APPLICATION OF BIG DATA TECHNOLOGY IN THE CONSTRUCTION INDUSTRY IN KENYA: A CASE STUDY OF NAIROBI COUNTY.

BY

WAGUTHII MUCHOKI JEDIDA

B53/12221/2018

A RESEARCH PROJECT SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS OF THE DEGREE IN MASTER OF ARTS IN CONSTRUTION MANAGEMENT

OF

UNIVERSITY OF NAIROBI

JULY 2021

DECLARATION

This research project is my original work and has not been presented for a degree or any other award in any other University.

Signature……..……. Date…2ND JULY 2021 …..

Waguthii Muchoki Jedida

B53/12221/2018

This research project has been submitted in partial fulfillment of the examination requirements with our approval as the University of Nairobi supervisors.

Signature……..…….. Date……..……..……..

Professor Robert Rukwaro.

Signature……..…… Date……..……..……..

Irene Nyokabi Wamuyu

DEDICATION

This research is dedicated to my dear four children, my parents, my siblings, and my spiritual mum and pastor for their unrivalled support, encouragement, prayers and believing in me throughout my studies and always urging me on whenever things seemed tough.

God bless you.

ACKNOWLEDGEMENT

Above all else, I thank the Almighty God for granting me the ability and resources to pursue my course. I would want to give my earnest thanks to each and every individual who added the acknowledgement of this project.

I owe my sincere gratitude to my two Supervisors, Professor Robert Rukwaro and Irene N. Wamuyu for their advanced insightful help, kindness, and giving guided direction towards the realisation of this research study. All time they availed themself as and when I required help. Their productive views and remarks formed the detailing of this research study.

Many thanks also go Professor Paul Syagga for his insightful comments, guidance and constructive criticism. His key notes believed in the fruition of this study. I would likewise impart the recognition of my work to Doctor Mary W. Kimani whose sentiments during the first proposal presentation of this study provoked and gave confidence and significant influences on this work.

To all of you, your astute and informative remarks and proposals on my draft propelled and supported an enormous aspect of my last corrections and commendably improved the standard of this research project.

Unique thanks and appreciation to my Dad and Mum for their comments and continuous encouragement throughout the study. To my four children I know it was not easy for you to understand why my several absences in the house but I say thank you very much for the support. To my spiritual Mum and Pastor, Josephine Wambui Nyamu thank you for the spiritual support and prayers. And to all my classmates and friends who in one way or another were instrumental in the success of this study, receive my sincere gratitude.

I also thank all the respondents' during the pilot and the main fieldwork study who were mainly the various construction professionals for allowing me time off their busy schedules and furnishing me with all the fundamental data that empowered me complete my hands-on work quickly.

Regardless of the insights and backing of the above-named people, any error(s) of commission and omission, as well as oversight remain exclusively my duty.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

ABBREVATION

BDT – Big Data Technology.

BD – Big Data.

QS - Quantity Surveyor.

BIMa - Building Information Management.

BIMo - Building Information Modelling.

ICT - Information and Communication Technologies.

IT - Information Technology.

ML – Machine Learning

AI - Artificial Intelligence

AEC – Architectural Engineering and Construction.

DPL - Defects Liability Period.

IoT - Internet of Things

BORAQS - Board of Registration of Architects and Quantity Surveyors of Kenya.

ICPMK - Institution of Construction Project Managers of Kenya.

NCA – National Construction Authority.

RoI - Return on Investment.

BSA - Building SMART Alliance.

SAIT - Southern Alberta Institute of Technology.

AWS - Amazon Web Services.

NBS - National BIM Standard.

MoTIHUD & PW - Ministry of Transport, Infrastructure, Housing Urban development and Public Works.

ABSTRACT

The mere ability to capture large amount of data digitally, store, clean, analyze, and retrieve insightful information from the comfort of one's keyboard in Kenya construction industry ignited this study. The focus is particularly on the finished and ongoing projects, each at its own capacity. Further this study was informed by the thought of a centralized digital retrieval system of information concerning all construction projects. As such, BDT platform creates the single most source of the truth through digital data handling and management. Hence the need for digitizing construction processes by integrating all BIMa processes, real time data, social media information, and internet of things to a BDT platform for central storage and retrieval. The purpose of this study was to investigate the application and the viability of BDT in the construction industry in Kenya. The overall objective was achieved by: examining the status of application of BDT in construction industry in Kenya; identifying the factors that determine the application of BDT in construction industry; and determining the frameworks to use in integrating application of BDT into the construction industry. The study was conducted through a cross-sectional descriptive survey method. Primary data was collected from construction consultants working at the NCA which is a contractors regulatory body, MoTIHUD & PW, AAK which is a professional body, private sector consulting firms, construction companies and universities, randomly selected using questionnaires for data collection. Data was analysed using Statistical Package for Social Sciences (SPSS). This study did point out on the general qualification information of the respondents, BDT extent of adoption, the education teaching on BDT, the benefits of and the limitations to BDT adoption, BDT and integrating platforms. Top findings were that BDT has not been adopted in Kenya Construction industry, Tertiary institutions in Kenya are not training students on BDT and thus the skill gap in the industry, social media ranked the top most known and used integrating platform in use in Kenya followed by BIMa. This calls for the integration of operations which are currently segmented. The study recommended that the tertiary institutions of the built environment schools to come up with a BDT curriculum. Practicing consultants can be enlightened of BDT through workshops and seminars. Government policy on BDT needs to be formulated so that all industry players are keen to wire project data to a central repository.

CHAPTER ONE

INTRODUCTION

1.0 Background and context of the study.

Construction processes produce voluminous, complicated data sets. Successfully handling this huge data sets to ensure its accessibility and at the same time retaining its true image is vital in the management discipline. Missing and poor data sets would easily make contracts land in time overruns, costly way forwards, and entirely failing to handover the planned final product (Hendrickson, 1998). This therefore calls for more effort and well thought approach in actualization a detailed digital data management system.

Information is normally analysed to produce knowledge, organized to give enlighten its consumers or used to yield intelligence of giving informed directions. From the technology perspective, huge data sets have acquired the 'new oil' terminology (Data in the Construction Industry, 2019). This can be achieved throughout the project cycle towards mitigating schedule breaches.

Further, data analysing has the potential to depict trends, evaluate production, enhance sector competition, and boost allocation of inputs, amongst many others. Even so, there are digital frameworks that are used to collect, store, analyse and distribute the rapidly increasing high volume amounts of data, which is termed as 'Big Data' (DitCI, 2019).

However, Hendrickson (1998) has highlighted that, while there may be expensive implications as a result of missing or inaccurate data sets, there are also considerable infrastructure and human resources expenses due to setting up data centres, storing, transferring, retrieving including manipulating other data sets. These commands for excessive amount of care from the construction project managers by employing such a qualified person. Such requirements may be scarce and expensive resources on any construction project.

Due to the huge data sets emanating from the construction processes, orderly arrangement of data sets management is important so as to eliminate noise from the information (Wilkinson, 1984). This calls for a centralized database management system. Therefore, the dire need of information management specialists (IMSs) in the construction industry.

1.1 Problem Statement.

Lack of one stop search engine for whatever construction information one may need pertaining a particular construction project, is a looming need that needs to be addressed in the wake of BIMa, cloud computing, IoT, real time data, social media and ICT adoption in the construction industry. In practice prior detailed research is normally done before design onset, which is very important because it guides any new design with important insights. Thus, information is needed from such like similar past and ongoing construction projects. However, this study imagines one has to do a lot of moving and phone calling to get data relating past construction projects. For instance, getting simple information like who were the consultants to a particular project translates to consumption of a lot of time that could be put to other relevant use.

Construction industry is diverse and has many players and has many activities happening thus, the need for digital integration for information sourcing and sharing. For instance, looking at the projects executed by foreigners, they normally go away with vital information which could help the Kenya's construction industry if it was stored at a central digital repository center. Hence, the need for digitizing construction processes by employing integrating frameworks and big data technology (BDT) for central storage and retrieval.

As such, project activities generate large amount of information, and mostly it is not in any presentable order and isolated. Information is often on hand written or on printed format which is later stored when the works are finished completed. Further, Hosey (2018) explains that, a lot of information is produced every day by all project participants and it is misplaced as the project gets executed, from inception stage to completion and handover. This information has very important details which have potential to safeguard developers, materials suppliers, contractors, and subcontractors. Data can also empower all construction consultants to remediate problems before they develop into bursting setbacks. Policy makers and the ruling government cannot be left out as vital consumers of information. It is at this point that integrating frameworks like BIMa comes in handy to collect and store all data in an unstructured form may be.

More so Hosey (2018) points out that, it is quite baffling to come to terms with that historical challenges are still the same ones encountered today by the project participants. A continual

observation is that the construction sector has uniformly dating back to historical dates, been facing schedule overruns, budget shooting, inadequate project management including imperfect document filing. For instance, BIMo is quite an achievement but still all BIMo projects needs to be brought under one platform thus forming a one stop shop so that anybody requiring information concerning whichever project mines it with ease and conveniently. This way BIMa becomes a subset or roots of feeding BDT and on vice versa BDT shall facilitate BIMa for any new project as a point of research and reference (a process here in referred as data mining).

The consistent poor performance is noted to have been improving in the recent time as construction organisations have realized that BDT in conjunction with the various frameworks in use are beneficial to them. Thus, making the potential opportunities and benefits of big data accessible. According to Hosey (2018) it is only in the past period that digital innovation has brought advancements of handling data while retrieving knowledge instantly as it is needed. However, there are key factors that cannot be ignored that affect the application of BDT and any other digital application like IT skills availability, project management skills, communication skills, hardware and software (H&S) initial and maintenance cost, H&S technology, statistical skills (Jadeja & Issa, 2017).

As such, one gets interest to know more on what exactly has been put in place for this extraction of information to be a reality. Hence, a dire need for a one stop shop concerning housing, commercial and industrial projects, roads projects, dam projects, water works and sewerage projects, among all others not mentioned in this study.

This way, insights generated from one project can be used to forewarn proposed future projects. The data collected and analysed from projects can highly improve productivity of any new project, and should be something for construction project managers to consider when starting a construction project (Akbar, 2018). From the foregoing, the main aim of this study is to examine the application & viability of BDT in construction industry in the Kenya context in order to better inform all the integrating frameworks that are already in use.

1.2 Research questions.

i. What is the status of application of BDT in construction industry in Kenya?
ii. What are the factors that determine the application of BDT in construction industry?
iii. What is the framework of integrating application of BDT into the construction industry?

1.3 Research objectives.

i. To examine the status of application of BDT in construction industry in Kenya.
ii. To identify the factors that determine the application of BDT in construction industry.
iii. To determine the frameworks to use in integrating application of BDT into the construction industry.

1.4 Hypothesis of the study.

H_o: BDT is not applied in the construction sector in Nairobi County.
$H_{1:}$ BDT is applied in the construction sector in Nairobi County.

1.5 Significance of the study.

To handle projects properly, construction project knowledge is needed so that it can be used by the following parties: construction project managers, contractors, policy makers, government, all construction consultants (design team), developers, material suppliers, in Kenya and globally. The need of project data by stakeholders shall help in embracing the importance of BDT in the construction sector which further curbs the challenge of geographical dispersion.

i. To the construction managers – Extraction of usable data assists project managers in managing the project team including resources for the works. They could also explore data from the budgeting software in coming up with the best decisions for the contract. This data has potential to be updated instantly and automatically, forthwith providing the project manager with real time knowledge that helps in making precise, quick decisions as required (Mark, 2013).

ii. To the Contractors - information makes works to be finished timely and within cost. This is possible since insights are provided from the details produced by

project participants, enterprise resource planning software, and various computerized project models. Accurate project data could assist contractors planning for inputs timely towards avoiding time lag (Mark, 2013).

iii. To the architects & quantity surveyors – big data will encourage more shared access of contract tools: designs and details, models, contracts and procurement schedules. These documents are further categorised and made accessible to all participants. This integration yields enhanced production including shortened processing period of documents and reports, which ultimately improves the core performance. Further, BDT will make sure all project participants are using uniform details to deliver their expected deliverables (Mark, 2013).

The lack of an efficient centralized digital retrieval system necessitates the need for this study. For instance, a) Suppose you wanted to know what really transpired during the design and construction of say the Telposta Towers, located along Kenyatta Avenue at the intersection of Koinange Street. Are you able to know without having to have a one on one sitting with project manager of this particular project? b) Again, after the end of the defects liability period it is difficult to even know who were the players of a project.

The findings of this study will therefore be beneficial to the above parties as well as other professionals, academicians and students with interest in the construction industry.

1.6 Scope of the study.

The centre of attention of this research is in studying the application and viability of BDT in the construction industry in Kenya. Nairobi City County was further chosen as the area to collect data from because it is the urban centre with -

i. The large firms that are likely to use BIMa and other integrating frameworks.
ii. Has concentration of information sources for interview areas.
iii. Again, Nairobi County is the most dynamic and fast-growing county in Kenya.
iv. Generally, has most organizations of the stakeholders in the construction sector.

These stakeholders are for instance the policy makers, construction project managers, architects, contractors, quantity surveyors, engineers amongst others.

From BORAQS website as on 23rd March 2020, there are a total of 791 registered QSs', 1354 registered architects, 246 registered QS firms, 411 registered architectural firms. ICPMK and AAK CPM chapters have 372 individual practicing construction managers, as at 25th March 2020 and 2447 registered engineers from ERB website. NCA has a total of 14,500 registered contractors, in all categories in Nairobi County as at 26th March 2020.

1.7 Definitions of terms.

Here below are the key terms utilized in this investigation:

Data; is raw, unorganized facts that need to be processed.

Big data; is a newly discovered concept describing huge amount of organised, partly-organised and unorganised data which can be extracted for knowledge and used in computerized software and models to perform high level analysis. BD is information originating from different applications. It can be structured, partly structure, or unstructured collected from people and machines such as sensors, computers, etc (Rouse, 2018).

Big Data Technology (BDT); BDT is an application made and used for analysing, processing and retrieving insights from huge data sets which conventional applications could not manage (Kiran, 2019).

Building Information Management (BIMa); is wide terminology defining the procedures of formulating and operating a framework that has digitized data of project (DitCI, 2019).

Building Information Modelling (BIMo); is a shared procedure enabled by different varied applications, where project details from inception to facility usage are shared digitally amongst project participants. This is model is a sound resource of knowledge that assists in making informed solutions to all arising queries (National BIM Standard, 2014).

Building Information Models (BIMs); are documents in no defined order that can be retrieved and shared digitally to act as baseline for giving solutions concerning a project (NBS, 2014).

Real time data (RTD); is information that is current and accessible instantly (Novotny, 2018).

Central digital repository platform; is a data centre which collects, stores, manages, further allowing accessibility of stored information (Xie, Krystyna, & Matusiak, 2016).

General purpose programming; is a coding language, because data processing machines cannot operate themselves. Human operators are needed to make them function by feeding them with coded instructions (Digital Technologies Hub, 2019).

Internet of things (IoT); refers to a system of linked gadgets that capture, wire information through installed sensors without the expert person input (Rose, 2019).

Data mining; refers to the exercise of obtaining current updates from large size of information. Also known as knowledge discovery in databases (KDD). Details retrieved from KDD are deemed latest and helpful (Data Mining, 2017).

Machine learning (ML); is whereby computers learn to make decisions instead of humans manually programming them.

Machine learning tools (MLT); are mathematical scientific software that enable models to be able to enhance, educate and perform in absence of adequate expert guide. Best MLT include: Google Cloud AI, Keras, and TensorFlow.

Artificial intelligence (AI); is the duplication of an expert knowledge in computers. They are conditioned to reason like people and imitate them. The terminology could be used digital devices that solve queries and educate just like a person does.

Structured Query Language (SQL); is knowledge-based commands acting as the foundation of the BDT generation. A warehouse cannot be complete without SQL (Verma, 2018).

NoSQL / non-SQL / non-relational; is a computerized system offering techniques of for storing and retrieving large collection of information in a computer. This data forthwith analysed using MLT and AI (Mohan, 2013).

Interoperability; is a feature on devices that enables them to share, access and exchange information across devices and models all the time with ease, at no limitations and reduced dependency (Definition of Interoperability, n.d).

Business Intelligence (BI) tools; are mechanisms that use sets of approaches and digital knowhow in preparing, presenting and assisting in analysing information. The analysed information enables managers and other equivalents in making most efficient information-based solutions. These are, enterprise resource planning system (ERP), online analytics processing (OLAP), cloud and open-source BI, predictive analytics (Jreport, n.d).

Smart city; is a city or municipality or town using of digital Internet of Things (IoT) sensors to capture information which is forthwith used in managing building facilities, amenities, systems and utilities well and profitably (Interesting Engineering, 2018).

1.8 Organisation of the study.

It was an applied research comprising a number of summative and formative components presented in the following six chapters: -

Chapter one covered the foundation of the investigation, the introduction to the study project, study queries and objectives, research hypothesis, research significance, investigation scope, and definition of terms.

Chapter two reviewed the literature on the status and extent of application of BDT worldwide. Factors that determine adoption of BDT and integrating frameworks in use that can feed the BDT system.

Chapter three outlined the research framework. Questions to examine the three investigation queries were formulated, and data collected using structured questionnaires.

Chapter four displayed the field results. The study hypothesis was tested in this chapter.

Chapter five discussed the findings from the main field study in relation to the literature review so as to highlight the contrasts and the supports.

Chapter six reported on the conclusions, recommendations of the study and suggestion of areas of further research.

CHAPTER TWO

LITERATURE REVIEW

2.0 Introduction.

This chapter reviews the available literature on BDT and on the possible integrating frameworks in use. There are not many researches works on BDT in the construction sector. Though, what is already documented is sufficient and has large effect in enhancing the sectors delivery and production. According to Deutsch (2015a) the built environment industry is a data intensive industry. The study will further develop a conceptual framework that shall guide the main field study. This research shall rely greatly on written books, peer reviewed journals articles, brochures, previously related research and various articles on the internet in undertaking the literature review.

2.1 Big data technology Overview.

2.1.1 Theoretical Foundation

As in other sectors, in construction sector too BD describes large datasets which are archived in a digital platform with a focus of being extracted later (Burger, 2019). In her explanation she states that, BD can come from people, computers, machines, sensors, mobile phone applications and any other data generating device or agent. Lacking in Burger's work is the mention that, gadgets like computers are powered by various software which is an important factor in upfront cost of BD infrastructure.

The size of data that can be termed as BD faces variance based on the experts' abilities and software and hardware in use. First experience of having large datasets in the organization will prompt upscaling towards innovative management of data because they are large sizes that demands notable attention (Magoulas & Lorica, 2009). However, in reference to Magoulas and Lorica (2009) overview, this study looks at possible integration of data from all projects in Kenya to one central platform.

