and databases. Now data comes in the form of emails, photos, videos, monitoring devices, PDFs, audio, etc. This variety of unstructured data creates problems for storage, mining and analyzing data.

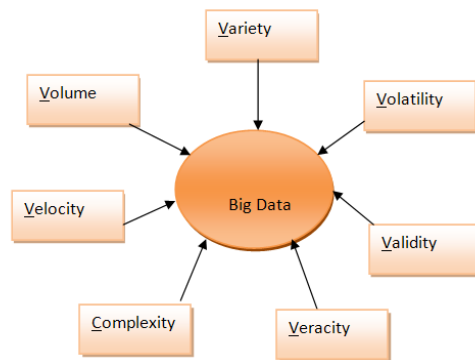


Fig.1 characteristics of Big data

c) Velocity

Big Data Velocity deals with the pace at which data flows in from sources like business processes, machines, networks and human interaction with things like social media sites, mobile devices, etc. The flow of data is massive and continuous. This real-time data can help researchers and businesses make valuable decisions that provide strategic competitive advantages and ROI if you are able to handle the velocity. Inderpal suggest that sampling data can help deal with issues like volume and velocity.

e) Veracity

Big Data Veracity refers to the biases, noise and abnormality in data. Is the data that is being stored, and mined meaningful to the problem being analyzed [5]. Veracity in data analysis is the biggest challenge when compares to things like volume and velocity. In scoping out your big data strategy you need to have your team and partners work to help keep your data clean and processes to keep 'dirty data' from accumulating in your systems.

f) Validity

Like big data veracity is the issue of validity meaning is the data correct and accurate for the intended use. Clearly valid data is key to making the right decisions

g) Volatility

Big data volatility refers to how long is data valid and how long should it be stored. In this world of real time data you need to determine at what point is data no longer relevant to the current analysis.

h) Complexity

Data must be able to travel through various physical as well as cloud.

From our analysis of big data in various domains, we reached a sense of what characteristics indicate higher or lower value potential from the use of big data to capture value from its use, as well as higher or lower barriers in realizing that value in different sectors. Using these insights, we created two indices: (1) an index on value potential (I_p), and (2) an index on the ease of capture (EoC). Each of these indices comprises multiple criteria which give us a relative sense of which sectors may be poised for greater gains and which sectors would face the toughest barriers [6]. We do not claim that these indices give a full picture, but we believe that they give a good sense of both the potential value available and the ease or otherwise of its capture across sectors.

II. VALUE POTENTIAL INDEX

The five criteria we use in this index act as a proxy for how The sectors can benefit from one of the five transformative opportunities we have discussed in this report:

A. Amount of data per firm

The larger the amount of data per firm, the more it indicates that a firm is likely to be able to benefit from increasing transparency in terms of data. We used the storage available per firm as a proxy. We built upon our data mapping analysis to estimate the available data storage, in bytes

B. Variability in performance

The higher the variability, the more it indicates a firm can benefit from the use of data and experimentation to expose variability and improve performance. We used the variability in earnings before interest tax depreciation and amortization as a proxy.

C. Customer and supplier intensity

The more customers and suppliers a firm has, the greater its potential to apply segmentation to tailor courses of action. We used the number of frontline employees (defined as those who interface with customers or suppliers) per firm as a proxy

D. Transaction intensity

The higher the transaction intensity, the more likely the sector can benefit from the use of automated algorithms to augment or replace human decision making. We used the amount of processing power of an average firm in a sector as a proxy.

E. Turbulence

Turbulence, or how frequently leaders and laggards in a sector change place, is a proxy for the amount of innovative disruptions to which a sector is susceptible. We hypothesize that the higher the turbulence, the greater the likelihood that a sector will benefit from the use of big data to innovate business models, products, and services.

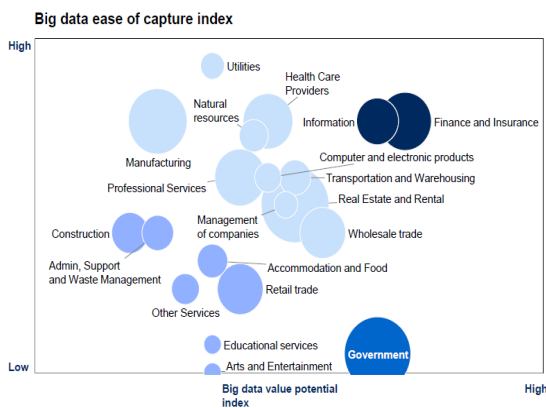


Fig. 2 Potential vs capture index

III. VALUE CAPTURE INDEX

This index is made up of four criteria, each of which aligns with a key barrier to the use of big data that we have identified

A. Talent

The more deep analytical talent a firm has, the better a position it is in to realize value from big data. Our research identifies three key types of talent required to capture value from big data:

a) deep analytical talent—people with technical skills in statistics and machine learning, for example, who are capable of analyzing large volumes of data to derive business insights;

b) Data-savvy managers and analysts who have the skills to be effective consumers of big data insights—i.e., capable of posing the right questions for analysis, interpreting and challenging the results, and making appropriate decisions;

And supporting technology personnel who develop, implement, and maintain the hardware and software tools such as databases and analytic programs needed to make use of big data.

