Big Data in Construction Perspective: Exploration of Google Cloud based Technologies and Offerings

Rajesh Saxena & S.K.Soni

Rabindranath Tagore University

ABSTRACT

This paper critically explores and reviews the literature to identify the data trends and how the construction industry can benefit from big data. The big data has brought the revolutionary changes in all the industries and construction industry is also not remaining untouched. Big data engineering (BDE) and statistics are among the most crucial steps for integrating big data technology in construction. We have reviewed related papers published in the various research and education institute across the world. The current application of Big Data in construction industry is already wide spread and future opportunities like big data research into construction safety, site management, heritage conservation, and project waste minimization and quality improvements are giving it totally new dimensions. Construction industry generates large amount of data every day but this data which has huge potential but not gets utilized effectively mainly due to lack of technology adoption. This paper will try to give different perspective of Big Data with the appropriate use of Big Data Engineering, Tools and Technologies.

Keywords: big data; big data engineering; big data analytics; big data storage; big data processing; machine learning; artificial intelligence; google cloud platform.

Classification: LCC: T58.64

Language: English

Great Britain Journals Press

LJP Copyright ID: 392953
Print ISSN: 2631-8474
Online ISSN: 2631-8482

London Journal of Engineering Research

Volume 23 | Issue 4 | Compilation 1.0

Big Data in Construction Perspective: Exploration of Google Cloud based Technologies and Offerings

Rajesh Saxena[α] & S. K. Soni[σ]

ABSTRACT

This paper critically explores and reviews the literature to identify the data trends and how the construction industry can benefit from big data. The big data has brought the revolutionary changes in all the industries and construction industry is also not remaining untouched. Big data engineering (BDE) and statistics are among the most crucial steps for integrating big data technology in construction. We have reviewed related papers published in the various research and education institute across the world. The current application of Big Data in construction industry is already wide spread and future opportunities like big data research into construction safety, site management, heritage conservation, and project waste minimization and quality improvements are giving it totally new dimensions. Construction industry generates large amount of data every day but this data which has huge potential but not gets utilized effectively mainly due to lack of technology adoption. This paper will try to give different perspective of Big Data with the appropriate use of Big Data Engineering, Tools and Technologies. We will also discuss the currently available tools such as computer-aided drawing (CAD) and building information modeling (BIM) and how they are providing tremendous opportunities to researcher in construction industry. We will also discuss how the rise of interest in big data is making it more effective due to the adoption of technology such as robotics, smartphones, computers and other gadgets. These gadgets helps in model development gathering data which gets used in the development of varioustechnique for Algorithm development, machine learning (ML), statistical analysis, and computational.

Keywords: big data; big data engineering; big data analytics; big data storage; big data processing; machine learning; artificial intelligence; google cloud platform.

Author α: Department of Civil Engineering Rabindranath Tagore University Bhopal, Madhya Pradesh.

σ: Department of Civil Engineering Rabindranath Tagore University Bhopal, Madhya Pradesh.

I. INTRODUCTION

Construction is a data-intensive sector where the bulk of data is generated and not capitalized on adequately due to slow technology adoption. Big data, a relatively new technology, are not properly adopted by construction. There are various steps involved in using big data, including data acquisition, storage, classification, and refining. These steps are handled through various software programs to refine the associated big data and make it usable for research and practical purposes In the case of construction, some barriers to big data adoption include latency, data privacy, data availability, data governance, poor broadband connectivity at construction sites, and cost implication for long-term use. For instance, big data adoption in construction may have latency issues with lower transfer rate and response time required due to software issues or network problems which may be a hurdle for some time-sensitive construction applications.

The construction industry is also benefiting from big data in a way that has revolutionized its traditional operational methods to a more automated process. The presence of digital tools and technologies for designing and executing construction projects has made the construction industry take enormous leaps in the last two decades.

The possibility of modeling building structures and identifying the functionality of those structures before they are built has led to industrial investments in big data and related technologies. Computer-aided design (CAD), such as building information modelling (BIM), is a term now synonymous with the construction industry.

The three-dimensional modeling of buildings and other construction infrastructures leads to the generation of digital files which can be stored in various formats, leading to a bulk of data generation. Other digital innovations such as digital twins, 3D laser scanning, and advanced wearable gadgets incorporated in hats, shoes, gloves, and other sensor-based tools have revolutionized the construction industry and helped generate useful big data.

Big data in the construction industry can accumulate quickly and become storage heavy due to the large size of the 3D modeling files and a huge amount of daily data generated by wearable gadgets. Management of such big data is a hectic but essential task as the usefulness of the models lies in ensuring that they are available for viewing and leveraging as and when needed. Apart from providing the ease of modeling infrastructure, big data also provide the opportunity to develop sustainable structures by using test models before actual constructions. These are made possible by using digital twins, geographical information systems (GIS)-based 3D point cloud structures, and other cloud-based scanning systems.

Furthermore, the software that enables CAD and BIM further feeds into the databases and contributes to big data.

The applications of big data in the construction industry are immense. Identifying how big data can be applied to the construction industry remains the real challenge.

Since each construction project leads to more data generation, it is crucial to analyze and sort the data accordingly. Some of the key features within the construction industry that can benefit from big data

- Construction safety
- Efficiency
- Waste minimization
- Productivity
- Competitive Advantage
- Pollution Management
- Project Management
- Claims Management
- Procurement

II. METHODOLOGY

The purpose of this study is to explore other nonconventional ways to control the cost and achieve the cost effectiveness in construction.

In order the achieve above stated purpose, following methodologies are outline:

- Step-1
  - Aim
- Step-2
  - Objectives
- Step-3
  - Literature Review
- Step-4
  - Google Clound Platform -Technology Exploration
- Step-5
  - Conclusion

III. AIM

- The core aim of using different technologies is to simplify how datasets can be used to guide future construction projects.
- The guiding principle here is to use modern technology to upgrade and update the ways in which information could be streamlined for the benefit of different projects.
- Identifying the materials that best suit a particular structure, developing project timelines, and streamlining the resources can become much more straightforward if the construction projects are developed with the help of big data technologies
- The overall purpose of this paper is to explore various tools and techniques for the effective use to Big Data in construction.