Bilal et al. (2016) opines that, using scientific analysis on data in the construction sector has started happening. However, application of BDT in construction sector is still at the budding

phase slowing the extensive BDT usage. (Martínez, Marin and Vila, 2015) state that, construction sector performance depends on accessing relevant well managed data so the contract can succeed. As such, project managers should be equipped with software that can handle information effectively.

For this reason, the construction sector has over time resulted to using unified and integrated communication digital systems because the building sector realised the ability and enhancements the technology has brought in otherwise industries. A lot of effort has been put in coming up with applications that can handle data of different activities across all the project phases (Martínez et al., 2015).

2.1.2 Evolution and extent of use of BD in the construction sector worldwide

Countries like Canada, United States, United Kingdom, Malaysia, Singapore have in place BDT application as outlined below. Meaning BDT infrastructures have been set up in some of these countries either privately owned or state owned as discussed in the subsequent sections:

i. United States

According to (Lynch, 2018), ten years back construction industry almost tailed in the listing of sectors that were using digital tools in their operations. Meaning the industry was resistance to adopt to change. Recently there has been major adoption in the sector as construction entities have realized that digital tools are helpful in their operations.

A construction company's tools kit should have technology as a valuable tool, because construction is about office and site operations. Such a consideration has helped companies win jobs, increase productivity, and deliver better projects. This is the shift construction industry needs to implement to achieve massive demand of high performance in the industry (Lynch, 2018).

For instance, Suffolk Construction Company (SCC) has successfully adapted to deploying automated technology processes, and the attitude of digitisation has yielded largely to predicting works procedures that have reduced errors. SCC top level management has invested heavily by taking the use digitisation very keenly. Suffolk has achieved quite a lot by using

technology in analyzing datasets, predicting risks, and delivering successful projects (Lynch, 2018).

Dodge Data & Analytics LLC Company is North America's leading provider of analysis tools and applications for tasks related to integration software that are used in the construction sector. Industry stakeholders, building product manufacturers, consultants, and contractors use Dodge in finding hidden development possibilities which are followed and implemented on for improved company success (Dodge, n.d). Elsewhere in Oregon, because of using BD insights, the Portland State University and the Oregon State University managed towards achieving large reduction of construction costs (We Build, 2018).

ii. Canada

In Canada, BDT is a digital revolution that is assisting in solving problems, saving costs, implementing safety rules, and coming up with productive procedures, using successful approach. Construction firms in Canada are focused to achieving these benefits by using data in an organized manner. BDT in Canada is playing a core role in managing works (We Build, 2018). This technology has extended to and has enriched the construction sector which is key in economic and social growth for it significantly enhances the livelihood of people. However, this sector was tailing in innovation.

The orderly and exhaustive capturing and storing of data is becoming an essential tool that is reducing time and costs overruns in construction sector. A research done by the Southern Alberta Institute of Technology (SAIT) in Calgary found that using Building Automation Modeling (BIM) during life cycle of the construction projects results in notable cost reduction (We Build, 2018).

iii. United Kingdom

According to Akbar (2018) there has been much excitement about BDT in past few years. Many entities have changed their business model in order to effectively utilize BDT. As such, BD has rapidly entered construction industry and has made notable achievements in implementing BDT.

Historically, 95% of data in the building sector is scattered in all over. This percentage is huge showing how big the challenge of data management is in the construction sector. However, BDT has brought integration and extraction of useful insights (Akbar, 2018).

iv. Malaysia

Maaz, Bandi and Amirudin (2018) in their study stated that: The Construction Industry Transformation Plan 2016-2020 of Malaysian construction industry is concerned and focused to multiplying its production by applying BD. Therefore, dire need for carrying out notable research in the new concept of BDT. Further they stated that the perception of BD in the Malaysian construction sector is relatively limited. In reference to the above this study deduces that Malaysia is also in the process of founding BD in the construction sector.

v. Singapore Data Centre

Singapore is at the moment experiencing very impactful evolution in data management. Data centres are being set up at a very high rate and all over the country. Big infrastructure providers in the world are currently investing heavily in Singapore (Wood, 2019).

Wood (2019) further outlines the Key Market Developments at Singapore, as follows -

a. That Singapore data centre market will continue to see notable growth over the next few years, whereas during late 2019 and 2020, a notable amount of new capacity is expected to enter the Singapore market.
b. Huge potential in supporting BDT growth and structuring is offered by renowned big infrastructure providers in the world such Equinix and Digital Realty (companies that are global giants in multimedia). Later, this evolution shall slacken because Singapore data centre market will face set up challenges.

Wood's highlights leave this study in a state of believing that the Singapore data centres includes the construction industry sector. On overall from all the above discussed areas, most countries are in the budding stage of laying big data infrastructure towards fruition.

vi. Sweden

Big data has increased the demand for information management specialists so much that: Software AG, IBM, Microsoft, SAP, EMC, HP, Oracle, Dell, and Corporation have spent

more than $15 billion on software firms specializing in data management and analytics (The Economist, 2012).

According to Frost & Sullivan (2018) BDT solutions are being used in the construction industry, courtesy of the increased connectivity and internet usage. This way, construction companies are even generating and collecting more data than ever before. In this evolution data collection techniques are now many and more advanced. Currently, smartphones, drones (see Figure 2.1), wearables, jobsite sensors, telematics, and GPS systems on heavy equipment and other mobile solutions are the gadgets in use for data collection.

For instance, a drone can fly over a work site and within 15 minutes gather millions of pieces of data. The data is then plugged into 3D design models to make work execution go faster and keep work places safer.

Figure 2. 1: Drone at work.
Source: Wasonga, 2019.

The collected and stored data becomes massive and organisations are trying finding out how to arrange, handle and analyze it so as to give informed solutions towards improving their operations. AI and ML digital fields are advancing rapidly in the construction sector helping organisations to use combined datasets from within and from external environment in forecasting possible results on contracts (Frost & Sullivan, 2018).

5G connectivity has filled a huge gap in the construction sector making many leading companies to research further on it. Digital connectivity of various devices shall be a going

concern since it shall improve on information sharing. For instance, Volvo Construction Equipment, Sweden, is the first in the globe to try 5G phone digital connectivity jointly with mobile operator Telia Company. Thus, one would be interested to know about the overall implications of 5G technology to the construction industry (Frost & Sullivan, 2018).

vii. Kenya Overview

From the foregoing worldwide overviews it is clear that construction companies have started relying on data to run their businesses. The benefits and value addition to BDT adoption in organisations are many. However, BDT is a new and a relatively unexplored concept. For this reason this study earmarked on the following studies from the Kenya sectors. None of BDT study has previously been carried before in construction industry in Kenya.

Mbaluka (2013) found out that the banking sector in Kenya was beginning to build out road maps of where BD would deliver the most value within the broad set of technology investment. At that time many financial investments were cautious of making broad based investments in BDT because it was a new and relatively nascent field. Some banks had implemented BDT which they were piloting and experimenting. Others were planning to adopt BDT, however they were in the process of developing adoption technology and knowledge gathering about BDT.

Data governance and security, large data volume, speed, flexibility of data deployment platforms, cloud based solutions, understanding and utilization of BD, lack of sharing capacity and integration of a wider variety of data, lack of business support from top management are most of the challenges that kept on recurring from Mbaluka (2013) survey.

Three years after, Ndambo (2016) in his study found out that some financial firms in Kenya have big data warehouses and business intelligence tools for reporting on and analysing customer behavior to better anticipate their needs, and optimize operations. At this time financial firms were at the early stages of deploying BD management initiatives. BD adoption by the financial firms was driven by the desire to differentiate themselves from their competitors through various methods such as developing, testing and delivering unique products and services to their clients.

Construction sector in Kenya should therefore borrow a leaf from the financial sector. Despite the low level of adoption, various integration software frameworks to use are in place and being used in some construction organisations for drawing and design, cost control, construction management and security checks (Nyaga, 2016). They produce raw data (operational big data) which is further supposed to be analysed to analytical big data. To start with construction firms should in-house expertly manage these raw data. However, the field study shall establish the status and level of BDT adoption in Kenya construction sector.

2.1.2 Characteristics of big data

As outlined by Laney (2001) BD was initially characterized with volume, variety, and velocity. Other characteristics associated with BD have been realized and are veracity (i.e. how much noise is in the data), value and variability. Big data does not amount to any specified amount of data. BD is used to refer to any amount of huge information captured within a particular period.

2.1.2.1 Breaking down the Vs of big data

i. Volume: Massive amount of data coming from many different sources. Mobile applications, internet clickstream logs, social networks, automatic sensors used in IoT set up and machine-generated information are some of the many sources of data. This data can be stored undefined or semi-defined using data mining software as it awaits detailed analysis (Rouse, 2018).

ii. Variety: Big data is composed of various data types, such as: defined information in SQL storages, undefined information in text and document files held in the Hadoop systems or in the NoSQL systems, and pre-defined information such as web server logs or streaming information from sensors. Additionally, BD incorporates various, concurrent information origins (Rouse, 2018).

iii. Velocity: Refers to the speed at which BD is produced, processed and analyzed. Whereby, BD sets are edited immediately as opposed to many traditional data warehouses which updates on daily, weekly or monthly intervals. Speed is also key as BD analyzing opens up to areas where analysis procedures instantly finds trends on the captured information and use these

fields to produce knowledge. Machine learning (ML) and artificial intelligence (AI) are some of these fields (Rouse, 2018).

Figure 2.2 shows the primary characteristics of big data and their growth levels. For instance, data volume grows from megabytes, to gigabytes, to terabytes, to petabytes and even to exabytes. Data velocity grows from batch to periodic, to near real time, and to real time transmission. Data sources could be from a mere table to socio media.

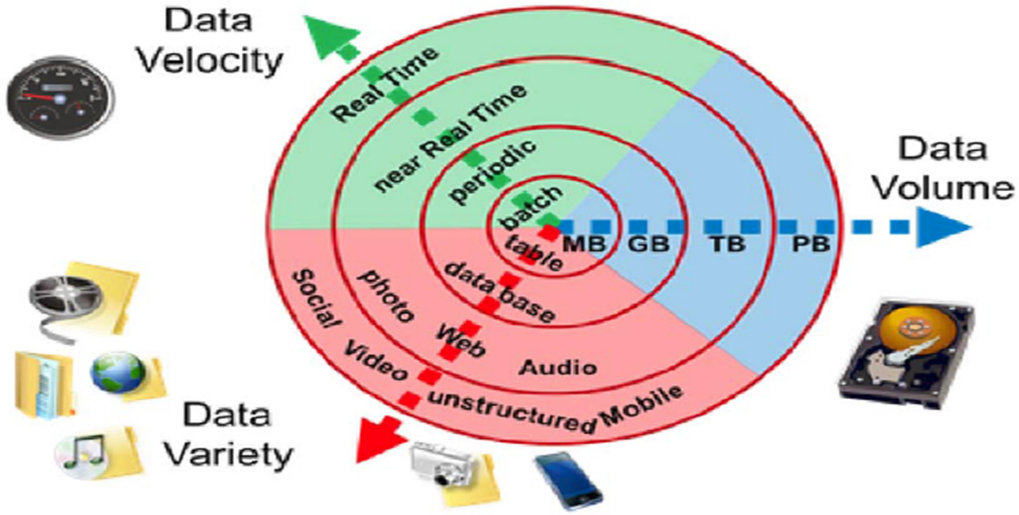

Figure 2. 2: Growth of BD's primary characteristics of volume, variety and velocity. Source: Kitchin & McArdle, 2016.

iv. Veracity: Data veracity refers to the degree of trust and accuracy that is in the captured data sets for it to produce valid results. Quality data is key to the users. It is good practice for IT and analytics teams to task companies to account for uncertain and inaccurate raw data before it is used in BD analytics applications. Use of inaccurate data can lead to critical quality issues that may be hard to box out later (Rouse, 2018).

v. Value: Data collected should relate to the relevant business issues for it to have real business value. Use of inaccurate data can weaken knowledge generated by BD analysing software. Data cleansing is therefore an important practice that companies should employ in verifying if collected data relates to the problem to be solved before experts channels it into BD analysis system (Rouse, 2018).

vi. Variability: Is data consistent or inconsistent. Consistency issues poses various ways in which BD can be used and formatted. Variability applies to BD sets which are not uniform as compared to the traditional transaction data. Variability may be having several definitions and

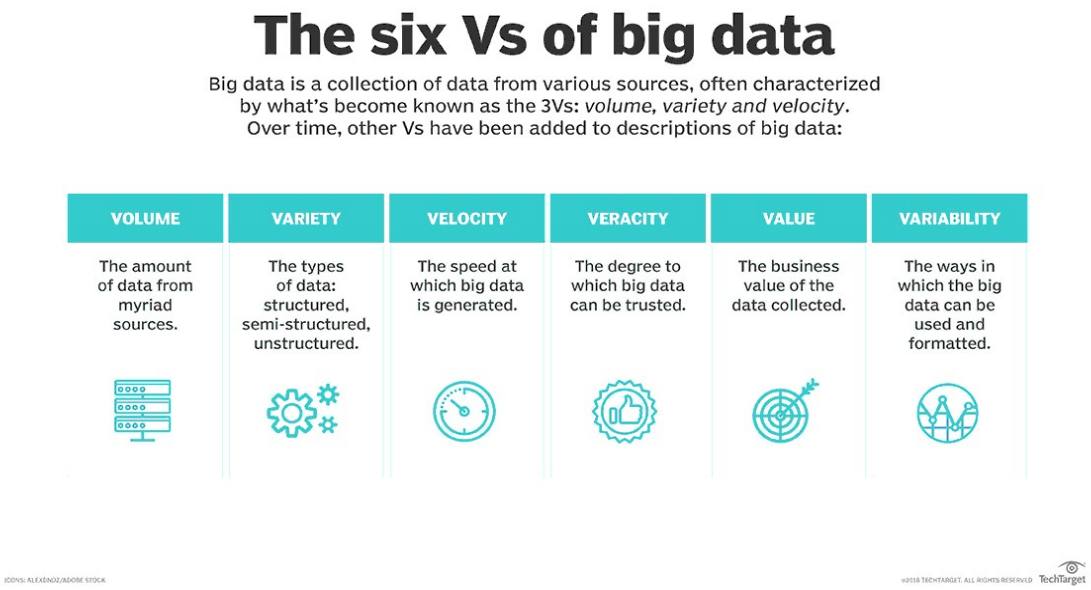

Figure 2. 3: The six Vs of big data
Source: Rouse, 2018.

can be organised in divergent forms depending on the various data sources. Data inconsistency complicates efforts to process and analyze the data (Rouse, 2018). Figure 2.3 shows the six Vs (characteristics) of big data.

However, data scientists and data consultants have come up with diverse lists of BD characteristics with between seven and ten Vs (Rouse, 2018). This study will find out if these characteristics are known in the Kenya construction industry.

2.1.3 Importance of big data

2.1.3.1 Usefulness of Big Data in the construction Industry

Akbar (2018) has outlined the following as the main ways in which big data (BD) has helped the construction industry so far –

i. Keeping track of time and better management.

ii. Brings knowledge thus better planning and eventually accurate improved timelines and budget estimates and superior understanding of costs. BD is therefore creating powerful construction managers that are using the tools at their disposal very skillfully.
iii. Lowers project risks that usually occur in projects due to many inefficiencies such as delays, litigations, and so no. Access to the relevant historical information lowers risks.
iv. Making the right choice. With the correct information at hand construction managers are well equipped to make the right choices. BD tools combined with the right qualified people can bring more insights to the projects. Eventually resulting to more information driven decisions for the construction projects.
v. Real time project management software have been developed. These software enable all executors to make decisions faster and advice the project participants in real time.
vi. Creates wealth of information. That needs to be analysed.
vii. Processes automation through machine learning helps find the root causes of persistent problems (from quality issues to safety risks)
viii. Helps in identifying construction trends and derive better outcomes on future projects.
ix. Solving problems, reducing costs, implementing safety rules, and organizing production processes in the most efficient way possible.
x. Enhances the productivity by shortening construction times, lowering costs, and reducing risks.
xi. Data analytics assists organizations to determine the most profitable projects to bid for and how to manage them efficiently (Frost & Sullivan, 2018).
xii. Helps to predict future outcomes on projects
xiii. Ability to promote the overall project delivery by improving processes.
xiv. Construction firms can benefit from heavy income by investing in BDT (Bilal et al, 2016).

Construction firms can harvest all the above benefits by using big data in a consistent way, because data is a junk of information that changes lives (Lynch, 2018). This study shall subject these benefits to field study to be rated by the respondents.

Megatrend and Innovations (2018) statistics further outlines why data is needed in construction industry by the players as follows -
57% of construction firms want “consistent, up to date financial and project information” whereas 48% want to be warned when particular occurrences happen. 41% want prior

forecasting so as to allow them to prepare well for best- and worst-case building events, while 14% want online analytics so that they see clearly the factors that are affecting revenue and by how much. All the foregoing reasons can be met by using big data which has the capability of data retrieval and derivation of insights.

2.2 Big data before, during and after construction.

The Construction Phases

According to We Build Value Digital Magazine (2018) big data is useful in all stages of the life of a construction project, which are: the design phase, the construction phase and after completion. Other in between important phases according to Koutsogiannis (2016) are the conception phase, pre-construction phase, procurement phase, post-construction phase. During these phases a lot of data is produced, that needs to be systematically captured for future retrieval as need arises.

2.2.1 The conception phase

Normally, the conception of the project starts when the client tables their dreams before the design team in leadership of a project manager. Ideally, research for the right project location and project specifications/standards are outlined (Koutsogiannis, 2016).

2.2.2 Design phase

There are usually four distinct interdependent steps in the planning stage of the project. They are programming and feasibility study, schematic plan, design detailing, and preparation of contract documents which are used in the construction field by those placing bids to work on the project. During the programming and feasibility step, all goals and objectives of the project must be clearly stated. Various decisions are made at this stage, including the scope of the project and how every space will be utilized (Koutsogiannis, 2016).

Data gathering in this phase is intense involving the project information, analysis of stakeholder indications, environmental information, and social media discussions between citizens impacted by the project. The gathered data is vital since it can influence the position and the orientation of the building (We build, 2018).

2.2.3 Construction phase

This stage of the construction project is preceded by two critical steps:

i. The pre-construction phase comes in first. Whereby, the project parties are required to visit the site in order to do a joint site investigation. The site inspection is supposed to allow the project parties to inspect all environmental challenges that may occur when project implementation is on course. It is usually a lengthy exercise since all concerns and opinions should be heard, documented and addressed (Koutsogiannis, 2016).

ii. Thereafter the procurement stage: of ordering and obtaining materials, equipment, and workforce (laborers) sets in. Most of the work is performed by the main contractor but sub-contractors are also in-charge of their areas of specialty (Koutsogiannis, 2016).

iii. Then the construction phase: Before the construction work begins, a pre-construction meeting is done to ensure that everyone is on the same page when the construction starts. This meeting normally includes minutes' information about the following topics: how to access the job site, the quality control of the project, how and where to store all the materials, the hours that everyone will be working. Construction stage demands for a lot of coordination of the roles and activities of the main contractors and subcontractors. It easily becomes evident that bad planning at this point can lead to serious delays and budget overruns (Koutsogiannis 2016).