B. IT intensity

The more IT assets a sector has on average, the lower the technology barriers to be overcome.

C. Data-driven mind-set

This indicates how receptive the organization is to using big data to create value

D. Data availability

We use the relative number of databases related to each sector in a proprietary corpus of data as a proxy for how accessible data is in a sector.

| | Big data savvy | Deep analytical | Big data infrastructure |
|--------------------|---|---|---|
| Definitions | Employees who can define key questions that data can answer and have basic knowledge of statistics | Specialists who have conduct data analysis and advanced training in statistics and/or machine learning and | IT personnel who serve as database administrators and programmers |
| Occupations | <ul style="list-style-type: none"> Business and functional managers Budget, credit and financial analysts Engineers Life scientists Market research analysts Survey researchers Industrial-organizational psychologists Sociologist | <ul style="list-style-type: none"> Actuaries Mathematicians Operations research analysts Statisticians Mathematical technicians Mathematical scientists Industrial engineers Epidemiologist Economists | <ul style="list-style-type: none"> Computer and information scientists Computer programmers Computer software engineers for applications Computer software engineers for system software Computer system analysts Database administrators |

Fig.3 Capture index

IV. CAPTURING CONSTRUCTION COST INDEX

The construction companies heed in all around to estimate the cost of each project they involve not only for this reason but also to focus on equality at all insights[8]. The construction companies should ensure its adoption of technologies properly and the guidelines in prepare workforce, infra structure, collecting data, processing and analyzing data, visualizing information's for making decisions.

A. Use project data on future work.

By collecting data on every project, design firms can eliminate rework and apply best practices on future projects.

B. Automate the data entry process.

Some firms are applying algorithm-based approaches to integrate the new data with existing historical data to avoid daunting situations without eliminate any information from the past project. This approach also minimizes rework, allowing designers to focus their time on solving new challenges.

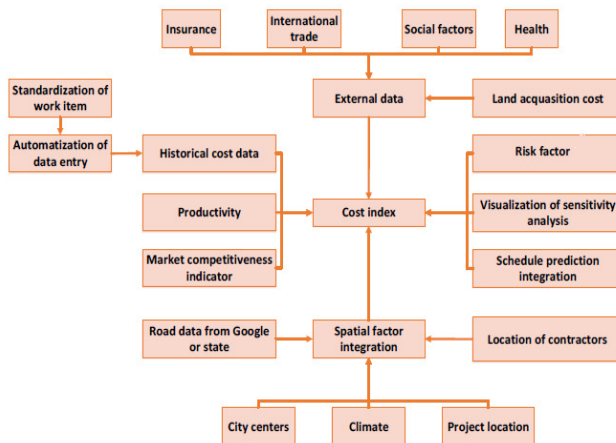


Fig 4. Application example: Construction cost index

V. BENEFITS OF BIG DATA OVER CONSTRUCTION INDUSTRY

C. Accurately budget for new project

First of all, big data allows projects to be kept on time and on budget by providing insights based on data generated by various sources, availability and prices of materials, suppliers, human resources systems, and communication. Site information can help constructors to better plan suppliers in real-time, create better routings and save money.

D. Reduces project Risks

One of the major problems for construction companies is the risk that the project is delayed, which can cost them a lot of money [8]. There are a lot of moments when a project can be delayed, especially with the rise of global projects. Example, the architect lives in one country and deals with contractors from another country to build a project in the some other country. On top of that, increased competition and skill shortages can be a threat to complete the project on time.

E. Creating transparency

Both external stakeholders such as citizens and businesses and internal stakeholders such as government employees and agencies can improve their efficiency when data from large public sector databases are made more accessible. Pre-filling would also have the advantage of reducing errors and speeding up processing time.

F. Replacing/supporting human decision making with automated algorithms

Some of the more sophisticated applications of big data use automated algorithms to analyze large datasets in order to help make better decisions. Algorithms can crawl through big data from a variety of sources, identifying inconsistencies, errors, and fraud. For instance, rule-based algorithms can flag suspicious correlations such as a person receiving unemployment benefits while filing for a work-related accident.