IV. OBJECTIVES

- The overall objective of this paper is to explore Tools and Technique which can be used to use Big Data for construction and its management based on the review of existing literature.
- The existing literature on big data does not provide Technology focus approach to use Big data for construction management and this is the very reason construction industry is lagging behind in the use of big data in the construction industry.
- The paper research work focus on following aspects
  - o How can we use Technologies for big data for research in construction engineering and management?
  - o How can we use Big Data Engineering to process and Store Big Data
  - o How can we use Big Data Analytics to use Big data in construction
  - o How can we use Machine Learning to effective use Big Data in construction
  - o How can we use Cloud Offerings (GCP) in construction?
  - o How to use 10 Vs of Big Data
  - o How can big data be used for planning construction projects in a futuristic way?

V. LITERATURE REVIEW

5.1 Big Data Trends

Following are sequence of trends Big Data have gone through since 2016

- 2016
  - Cloud computing and Big Data for construction
- 2017
  - Development of Big data modes for construction
- 2018
  - Implementing Big Data in construction engineering
- 2019
  - Construction Management through Big Data
- 2020
  - Public awareness of construction using Big Data
- 2021
  - Harnessing ML and Big Data for futuristic construction Project
- 2022
  - Artificial Intelligence (AI),Deep learning
- 2023
  - Block chain, Chat GPT

5.2 Big Data Research Techniques

Following are the various research techniques getting used in Big Data

- Big Data Analysis
- Block Chain
- Data Bases ,Data Mining & Statistics
- KDD & Pattern Recognition
- 3D & Computational Neuroscience
- Artificial Intelligence
- Chat Bots & Chat GPT

5.3 Big Data Domains

☐ Big Data is classified in to Major Domain BDE & BDA. These two main domains are further divided in to many classes and subclasses. The third domain that comes under the canopy of Big Data is ML.

- Big Data Domains
  - Bata Data Engineering
  - Big Data Analystics
  - Machin Learning

5.4 Big Data Engineering

1. Big data analytics (BDA) is supported by BDE that provides a framework to conduct it. BDE has tremendous applications in construction.
2. It has been used for BIM to improve project management.
3. It has also been used to improve building design and for effective performance monitoring, project management, safety, energy management, decision-making design framework, resource management, quality management, waste management and others.
4. To understand BDE, it is important to discuss big data platforms. These platforms are divided into two groups based on variations in their inherent characteristics.

   4.1 Horizontal Scaling Platforms (HSP)
   HSP utilizes multiple servers by distributing processing across them and bringing new machines into the cluster. In construction, HSPs have been used for waste management, profitability performance, smart road construction, and others.

   4.2 Vertical Scaling Platforms (VSP) VSPs are single-server-based configura- tions that achieve the scaling by upgrading the hardware of the related server. VSPs have been reported in one-off construction projects, transportation and others.
5. BDAS (Berkeley Data Analytics Stack)has been in the limelight since it has greater performance gains over Hadoop.
6. Hadoop has been widely utilized in big data applications. The tools offered by these platforms are useful in the storage and processing of big data.
7. Classification of Big Data in to its key domains
   I. Big Data Engineering
   i. Big Data Processing
      a. Map Reduce
      b. DAG
   ii. Big Data Storage
      a. NoSQL
      b. Distributed File System
         a) HDFS
         b) Tachyon
8. Big Data Processing-Distributed and parallel computation is present in the core of BDE. In construction, big data processing has been utilized for waste management, prefabricated construction project management, profitability analyses, and other construction management applications.

   8.1. MapReduce (MR)

- MapReduce was developed for the handling of big data. It utilizes a distributed processing model in which two functions, as indicated by the name itself, map and reduce, are employed to write analytical tasks. Mappers and reducers are the processes that collect Big Data from these functions for further processing. Initially, mappers collect and read the input information to process it for subsequent results generation. The output of mappers is used by reducers which give the results that are ultimately stored in the file system
- MR has been used by Jiao et al. to develop an augmented framework for BIM. Similarly, it has also been used in construction knowledge maps and other big data applications.
- The overall features and application of MapReduce is the reduction of data into

manageable chunks. The use of MapReduce not only distributes data into smaller chunks but also helps develop datasets that present a more analytic view of big data.

- Having organized datasets within the construction industry is of key importance as it can greatly increase the efficiency of data management and decision making based on data analysis.
- Hadoop was the popular and first big data platform that introduced and made it easy for people to work on MR by executing its programs successfully. For tasks requiring batch processing, MR proved itself to be an effective tool as a typical cluster contains interlinked mappers and reducers that assist by running MR programs side by side at the same time
- Yet another resource negotiator (YARN) has also been introduced, which functions by providing resource management and scheduling related functions of MR and has made it easy to implement innovative applications by Hadoop.
- Hadoop models have been used in construction for smart buildings and disaster management , failure prediction of construction firms [56], workers' safe behaviors in a metro construction project and other relevant applications.

8.2. Directed Acyclic Graph

- Big data platforms also use Directed Acyclic Graph (DAG) which is an alternative processing model. In comparison with MR, DAG works by relaxing map-then-reduce, the style of MR, which is supported by Spark.
- Spark is widely accepted for reactive and iterative applications due to its supremacy over MR in high expressiveness and in-memory computation
- Disk-resident and memory-resident tasks are conducted ten and one hundred times faster using Spark than MR.
- DAGs provide major advantages that enable experts and researchers to construct complex causal relationships in which nodes represent stochastic variables, and directed edges (arrows) indicate direct probabilistic dependencies among the relevant variables.
- DAGs are also able to encode deterministic as well as probabilistic relationships among the variables.
- The usage of Spark and associated DAGs has been reported for construction profitability analysis, waste management, energy monitoring service on smart campuses , and others.
- Spark and Hadoop are among the ML tools with enormous potential in cons- truction engineering and management.
- The speed of both these systems is better than other algorithms and ML tools currently in use in the construction industry.
- Fault tolerance in both these systems is also high and has greater scalability than existing models.
- The data storage in these systems is slightly different in that Spark uses a memory system while Hadoop utilizes a disk for data storage.
- The language for both these tools is also different since Spark is written in Scala while Hadoop has been developed using JavaScript.
- Despite the slight differences, both these tools provide the opportunity to process data in the form of batches and at a higher speed than previously existing models, making them potential tools for futuristic model developments in construction engineering and management.
- JavaScript has been used in construction to anticipate building material reuse, automated progress control coupled with laser scanning, shared virtual reality for design and management, construction information mining
- Scala has been used for the process information modeling concept for on-site construction management.

8.3. Big Data Processing in Construction

- Big data processing has been effectively utilized in the construction industry for failure prediction data, construction waste analytics, profitability data, modular and prefabricated construction, fire incident management, smart campus energy monitoring, healthier cities management, smart road management, and others.