The expectation at the construction phase is to have planned everything so carefully that everything goes off without a hitch. Which rarely happens, as something always goes wrong during a construction project. Thus, to avoid problems while planning a construction project, some kind of digital solution should be used (Koutsogiannis, 2016). During execution stage the construction project manager and all other stakeholders are concerned on all factors that can contribute to making the job execution quick and efficient (We Build, 2018).

2.2.4 Post-construction phase

This phase of the construction project is divided into three important steps:

i. Commissioning of the new construction building: Inspection and snagging of the whole building is done. The project team also trains the client in operating and maintaining the newly built structure which is an important exercise for it will contribute to increasing the lifecycle of the project (Koutsogiannis, 2016).

ii. Owner occupancy: Client takes over the building and forthwith the warranty / defects liability period commences. A period when the client uses the facility and feels safe that there is enough time to use and test all the different systems, equipment, and materials that have been installed (Koutsogiannis, 2016).

iii. Closure: Is the last step in the lengthy process of a construction project. When the project team has the overall contractual agreements to make sure that the project is free from any type of legal burden. At closure, it is good practice to carry out a monitoring and evaluation exercise which helps the all project participants to detect any tasks that were not completed, analyse the cause and put together a list of insights for the future refers. A post project evaluation review is the baseline for writing a thorough project completion report (Koutsogiannis, 2016).

2.2.5 Operation phase

In this phase the quality of the work carried out is tested, and any future need for repairs are predicted. For instance, sensors networks are being placed inside completed structures to monitor the functioning and performance level of the structure. This level of information gathering can help avoid future risks and enhance the safety of its users (We Build, 2018).

Overall, every phase of a construction project is a chain of tasks, decisions, and use of various bodies of knowledge. Management needs depend largely on the size and type of the project but there are always some major steps that cannot be skipped. Flawless communication between stakeholders and fact-based decisions are two main principles in executing a construction project towards making sure that all project phases are implemented within the agreed period and cost (Koutsogiannis, 2016).

As such, adopting construction management technology is a good move towards managing all construction related risks. Project related documents should be captured digitally so that significant information is amassed and grows to notable volume with every new project. Thereafter, machine learning tools looks through to find and report trends in project level of

performance and flag out the risks from the amassed project data. Meaning, the more most relevant the information is put into the system, the more predictive the results (Lynch, 2018). This study will find out from the respondents at which phase would they introduce BDT.

2.3 Big data challenges and how to manage.

According to Fujikawa (n.d) the handling and integration of large data sets is an involving exercise. There are several challenges that one faces during the integration, such as: data capturing, data analysis, content curation, data sharing, data search, data visualization, information privacy, data transfer, data querying, data updating and data storage.

The central components of the are to handle the data in new ways from how the traditional relational database handled. Confident decision making will be arrived at through managing big data in an accurate approach. In a Fujikawa (n.d) the six challenges that can be faced during the implementation of integration process, are discussed here below. These challenges must be considered and should be taken care of if one is going to manage any big data platforms.

i. Uncertainty with data management tools: Big data management has options of using many types of innovative data management tools and frameworks whose designs are only for supporting operational and analytical processing. The many types of NoSQL tools, developers and the disruptive status of the market causes uncertainty with the data management (Fujikawa, n.d).

ii. Skills gap in big data: The truth is that there is lack of skills in the market for BD technologies. The available experts are few and have gained experience on job via sheer luck of a platform that needed to be structured. The big data management discipline demands for high value in analysis and stable technologies that can be relied on (Fujikawa, n.d).

iii. Feeding data into a big data platform: Big data hoisting involves transmission, access, loading and delivery of data from many diverse sources into a big data structure. This is a complicated exercise where huge amount of data is feed to the platform for it to be managed through analyzing and processing (Fujikawa, n.d).

iv. Syncing risks across many data sources: Once data is imported into a big data platform; data copies are migrated from many sources on different sizes and speed. As such, data

schedules can quickly get out of the synchronization with the original source. The traditional data management and data warehouses, the order of data transformation, extraction and data migrations can birth situations that are risky by making data to become unsynchronized (Fujikawa, n.d).

v. Ensuring data availability to the data consumers as needed: Big data involves availing of data to users, increasing existing storage of data and allowing access to end users by employing business intelligence tools to connect to different big data platforms for the purpose of the discovery of data and extracting. Number of data consumers is deemed to grow over time demanding the need to support increased data capturing for many concurrent user accesses depending on different aspects of business process cycles needs (Fujikawa, n.d).

vi. Miscellaneous challenges while hoisting a big data structure are such as - upfront cost, processing of large amount of data at a reasonable speed so that information is available for data consumers without delay, the rate of transformation of data, veracity and validity of data (Fujikawa, n.d)

2.4 Skills required in BDT and Job roles.

According to Manuel (2017) herein below is a glimpse Figure 2.3 of the top BD skills that are most sought after by employers of analytics. Apache Hadoop skills has almost become synonymous to big data. Other BDT skills according to Verma (2018) are Quantitative analysis skills., Problem solving skills., and SQL skills.

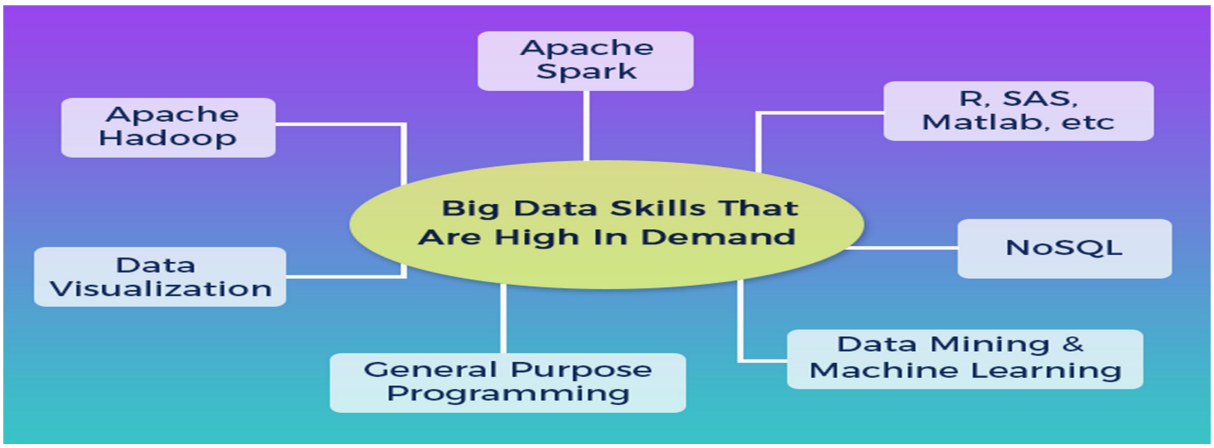

Figure 2.4: Big data skills
Source: Manuel, 2017.

It is becoming necessary for all professionals in whichever industry to have knowhow of

general-purpose programming. Since they are the people that identify a challenge in their line of operation and come up with a solution thereto. A case of the Australian Curriculum: Digital Technologies for secondary schools, has been prescribed (Digital Technologies Hub, 2019).

Job roles

Job roles in the big data field are as in herein below listed and in the subsequent illustration Figure 2.5. Data Analyst., Data Administrator., Data Scientist., Data Architect., Database manager., Data Engineer., and Data Consultants. Big data field is evolving all round birthing big data jobs of various shapes and sizes across all sectors. One just need to be ready with the skills to get the job that best tallies with their expertise and interests (Manuel, 2017).

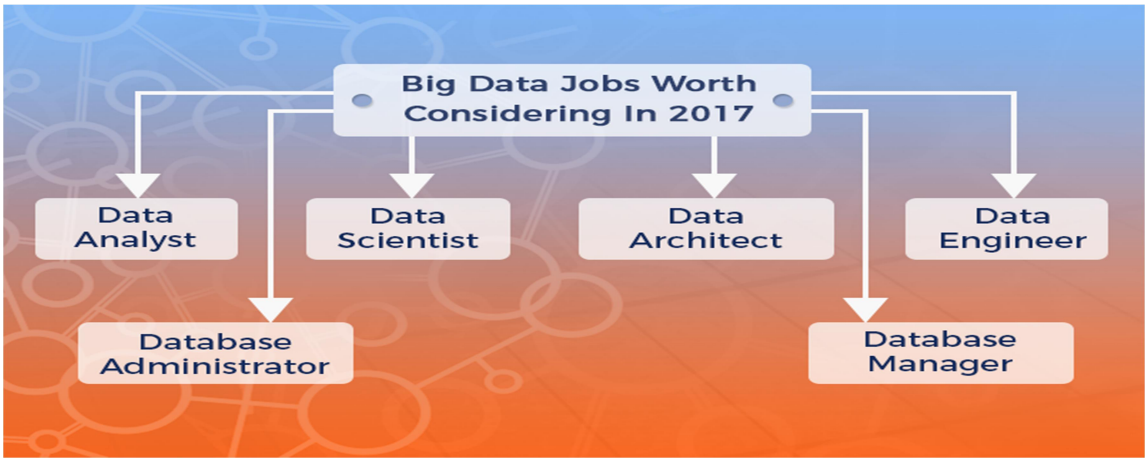

Figure 2.5: Big data job roles.
Source: Manuel, 2017.

2.5 Factors that determine the application of BDT in construction industry.

These factors are the related software and hardware initial costs, maintenance costs, and technology affect the implementation of big data. Further, skills in statistics and in IT, project management skills and communication skills are vital for the persons who work with big data platforms (Jadeja & Issa, 2017).

i. Relative advantage over the superseded technology

A new technology is considered to be advantageous if it is predominant regarding cost and functionality than the innovation it supersedes (Fichman & Kemerer, 1993). Big data innovations are open-source, less expensive to store, cheaper to handle, process detailed and huge volume of data. It is scalable and adaptable.

ii. Compatibility

New technology should match well with existing "values, skills, and work practices of potential adopters" (Fichman & Kemerer, 1993). Big data experts need to be trained on data management devices, and equivalent skills since BD is a new norm.

iii. Ease of use

The concept of big data takes place by storing data in Hadoop Distributed File System (HDFS) in raw form (Liu, Byna & Chen, 2013). At the HDFS there is reduced data movement such that once data arrives in HDFS from heterogeneous sources. All analytical tools, machine learning tools, and reporting tools run against HDFS.

iv. Flexibility to consolidate data from various sources

BD devices and innovations gives more flexibility to merge data from different sources into a central place (HDFS). The data that is merged into a central platform gives improved data mining and business insight capacities.

v. Capability to make instant decisions

One of the key abilities of big data technologies is analyzing data in real-time, or close to real time (Goes, 2014). Data is captured in real-time with streaming technologies transmitting it to the big data platform. It is then processed and predictive analytics performed with machine leaning tools.

vi. Security and Privacy Considerations

Data privacy and security are concerns with any technology. Richards and King (2014) stated that big data has issues in terms of security, privacy, confidentiality, and identity.

vii. Storage scalability and data processing

Scalability is one of the main competent characteristics for a successful data warehousing (Sen & Jacob, 1998; Sen & Sinha, 2005). The main advantage of Hadoop system is the ability to scale-out storage capacity (Aye & Thein, 2015). Big data technologies provide scalability in relation to capacity to store, information processing, and scalable mining platforms.

viii. Seller support

Seller support is significant because most of the big data devices and innovations are open source. Entities might want to make sure that seller support is available to assist them get their BD platforms running even when new versions of devices and innovations are rolled out

(Rahman, 2016). Sellers' attributes such as after training service, maintenance service, financial offer, after sales service, guarantee, warranty, relationship & motivation program, vendor size of organization.

ix. Entity size of skills workforce

Entity size in terms of workforce in IT skills and business size may affect the acceptance of the new technology. Big data comes with several devices and innovations and learning plus maintaining these devices and data requires ready skills in the workforce and other resources. An organization should have resources capacity, specialized technical experts and readiness to keep up another innovation (Rahman, 2016).

x. Managerial and individual users' factors

Dong, Neufeld, and Higgins (2009) argue that the organization management should encourage use of the new technology by influencing positive attitude and bettering the overall technology adoption uptake. Likewise, (Gambatese & Hallowell, 2011) found that top management support had huge effect of enabling implementation of innovation in construction firms.

Management support is the main factor that workers look on when adopting a new technology. Company's internal facilitating conditions and efforts positively encourages the intention to use the technology. Positive perceptions by the users on ease to use the technology and advantages thereto are some of the most important signals for behavioral intention to adopt and implement a new technology (Lu, Yu, & Liu, 2005).

xi. Internet connectivity

To monitor project site activities at real-time, prompt data transmission between the project sites and the central big data repository should be enabled. However, project sites usually have low bandwidth because they are not equipped with strong networking infrastructure. Use of advanced wireless sensor networks can solve internet connectivity issues in these types of big data applications. Making decisions from instant online data is more helpful for effective project monitoring than from old offline data (Bilal et al., 2016).

xii. Upfront costs

BDT is a cost demanding concept because entities are supposed to set up data centers, purchase software licenses, which demands huge initial capital outlay. Another cost is on the

skilled IT personnel whom must be engaged to keep the entire system running (Bilal et al., 2016).

xiii. Time required for training on big data model (Omran, 2016).

xiv. Exposure during under graduate education and several mentions in academic, otherwise lack of.

xv. Acceptance to change or resistance to change.

xvi. Government rules, regulations and requirements: For instance, on licenses.

2.6 Frameworks to use in integrating the application of BDT.

2.6.1 Use of the BD analytics in BIM processes

According to Building SMART Alliance (2017) in the most recent version released in July 2015, the National BIM Standard' definition of BIM includes three separate but linked functions, namely:

i. BIM is a business process function. For creating and managing information on a construction project across the project lifecycle.
ii. BIM is a digital representation function: of physical and functional characteristics of a built facility.
iii. Organization and control function. Where by the business value aspect of BIM on the company comes up referring to measurable economic effects of BIM on company performance, such as productivity enhancement, cost reductions, reduced errors and rework, time saving and so on. Also, BIM non-economic effects which are intangible and semi-tangible such as improved communication, better decision making, improved coordination among team members, and more efficient information management are realized (Cao, 2010).

As such, BD analytics can play an important role in BIM by introducing computational expertise in the construction industry by employing knowhow and applications from various computer aided engineered scientific models which have ability to analyse large data sets. The essence aim being to improve information, to find insights, to effectively forecast results, to automate procedures and analyses, and to perfect solutions (Baesens, 2014).

BD analytics can be applied to the construction industry business because of its dependency on information and its value for BIM processes (Heger, 2014; Mencarini, 2014). It is however important to understand that, implementation of BD analytics would demand many resources such as machines, cloud services, software, and expert people, which would cost as much to set up a BD infrastructure (CRC Construction Innovation, 2007).

2.6.2 Real Time Data (RTD) Application in BDT

RTD is data that is up to date and viewable the moment it is available. RTD is effective in the field of drawing tools and BIM. Meaning project managers can instantly access and guide by marking out on the design plans and drawings. Main contractor, subcontractors, design team and key stakeholders can easily communicate thus creating a more integrated project and an integrated labor delivery (Novotny, 2018).

RTD uses in construction projects are countless because it can be used to track material and fleet, show when workers clocked in and took their breaks, and how productive the site was. (Novotny, 2018). RTD can make schedule breaches and cost overruns a thing of the past.

Novotny (2018) further explains that, the construction industry is continuously learning the benefits of RTD in the industry. RTD in work place enables project managers to have consistent, accurate data, and reduces the guesswork typically involved. Guesswork leads to poor decision making thus making mistakes that reduce projects profit. With the use of RTD, outdated information which sets back construction projects is done away with. RTD is a vital input for all companies to adopt. Its adoption by the company's management ensures the decision-making parties make effective and fruitful resolves to the advantage of the planned construction works.

2.6.3 Cloud Computing Application in BDT

Cloud computing is a framework support that by request, gives access to the integration of all captured data (Chui & Manyika, 2010). Core aim being to give several users ability to use the data computing platform. This enables users to operate on one overall license. Cloud computing technology usage by many people has resulted to the emergence of BD (Qubole, 2017). Cloud computing is also enabling effective relationship of tasks in the usage of BIM models; hence it is being widely used in the sector of construction. BD is therefore set to

perform excellently during this time of intense innovation (Bilal et al., 2016). To add on, cloud computing and BD are the best combination that births a cost effective and a scalable platform in enabling BD and Business Analytics (Ferkoun, 2014).

2.6.4 Internet of Things (IoT) Application in BDT

The Internet of Things (IoT) has been the main support that has been initiating the big data phase, because IoT and BD are interdependent trends. In principle, IoT is a network of inter-connected gadgets that collect, wire data through installed sensors (Meola, 2016). Substantial number of sensors devices for data collection must be in place for IoT application to work. This way large amount of data is created, accessed and analysed in real-time in construction applications (Bilal et al., 2016). To add on, Pal (2015) at one time during the selection of a BD processing technology suggested that, huge amount of information produced by IoT triggers BD on a vice versa basis.

2.6.5 Social Networking Services Framework Application in BDT

Social media (like Twitter, LinkedIn, Facebook) is an amazing framework that should enable the building sector to enhance transmission of information amongst parties to the project (Jiao, Wang, et al., 2013). Nonetheless, a major challenge is to accept the value of the social media information and finding out how to analyse it (Chen, Chiang, & Storey, 2012). This is because social media networks produce large amount of diverse data.

As such, to thoroughly analyse data from social media, the analytical techniques of data analysis require improvements and inclusion of the all collected voluminous information into the huge data processor (Bello-Orgaz, Jung, & Camacho, 2016). In connection to this, BD can be used to create a performing field of software in respect to the data characteristics (or else the Vs) of the social media data towards improving the users production.

Other frameworks in use to feeding BDT are smart building technology and augmented reality (AR) technology coming from virtual reality (VR)

2.7 Big data platforms and Big data analytics software.

2.7.1 What is Big data platforms and Big data analytics software

PAT Research (2019) describes big data platforms as technology infrastructures with characteristics and operations of BD software in a central platform that is capable of creating, installing, functioning and administering BD. BD analysing application has ability of revealing unseen trends, not known relationships, vending patterns, user priorities including helpful data tips, collected from various varied sources.