G. No benefits from using big data technology without Privacy

The cloud service provide the basis for big data because there will be a vast amount of data move from system to system in the various processes according to the defined standards and regulations but it is difficult job to spot where these data from. The cloud service enables access to the required system and service in more secure manner often it provide services in two essence: one is providing “data as service” and another one “algorithm as service”.

Choosing right platform also has significant affects in the growth of any domain to meet the end user requirements. Some of the factors which shift companies to real time data platform adoption are follows:

a) *Response*: The response rates are key component to the new platforms in the area of big data analytics can deliver. Data delivery is not necessary for all applications and data-driven scenarios, but real time use cases such as fraud detection, predictive maintenance, asset tracking and real time customer engagement are growing in importance and becoming more critical to many companies.

b) *Analytics complexity*: Addressing the requirements of complexity within analytic environments is getting more challenging while running highly complex analytic models over massive data stores is becoming more commonplace.

c) *Economics*: The economics of technology is the great equalizer and often can attribute to an early majority adoption of the technology.

d) *Structure*: Flexibility of structure is a growing decision point for selecting Big Data platforms.

e) *Load*: Data loads are growing more complex and the sources more diverse. Driven by greater complexity and demand, Big Data adoption is driven by the need to provide flexibility. The power of Big Data platforms to load a mixture of data creates an opportunity to address both analytic and operational scenarios.

Including these five requirements in your planning for Big Data platforms will ensure that the companies can select the best possible solution or combination to execute your Big Data strategy.

CONCLUSION

This paper discussed so far has emphasized the insights available on construction sites as either we take as globally or in small construction sites. Construction sectors have already hiring for awaking knowledge about the big data technology benefits, models, it usage and how extend it improve the businesses by using tools. The potential and capture index value well shows how one sector driven along the big data technology among other sectors and where the improvement needed today to map the technology. The need of data map methodologies further strengthens the growth of individual sectors. Finally, this paper signifies some of the requirements or need for choosing platforms in any sectors to respond at new speeds, the new ways to leverage data and provides insights to the new users.

REFERENCES

- [1] Big Data and the Construction Industry, leopard construction companies leveraging big data analytics, February 5, 2015.
- [2] Big data infrastructure covers guiyang, chinadaily.com.cn, 2014-12-17.
- [3] Big Data, Predictive Analytics, and Data Visualization in The Construction Engineering December 2013
- [4] Perdomo, J., Cavallin, H. (2014) Transforming Building Design Through Integrated Project Delivery in Architectural and Engineering Education, to appear in Proceedings of the ASCE 2014 Construction Research Congress
- [5] Efficient document handling for saving time and money, view point.com, October 7, 2014
- [6] Construction Industry Trends, Technology for Construction, www.stratusvue.com, 27 April 2014
- [7] How Big Data is Transforming Construction, construction global, Aug 19, 2014
- [8] Big data insight, Bigdata.cioreview.com, Aug 19, 2014
- [9] Stefan Heck & Matt Rogers, "Resource revolution", Amazon publishers ISBN 978-0-544-11456-2
- [10] 5-ways-big-data-is-changing-healthcare-the-latest-from-mckinsey, http://rockhealth.com, April 04, 2013
- [11] A. Vasenev, F. Bijleveld, T. Hartmann, A.G. Dorée, Visualization workflow and its implementation at asphalt paving construction sites, joint CIB W078-W102 2011 conference, France, 2012

- [12] Constructions Big data problem, randy_deutsch, January 14, 2014
- [13] WHAT YOU DON'T KNOW ABOUT YOUR (BIG) DATA CAN HURT YOU, ENR.CONSTRUCTION.COM, MAR 12 2012
- [14] Construction's Big Data Mistake, stratusvue.com, July 9, 2013
- [15] Three Reasons Why Construction Companies Fail When Trying to Predict Injuries, Griffin Schultz, and August 2014
- [16] Three Ways to Use Big Data Analytics to Your Advantage, proest.com, Mar 12, 2015

AUTHORS PROFILE

Lokesh studied computer science and engineering from Anna university, chennai obtained his Under graduate degree in 2007 and master degree in computer science and Engineering in 2011.he started working as Assistant prof. since 2008.he has been working in Real time projects related and hands on experience related to the fields of wireless networking, data mining and Big data. His area of specialization is Big data.



Sakthivel obtained his degree in 2005 in computer science and Engineering from Anna university.he started working as Assistant prof. since 2006.he has been working in developing models for big data projects related to data management.



Rajesh received his degree in 2007 from Anna university, Chennai. He has been working as Assistant professor since 2008.He works for data mining and information management projects. His area of specialization is Information systems.