- Though MR and Spark have their own significance, these are less frequently employed in the construction industry to process big data such as BIM-associated data. Partial BIM models' retrieval was optimized by MR by Bilal, et al. and Chang and Tsai The authors found a loop in the Hadoop MR logic of data distribution.
- Another research group worked for naive and expert BIM users by developing a system for BIM data storage and retrieval. The authors developed a system for cloud BIM to retrieve and represent big data intelligently. This system helped develop an interactive interface to maximize the usability and utility of construction big data. Complex BIM data are retrieved by processing proposed natural languages after reformulating user queries. This data are then visualized by mapping on various visualizations. Before query evaluation, two BIM collections are merged to optimize the process of query execution. Using this technology, a 40% reduction in response time has been witnessed compared to other traditional technologies.
- Currently, the utilization of BIM is limited across the construction and facilities management stages. The real intent of BIM could only be achieved once applied at each stage of the building lifecycle.

9. Big Data Storage-

- Big data storage is also an important aspect of BDE. In construction, big data storage has been explored for forecasting the success of construction projects, smart buildings data storage, tender price evaluation, and others.
- Despite the availability of BIM data storage, the current applications in construction still require successful implementation. Social BIM, proposed by Das et al., captures building models and the social interactions among the users. The authors developed BIM Cloud based on the distributed BIM framework.
- Similarly, a two-tiered hybrid data infrastructure was proposed by Jeong et al. for data management and monitoring of bridges. In this model, the client tier efficiently completes some analytical tasks by storing structured data momentarily using MongoDB, while the central tier stores sensor data permanently using Apache Cassandra. Lin et al. also used MongoDB to store BIM data obtained through building models.
- Overall big data storage is provided by either emerging NoSQL databases or distributed file systems, as explained subsequently

9.1. Distributed File Systems

- The distributed file systems consist of Hadoop Distributed File System (HDFS) and Tachyon.
- HDFS is designed to deal with large and complex databases such as those related to BIM, waste, and other construction big data sources.
- It operates with the commodity servers grouped together in a cluster.
- As it utilizes several servers, the probability of hardware failure also increases. To overcome this problem, HDFS introduces fault tolerance achieved through the distribution of data and their replication. However, in situations where low-latency data access is required, HDFS is not a suitable option as it shows inferior performance.
- HDFS has been utilized by construction researchers for observing construction workers' behavior, improving road performance, and investigating profita- bility performance.
- Furthermore, based on the distributed input from HDFS, it facilitates building predictive models for conducting building simulations that give output in a predictive model markup language.
- Tachyon is a distributed file system designed to extend HDFS benefits by providing access to the distributed data across the cluster at memory speed.
- Tachyon has been utilized in construction for handling unstructured documents and file storage.
- The Tachyon performs better than HDFS, is backward compatible and can handle the MapReduce jobs without any further modifications.

9.2. NoSQL Databases

- Relational databases have been common for data management in past decades. However, new applications were designed for better performance, scalability, and flexibility as the technology emerged. Relational databases lag because of their special processing and storage needs.

- As a result, new systems were devised to fill this technology gap. One such system is the "Not only SQL" system that has optimized data management in several ways.
- NoSQL has been widely used in different industries, including construction, dueto its fragmented nature. Some examples of NoSQL in construction include integration of lessons learned knowledge in BIM], web service framework for construction supply chain collaboration and management, and Social BIM Cloud implementation .
- NoSQL systems store schema less data in a non-relational model. It does not set too many restrictions on value and allows easy product determination. Generally, when NoSQL databases are set to key values, they carry out only specific tasks without evaluating specific values. The key-value database is mainly tailored to the business accessed through the primary key. These systems have four data models that are briefly discussed below.

i. Key-value: This is the simplest data model used for unstructured data storage. However, the data lack self-description. It has been used for knowledge management in construction and integration of lessons learned knowledge in BIM. Big data utilization in BIM can be beneficial to discover root causes of poor building performance, perform real-time data queries, improve the decision-making process, improve productivity, and reveal new designs and services in the construction industry, as is the case in every industry.

ii. Document: This model can store self-describing data. However, this model can lag in terms of efficiency. It has been used for unified lifecycle data management in architecture, engineering, construction, and facilities management through BIM integration.

iii. Columnar: Aggregated columns, grouped sub-columns, and sparse data can be stored by using this model. It has been used for integrating digital construction through the internet of things and smart archiving of energy and petroleum construction projects.

iv. Graph: This model works well for property-graph-based huge datasets in relationship traversal. It has been used for the 4D construction management information model of prefabricated buildings and the development of a BIM-enabled software tool for facility management.

1. *Big Data Analytics:*
   a. Big Data Analytics
      i. Statistics
         1. Bayesian Analysis
         2. Resampling
         3. Shrinkage
         4. Tree-based Analysis
      ii. Data Mining

- BDA gathers information from a variety of disciplines. All these disciplines have one thing in common: to find out data patterns. Some of these related disciplines are data mining, statistics, business analytics, predictive analytics, data analytics, knowledge discovery from data, and the most recent one, big data.
- Big data use the previous techniques to broaden the field of data analytics. For BDA, some of the ML-based tools are developed.
- In construction projects, BDA has been used for improving building design and effective performance monitoring, project safety, energy, resource, overall management and decision-making frameworks, and quality and waste management.
- Big data analytics has been taken a step further by developing predictive analysis techniques.
- Ngo et al. used a factor-based big data predictive analytic tool for analyzing the capacity of construction industries to deal with big data. This tool was tested and validated on four different construction organizations to ensure that the predictive analytic method could improve how the construction industry can use big data.

- The integration of big data in the construction industry remains an avenue that requires further research in terms of big data analytics.
- Overall, data analytics is conducted through statistical, data mining, and regression techniques, as explained below

1.1. Statistics

- Statistics has wide applications in the construction industry.
- Statistical techniques including Monte Carlo simulation, Gaussian distribution, non-Bayesian methods, correlation analysis, factor analysis, decision trees, Naïve Bayes, and others have been reported by various studies in construction.
- Some of the areas that benefitted from statistics include learning from post-project reviews, identifying causes of construction delays, analyzing buildings for structural damages, construction litigation, and identifying and recognizing heavy machinery and workers.
- Other examples of statistics in construction are those of bidding statistics to predict completed construction cost, accidents statistics, quality control, and six sigma for project success]. From measuring the bid-to-win ratio to how much a project is over budget or schedule, and KPIs, the more numbers you can put behind your work, the better. Data not only allow for more visibility into the state of a particular project, but relevant industry statistics and facts can provide valuable information needed to make important future decisions regarding preconstruction and planning, productivity tools, risk assessment, and workforce and operational efficiency.