BD infrastructures and BD analysing applications aims at offering effective analysis on voluminous captured data. The entire analysing exercise assists companies in getting enlightened. This is by converting collected information to valuable data, giving detailed tips and enlightens the entrepreneurial ventures. BD analytics assists enterprises in having an upper hand in the world technology and innovations (PAT, 2019).

PAT (2019) states that the conventional information handling methods and depository have no ability to effectively analyse large datasets. Therefore, organisations have resulted to fully using BD infrastructure which is a tool that has been developed by companies that specialize in setting up data centers. Main objective being to increase expandability, accessibility, production, and safety on the companies which are powered by the huge amount of data.

PAT (2019) further explains that, a BD infrastructure is created to hold huge amount of information which is collected instantly as it is produced. The BD system has the potential of performing various planned activities by various experts. For instance, data engineers use the infrastructure for sorting, merging, while eventually making analysing of datasets ready. Entrepreneurs use it to solve problems using generated insights, whereas scientists use BD to analyse trends from voluminous captured data by the aid of the programmed computer-based applications.

Figure 2.6 illustrates the general components of the big data technology.

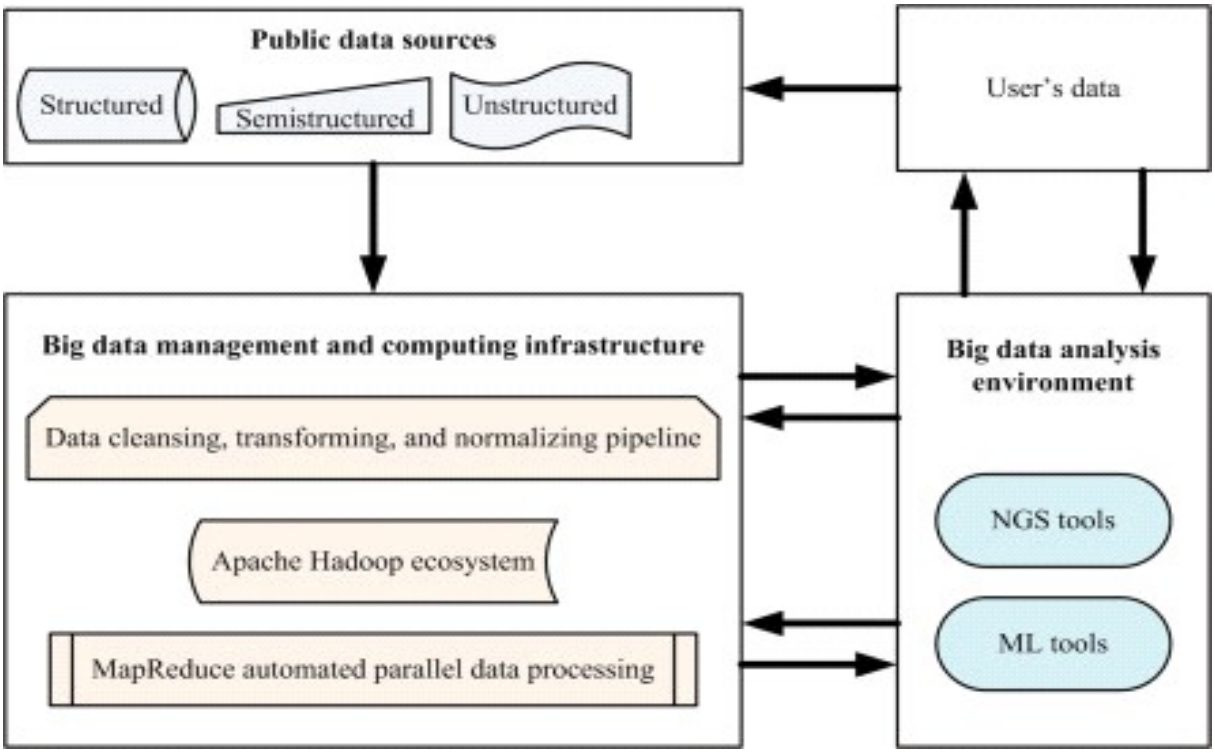

Figure 2.6: Describes the general components of the big data technology.
Source: Gharajeh, 2018.

In Kenya a national central BDT platform for the construction industry can be hosted by Architectural Association of Kenya (AAK), National Construction Authority (NCA), Board of Registration of Architects and Quantity Surveyors of Kenya (BORAQS), Institution of Construction Project Managers of Kenya (ICPMK), Engineers Board of Kenya (EBK) not limited to the emergence of a private owned entities (Author, 2020)

2.7.2 Advantages of big data systems and big data analysing applications

Further PAT (2019) outlines the benefits of big data platform as follows:

i. Offers precise information.

BD provides precise information to the enterprises and assists in settling on the correct way forward in an organization. The analysing device lessens the danger of getting incorrect information that happens because of managing raw data that has not been investigated. This causes the organization to settle on well searched decisions that are based on the correct realities. Settling on decisions while relying on inaccurate information is not safe. It negatively influences the growth and development of the organization.

ii. Productivity is enhanced.

Currently the number of millennials in the work place is high. This is the age group that is digital smart. This implies they grasp innovation and the greater part of their connection is digital based. They can get to any sort of information by a click of a button. A big data infrastructure can make it simple for them to hotspot for the data they would require. BD infrastructure acts in advance preparing for what lies ahead as it will generally provide data that will be required afterward hence time saving and increasing the productivity of the workers.

iii. Provides solution to difficult queries.

To control the projects, one should have the ability to address basic inquiries influencing the projects activities. Having a BD infrastructure set up would assist in addressing queries that project participants may raise, in a more informed way. Basic inquiries that would take a long time to give solutions can take shortest time to address.

iv. It is secure.

A big data infrastructure is a protected platform that ensures the wellbeing of your information. Security break on information is a risk that all organization fears. The BD infrastructure guarantees your information is secure. Thus, assuring business set ups safety from loss of income that would emerge from security violation.

2.8 Legal and risk issues that come with any digital framework.

Across industry sectors these are issues concerning: data rights, intellectual property rights, product liability, regulatory compliance, privacy vulnerabilities, high-profile data security incidents, weak or absent data rights, and cybersecurity threats are paramount and less well understood. These risks can be expertly managed (Sharron, Bosman, Do, & Lyon, n.d.). Further, (Mahajan, Parthasarathy, & Jain, 2018) outlines the below risks areas.

i. Technology failures / obsolesce risks

Technology failures and obsolesce can cause a company huge loss. Risks related with technology have great effect on platforms, users, and on procedures. Major threat is on expandability, interoperability, including performance precision of the digital system in use.

ii. Cyber risks

Technology environment need proper protection from unapproved access. Guarantee on honesty and confidentiality of the digital frameworks should assured. Key controls should incorporate platform hardening, network design and redesigning, application of security, management of weak points, and security monitoring.

iii. Privacy risks

Are risks emerging because of inappropriate handling of one's sensitive information, that may affect the privacy of a person. Incorporation of notice, one's decision to consent, accuracy, and other privacy standards should be the key control measures.

iv. Data leakage risks

The authors of data need assurance on data protection across the technology platform at various phases of information life-cycle. Protection on information on use, data in movement and on idle data. Control measures of this risk should revolve around data classification, data retention, data processing, data encryption, but not limited to.

v. Third-party control risks

Improper controls at the vendors and third-party operating environment can cause this risk. Data sharing, technology integration, operations dependency, vendor resiliency, and so on would be the best control solutions.

vi. Forensics risks

The capability of the technology's environment to enable investigation when security fraud occurs is vital. Absence of the enabling environment can prevent the investigators to capture data evidences that are presentable in the court of law.

vii. Organization's improper working processes

It is about the occurrence of internal or external happenings, that impacts the company's capacity to accomplish the business targets through its programmed activities. Some of these risks emerge because of lack of controls in the working methodology.

viii. Risk of disruption of the services

Is about disturbance of service accessibility, because of high reliance of the most demanded technology platform. To control service disruptions the host should consider scalability, IT/Network disaster recovery, cyber resiliency, and crisis management.

ix. Legal requirements

This risk demands adherence to the legal requirements including guidelines and laws of technology, and sectoral laws. Failure to adhere comes in handy with fines.

2.9 BDT and Construction Industry Key Performance Indicators (KPIs).

KPIs are quantifiable yardsticks which assist project managers and the team to measure and rule out if the works are being executed as anticipated. For the works to be executed within program, project managers should track, evaluate and comprehend the execution processes so that projects can achieve their inception stated objectives (Karola, 2015).

Over-promising and under-delivering is common in the construction industry. At times we have been involved in projects where all the outputs of performance were troubled. Contract sum was not making sense, programmed time was out of hand, making all participants to be stranded. Fortunately, KPIs, are helping negate these challenges. KPIs determines project goals better (Olive, 2019). KPIs assist in improving quality, reducing labour and material costs, and identifying strengths and weaknesses in the construction processes. Also, KPIs aim to track performances of construction project manager and supervisor (Hosey, 2018).

The variables of measuring KPIs include the contractual works program, construction cost, profitability, project management, material ordering, handling and management, risk management, quality assurance, client satisfaction (product and service perspective), site health and safety, time predictability (project, design, construction), productivity. Towards avoiding under delivering and over promising (Sibiya, Aigbavboa, & Thwala 2015).

These KPIs coupled with the application of BDT can make schedule breaches and cost overruns a thing of the past. This can be achieved throughout the project cycle towards mitigating time schedule breaches. BD assists works to be completed within program, within the signed contract sum. Big data offers guidance from information captured from assorted

origins, such as work force numbers and qualifications all got from existing workplaces; inputs cost implications and supplies logistics by the suppliers, organization workforce software, written interactions and enterprise resource planning system (Mark, 2013). Hence increased productivity and reduced document turnaround time.

2.10 Conceptual Model.

A conceptual model represents the researcher's understanding of the reviewed literature by outlining through a diagramed table or model for instance the connection of the all the identified variables based on the study objectives. It guides the author on the actions to take in their field study based on the previous knowledge of other researchers' point of view and findings on the study. In other words, the conceptual model is the researcher's guide of how variables in the research connect with each other (Regoniel, 2015).

The conceptual model outlined

This is guided by the study objective of finding out the extent of application of big data in Kenya construction industry. Whereby application of BDT is the dependent variable of this study. A close intertwined primary independent variable (IV) is the integrating frameworks available and in use for instance BIM processes which produces a lot of data that requires to the analysed (Heger, 2014; Mencarini, 2014). Figure 2.7 illustrates the conceptual understanding of this study around the research objectives.

Further, secondary independent variables of this study are as outlined herein below.

i. Initial capital outlay as outlined by (Bilal et al., 2016).
ii. Management support as explained by (Gambatese & Hallowell, 2011).
iii. Legal and risk issues as detailed by (Mahajan, Parthasarathy, & Jain, 2018).
iv. Availability of data analytic skills as explained by (Baesens, 2014).
v. Construction technology knowhow as outlined by (Omran, 2016). From the education curriculum and through the software manufacturers / sellers.
vi. Awareness of BDT through its characteristics, technologies, skills and fields (Rouse, 2018) and (Manuel, 2017).
vii. Nature of construction industry (Mark, 2013), (Magoulas and Lorica, 2009) and (Akbar, 2018). The construction industry processes needs to be integrated.

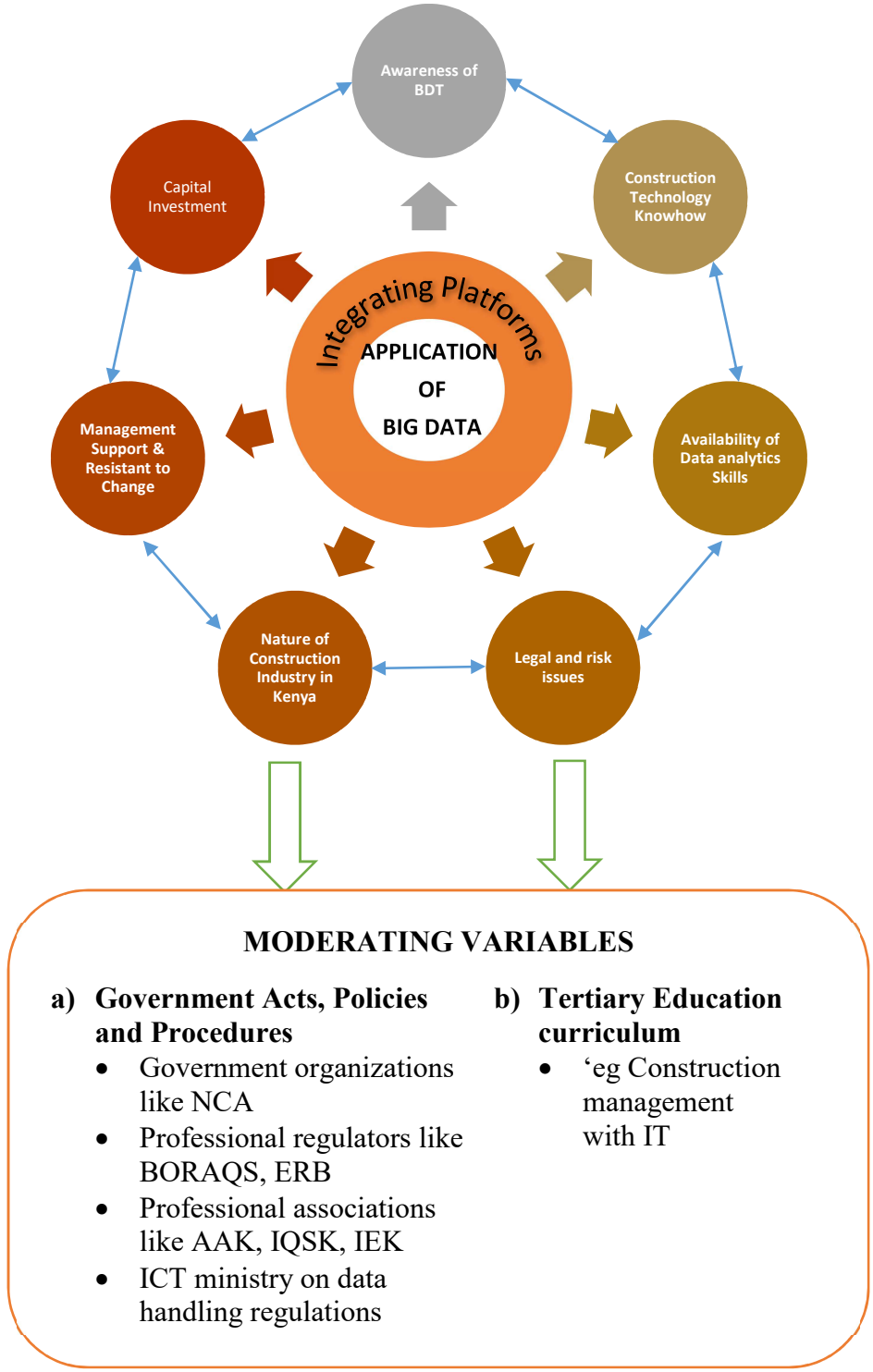

Figure 2.7: Conceptual Model of the Study

Source: Author, 2021.

The foregoing independent variables are closely interrelated, and are moderated by government policies, acts, procedures hand in hand with tertiary education curriculum that

starts to shed light on BDT knowhow.

The relationship of the conceptual model and the study hypothesis which is "big data is not applied in the construction sector in the county of Nairobi", is that the primary and all the intermediate independent variables could highly hinder the application of BDT in Kenya. The field study will be able to establish the true position.

BDT calls for change of business models in the construction industry, towards managing projects better. Regular performance checks should be done to the BDT platforms so as to allow for a stop and re-tool exercise.

2.10.1 Concept Definition

Table 2 1: Operational Definition

Variables	**Conceptual Definition. (Parameters to Measure)**	**Operational Definition. (How to Measure)**
Application of Big Data	- Is BDT known by consultants in Kenya? - Duration of use by the consultants.	- Selection options. Yes / No/ Planning to adopt/ Don't plan to adopt. - Ratio scale e.g. 1 – 5 years: Range choices
Awareness of BDT	- Consultants awareness of BDT characteristics, fields, skills, and technologies.	- Ticking listed choices.
Construction Technology Knowhow	- Is training happening at the education institutions and CPDs. - Use of BDT platforms like BIM, cloud computing. - Consultants awareness of BDT benefits. Improves project planning Quickens construction time Handles huge amount of data	- Multiple choices e.g. Yes / No - Ticking listed choices. - A Likert scale of five points will be used to measure the constructs under the variable.

Availability of Data Analytics Skills.	- Availability of computational expertise on ML, AI, Data Analysis, Data Mining, Data Science. Data management. - Are the experts listed above employed in construction companies?	- Selection options. Yes / No/ Planning to adopt/ Don't plan to adopt.
Integrating Platforms Availability and Use	- BIM Models, - BIM software, - Real Time Data, - Social media, - Cloud computing, - IoT. - Requirements and benefits that come with BDT platforms.	- Respondents are expected to select all the platforms they are aware of. - A Likert scale of five points will be used to rate the constructs under the variable.
Legal and risk issues	How for instance - intellectual property rights, information privacy, data security and legal requirements are weighted by the respondents.	- A Likert scale of five points will be used to measure the constructs under the variable. - Each respondent is to rate legal risks that come with BDT.
Nature of Construction Industry in Kenya	- Are consultants using integrating platforms? – like BIM, cloud computing. - Benefits of integrating various sources of data, challenges thereto.	- A Yes / No question will be asked. On advocacy of use of BIM in all projects. - A Likert scale of five points will be used to measure the constructs under the variable.
General Factors	- Managerial support - Strong internet connectivity - Skill gap - Formal education - Resistant to change	- A Likert scale of five points will be used to measure constructs under the variable.
Capital investment	- Initial / upfront investment and hosting costs thereto.	- A Likert scale of five points will be used to measure constructs

	E.g. on Computers, Phones, Drones, Sensor Networks, Software, Models, strong internet connectivity, employ BDT experts.	under the variable. - Each respondent is to rate their perception on upfront costs as a limiting factor.

Source: Author, 2021.

CHAPTER THREE

RESEARCH METHODOLOGY

3.0 Introduction.

This chapter covered the research design, data sources, sampling design, data collection tools and techniques, data analysis and presentation, and expected results (Rukwaro, 2016). All the highlighted sections were well framed so as to answer the research questions and achieve the set research objectives (Mugenda & Mugenda, 1999). Objectives of this study are –

i. To examine the status of application of BDT in construction industry.

ii. To identify the factors that determine the application of BDT in construction industry.

iii. To determine the frameworks to use in integrating application of BDT into the construction industry.