1.2. Data Mining

- Data mining is used to extract meaningful patterns in the data. It has been an integral part of all big data management systems. It employs the techniques used in pattern recognition, ML, and statistics. Several models are assessed, and the ones with the best tolerance and high accuracy are selected and used for obtaining predictive results.
- In construction, data mining has been reported in waste management, BIM-based construction engineering quality management, and other relevant areas.
- Data mining detects useful regularities and information necessary for decision making for construction management projects.
- A data mining method such as cluster analysis is important for the construction industry, as it combines different construction objects into homogeneous groups and investigates them.
- Data mining is supported through data warehousing. Specially structured data is stored in data warehousing for querying and analysis. Extract, transform and load (ETL) is a program that allows the collation of transactional data and operational data.
- Warehouse Data analysis is conducted using Online Analytical Processing (OLAP), which performs better than SQL in computing breakdown and data summaries. OLAP has been used for cost data management in construction cost estimates by Moon et al. OLAP technology deals with the operational data and data obtained using big data technology. OLAP is presented as a multidimensional cube that rapidly processes datasets.
- Similarly, different data mining techniques have been used to identify construction delays. For analyzing construction datasets, Kim et al. presented a framework of knowledge discovery in databases (KDD).
- In the KDD, the most time-consuming and challenging step is data preprocessing. Nevertheless, KDD is a powerful tool for identifying casual relationships in construction projects and reducing construction variability by identifying and eliminating causes for possible deviations. With the application of KDD, randomness of construction projects and novel patterns can be determined. Other techniques include dimensional matrix analysis, link analysis, and text analysis . Other datasets with information related to delay causes, BIM-based knowledge discovery, intelligent learning, and the prevention of occupational injuries can be easily extended in the domain of data mining.

1.3. Regression Techniques

- Based on an input variable, regression predicts the value of the target variable. It is a supervised ML method.
- Regression is categorized into simple linear and multiple linear regression based on explanatory variables.
- In simple linear regression, the relationship between two variables (an explanatory variable x and a dependent variable y) is modeled using ML.
- While in multiple linear regression, two or more explanatory variables are used and their relationship with the dependent variable is modeled. The more common regression technique is multiple linear regressions.
- Regression has been extensively used in construction research. For example, it has been used to predict properties of concrete cured under hot weather , predicting final cost for competitive bids on construction projects , determining contingency in international construction projects, estimating performance time for construction projects and others.
- Moreover, regression has been used for cost estimation, which is a difficult task in the early stages of the project. Adoption of parametric methods such as regression and multiple regressions can be applied as both analytical and predictive techniques to estimate the overall reliability of the cost estimation.

2. The 10 vs. of Big Data-

- The most crucial properties of big data include their value, volume, velocity, variety, veracity, volatility, validity, variability, vulnerability, and visualization, also known as the 10 vs. of big data
- In terms of the use of big data in the field of construction, analyzing the vs. can help explore how big data can be used for developing better construction models in the future.
- Overall, multiple construction-related studies have reported the usage of vs. of big data. For example, velocity has been reported for high-speed construction data processing. Value has been reported for smarter universities and campuses. Volume has been reported for mass level offsite construction material and component production. Variety has been reported for investigating the profitability performance of construction projects. Veracity has been reported for forecasting the success of construction projects. Similarly, variability has been reported for modeling occupational accidents in construction projects.

2. *Machine Learning:*

   a. Machine Learning
      i. Regression
         1. Linear Regression
         2. Ridge Regression
         3. Neutral Network Regression
         4. Lasoo Regression
         5. Decision Tree Regression
         6. Random Forest
         7. KNN Model
         8. Support Vector Machines
      ii. Classification
      iii. Clustering
         1. Centroid Based Clustering
         2. Density Based Clustering
         3. Distribution Based Clustering
         4. Hierarchical Clustering
      iv. Natural Language Processing
      v. Information Retrieved(IR)

- One AI subdomain is ML which can be used to learn from the data using computational systems.
- ML is further categorized into:

(i) supervised learning;
(ii) unsupervised learning;
(iii) association
(iv) numeric prediction.

- ML has several applications in the construction industry. It uses different approaches, including rule-based learning approaches, case-based reasoning techniques, artificial neural networks, and hybrid methodologies.
- ML has immense potential as a tool in the field of construction. Over the last two decades, several ML algorithms have been proposed to aid and improve the overall process of construction. For example, ML has been used to predict properties of concrete,

contract management, site safety and injury prediction, delay risks management, BIM integrated on-demand site monitoring, and other areas of construction engineering and management.

- Various ML tools are integrated at different steps along with the construction management processes. Different ML interfaces such as PyTorch and Keras.io help develop computational models based on existing data for building futuristic construction models.
- BIM can also be improved by using big data and ML tools, as these technologies allow the opportunity to explore how technology could be applied to the construction industry .
- Over the last few years, different algorithms have been explored to predict various project phases and guide construction projects from inception to closure.
- Firstly, decision trees and similar tools are used for developing an overall project timeline to predict or determine construction project performance in various phases.
- Secondly, statistical analysis tools are used for analyzing previous projects and choosing guiding principles for future projects .
- Finally, design tools are integrated with ML algorithms to build 3D construction models and graphics for building models.
- These computational models enable analyzing construction projects by planning through look-up schedules and looking for ways to improve buildings and other structures. The combined use of big data, ML, and AI holds the potential to develop seamless construction projects and enable the development of structures that can withhold severe weather conditions and disasters.
- For example, one of the key uses of ML tools in futuristic construction projects can be the development of structures that can stand through natural disasters and provide safety nets to communities during floods and other disasters.
- Similarly, post-disaster evacuation and rescue of individuals can also be carried out more easily if the area contains structures such as roads and buildings built through the use of statistical modeling, thus providing safe routes for people.
- Although the automation of construction projects remains a future goal, the integration of different ML algorithms is already underway. Managing costs, timelines, and human resources on a construction project are areas guided by various algorithms and computational models.
- The ML approach can also be applied to develop leading indicators to classify sites according to their safety risk in construction projects

Future Opportunities of Big Data in Construction-

- There is immense potential for the use of big data in the construction industry. The use of big data and ML can enable construction automation.
- The construction industry is quite dynamic and demanding, with the need for labor strength and human resources to ensure the smooth running of projects. The constant challenge of keeping projects on track and ensuring that new buildings and structures are made up to modern standards puts much strain on the project management teams. These roadblocks can greatly be reduced with the use of big data and ML.
- The core aim of using big data in the construction industry is to enhance the project planning phases and speed up the overall construction process by predicting the possible timelinesfor particular projects and identifying what factors can be worked on to improve the overall process
- The automation of the construction projects will require the combined use of big data, deep learning, and ML tools.
- The use of big data and related tools can ensure that existing data and information can be used for drafting guiding principles and then building computational models accordingly. For example, using sensor-based wearable personal protective equipment, the big data of near misses, onsite accidents, hazards, and other issues can be generated for

developing safety plans and management techniques.