3.1 Research design.

According to Mugenda and Mugenda (2012) a research design is a procedural outline, a set framework and approach for investigating respondents to get answers to the study questions validly, objectively, accurately, and economically. The study adopted a cross-sectional survey research design which combined both qualitative study and quantitative strategies, so as to find out the extent of application of BDT in the Kenya's construction industry. Questionnaires and interviews during the pilot study were used to collect the primary research data.

A cross-sectional study (also known as a cross-sectional analysis, transverse study, or prevalence study) is a research study that is focused in establishing the true position by analysing data from an identified populace based on a given particular time. Cross-sectional studies are descriptive studies 'neither longitudinal nor experimental' It is a study that describes and expresses a concept (Schmidt & Kohlmann, 2008; & Lee, 1994).

3.2 Data sources.

The study used both primary and secondary data in an attempt to focus on the stated problem and address the objectives. The primary data used in this study was sourced from various consultants that are trained and experienced in the construction industry. This study targeted

response from people well versed with the construction work and in particular construction project management discipline in the construction sector such as the construction project managers, architects, quantity surveyors, and engineers. It involved use of structured and closed ended questions. Some questions though, had space for the respondent to comment further.

In addition to the questionnaires, data was obtained throughout the study from papers, other material subjects and ongoing and previous literature review. Secondary research was widely and additionally used for data collection during the study. As (Church, n.d.) states, secondary research is founded on existing published literature sourced from similar carried researches.

This study entirely used several internet search engines such as Emerald, Science Direct, Wiley Online Library, and Google Scholar among others in order to form an in-depth comprehension of BDT in the construction industry. Several published and unpublished PDF documents were sourced and downloaded.

Information collected using secondary research included construction projects that have employed BDT and organizations using BDT plus how BDT has been perceived elsewhere. According to (Bosch, n.d.; Ormsby, n.d.) secondary data can be sourced from the textbooks, encyclopedias, news articles, reviewed papers, meta analyses, authored scholarly articles, government releases, statistical databases, historical records, and that published and uploaded on the internet.

3.3 Sampling design.

3.3.1 Location of study

Nairobi City County was chosen as the area of the study because it is the urban centre with
i. Firms that are most likely to be using BIMa and equivalent frameworks.
ii. Has concentration of information sources for interview areas.
iii. Generally, has the most organisations employing or owned the stakeholders in the construction industry. These stakeholders are the likes of the construction project managers, architects, contractors, quantity surveyors, engineers, policy makers amongst others.

vi. Again, Nairobi City County was chosen because it is the most dynamic and fast-growing city in Kenya. This alone makes Nairobi City County a hub of construction activities. Most residents have relocated from elsewhere to Nairobi in pursuit of education and career. It is also Kenya's largest city with a population of 4,396,828 as at 2019 census (KNBS, 2019). The cities rate of growth is four percent (4%) annually and this is due to rural to urban migration in search of employment opportunities.

3.3.2 Unit of analysis

Unit of analysis is defined as "the element whose data is aggregated and analyzed in the study to make conclusions, decisions or inferences" (Mugenda & Mugenda, 2012). The unit of analysis is normally informed by the research questions. It is therefore crucial to determine the unit analysis before undertaking sampling process because, the units constitute measured variables that form the analyzed data (Babbie, 2001). Unit of analysis in this study was the individual consultants from a cross sectional of various organizations as explained in 3.3.3 here in next.

3.3.3 Population frame

The target population of this study was made up of the various trained construction consultants who have ventured into the various sub sectors within the construction industry in Nairobi County. These consisted of consultant respondents practicing in the fields of construction management, architecture, quantity surveying and engineering and are deemed to be rich source of information in the construction sector. Population frame is a complete listing of all the units of the target population from which the sample is selected (Mugenda & Mugenda, 2012).

The total population for this study was 4320 practicing consultants. Who were sampled from various organisations where they are employed. Which included 1233 NCA 1 to NCA 4 registered construction companies in the Nairobi County; this was obtained from the list of companies from the NCA register. At the same time this study sampled and interviewed some technical staff of NCA from a total of 97. Registered Architectural firms are 411, and 246 registered Quantity surveying firms both obtained from BORAQS. 186 individual registered construction managers with ICPMK. 119 registered engineering consulting firms obtained

from Engineering Board of Kenya (EBK). 530 Teaching staff of construction related courses from the 3 universities in the Nairobi County, namely University of Nairobi with about 200, Technical University of Kenya with about 300, and Kenyatta University with about 30. Department of Public Works which had about 300 technical staff and about 52 AAK main & chapters officials.

3.3.4 Sampling size

Mugenda and Mugenda (2003) states that a sample size needs to be not less than 30 persons. Further, according to (Warren, 2002; & Bryman, 2012), for a qualitative research to be publicized, it needs not less than 30 samples. This study shall focus to fulfill this threshold.

A number of consultants were sampled to be able to achieve the study objectives. The consultants were expected to give their insights and opinions on BDT use in executing their projects, parameters to be considered in acceptance and strategies that can be put in place to increase the uptake of BDT. Due to the large number of consultants, random sampling of organisations was carefully done such that the questionnaires reached the target consultants. Follow up for responses was carefully and consistently done.

The simple random sampling aimed at harvesting responses that were needed to give this study a general informed feel across the public set ups, regulating bodies, teaching institutions and private set ups on BDT use. This way broad coverage was achieved.

As such, sample size determination where population size is known, the following formula modified by (Syagga, 2018) from Mugenda and Mugenda (2003) formula, where Syagga proposed p > 50%. Meaning that the higher the target population above 50% with similar characteristics, the higher the reliability to arrive at a sample size from a given population.

$$n = (z*z)(p*q)N / e*e(N-1)+(z*z)(p*q)$$

Where n = sample size

z = standard deviation (1.96)

p = % target population assumed to have similar characteristics (say taken as 90%, the higher the % above 50% the higher the reliability)

q = 1-p (0.1)

N = population size

e = confidence interval (margin of error (say, 0.05)

As such out of a total population of 4320, and in reference to the above formula, a sample size of 270 consultants was calculated. This study targeted to interview two consultants from the randomly selected organisations. This decision was made due to time constraints. This study just required a representative enough from the sample size. This study targeted more than one respondent as a sensitivity caution sought by this study from every organisation. This means that a total of 270 questionnaires were sent out to collect primary data.

Adopting proportional allocation' of the sample, gives the researcher satisfaction in targeting a well allocated framework. Each category of the respondents is sampled through an allocation formula based on sample size. This way each category of the sample is accorded the same measure scale so as to come up with accurate populace representation towards collecting valued feedback (Mugenda and Mugenda, 2003).

Getting the sample size under different strata, using the formula; n1 2 3…. K = S (Pi / N 1 2 3 …. K) where;

n123…k = number of items to be selected from each stratum.
S = sample size required.
Pi = proportion of population included in stratum
N = target population

With the sample size of 270 denoted as 'n' calculated from a populace of size N = 4320 which is allocated to four categories: n1 for registered architects; n2 for registered quantity surveyors; n3 for registered engineers; and n4 for practicing construction managers. Therefore, the sample size under the four strata was as follows;

n1 = 270 (1286/ 4320) = 80, (registered architects) 29.78%
n2 = 270 (752/ 4320) = 47, (registered quantity surveyors) 17.39%
n3 = 270 (1930/ 4320) = 119, (registered engineers) 44.64%
n4 = 270 (353/ 4320) = 24, (construction managers) 8.19%

3.3.5 Sampling method

A sampling design is the method of selecting items to be discovered for a particular study from a populace (Kothari, 2004). The study employed simple random sampling method. Simple random sampling is used to make factual derivations about a populace (Thomas, 2020). The study identified all the consultants that would help in achieving the research objectives and randomly picked them. These respondents had to be practicing and based in County of Nairobi. The study was sure that the participants chosen had information on the big data technology.

The primary objective of choosing simple random sampling was to assist this study to zero in on specific attributes of a populace that were of interest, which would best empower the study to get responses to the research questions. Further, the type of the random sampling technique used by this study was expert sampling because the research needed to obtain feedback only from the practicing consultants in the construction sector.

The target expert practicing consultants comprised of a four strata as follows: the construction managers, engineers, quantity surveyors and the architects working in various entities based in the county of Nairobi as outlined in 3.3.3 here above. The research was careful to formulate and allocate the sample size against the four strata as calculated in 3.3.4 above. According to (Thomas, 2020) benefits of simple sampling is that it guarantees high internal validity. Randomization is the best strategy to reduce the effect of potential biased factors. Also, with a huge enough sample size; simple random sample has high external validity and as such it addresses the attributes of big populace.

3.4 Data collection tools and techniques.

Data collection was conducted using questionnaires and interviews. The main tool of data collection was questionnaires, which were further prepared to a google format document that enabled forwarding them to the respondents via email or whatsapp. The interviews were guided by the structured questionnaires especially during the pilot study which did target every category of sampled pull. Thus, it was closed ended interviews, because this study just required straight forward answers. According to (SurveyAnyPlace website, n.d) benefits of closed ended interview are: It is uncomplicated and interviewee respond fast to questions.

Comparison of the responses from various study participants is easily done. Responses are easily coded and analytically analyzed.

Face to face interviews and online google meets were employed. This allowed for observation of any non-verbal communication but also allowed for seeking any necessary clarification. The interviews were carried out in a conversational style. The questionnaire was divided into five categories; part one addressing the general information of the interviewees, part two was on BDT extent of adoption, third part was on the important components of big data, fourth part was on construction project data, whereas the fifth part was on the BDT and integrating platforms (Appendix B – the Questionnaire).

The questions in the questionnaire were close ended multiple-choice format as well as short answer questions especially where this study needed to seek further the understanding of the sampled respondents on the various BDT variables for ease of analysis and interpretation. The structured questions were rated using a two version of 5-point Likert scale assigning [5] Very important [4] Important [3] Uncertain [2] Slightly important [1] Not important. The second version was assigned as follows; 5-point Likert scale of [5] Strongly agree [4] Agree [3] Uncertain [2] Disagree [1] Strongly Disagree. In some instances, specifically Yes or No responses was sought. Further, more diverse answers like Planning to adopt, Don't plan to adopt, Other (please elaborate) were used towards gauging the consultants feel on BDT.

3.4.1 Pilot Study

A pilot study is an important exercise carried out by collecting data from selected few respondents to test if the questionnaire and other tools are ok or need improvements. Target response can be 10% of the sample size. This is done before going out to collect data for the main field study (Lancaster et al., 2004 & Kraemer et al., 2006). As such, a pilot study has been defined as a study prior the big study with a focus to check study procedures as documented, questionnaires, selection of respondents, and else study methods towards preparing for the main field data collection (Stewart, n.d).

Main field study was targeting to collect feedback from 270 respondents. However, in the pilot study 10% of the total respondents was targeted. Translating to about 27 respondents. As such, a total of 27 respondents were reached during the pilot study and 15 responded.

Experiences from the pilot study

i. Day one the study dropped the questionnaire via hand delivery and interviewed using the questionnaire one on one, however by the end of the day, the study resolved to the online google meet. This made this study formulate the questionnaire into a google document, that was sent to the respondents via email and whatsapp. Google meets were quite efficient in time management, and travelling cost.
ii. The pilot study observed that the respondents' working in contractors' companies and lecturers did not highlight that in their responses. This made this information to be added for the main study questionnaire.
iii. Average response time by the respondents was about 45 minutes, least time taken was 30 minutes, highest time taken was 60 minutes.
iv. Some keen respondents highlighted repetition or similarity of some of the questions. This did help this study in polishing up the questionnaire for the main field study.

3.5 Data analysis and presentation.

Data to this study was collected in ordinal scales. Meaning sequence of the responses is what is key and crucial, however the variances between two ordinal sets is not clearly recognised. In general, ordinal scales are estimations of nonnumerical terms like satisfaction, understanding, happiness, discomfort, to mention a some (My Market Research Methods, 2018). Further, the acceptable approach of calculating central tendency on a set of ordinal data is by using the mode or median; since the mean is indefinable from an ordinal set.

The primary data collected was systematically organized and analysed using descriptive statistics of weighted averages. This was done by using means, standard deviation, variance and percentages. The final data was thereafter analysed using two computer applications i.e. Microsoft Excel and Statistical Package for Social Sciences. The easiest approach of analysing ordinal data is using displaying methods such as tables, charts. Distribution tables of frequencies, percentages, standard deviation and variances, were summarized where diagrams such as multiple bar charts and pie charts were used during the analysis to present the results.

Tables refer to group of figures systematically presented in the format of rows and columns. Analytical tables were used to interpret figures as they are suitable for comparison. The data

was analysed and statistical inferences drawn.

Charts like bar charts were used to represent the data and analyse it logically. A bar chart comprises a number of spaced rectangles, which generally has their major axis vertical. They can be used to represent a large number of statistical data. Multiple bar charts were used too. This is more useful when there are more sets of comparable data to be compared and contrasted. Multiple bar charts not only make the actual number of items involved clear but also make it easy for the eyes to gauze the relative position of each group. This therefore makes interpretation easier to understand.

3.6 Hypothesis testing.

The hypothesis of this study is "BDT is not applied in the construction sector in Kenya: case study of Nairobi County". The analytically significant evidence of at $\alpha = 0.05$ to reject or accept this hypothesis, is what this study used. The main field research was structured in such a way to test the clear understanding of the various construction consultants on BDT concept.

The hypothesis testing of ordinal data can be carried out only using non-parametric tests such as the Mann-Whitney U test (Wilcoxon Matched-Pairs test). This test is used in testing only what two samples (groups) are possibly able to voice out concerning a particular study. It has been understood as a test for contrasting results between two populaces (Larmorte, 2017).

A one-sided research hypothesis was used because the interest of this study was based in finding out the observations of one populace sample as matched to the other sample. The process for the test constitutes comparing responses from the two samples to one combined sample, as well as noting clearly from which sample each response belong to. Responses are thereafter ranked from the lowest to highest from 1 to n1+n2, respectively (Larmorte, 2017).

Further tests were conducted using the Kruskal-Wallis H test which an extension of the Mann-Whitney U test (Wilcoxon Matched-Pairs test). It is used for contrasting two or more independent samples of the same or varying sample sizes (Kruskal & Wallis, 1952).

CHAPTER FOUR

RESULTS

4.0 Introduction.

This chapter presents the data results presentations obtained from questionnaires and conducted interviews using the questionnaires with the aim to drawing conclusion to the main study. The specific areas of interest covered in the study included examining the status of application of BDT in construction sector; identifying the factors that determine the application of BDT in construction sector; and determining the frameworks to use in integrating application of BDT into the construction industry.

4.0.1 Response rate

The researcher distributed 270 questionnaires and received 161 back, this represents a response rate of 60% as follows: 21 Construction Project Managers, 42 architects, 50 quantity surveyors, 40 from engineers, 3 contractors, and 5 others (who were a BIM consultant, Real estate consultancy/Valuation, Contracts Administrator, Facility Manager, and a Data Analyst) from Nairobi County. A response rate of over 60 % is considered good while 70% and above is considered very good according to Mugenda and Mugenda (2003), thus the response rate in this study was adequate for analysis as shown in Table 4.1 and Figure 4.1.

Table 4.1: Showing response rate: of the approached professionals.

Profession who responded	Frequency	Percent	Cumulative Percent
Construction Project Manager	21	13.0	13.0
Architect	42	26.1	39.1
Quantity Surveyor	50	31.1	70.2
Contractor	3	1.9	72.0
Engineer	40	24.8	96.9
Other	5	3.1	100.0
Total	161	100.0	

Source: Field Survey, 2021.

Table 4.1 clearly shows that quantity surveyors were the most responsive at over 31% of the total respondents. It was important for the respondent to indicate their profession so as to help this study analyse how the total target sample competed and behaved in responding. The

category other was not in the sampling, however in the in the questionnaire it was. Where some respondents recognized themselves as contracts administrator, data analyst, bim consultant, real estate consultant and facility manager.

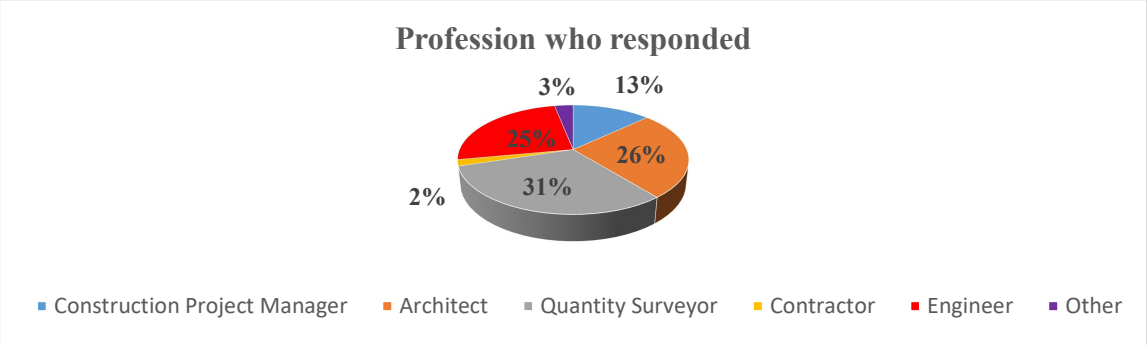

Figure 4. 1: The response rate of the respondents: by profession
Source: Field Survey, 2021.

4.0.2 Doubling up Professionals

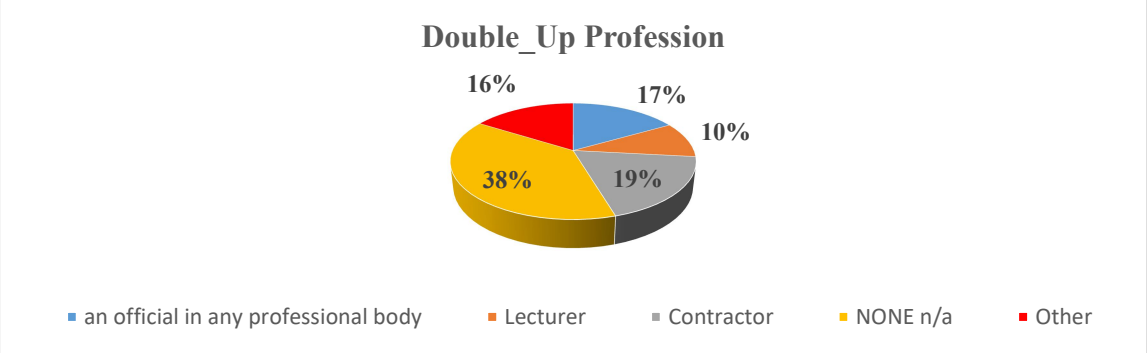

Figure 4.2: Doubling up professionals
Source: Field Survey, 2021.

Findings from Figure 4.2 shows that 19% of the respondents double up as contractors, 16% as in officials in professional or regulating bodies, and 10% as lecturers. However, 16% of the respondents double up in other fields such as (Arbitrators, Mortgage Finance expert) while 38% of the respondents do not double up, they are only practicing professionals. This feedback was important in helping this study realise how the respondents have diversified in the construction sector, translating to more data.