- Big data, BIM, and cloud-powered simulations can help minimize project waste and help produce superior quality constructed facilities.
- Big data artifacts generated by 3D scanners for as-built drawing development are another key advantage whereby the rehabilitation plans of ancient heritage sites can be developed.
- The future holds great potential for the construction industry through big data integration. Some of the key opportunities for the construction industries lie in using big data for business and environmental sustainability. The current roadblocks faced by the construction industry can be overcome in the future through the integration of information extracted through big data.
- The use of information gathered from past and present projects can help develop sustainable infrastructure in the long term. It is possible to avoid past mistakes and use better quality products guided by the information found through big data in construction.
- Future research directions in the field of construction rely heavily on big data as the presence of information sources can help in building better infrastructure and greatly improve building designs and the overall construction business. The construction industry must move towards automation and build upon the integration of technology to make the future use of big data seamless and hassle-free.
- The use of big data tools, BIM, and CAD can only be possible if the relevant support and integration systems are present. Hence, the future of the construction industry depends on upgrading the present environment gradually.
- Overall, the role of big data in enabling the entire process of futuristic construction projects is undeniable.
- Data play a crucial role in developing training models and smoothly enabling the process of construction.
- Future developments in this field will also include the generation and use of more algorithms and models that rely on big data, owing to the need to train the models reliably.
- All above highlighted benefit of Big Data can be taken to the next level by utilizing Cloud Offering in the Market.
- Now a days Cloud provider like Google, Amazon and Microsoft are offering various platforms and services for the effective utilization of Big Data.
- Following are some benefit of using cloud services in conduction industry:
  - o Data generated by Construction industry is huge and storing and processing data can be very expensive.
  - o Cloud provides data storage in cheap rate which can be accessed from anywhere easily. No need to physically score data on your premises.
  - o Cloud Storage is scalable so can be optimized based construction site requirement.
  - o Cloud provides various tools and technique for effectively managing construction data.
  - o Cloud Big Data offering helps in effectively storing and managing huge volume of construction data.
  - o Wide range of data analytical tools available on Cloud platforms can be used to process and effectively utilize the huge volume of Construction data.
  - o Cloud providers various Machine Learning and Artificial Intelligence capability which can be ,if effectively used , can be game changer to address following concerns associated with construction industry
    - ▪ Effective planning specifically when construction industry has to deal with lot of unknowns
    - ▪ Effectively managing day to day construction activities by giving due consideration of dependent tasks and activities
    - ▪ AI can be helpful with forecast and planning the construction activities accordingly.
    - ▪ Safety and Security concern of construction industry can also be effectively address by using various Cloud offering
    - ▪ One construction project experience and data can be used effectively planning for future projects.

Google Cloud Platform(GCP) -Technology Exploration

- Google Cloud Platform Domains
  - GCP Platform Services
  - GCP Data Analytics
  - GCP AI/ML Services

6.1 GCP Platform Services

Compute:

App Engine: App Engine enables you to build and host applications on the same systems that power Google applications. App Engine offers fast development and deployment; simple administration, with no need to worry about hardware, patches or backups; and effortless scalability.

Compute Engine: Compute Engine offers scalable and flexible virtual machine computing capabilities in the cloud, with options to utilize certain CPUs, GPUs, or Cloud TPUs. You can use Compute Engine to solve large-scale processing and analytic problems on Google's computing, storage, and networking infrastructure.

App Engine & Compute Engine offerings can be very useful for processing the big data in construction industry where data demand in dynamic in nature require flexibility in terms of processing power and scalability.

Workload Manager: Workload Manager is a rule-based validation service for evaluating workloads running on Google Cloud. If enabled, Workload Manager scans application workloads to detect deviations from standards, rules, and best practices that improve system quality, reliability, and performance.
Work load Manager can provide great amount of help to manage the dynamic workload of construction industry which required to effective management the demand deviations.

Storage

Cloud Storage: Cloud Storage is a RESTful service for storing and accessing your data on Google's infrastructure. The service combines the performance and scalability of Google's cloud with advanced security and sharing capabilities.

Persistent Disk: Persistent Disk is a durable and high performance block storage service for Google Cloud Platform. Persistent Disk provides SSD and HDD storage that can be attached to instances running in either Compute Engine or Google Kubernetes Engine.

Cloud Filestore: Cloud Filestore is a scalable and highly available shared file service fully-managed by Google. Cloud Filestore provides persistent storage ideal for shared workloads. Above mentioned Google Storage options should be explored to establish best suited storage mechanism to Store Big Data in construction industry which is normally very huge in volume.

Databases

AlloyDB: AlloyDB is a fully-managed, Postgre SQL-compatible database for demanding transactional and analytical workloads. It is designed to provide enterprise-grade performance and availability while maintaining compatibility with open-source PostgreSQL.

Cloud Bigtable: Cloud Bigtable is a fast, fully-managed, highly-scalable NoSQL database service. It is designed for the collection and retention of data from 1TB to hundreds of PB.

Datastore: Datastore is a fully-managed, schemaless, non-relational datastore. It provides a rich set of query capabilities, supports atomic transactions, and automatically scales up and down in response to load. It can scale to support an application with 1,000 users or 10 million users with no code changes.

Firestore: Firestore is a NoSQL document database for storing, syncing, and querying data for mobile and web apps. Its client libraries provide live synchronization and offline support, while its security features and integrations with Firebase and Google Cloud Platform accelerate building serverless apps.

Memorystore: Memorystore, which includes Memorystore for Redis and Memorystore for Memcached, provides a fully-managed in-memory data store service that allows customers to deploy distributed caches that provide sub-millisecond data access.

Cloud Spanner: Cloud Spanner is a fully-managed, mission-critical relational database service. It is designed to provide a scalable online transaction processing (OLTP) database with high availability and strong consistency at global scale.

Cloud SQL: Cloud SQL is a web service that allows you to create, configure, and use relational databases that live in Google's cloud. It is a fully-managed service that maintains, manages, and administers your databases, allowing you to focus on your applications and services.

Construction Industry generate huge amount of data which is very diverse in terms of all 10 V's which include variety, value, volume, velocity, variety, veracity, volatility, validity, variability, vulnerability, and visualization. Above available Data bases those are available on GCP can be effectively use to store the construction data based on the data demand.

6.2 GCP Data Analytics

BigQuery: BigQuery is a fully-managed data analysis service that enables businesses to analyze Big Data. It features highly scalable data storage that accommodates up to hundreds of terabytes, the ability to perform ad hoc queries on multi-terabyte datasets, and the ability to share data insights via the web.

Cloud Composer: Cloud Composer is a managed workflow orchestration service that can be used to author, schedule, and monitor pipelines that span across clouds and on-premises data centers. Cloud Composer allows you to use Apache Airflow without the hassle of creating and managing complex Airflow infrastructure.

Cloud DataFusion: Cloud Data Fusion is a fully-managed, cloud native, enterprise data integration service for quickly building and managing data pipelines. Cloud Data Fusion provides a graphical interface to help increase time efficiency and reduce complexity and allows business users, developers, and data scientists to easily and reliably build scalable data integration solutions to cleanse, prepare, blend, transfer, and transform data without having to wrestle with infrastructure.

Cloud Life Sciences (formerly Google Genomics): Cloud Life Sciences provides services and tools for managing, processing, and transforming life sciences data.

Data Catalog: Data Catalog is a fully-managed and scalable metadata management service that empowers organizations to quickly discover, manage, and understand their data in Google Cloud. It offers a central data catalog across certain Google Cloud Services that allows organizations to have a unified view of their data assets.

Dataform: Dataform provides an end-to-end experience for data analysts to develop, test, version control, and schedule complex SQL workflows.

Dataplex: Dataplex is an intelligent data fabric that helps customers unify distributed data and automate management and governance across that data to power analytics at scale.

Dataflow: Dataflow is a fully-managed service for strongly consistent, parallel data-processing pipelines. It provides an SDK for Java with composable primitives for building data-processing pipelines for batch or continuous processing. This service manages the life cycle of Compute Engine resources of the processing pipeline(s). It also provides a monitoring user interface for understanding pipeline health.

Datalab: Datalab is an interactive tool for exploration, transformation, analysis and visualization of your data on Google Cloud Platform. It runs in your cloud project and enables you to write code to use other Big Data and storage services using a rich set of Google-authored and third party libraries.

Dataproc: Dataproc is a fast, easy to use, managed Spark and Hadoop service for distributed data processing. It provides management, integration, and development tools for unlocking the power of rich open source data processing tools. With Dataproc, you can create Spark/Hadoop clusters sized for your workloads precisely when you need them.

DataprocMetastore: DataprocMetastore provides a fully-managed metastore service that simplifies technical metadata management and is based on a fully-featured Apache Hive metastore. DataprocMetastore can be used as a metadata storage service component for data lakes built on open source processing frameworks like Apache Hadoop, Apache Spark, Apache Hive, Presto, and others.

Datastream: Datastream is a serverless change data capture (CDC) and replication service that enables data synchronization across heterogeneous databases, storage systems, and applications with minimal latency.

Google Earth Engine: Google Earth Engine is a platform for global-scale analysis and visualization of geospatial datasets. Google Earth Engine can be used with custom datasets, or with any of the publicly available satellite imagery hosted (and ingested on a regular basis) by Earth Engine Data Catalog.

**Looker (Google Cloud core):* Looker (Google Cloud core) is a business intelligence and embedded analytics solution hosted on Google infrastructure. With Looker (Google Cloud core), customers can build semantic models using various data sources, develop customized insights from the models, and share those insights for collaboration via dashboards and other services.

**Looker Studio:* Looker Studio is a data visualization and business intelligence product. It enables customers to connect to their data stored in other systems, create reports and dashboards using that data, and share them throughout their organization.

- Looker Studio Pro: Looker Studio Pro is a paid edition of Looker Studio that adds enterprise governance, team management features, and other features listed at https://cloud.google.com/looker-studio/ or a successor URL. Unlike Looker Studio, Looker Studio Pro is eligible for partner resale.

Pub/Sub: Pub/Sub is designed to provide reliable, many-to-many, asynchronous messaging between applications. Publisher applications can send messages to a "topic" and other applications can subscribe to that topic to receive the messages. By decoupling senders and receivers, Pub/Sub allows developers to communicate between independently written applications.

Data Analysis is one of important aspect which is required to effectively use the Big Data in construction industry.It is required for effective planning and managing the construction activities. Above mentioned Data Analytics tools available on GCP can be proved very effective mechanism to explore the various dimensions of construction data.

6.3 GCP AL/ML Services

AI Solutions

Anti Money Laundering AI (AML AI): AML AI enhances financial institutions' legacy transaction monitoring systems with an AI-powered risk score to improve financial crime risk detection.

Contact Center AI (CCAI): CCAI uses AI to improve the customer experience in contact centers. It includes Agent Assist, Dialogflow Essentials, Dialogflow Customer Experience Edition (CX), Insights, Speech-to-Text, Text-to-Speech, and Speaker ID.

Contact Center AI Insights: Contact Center AI Insights helps customers extract value from their contact center data by identifying sentiment and topics and highlighting key insights in the data.

Contact Center AI (CCAI) Platform: CCAI Platform is a contact-center-as-a-service (CCaaS) platform leveraging CCAI. It integrates directly with CRMs and queues and routes customer interactions across voice and digital channels to resource pools, including human agents.

DialogflowEssentials(ES): Dialog flow ES is a development suite for voice and text conversational apps that can connect to customer applications and telephony and digital platforms.

Dialogflow Customer Experience Edition (CX): Dialogflow CX is a development suite for creating conversational AI applications including chatbots and voicebots. It includes a visual bot building platform, collaboration and versioning tools, bot modularization tools, and advanced IVR feature support.

Discovery Solutions: Discovery Solutions enable customers in retail, media, and other verticals to deliver Google-quality search results and recommendations.