4.0.3 General Gender Response

Table 4. 2: Showing Gender Response

Gender	Frequency	Valid Percent	Cumulative Percent
Male	115	71.4	71.4
Female	46	28.6	100.0
Total	161	100.0	

Source: Field Survey, 2021.

According to the findings from Table 4.2, 71% of the respondents were males while 29% were females. A close indicator of the gender balance rule achievement. This was for general observation by this study.

4.0.4 Education Level of the Respondents

Table 4.3: Showing Education Level of the Respondents

Education Qualification	Frequency	Valid Percent	Cumulative Percent
Diploma	3	1.9	1.9
Bachelor's degree	89	55.3	57.1
Master's degree	62	38.5	95.7
Doctorate	7	4.3	100.0
Total	161	100.0	

Source: Field Survey, 2021.

The findings from Table 4.3 revealed that over 4% of the respondents have completed Doctorate, 38.5% Master's degree while over 55% have completed Bachelors' degree levels. 1.9% are noted to have diploma. This feedback was important in helping this study note the literacy qualifications of the respondents.

4.1 Status of application of BDT in construction industry in Nairobi County.

The first objective was to examine the status of application of BDT in construction industry in the Nairobi County; and in adherence to this study questionnaire, several indicators of the status of BDT application in the construction industry were tested, such as: number of years a consultant has used BDT, whether they have adopted BDT in their company, whether they had used BDT in any construction project, the characteristics of BDT they are familiar with – or else the big data Vs, and whether they are familiar with the BDT fields, technologies and

skills. The findings revealed the following, based on the indicators highlighted in the foregoing paragraph.

4.1.1 The duration the respondents have used BDT

Table 4.4: Displaying duration the respondents have used BDT

Years of BDT use	Frequency	Percent	Cumulative Percent
0	102	63.4	63.4
1-5	38	23.6	87.0
6-10	16	9.9	96.9
11-15	5	3.1	100.0
Total	161	100.0	

Source: Field Survey, 2021.

This feedback was important to this study towards establishing BDT use by the respondents at this particular point of time. From Table 4.4, 63% of the respondents have never used BDT.

4.1.2 Adoption of BDT in organisations

Table 4.5: Adoption of BDT in organisations

	Frequency	Percent	Cumulative Percent
Yes	49	30.4	30.4
No	84	52.2	82.6
Planning to adopt	28	17.4	100.0
Total	161	100.0	

Source: Field Survey, 2021.

From Table 4.5, only 30% of the organisations indicated to have adopted BDT.

4.1.3 Awareness of BDT characteristics (the Vs) in Nairobi County

Table 4.6: Showing awareness of BDT characteristics.

BDT Vs	Valid	Responses	Percent	Mean	Std. Dev	Variance
Volume	161	91	56.5%	0.57	0.497	0.247
Value	161	66	41.0%	0.41	0.493	0.243
Variety	161	54	33.5%	0.34	0.474	0.224
NONE	161	52	32.3%	0.32	0.469	0.220
Variability	161	38	23.6%	0.24	0.426	0.181
Velocity	161	34	21.1%	0.21	0.409	0.168
Veracity	161	17	10.6%	0.11	0.308	0.095

Source: Field Survey, 2021.

The respondents were tasked to highlight all the characteristics they were aware of. The results

are as shown in Table 4.6. This was important in finding out the respondents' knowhow on BDT characteristics, because when one understands a particular knowledge it is critical to have detailed know how of the related concepts thereto.

4. 1.4 On awareness of BDT fields in Nairobi County

Table 4.7: Awareness of BDT fields

CONVERSANT WITH BDT FIELDS					
BDT fields	**N**	**Yes - % age**	**Mean**	**Std. Dev**	**Variance**
Data storage	161	66.7	0.66	0.476	0.226
Data analytics	161	42.8	0.42	0.495	0.245
Data mining	161	40.3	0.40	0.491	0.241
Data visualization	161	33.3	0.33	0.471	0.222
NONE	161	26.4	0.26	0.440	0.194

Source: Field Survey, 2021.

The respondents were tasked to highlight all the BDT fields they were aware of from the above four (4) listed fields. This information is critically important in enabling this study find out how well versant is the respondent with the BDT terms. These fields are basically the BDT terminologies that enable the user of the technology to have a clear understanding of the technology scope and coverage. The results are tabulated in Table 4.7.

4.1.5 On awareness of BDT skills

Table 4.8: Showing respondents awareness of BDT skills.

Skills	**Valid**	**Std Dev**	**Variance**	**Sum**	**%age**
NONE	161	0.497	0.247	70	43.8%
Quantitative Analysis	161	0.487	0.237	61	38.1%
Problem solving skills	161	0.467	0.218	51	31.9%
Data visualization	161	0.447	0.200	44	27.5%
Data mining & Machine Learning	161	0.418	0.175	36	22.5%
General purpose programming	161	0.375	0.140	27	16.9%
SQL	161	0.351	0.123	23	14.4%
NoSQL	161	0.273	0.075	13	8.1%
Apache Hadoop	161	0.174	0.030	5	3.1%
Apache Spark	161	0.136	0.018	3	1.9%

Source: Field Survey, 2021.

The respondents were tasked to highlight all the BDT skills they were aware of from the nine (9) listed skills. From Table 4.8, 44% of the respondents were not aware of any BDT skills.

The respondents' feedback was vital in enabling this study establish the exposure they had on BD technology.

4.1.6 On awareness of BD Technologies

Nineteen (19) technologies were listed and each respondent was required to select all the technologies they are aware of.

Table 4.9: on Awareness of BD technologies

Big Data Technologies	N	Std. Dev	Variance	Sum	Percentage
NONE	161	0.497	0.247	91	56.9%
Blockchain	161	0.469	0.220	52	32.5%
Tableau: & Airflow: each	161	0.242	0.059	10	6.21%
RapidMiner: & R_Language: each	161	0.218	0.048	8	4.97%
Hadoop	161	0.190	0.036	6	3.8%
ApacheKafka	161	0.174	0.030	5	3.1%
Elasticsearch: & ApacheBeam: each	161	0.156	0.024	4	2.5%
MongoDB: & Presto: & Spark: each	161	0.136	0.018	3	1.9%
TensorFlow	161	0.136	0.018	3	1.9%
RainStor: & Splunk: & Plotly	161	0.111	0.012	2	1.3%
Hunk: & Docker: & Kubernetes	161	0.079	0.006	1	0.6%
KNIME	161	0.000	0.000	0	0%

Source: Field Survey, 2021.

The respondents were tasked to highlight all the BDT technologies they were aware of. The findings tabulated in Table 4.9, established that 57% of the respondents are not aware of any BD technologies. With the above analytical big data technologies, actual performance part comes into the picture and the crucial real-time business decisions are made by analyzing all the operational big data from socia media, BIMa, cloud computing and many more.

The response from Table 4.4 to Table 4.9 was vital because it formed the baseline for hypothesis testing.

4.1.7 Project Life Cycle Stages to introduce BDT

Table 4.10: on Stages at which to introduce BDT to the project.

Project Phases	Frequency	Valid Percent	Cumulative Percent
Conception	78	48.4	48.4
Design Stage	42	26.1	74.5
Pre-Construction	10	6.2	80.7

Construction	21	13.0	93.8
Post-Construction	4	2.5	96.3
Operation	4	2.5	98.8
Maintenance Stage	2	1.2	100.0
Total	161	100.0	

Source: Field Survey, 2021.

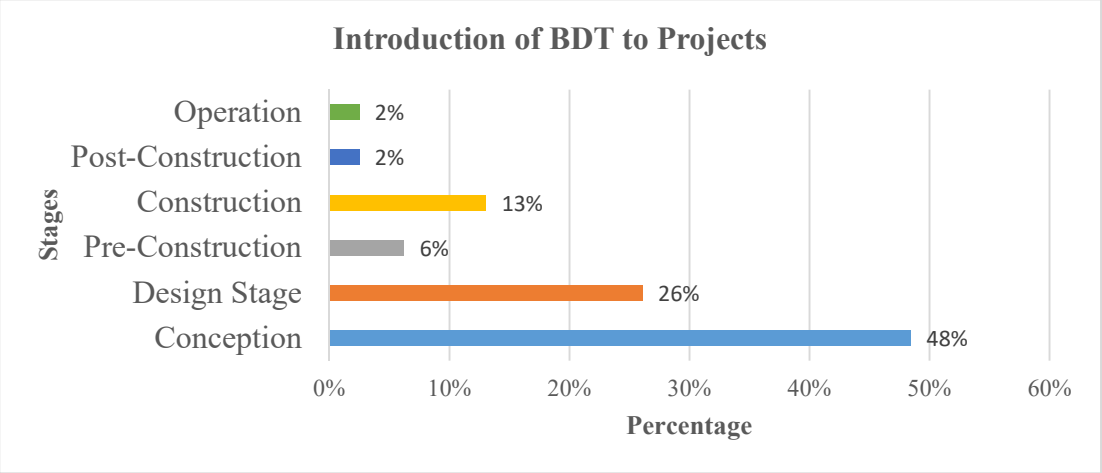

Figure 4.3: Bar chart showing Stages at which to introduce BDT
Source: Field Survey, 2021.

This feedback was important because it enabled this study to find out which project phase the respondents' deemed as very critical for introducing BDT to the project. Each respondent was clearly instructed to select only one phase. As tabulated in Table 4.10 and shown in Figure 4.4, the conception phase ranked top at 48%.

4.2 Factors that determine the application of BDT in the construction industry.

This was the second objective of this study. Where reasons of why BDT was in use or else why not in use in many organisations were sought from the respondents in the Nairobi County. The key indicators that guided collection of data in this objective were: benefits of BDT in construction projects, and the limiting factors in the adoption of BDT.

4.2.1 BDT benefits in construction projects

In finding out how the respondents rate the various BDT benefits; they were given questionnaires listing twenty-seven (27) benefits.

Table 4.11: on BDT benefits in construction - Descriptive Statistics

Benefits of BDT	N	Mean	Std. Dev	Variance
Improved project planning & scheduling.	161	1.44	.669	.448
Saves on construction time (quickens execution time)	161	1.55	.733	.537
Capability to handles huge amounts of data.	161	1.56	.740	.548
Allows better project management	161	1.56	.757	.573
Improved research on new proposed projects	161	1.57	.705	.497
Provides accurate data.	161	1.57	.723	.522
Increases efficiency in the workplace.	161	1.57	.696	.485
Consistent, up to date project information	161	1.60	.710	.505
Brings knowledge thus better planning.	161	1.60	.692	.478
Risk management	161	1.61	.735	.540
Quick response times is realised.	161	1.62	.733	.537
Real time capabilities.	161	1.66	.759	.576
Information is accessible in a transparent way	161	1.67	.756	.572
Integration and collaboration of platforms.	161	1.68	.787	.620
Opens up for central repository storage & retrieval	161	1.68	.825	.680
Enhances productivity: Improved resource allocations	161	1.71	.788 / .841	.620 / .708
Allows predictions to reduce the risks in future.	161	1.71	.764	.583
Informed decision making. Brings more insights	161	1.71	.737	.543
Improved budget estimates and superior understanding of costs.	161	1.73	.806	.650
Cost saving (lowering costs) curbs cost overruns.	161	1.73	.806	.650
Increased data collection techniques	161	1.75	.793	.628
Quality improvement	161	1.81	.853	.728
Improved return on investment – ROI	161	1.83	.846	.715
Determines the most profitable projects to pursue.	161	1.98	.884	.781
Allows improved safety and health management.	161	1.98	.848	.718
Gives answers complex questions	161	2.01	.877	.769

Source: Field Survey, 2021.

The respondents were required to rate each of the importance of BDT benefits using a five-point Likert scale where 1= Strongly Agree, 2= Agree, 3= Uncertain, 4= Disagree and 5= Strongly Disagree. From Table 4.11 presentation above, the lower the (mean, standard deviation, and variance) the greater the benefit of adopting BDT and vice versa.
This response was key because it helped this study to rank the BDT benefits in the construction industry from Nairobi County respondents which double up as the KPIs in construction sector.

4.2.1.1 Data consumers in order of priority

Table 4.12: Ranking of the consumers of data in order of priorities - Descriptive Statistics

Consumers	N	yes- Freq	YES - %	Mean	Std Dev	Variance
Construction Project Managers	161	152	94.4%	1.06	.230	.053
Quantity Surveyors	161	144	89.4%	1.11	.308	.095
Architects	161	142	88.2%	1.12	.324	.105
Engineers	161	141	87.6%	1.12	.331	.109
Contractors	161	130	80.7%	1.19	.396	.156
Students	161	126	78.3%	1.22	.414	.171
Policy Makers	161	117	72.7%	1.27	.447	.200
Site Supervisors	161	101	62.7%	1.37	.485	.235
Ruling Government	161	86	53.4%	1.47	.500	.250

Source: Field Survey, 2021.

The respondents were tasked to highlight all the data consumers they deemed critical. CPM ranked highest with a mean of 1.06, standard deviation of 0.230 and variance of 0.053. From Table 4.12 presentation the lower the (mean, standard deviation, and variance) the greater the demand of usage of data by the listed stakeholders and vice versa.

4.2.2 Limiting factors in BDT adoption

In finding out about the limiting factors in BDT adoption, that respondents would consider as hindrance in accepting BDT adoption, they were given questionnaires listing nineteen (19) factors. They were asked to indicate how they felt each of the limiting factors would influence the adoption of BDT using a five-point Likert scale where 1= Strongly Agree, 2= Agree, 3= Uncertain, 4= Disagree and 5= Strongly Disagree. The results are shown on Table 4.13.

Table 4.13: Limiting factors of BDT - Descriptive Statistics

Limiting factors in BDT adoption	N	Mean	Std Dev	Variance
Skill gap in BDT	161	1.61	.783	.613
Formal training is lacking	161	1.71	.847	.718
Upfront cost of adoption	161	1.75	.808	.653
Resistance to change	161	1.84	.905	.819
Lack of top management support.	161	1.98	.844	.712
The complexity behind getting data in the BD structure.	161	1.99	.887	.787
Lack of standards and guidelines	161	2.00	.873	.763
Technological capabilities.	161	2.02	.829	.687

Trusting data integrated from various sources.	161	2.07	.905	.819
Lack of government policy	161	2.09	1.030	1.060
Information privacy.	161	2.10	.989	.978
Information security.	161	2.12	1.002	1.005
No government / client requirement to use it	161	2.13	1.001	1.002
Size of a project	161	2.19	.965	.931
Ability to ensure that you are getting the right data	161	2.20	.975	.951
Data un-synchronizing risks.	161	2.22	.913	.834
Satisfied with existing system	161	2.25	1.013	1.025
Uncertainty created by wide range of BD tools	161	2.26	.959	.919
Pressure to remain competitive.	161	2.47	1.073	1.151
I do not understand BDT	161	2.76	1.302	1.694

Source: Field Survey, 2021.

Frequencies were thereafter computed and the lower the mean, the standard deviation and the variance the greater the affect the factor would have in BDT adoption and vice versa.

Further on, related to skill gap and lack of training the respondents had this to say based on:

4.2.1.1 BDT education and training in universities and colleges

This study sought to know more on education and training on BDT in our tertiary institutions.

Table 4.14: Educated on BDT at the University / College

Response Y/N	Frequency	Percent	Cumulative Percent
Yes	21	13.0	13.0
No	140	87.0	100.0
Total	161	100.0	

Source: Field Survey, 2021.

As tabulated in Table 4.14, the study established that 87% of the respondents had not been educated on BDT.

4.2.1.2 Otherwise source of initial BDT knowhow/ training

Table 4. 15: Otherwise initial BDT knowhow source.

Other BDT knowhow sources	Frequency	Percent	Cumulative Percent
Software companies	3	1.9	1.9

Software sellers	9	5.6	7.5
BDT champion	5	3.1	10.6
My employer	28	17.4	28.0
Inner Self-Drive	87	54.0	82.0
Other e.g workshops	3	2.0	84.0
None of the Listed	26	16.0	100.0
Total	161	100.0	

Source: Field Survey, 2021.

The respondents were tasked to highlight only one other source of BDT knowhow. From Table 4.15, inner self-drive ranked top at 54% whereas 16% had not had any exposure on BDT. Software companies tailed in the ranking.

4.2.1.3 Necessity of general-purpose programming (GPP)

Table 4.16: GPP becoming a necessity

Response Y/N	Frequency	Percent	Cumulative Percent
Yes	143	88.8	88.8
No	18	11.2	100.0
Total	161	100.0	

Source: Field Survey, 2021.

Results in Table 4.16 outlays that 89% of the respondents are for the opinion that GPP is becoming a basic necessity skill that all consultants should have.

4.2.1.4 Advocating for inclusion of GPP in curriculum

Table 4.17: Inclusion of GPP in the curriculum

Response Y/N	Frequency	Percent	Cumulative Percent
Yes	157	97.5	97.5
No	4	2.5	100.0
Total	161	100.0	

Source: Field Survey, 2021.

98% of the respondents advocated for inclusion of IT (GPP) in the construction related courses in the tertiary curriculum e.g. construction management with IT.

4.3 Frameworks to use to integrate to a BDT platform in Kenya construction industry.

This was the third objective of this study. Which first sought the know-how of various integrating platforms from the respondents. Then sort to determine the respondents understanding on what it entails to integrate. Respondents views on legal risks that relate to BDT were also interrogated in this objective. Awareness of new job roles and views of the ideal body that can host a national data repository centre were sought by this study.

4.3.1 Views on the Integrating Principles

The study sought to determine how the respondents understood and considered the listed requirements and statements for the BDT integration platform to function.

Table 4.18: Integrating principles to be considered in BDT adoption.

Integrating Principles (ranked)	N	Mean	Std Dev	Variance
Internet usage and strong connectivity is a necessity.	161	1.30	.513	.263
Optimizes decision making.	161	1.61	.690	.476
Integration aids in early problem detection.	161	1.63	.629	.396
Better prediction of outcomes.	161	1.65	.654	.428
Management information' factors (such as project schedule, safety, cost, quality) are integrated.	161	1.66	.689	.474
Integration enhances communication.	161	1.73	.733	.537
Processes and analyses are automated.	161	1.73	.756	.572
Enhanced interoperability support integration.	161	1.73	.678	.459
Integration improves good working relationship.	161	1.76	.763	.581
Knowledge is discovered.	161	1.77	.735	.541
Integration helps to reduce variations.	161	1.83	.755	.570
Integration births full information model throughout full project lifecycle.	161	1.83	.673	.453
Implementation of integration demands many resources (machines, cloud services, software, and expert people).	161	1.84	.914	.836
Integration births transparency and a more proactive approach.	161	1.84	.738	.544
Integration helps to reduce claims.	161	1.85	.768	.590
Promotes positive return on investment (ROI).	161	1.86	.749	.561
Integrations demands computational expertise.	161	1.86	.813	.661
Integration improves work quality.	161	1.88	.820	.672
Integration reduces construction cost.	161	1.91	.794	.630
BDT is fed by the integration platforms while integrating platforms extract information from BDT.	161	1.94	.764	.584

Information is transformed.	161	1.97	.794	.630
Integration reduces construction time.	161	1.98	.855	.731

Source: Field Survey, 2021.

In finding out the understanding of the integrating principles to be considered in BDT adoption, that the respondents considered important, they were given questionnaires based on twenty-two (22) listed considerations. They were required to rate all the listed integrating principles of BDT using a five-point Likert scale where 1= Strongly Agree, 2= Agree, 3= Uncertain, 4= Disagree and 5= Strongly Disagree.

From Table 4.19, computations with the lower mean, standard deviation and variance are the ones that are most critical in BDT hosting and vice versa. Internet usage and strong connectivity as a basic necessity ranked top at 1.30 mean, 0.513 standard deviation, 0.263 variance.

4.3.2 Legal Risks that come with BDT

The study sort to understand how the respondents perceive BDT integration from the legal perspective. Respondents were required to rate the fifteen (15) listed legal risks that come with BDT.

Table 4.19: Legal risks that come with BDT

Legal risks that come with BDT	**N**	**Mean**	**Std. Dev**	**Variance**
Data security	161	1.42	.628	.394
Sharing of copyright data	161	1.46	.602	.362
Intellectual property rights	161	1.46	.661	.437
Information privacy	161	1.47	.671	.451
Professional liability	161	1.65	.728	.530
Lack of BDT standards.	161	1.71	.747	.558
Standard of care and professional negligence.	161	1.84	.755	.569
Legal validation of design (NCC submissions).	161	1.89	.814	.662
Condition of contracts	161	1.90	.800	.640
Processes and responsibilities	161	1.92	.774	.600
Data interoperability	161	1.93	.784	.614
Legislation and judicial precedence.	161	1.95	.835	.698
Cost compensation	161	1.96	.904	.817
Admissibility of electronic based documents.	161	1.99	.915	.837
Model management	161	1.99	.802	.644

Source: Field Survey, 2021.

In finding out the respondents' views they were given questionnaires listing fifteen legal factors. They were asked to indicate how they perceived each of the legal factors in integrating BDT using a five-point Likert scale where 1= Strongly Agree, 2= Agree, 3= Uncertain, 4= Disagree and 5= Strongly Disagree.

From Table 4.20, computations with the lower mean, standard deviation and variance are the risks that are most critical and vice versa. Data security ranked top at 1.42 mean, 0.628 standard deviation, 0.394 variance. Figure 4.5 shows further presentation in bar chart format.

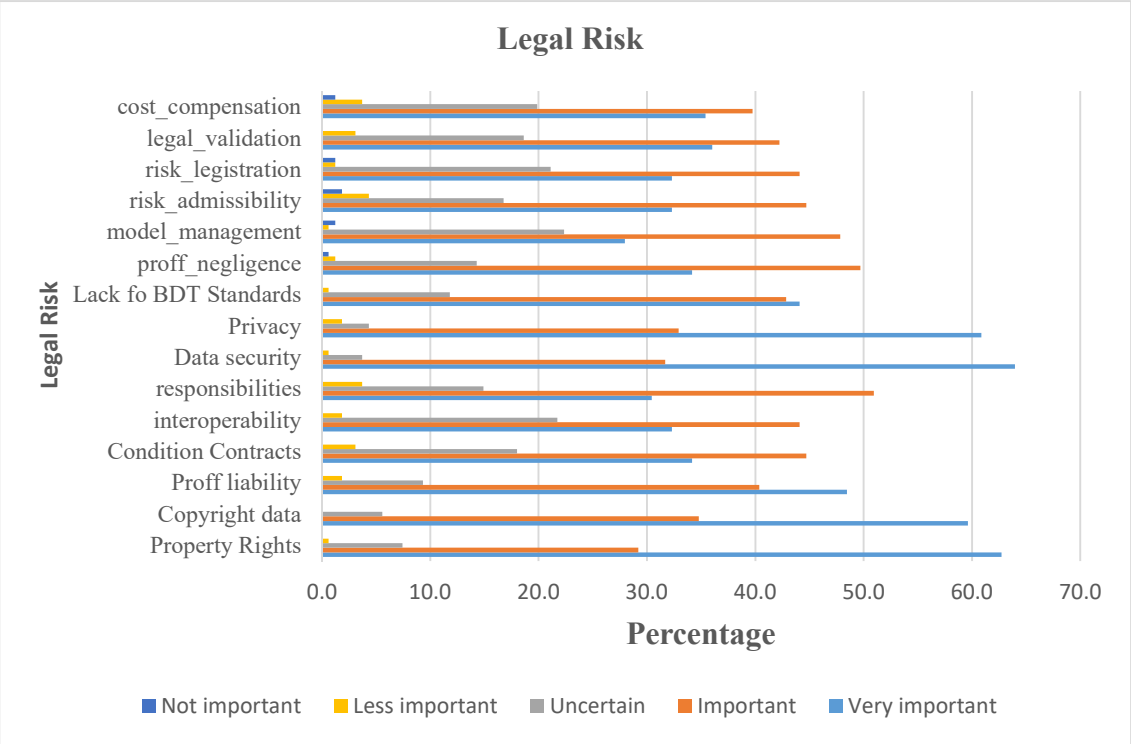

Figure 4. 4: Legal risks that come with BDT
Source: Field Survey, 2021.

4.3.3 Awareness of integrating frameworks

Table 4.20: Five integrating frameworks: Descriptive Statistics

Operational BDTs	N	Y- %	Mean	Std Dev	Variance
Social Media: WhatsApp, Facebook,	161	77.6%	1.22	.418	.175
Building Information Management	161	70.2%	1.30	.459	.211
Cloud Computing	161	45.3%	1.55	.499	.249
Internet of Things / Sensor Networks / Telematics (GPS devices)	161	29.8%	1.70	.459	.211

Real Time Data	161	28%	1.72	.450	.203
NONE	161	5.6%	1.94	.230	.053

Source: Field Survey, 2021.

The respondents were tasked to highlight all the frameworks they were aware of from the listed.

These frameworks are also known as the Operational Big Data Technologies. Which are all about the normal day to day data that we generate. From Table 4.18, computations with the lower mean, standard deviation and variance are the ones that are top in being known and used and vice versa. Social media ranked top at 1.22 mean, 0.418 standard deviation, 0.175 variance and 78% aggregated score.

4.3.4 Knowhow of new job data roles in BDT integration

Every respondent was tasked to highlight all the new job roles they were aware of. Figure 4.6 and Table 4.21 shows the results. Computations with the lower standard deviation and variance are the ones that are top in being known and vice versa. The role of data analyst ranked top at 0.324 standard deviation, 0.105 variance and 88% aggregated percentage score.

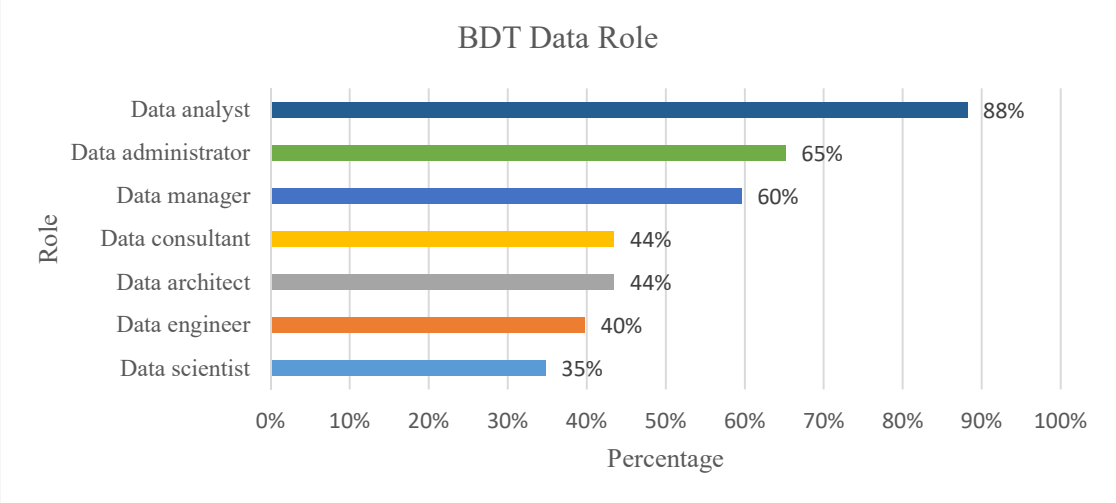

Figure 4.5: New job roles
Source: Field Survey, 2021.

Table 4. 21: On awareness of new job roles

Job roles	N	y- Freq	Y - %	Std. Dev	Variance
Data scientist	161	56	34.8%	.478	.228
Data engineer	161	64	39.8%	.491	.241
Data architect	161	70	43.5%	.497	.247

Data consultant	161	70	43.5%	.497	.247
Data manager	161	96	59.6%	.492	.242
Data administrator	161	105	65.2%	.478	.228
Data analyst	161	142	88.2%	.324	.105

Source: Field Survey, 2021.

4.3.5 Regulating Body for hosting a Central Data Repository Centre

In this sub subject, the study sought the views from the respondents on their preferred body to host a national central data repository centre.

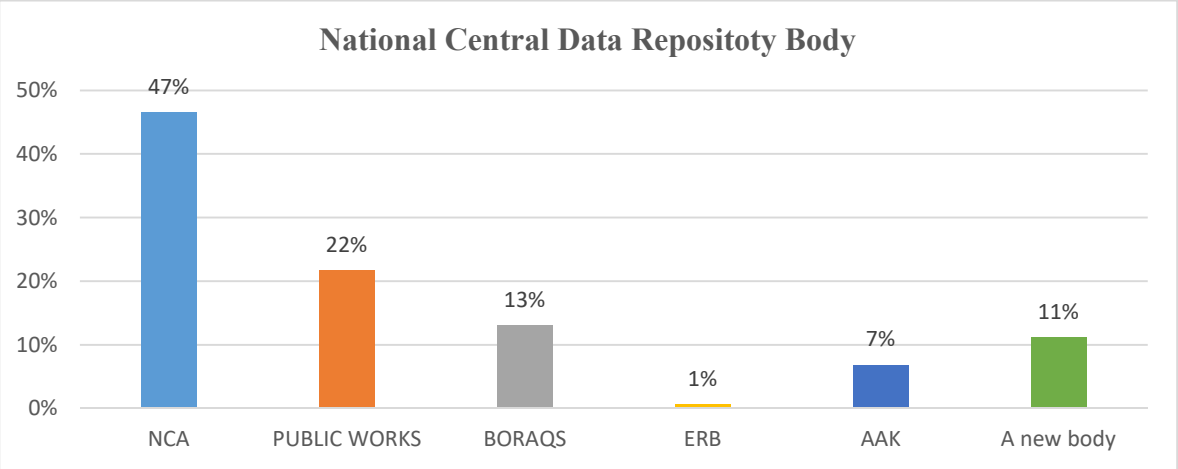

Figure 4.6: Preferred body for a national repository centre
Source: Field Survey, 2021.

Every respondent was tasked to highlight only one body. From Figure 4.7 National Construction Authority (NCA) ranked top preferred at 47 percent score.

Table 4. 22: Possibility of emergence of a private body

	N	Y- %	Mean	Std Dev	Variance
Foresee a private owned repository in Kenya	161	75%	1.25	.433	.188

Source: Field Survey, 2021.

From Table 4.22, 75% of the respondents' foresee the emergence of a private owned entity.

4.4 Hypothesis Testing.

This research is examined on the speculation that BDT application is not applied in the construction sector in Kenya for projects within Nairobi County. Application of BDT is measured through time the duration the respondents had used BDT, adoption of BDT in the organisation, respondents' awareness of BDT characteristics, fields, skills and terminologies, but not limited to.

The analysis compares the adoption of BDT in different professions and categories of employers. The multiple professions assessed include construction project managers, architects, quantity surveyors, contractors, and engineers. The main categories of employment include government entities and private entities.

Therefore, the hypothesis that will address the research hypothesis will be as follows.

Hypothesis 1

H0: There is no significant difference in application of BDT in construction projects citing response by government employees and those in private entities.

H1: There is a significant difference in application of BDT in construction projects citing response by government employees and those in private entities.

Hypothesis 2

H0: There is no significant difference in application of BDT in construction projects by various professionals having used BDT for more years than the other.

H1: There is a significant difference in application of BDT in construction projects by various professionals having used BDT for more years than the other.

Each hypothesis requires the use of a one-Way Analysis of Variance. The method is applicable when testing for variation in means for groups with more than two levels. However, for such to happen, testing for the assumptions of the test is vital. One key to the key assumption of ANOVA is that the dependent variable should be normally distributed. Testing the assumption ensures the right method of analysis is adopted to generate results with minimal errors.

4.4.1 Test of Normality

The results in Table 4.23 provide the results of the normality test. From the findings, the p-value is 0.000 which is less than 0.05; hence the assumption that the data is normally distributed is rejected.

Table 4.23: Normality test results

Tests of Normality						
	Kolmogorov-Smirnova			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
BDT Application	.117	161	.000	.896	161	.000
a. Lilliefors Significance Correction						

Source: Field Survey, 2021.

The histogram presented in Figure 4.7 further explains the distribution of the dependent variable. The data is not normal, and it's skewed to the left, as shown in the figure below. From the normality results, the analysis will adopt a non-parametric approach of data analysis to address the research hypothesis adequately. The method adopted is an alternative for ANOVA, which is the Kruskal Wallis Test.

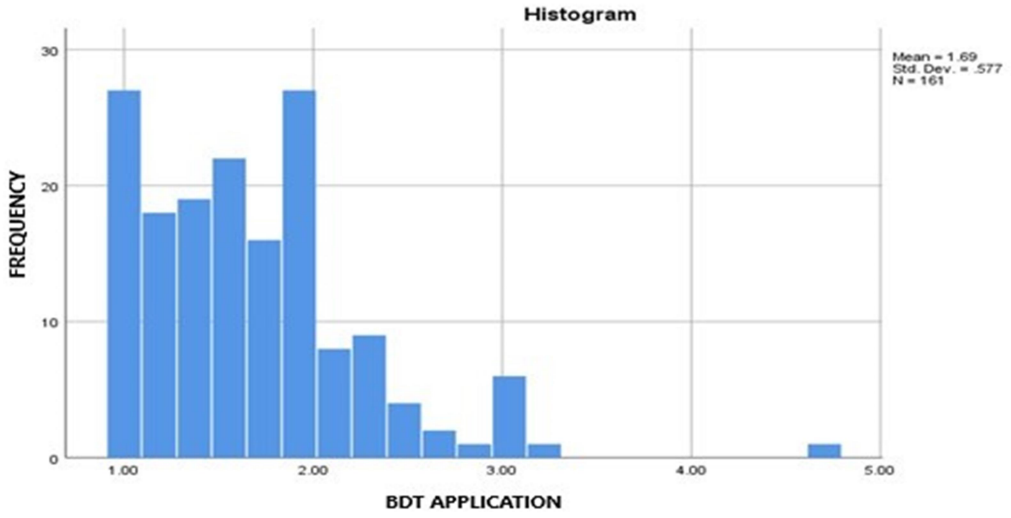

Figure 4.7: Distribution of the independent variables
Mean = 1.69, Standard Deviation 0.577
Source: Field Survey, 2021.

4.4.2 Kruskal Wallis & Mann-Whitney U Test Results

BDT application versus Different Employing Entities.

Table 4.24: Mean rank based on the type of employer

Ranks			
	Employer	**N**	**Mean Rank**
BDT Application	Government Agency / Department	29	71.72
	Private Entity	117	81.39
	Non-Governmental	4	74.25
	Self Employed	8	95.75
	Both Government & Private	2	145.75
	Retired	1	83.50
	Total	161	

Source: Field Survey, 2021.

Referring to the ranking of adoption of BDT in the construction sector from Table 4.4, 4,5, 4.6, 4.7, 4.8 and 4.9 by the government agency employees versus the private entity employees which are the independent variables in this study and who registered many respondents, Table 4.24 results above shows that profession in private entities have the highest mean rank of (81.39) whereas the profession in government entity had a mean rank of (71.72) on rating the BDT application. Thus, the variance in rating the results of BDT application by private and government entity profession is little.

Table 4. 25: Kruskal–Wallis H test results

Test Statistics[b]	
	BDT Application
Kruskal-Wallis H	5.924
df	5
Asymp. Sig.	.314
a. Kruskal-Wallis Test	
b. Grouping Variable: Employer	

Source: Field Survey, 2021.

The Kruskal – Wallis H test results in Table 4.25 shows that the results are non-significant, H (5) =5.924, p = 0.314. The results of the Kruskal-Wallis test show no difference in BDT application between employees of different employers. This is as per the results from Table

4.25 herein above and Table 4.4 to 4.9. The p-value for the test is 0.314, which is higher than the significance level of 0.05. For this reason, the null hypothesis is retained.

Checking for the variance between the government and the private entity employees, an alternative test for an independent sample t-test was adopted. The alternative test is known as the Mann-Whitney U test.

Table 4. 26: Mann-Whitney U test results

	Null Hypothesis	**Test**	**Sig.**	**Decision**
1.	The distribution of BDT application is the same across categories of employer - employee	Independent Samples Mann-Whitney U Test	0.308	Retain the null hypothesis

Source: Field Survey, 2021.

In this test, whose results are shown in Table 4.26, the significance value obtained is 0.308, which is higher than 0.05. The results imply that there is no difference in BDT application considering responses between by government and private entity employees.

BDT application versus various Professionals

Table 4. 27: Profession data to test

Ranks			
	Profession	**N**	**Mean Rank**
BDT Application	Construction Project Manager	21	76.85
	Architect	42	71.51
	Quantity Surveyor	50	78.54
	Contractor	3	53.17
	Engineer	40	72.53
	Other	5	46.00
	Total	161	

Source: Field Survey, 2021.

Table 4.27 results shows that the profession of quantity surveyor has the highest mean rank (78.54). Meaning the application of BDT indicators like duration that the respondents have used BDT, awareness of BDT technologies and skills: were rated higher by the quantity surveyors. Followed by the construction managers 76.85, engineers 72.53 and architects ranked fourth at 71.51.

Table 4. 28: Kruskal – Wallis H test results

Test Statistics, b	
	BDT Application
Kruskal-Wallis H	3.717
df	5
Asymp. Sig.	.591
a. Kruskal-Wallis Test	
b. Grouping Variable: Profession	

Source: Field Survey, 2021.