- *Recommendations AI: Recommendations AI enables customers to build a personalized recommendation system using ML models.
- *Recommendation Engine API: Recommendation engine API is the Version 1 API of Recommendations AI.
- *Retail Search: Retail Search, powered by Google's Retail API, allows retailers to leverage Google's search capabilities on their retail websites and applications.

Document AI: Document AI is a unified console for document processing that lets you quickly access all document processing models and tools. Customers can use Document AI's pre-trained models for document extraction, including OCR, Form Parser and specialized models.

- Document Workbench: Document Workbench allows you to build a custom classification, extraction or splitting model.
- *Human-in-the-Loop AI: Human-in-the-Loop AI uses Document AI to provide workflow tools for human verification of data extracted from documents.
- Document AI Warehouse: Document AI Warehouse is a highly-scalable, fully managed data management and governance platform that integrates with enterprise document workflows to store, search, and organize documents and their metadata.

**Talent Solution:* Talent Solution offers access to Google's machine learning, enabling company career sites, job boards, ATS, staffing agencies, and other recruitment technology platforms to improve the talent acquisition experience.
Pre-Trained APIs

Cloud Natural Language API: Cloud Natural Language API analyzes text to identify entities, sentiment, languages, and syntax.
Cloud Translation API: Cloud Translation API automatically translates text from one language to another language.

Cloud Vision: Cloud Vision classifies images into categories, detects individual objects and faces, and finds and reads printed words.

Media Translation API: Media Translation API is a gRPC API that automatically translates audio from one language to another language (e.g., French to English) and supports streaming real time.

**Speaker ID:* Speaker ID allows customers to enroll user voice prints and later verify users against a previously enrolled voice print.

Speech On Device: Speech On Device deploys speech-to-text and text-to-speech services locally on embedded hardware and operating systems.

Speech-to-Text: Speech-to-Text converts audio to text by applying neural network models.

Text-to-Speech: Text-to-Speech synthesizes human-like speech based on input text in a variety of voices and languages.

Timeseries Insights API: Timeseries Insights API enables large-scale time series forecasting and anomaly detection in real time.

Video Intelligence API: Video Intelligence API analyzes videos to extract metadata, add annotations, and identify entities in a video.

Visual Inspection AI: Visual Inspection AI automatically detects, classifies, and localizes abnormalities foundin images to improve production quality and develop analytics.
AI Platform/Vertex AI

AI Platform Data Labeling: AI Platform Data Labeling helps developers label data and centrally manage labels for training and evaluating machine learning models.

AI Platform Training and Prediction: AI Platform Training and Prediction enables customers to easily train and deploy machine learning models.

AutoML: AutoML enables customers to leverage Google's transfer learning and Neural Architecture Search to build custom models using a variety of data types. AutoML Services include AutoML Natural Language, AutoML Tables, AutoML Translation, AutoML Video, and AutoML Vision.

Deep Learning VM and Container: Deep Learning VM and Container provides virtual machine and Docker images with AI frameworks that can be customized and used with Google Kubernetes Engine (GKE), Vertex AI, Cloud Run, Compute Engine, Kubernetes, and Docker Swarm.

Vertex AI Platform: Vertex AI Platform is a service for managing the AI and machine learning development lifecycle. Customers can (i) manage datasets and associated labels; (ii) build pipelines to train and evaluate models using Google Cloud algorithms or custom training code; (iii) deploy models for online or batch use cases; (iv) manage data science workflow using Vertex AI Workbench (also known as Notebooks); and (v) create business optimization plans with Optimization AI.

Vertex AI Neural Architecture Search (NAS): Vertex AI NAS leverages Google's neural architecture search technology to generate, evaluate, and train model architectures for a customer's application.

Vertex AI Vision: Vertex AI Vision is a service that allows you to easily build, deploy, and manage computer vision applications with a fully managed, end-to-end application development environment.

Generative AI Services

Generative AI Services include: (i) Services listed in this subsection, (ii) any Service identified as a "Generative AI Service" or similar designation in the Documentation, and (iii) any feature of a Service identified as a "Generative AI Feature" or similar designation in the Documentation.

Generative AI App Builder: Generative AI App Builder allows customers to leverage foundational models, conversational AI, and search technologies to create multimodal chatbots and search experiences.

Generative AI Supporton Vertex AI: Generative AI Support on Vertex AI includes features for generative AI use cases, including large language, text-to-image, and image-to-text models that are available in Model Garden and Generative AI Studio.

AI/ML like any other industry has immense potential as a tool in the field of construction. Various AI and ML services available on GCP, as mentioned above, offers tremendous opportunities to process Big Data in construction industry.

These can be to predict properties of concrete, contract management, site safety and injury prediction, delay risks management, BIM integrated on-demand site monitoring, and other areas of construction engineering and management.

VII. CONCLUSIONS

- The construction industry is yet to reap the true benefits of using big data aptly. Over the last two decades, the rapid growth of big data technologies has caused a spike in the number of models and platforms that have been developed for increasing digitalization across different fields. However, the same level of digitalization has not truly been harnessed or integrated by the construction industry.
- However, the state of implantation of adoption in construction is below par. Therefore, the utilization and commercialization of big data to benefit the construction industry are crucial. An extensive literature review enabled us to identify the potential of big data in construction as the industry generates huge amounts of data daily and can greatly improve using the latest technologies.
- The development of online tools and software which enable infrastructure modeling and CAD is a crucial step in the right direction for futuristic constructions.

- Having explored the existing ML tools, we found that these tools, coupled with big data, can be applied in the construction industry.
- There are currently various gaps and pitfalls that act as barriers to using big data to its full potential.
- Firstly, data generation is much faster than the tools available for processing it. Moreover, big data integration into the construction industry is quite an uphill task even with the existing data processing tools.
- After going through this paper along with other papers which I reviewed, I want to further work on following areas which I still think that need more focus and research .
  - Big Data use to Improvement the Productivity and Performance Improvement
  - Big Date use for effective construction management
- In order to meet the above 2 objectives , my approach would be to develop some processes and protocol by using various analytical tool, storage mechanism and ML/AI processes available on Google Cloud Platform (GCP)
- Various Capability like Compute, Storage, Data bases, Data Analytics and AI/ML services available on GCP can be proved as game changer for the effective utilization of Big Data in construction Industry.

REFERENCES

1. Big Data in Construction: Current Applications and Future Opportunities https://www.mdpi.com/2504-2289/6/1/18 By Hafiz SulimanMunawar; Fahim Ullah; Siddra Qayyum;DanishShahzad
2. Application of Big Data Technology in Construction Organization and Management of Engineering Projects https://iopscience.iop.org/article/10.1088/1742-6596/1616/1/012002 Lijian Pei[1]Published under licence by IOP Publishing Ltd. Journal of Physics: Conference Series, Volume 1616, 3rd International Symposium on Big Data and Applied Statistics 10-12 July 2020, Kunming, ChinaCitation Lijian Pei 2020 J. Phys.: Conf. Ser. 1616 012002 DOI10.1088/1742-6596/1616/1/012002.
3. Review of Big Data Integration in Construction Industry Digitalization https://www.frontiersin.org/articles/10.3389/fbuil.2021.770496/full Omar Sedeeq Yousif1* RozanaBinti Zakaria1 Eeydzah Aminudin1 Khairulzan Yahya1 Abdul RahmanMohd Sam1 Loganathan Singaram1 Vikneswaran Munikanan[2] MuhamadAzani Yahya2 Noraziah Wahi3 SitiMazzuana Shamsuddin3.
4. Big Data in the construction industry: A review of present status, opportunities, and future trends https://pure.coventry.ac.uk/ws/portalfiles/portal/13749152/Big_Data_in_construction_Industry.pdf Bilal, M, Oyedele, LO, Qadir, J, Munir, K, Ajayi, SO, Akinade, OO, Owolabi, HA, Alaka, HA & Pasha, M.
5. Big data technology in construction https://www.e3s-conferences.org/articles/e3sconf/pdf/2019/23/e3sconf_form2018_01032.pdf Nikolay Garyaev* and VeneraGaryaeva.
6. Building Information Modeling (BIM) and BIG Data Analytics for Construction Industry https://www.academia.edu/70816308/Building_Information_Modeling_BIM_and_BIG_Data_Analytics_for_Construction_Industry Profile image of Yusuf ArayiciYusufArayici 2019.
7. Investigating profitability performance of construction projects using big data: A project analytics approach July 2019. https://www.researchgate.net/publication/334201799_Investigating_profitability_performance_of_construction_projects_using_big_data_A_project_analytics_approach Journal of Building Engineering 26:100850, DOI:10.1016/j.jobe.2019.100850 License;CC BY 4.0 Authors: Muhammad Bilal; Lukumon O. Oyedele; Habeeb O. Kusimo; Hakeem A. Owolabi;; Anuoluwapo O. Ajayi; Olúgbénga O Akinadé; Manuel Davila Delgado
8. Cloud computing in construction industry: Use cases, benefits and challenges https://www.researchgate.net/publication/347679916_Cloud_computing_in_construction_industry_Use_cases_benefits_and_challenges December 2020; Automation in Construction 122(1):103441 DOI:10.1016/j.autcon.2020.10

3441 ; License CC BY-NC-ND 4.0 Authors: Sururah Apinke Bello; Lukumon O. Oyedele; Olúgbénga O Akinadé; Muhammad Bilal;Manuel Davila Delgado; Lukman Akanbi; Anuoluwapo O. Ajayi; Hakeem A. Owolabi

9. Transformative role of big data through enabling capability recognition in construction https://www.tandfonline.com/doi/full/10.1080/01446193.2022.2132523. Bernard Tuffour Atuahene; Sittimont Kanjanabootra & Thayaparan Gajendran Pages 208-231 | Received 01 Aug 2021, Accepted 30 Sep 2022, Published online: 19 Oct 2022.
10. Big Data and Management https://journals.aom.org/doi/abs/10.5465/amj.2014.4002?journalCode=amj Gerard George; Martine R. Haas and Alex Pentland Published Online: 4 Apr 2014 https://doi.org/10.5465/amj.2014.4002.
11. An Appraisal into the Potential Application of Big Data in the Construction Industry https://www.researchgate.net/publication/325993944_An_Appraisal_into_the_Potential_Application_of_Big_Data_in_the_Construction_Industry/fulltext/5b323ea74585150d23d55f01/An-Appraisal-into-the-Potential-Application-of-Big-Data-in-the-Construction-Industry.pdf SitiAisyah Ismail; ShamsulhadiBandi and Zafira Nadia Maaz Department of Quantity Surveying, Faculty of Built Environment, Universiti Teknologi Malaysia Email: aisyaaahismail@gmail.com
12. Big Data Technology in Construction Safety Management: Application Status, Trend and Challenge https://www.mdpi.com/2075-5309/12/5/533 by QingfengMeng; QiyuanPeng; Zhen Li;andXin Hu. School of Management, Jiangsu University, 301 Xuefu Road, Zhenjiang 212013, China. School of Architecture and Built Environment, Deakin University, 1 Gheringhap Street, Geelong, VIC 3220, Australi.
13. Readiness of Applying Big Data Technology for Construction Management in Thai Public Sector http://www.jait.us/uploadfile/2020/1216/20201216030941159.pdf
Korb Srinavin1, Wuttipong Kusonkhum1, Boonyarit Chonpitakwong1, Tanayut Chaitongrat2, Narong Leungbootnak1, and Phatsaphan Charnwasununth3 1 Department of Civil Engineering, KhonKaen University, Khon Kaen, Thailand 2 Faculty of Architecture, Urban Design and Creative Arts, Mahasarakham University, Mahasarakham, Thailand 3 Department of Housing, Faculty of Architecture, Chulalongkorn University, Bangkok, Thailand Email: korbsri@kku.ac.th, w.kusonkhum@gmail.com, boonyarit_por@kkumail.com, Tanayut.c@msu.ac.th, Narongl.fec@gmail.com, Phatsaphan.C@chula.ac.th
14. Big Data is a powerful tool for environmental improvements in the construction business https://iopscience.iop.org/article/10.1088/1755-1315/90/1/012184/pdf To cite this article: AleksandrKonikov and Gregory Konikov 2017 IOP Conf. Ser.: Earth Environ. Sci. 90 012184
15. Research on Construction Methods of Big Data Semantic Model https://www.iaeng.org/publication/WCE2014/WCE2014_pp213-218.pdf Li Kang, Li Yi, LIU Dong.
16. Big data in building energy efficiency: understanding of big data and main challenges https://cyberleninka.org/article/n/722385.pdf NatalijaKoseleva*, Guoda Ropaite Faculty of Civil Engineering, Department of Construction Technology and Management, Vilnius Gediminas Technical University, Sauletekio al. 11, 10223, Vilnius