The Kruskal–Wallis H test displayed in Table 4.28 shows that the results are non-significant, H (5) = 3.717, p = 0.591. These results show no difference in BDT application between different professions in consideration of the duration they have used BDT. The p-value for the test is 0.591, which is higher than the significance level of 0.05. For this reason, the null hypothesis is retained.

4.4.3 Decision – to Accept the null hypothesis.

The null hypothesis for the two hypotheses is not rejected. The non-significance of the results is a prove that professions and different entities in the construction industry have not embraced the application of BDT in their operations. The value for the BDT application is the same across the different groups.

Although BDT contributes massively to improved project planning & scheduling, saves on construction time (quickens execution time), and contributes to accurate decision making: the construction industry has not utilized the application of the BDT. Low BDT uptake can be attributed to some of the factors such as skill gap, lack of formal training, upfront cost of adoption, resistance to change and lack of support from the top management.

The massive contribution of BDT application is given weight by the research findings that 87% of the interviewee consented that digital data mining can solve the consistent and historical problems of schedule overruns, ineffective construction management, poor record keeping, and cost overruns in the construction industry. However, due to the current low and slow adoption of BDT across various employing entities and profession groups, future analysis should therefore, be conducted to establish the variation after a greater adoption of BDT by several professions and entities in the construction sector.

CHAPTER FIVE

DISCUSSION OF THE FINDINGS

5.0 Introduction.

The results of the analysis are well illustrated by both descriptive and inferential statistics. Results of descriptive statistics support the findings of the inferential statistics. The key area addressed in descriptive statistics entails the extent of BDT application, the factors to consider in adopting BDT and the available frameworks in use for integrating to BDT platform.

5. 1 Status of application of BDT in construction industry in Kenya.

BDT use in County of Nairobi and adoption rate in companies

The focus of the study was mainly on the application of BDT. Findings at table 4.4 and figure 4.3 show that 63% of the participants have never heard of or used BDT, whereas some of the remaining 37% contradicted themselves against number of years they have used BDT and at the same time saying they are unsure of having used BDT in previous projects. The results provide a key insight of how majority of professions in the construction industry have never interacted with BDT hence a greater percentage in the non-use of the technology.

Further, findings in Table 4.5 revealed that 30% of the organisation have adopted BDT, meaning 70% of the respondents have not, though 17% are planning to adopt. However, from the 30%, this could be use of servers and cloud computing services. Further, some of the 30% contradicted themselves when they agreed to have adopted BDT but have never used BDT posing lack of understanding of the questions.

BDT characteristics awareness

Findings in Table 4.6 revealed that data volume as the most known BDT characteristic by the respondents at a standard deviation of 0.497. This is largely supported by its definition that, BD is a new concept that stands for massive amount of defined, pre-defined and undefined data (Rouse, 2018). Data value ranked second at a standard deviation of 0.493 followed by data variety at 0.474. However, 32% of the respondents were not conversant with any BDT characteristics, while from the remaining 68% distribution of knowhow across all

characteristics was sparsely known save for volume and value and variety. The intrigues of the BDT characteristics call the expertise of the data experts.

BDT fields awareness

Findings in Table 4.7 revealed that the field of data storage is widely known at a mean of 0.66 and a standard deviation of 0.476. The field of data analytics ranked second at a mean of 0.42. Data mining field ranked closely as third at a mean of 0.40 and a standard deviation of 0.491. However, 26% of the respondents were not aware of any of the big data fields. Therefore, the results are an indication that professions in the construction are not well acquainted with knowledge related to the key fields of BDT.

BDT skills awareness

Findings in Table 4.8 revealed quantitative analysis as the most know BDT skill, at a standard deviation of 0.487 and a variance of 0.237. Problem solving skills ranked second at a standard deviation of 0.467. However, the top ranking was on those who were not aware of any BDT skills at a standard deviation of 0.497 analyzed at 44%. The rest 56% sparingly indicated to have some knowhow on quantitative analysis skills, problem solving skills, data visualisation skills, data mining & machine learning skills, general purpose programming skills, SQL & NoSQL skills, Apache Hadoop & Apache Spark skills.

BD technologies awareness

Findings in Table 4.9 revealed that most respondents are aware of block chain technology at a standard deviation of 0.469. Tableau and Airflow ranked second at a standard deviation of 0.242 respectively. A clear indication that the rest of 18 listed technologies are sparingly known. However, 57% of the respondents were not aware of any BDT technologies.

Project stages to introduce BDT

Findings in Table 4.10 and Figure 4.4 revealed that most respondents (at percentage score of 48%) would introduce BDT at the conception stage. This finding supported the importance of collecting and storing construction information across all construction stages, which will be forthwith wired to a national central place for ease of mining. Introduction of BDT at design stage ranked second at 26%.

5.2 Factors that determine application of BDT in the construction industry in Kenya.

Benefits of BDT adoption

Through the mean item rating scale, standard deviation and variance subjection, the ten most benefits of BDT in the construction industry ranked as follows, from Table 4.11. Improves project planning & scheduling, saves on construction time (quickens execution time), capability to handle huge amounts of data, allows better project management, improved research on new proposed projects, provides accurate data, increases efficiency in the workplace, availability of consistent, up to date project information, brings knowledge thus better planning, helps in risk management towards lowering projects risks.

The findings in this subject are supported by Akbar (2018) who outlined keeping track of project time and better management, brings knowledge thus better planning and eventually accurate improved timelines and budget estimates and superior understanding of costs, lowers project risks (due to access to the relevant historical information), and development of project management software, as the usefulness of BD in the construction sector.

However, from this research findings improved budget estimates, superior understanding of costs; cost saving/lowering towards controlling cost overruns ranked low at position 19 and 20 respectively. During execution stage the construction project manager and all other stakeholders are concerned on all factors that can make work implementation fast and successful (We Build, 2018). Quality improvement, Improved return on investment – ROI, improved facility management, Determines the most profitable projects to pursue, allows improved safety and health management, and giving answers to complex questions, ranked as the least five rated benefits.

From the findings displayed in Table 4.12, it is revealed that construction project managers are the pertinent consumers of data at a mean of 1.06, standard deviation of 0.230, and variance of 0.053, followed by quantity surveyors, architects, engineers, fifth ranked is contractors, followed by students, then policy makers, whereas site supervisors and ruling government were the least in the rank.

Limitations to adoption of BDT

Through mean rating scale, standard deviation and variance of the results in Table 4.13, skill gap, lack of formal training, upfront cost of adoption, resistance to change, lack of top management support, ranked as the top five leading limiting factors. The results show that the most critical factor for BDT not being used is the skill gap with a mean of 1.61 and a standard deviation 0.783 while the least critical reason is, I do not understand BDT rated at a mean of 2.76 and a standard deviation of 1.302 as shown in Table 4.13. Skill gap as the most critical limiting factor in the BDT adoption in this study, has already been supported by Fujikawa (n.d) when they highlighted that, "the field of BDT is lacking diverse skilled personnel''.

This study revealed upfront set up cost as the third most reason as to why BDT is not used. From the literature review Bilal highlighted that every technology incurs cost. These are costs in: setting up of data centres, purchasing of software licenses, acquiring skilled IT personnel or the human resources for that matter (Bilal etal., 2016).

Lack of training at a mean of 1.71 ranked second and can be attributed to the fact that tertiary colleges and universities in Kenya are not offering BDT training, neither the vendors. The training agenda demands for a curriculum and time thereto. Training needs to be considered so as to aid in penetration of BDT and extent of use.

The findings have also revealed that lack of top management support was the fifth most critical factor under this theme. This finding is further supported by Gambatese and Hallowell (2011) who opines that top management support had huge effect of enabling implementation of innovation in construction firms. Towards improved technology adoption uptake. Ability to ensure that you are getting the right data, Data un-synchronizing risks, satisfied with existing system, Uncertainty created by wide range of BD tools, and pressure to remain competitive ranked as lesser limiting factors.

In relation to skill gap and lack of training: From Table 4.14 the investigation disclosed that only 13% of the respondents were trained formally on BDT during their tertiary education. A huge gap in this indicator because application and awareness of whatever technology aspect goes handy with education and training beforehand. This study suspects that the educated-on BDT were trained outside Kenya.

From Table 4.15 inner self-drive rated highest at 54% as the otherwise how respondents came to know about BDT. Employers rated second at 17% and software sellers followed at 6%. However, BDT champion and software companies ranked low in creating awareness of BDT in detailed perspective in Nairobi County which is a hub of most construction consultants. Other sources of knowhow at 2% where only 3 respondents who indicated as follows: I got exposed to BDT when I started working with a global company that has applied this before; via Seminar and workshops; via this questionnaire. 16% of the respondents had never known of BDT.

Inclusion of general-purpose programming (GPP) training in the tertiary education was vouched for by 98% of the respondents, while 89% agreed that GPP was becoming a basic necessity skill that all professionals should have.

5.3 Frameworks to use in integrating a BDT platform in Kenya construction industry.

Frameworks already in use

The findings as presented in Table 4:18 revealed that social media was the most used and known BDT integrating framework with a mean item rating of 1.22 and standard deviation of 0.418. Overall awareness of 78% and use. Example of social media platforms are the likes of WhatsApp, Facebook, LinkedIn, and Twitter just to mention some. This study believes that vast project management information is in the various WhatsApp groups. Social media is rich with management information and feedback from the audience / followers, that needs to be tapped and utilised well.

BIMa ranked the second most used and known to the respondents as a potential BDT integrating platform, with a mean item rating of 1.30 and standard deviation of 0.459. Overall awareness of 70% and use. This can be attributed to the level of awareness and use of BIMa amongst the consultants in the construction industry in Kenya.

The survey further deduced that cloud computing ranked third, with a mean item rating of 1.55 and standard deviation of 0.499. Overall awareness of 45% and use. Cloud computing is thus gaining momentum in the construction industry. Internet of Things ranked fourth with a mean item rating of 1.70 and standard deviation of 0.459, at overall awareness of 30%.

Being conversant with RTD ranked the least known and used, at 28% of awareness. Going by the definition of RTD (is information that is current and accessible instantly) this study deduces that the respondents do understand this and the applicability of RTD. This is as Novotny (2018) has explained that, the construction sector is continuously appreciating the contributions of RTD in the sector. However, RTD is effective in the field of drawing tools and BIM because project managers can instantly access and guide by marking out on the design plans and drawings

In summary, integrating frameworks are supposed to feed the big data platform with the voluminous information. From the literature review all the above frameworks are interrelated. However, 6% of the respondents were totally not aware of any of the listed integrating frameworks.

Integrating considerations

From Table 4:19, with a mean item rating scale of 1.30, standard deviation of 0.513 and variance of 0.263, the study revealed that internet usage, and strong internet connectivity to be the most critical necessity that should be in place for BDT integration to take place. Followed closely by the benefits to BDT integration as follows: it optimizes decision making, integration aids in early problem detection, better prediction of outcomes, management information' factors (such as project schedule, safety, cost, quality) are integrated, integration enhances communication, processes and analyses are automated, enhanced interoperability support integration, integration improves good working relationship, knowledge is discovered.

Internet usage, and strong connectivity necessity finding considered as the most important requirement is further supported by Bilal et al. (2016) when he states that "to keep track of instant work place tasks as they happen, then prompt information communication among project sites and integrated BD depository should be enabled. This is nonetheless an expense just like that of implementation of integration platform demands many resources (machines, cloud services, software, and expert people).

Integration optimizes decision making ranked second. This finding is supported by Baesens (2014) when he stated that BDT introduces computational expertise in the construction

industry, with the principle aim being to improve information, to find insights, to effectively forecast results, to automate procedures and analyses, and to perfect solutions.

Further, management information' factors are integrated ranked fifth. This is supported by Lynch (2018) on analytical applications, which looks at the contract datasets to discover and reveal patterns on the key performance indicators, as well as highlight the dangers. Meaning, the greater a lot of crucial information is fed into the digital set up, the greater that the forecasted insights are realised in regard to the key performance indicators.

Legal risks factors

Through the mean item rating scale and as shown in Table 4.20, the study revealed the top five critical legal factors as follows: data security, sharing of copyright data, intellectual property rights, information privacy, and professional liability. Table 4.20 is further presented through Figure 4.5. While the bottom five were: data interoperability, legislation and judicial precedence, cost compensation, admissibility of electronic based documents, and model management. This was from the fifteen listed legal fears/risks.

These results are strengthened by Sharron et al. (n.d) who states that, across industry sectors issues concerning: data rights, intellectual property rights, product liability, regulatory compliance, privacy vulnerabilities, high-profile data security incidents, weak or absent data rights, and cyber security threats are paramount and less well understood. Further, she states that, these risks can be expertly managed.

New job roles

This study further through literature review learnt of new job roles that have emerged to aid in integrating various frameworks to the BDT platform. The field study further, through the findings displayed in Figure 4.6 and Table 4.21, revealed the popularity of them as listed: data analyst, data administrator, data manager, data consultant, data architect, data engineer and data scientist. The gist was to have a feel of whether the respondents were aware of these new emerging digital data roles.

National central hosting bodies

Firstly, this study established that 80% of the respondents have ever heard of a centralized

database management system (CDMS) or one stop search engine. From the findings presented in Figure 4.7, NCA at 47% was the most preferred ideal hosting body for a national central data repository centre in Kenya construction industry. Followed by public works at 22%, BORAQS at 13%, AAK at 7%, while ERB at 1%. However, it is important to note that 11% of the respondents advocated for formation of a new body, and reason was that the current bodies have no capacity to handle a BDT system.

Further, all respondents indicated the need of a one stop search engine/shop of various BIM projects. These findings form results base for the construction sector participants and players towards vetting and settling on which body can host the national BDT platform. Other existing bodies were mentioned once respectfully are JBCC, IQSK, KNBS, Ministry of ICT, KENNET.

Private owned central data repository centre

Interesting from the above findings, possibility of emergence of a private owned repository centre for the construction industry in Kenya was foreseen by 75% of the respondents. A central data repository centre (CDRC) is a data system which collects, stores, manages, and allows accessibility of stored information (Xie, Krystyna, & Matusiak, 2016).

Further allusion by Bilal et al. (2016) in a statement that to help supervise workplace tasks at real-time, a centralised big data repository should be supported, is well supported through the findings of this study. 98% of the respondents concurred that a well thought approach in actualization of a detailed digital data management system is a huge task that lies ahead of the Kenya's construction industry.

5.4 Summary

i) Firstly, extent of adoption of BDT in the construction sector in Nairobi County is yet to find its roots.
ii) Training of BDT in the universities and colleges need to be considered.
iii) Joint effort is needed in creating BDT awareness by BDT champions, employers, software companies, software sellers, and of course employees self-drive.
iv) All respondents agreed well managed data protects the many construction industry players.

v) 98.8% of the respondents agreed that effectively managing construction voluminous data is an important managerial task.

vi) 92% of the respondents' capture data electronically in their organisation.

vii) 77% of the respondents use real time data exchange platforms in executing their projects, and that is WhatsApp.

viii) 51% of the respondents use building information modelling (BIMo) in executing their projects.

ix) 15% of the respondents have used drones in their projects.

x) 50% submit data to a central digital repository platform, which this study feels are the office (servers & cloud platforms) which are configured such that one can access information from where ever they are.

CHAPTER SIX

CONCLUSION AND RECOMMENDATIONS

6.0 Introduction.

This chapter presents the conclusion and the recommendations with respect to the objectives of the study. The main objective of the study was to examine the application and viability of BDT in construction sector in the Kenya in order to inform all the integrating frameworks. The specific objectives were to: examine the status of application of BDT in construction sector, identify the factors that determine the application of BDT in construction sector, and determine the frameworks to use in integrating application of BDT into the construction sector.

6.2 Conclusion.

The study was based in investigating the application of BDT in the construction industry in Kenya against three research questions, as follows: what is the status and the extent of penetration of BDT in the construction industry?, to find out what are the factors that determine application of BDT?, and to find out what are the frameworks in use towards feeding the BDT platform in Kenya?.

On the status and extent of penetration of BDT in construction in Kenya, the construction industry has not embraced the application of BDT. On analysis 63% of the respondents had never heard of or used BDT. 27% level of awareness of BDT among the practicing consultants was very low. 32% were not conversant of any BDT characteristics while the remaining 68% sparsely knew the six BDT characteristics. On BDT skills 56% of the respondents had idea of them, 44% were clueless. On BD Technologies 43% of the respondents had idea of them, 57% were clueless. The foregoing findings revealed that the extent of penetration of BDT in the construction industry in Nairobi County is yet to find its roots.

From the literature review, the study adopted parameters that were subjected to field study and the respondents identified and ranked them in order of criticality. As such in identifying the factors that determine the application of BDT, the top critical ones were ranked as follows: skill gap that need to be addressed through formal training, upfront cost of setting up a BDT

infrastructure (machines, cloud services, software, expert people and strong internet connectivity), resistance to change and top management support. BDT benefits were also considered as factors that could determine the uptake of BDT.

According to this study the benefits of BDT use improves project planning & scheduling, saves on construction time (quickens execution time), huge amounts of data is handled, allows better project management, research on new proposed projects is improved. Lastly, from the BDT platform decision making is optimized, early problem detection is aided, better prediction of outcomes is realised, management information' factors (such as project schedule, safety, cost, quality) are integrated, and communication is enhanced.

On determining the frameworks to use in integrating application of BDT into the construction sector, social media platforms were the most used and known at 78%. Building information management was second at 70% level of awareness. Cloud computing was third at 45% of awareness. These frameworks are for feeding the central BD platform and require strong internet connectivity for BDT integration to take place. However, all legal and risks issues for instance data security, information privacy, data rights: needs to be legally and expertly managed.

In summary, the many diverse players in the construction industry need to integrate their operations towards BDT adoption and application. Integration will birth and encourage shared access to contract tools which include - designs and details, contract management software, models and procurement schedules. Currently in Kenya the operations of the construction industry are segmented. This can only be addressed by first mitigating the prevalent skill gap, readiness to meeting the capital cost, curbing resistance to change and top management being in the front line to support data integration.

6.3 Limitation of findings.

There are various limitations associated with the study. This research was mostly concerned with investigating the application of BDT in the construction industry in Kenya. As such, this study was confined to finding out whether BDT is known and applied in Kenya construction industry, the factorial benefits and limitations to BDT adoption in Kenya construction industry and finding out possible integrating platforms in use in Kenya. This is to say that, various

many other aspects were not studied, for instance (which data should be hosted on BDT platform).

In addition, the extent of BDT awareness in Kenya construction industry identified in this research is limited to this point in time of research. However, individuals' level of awareness of this new concept may change with time. Thus, the findings of this study will require regular updates.

The target response was principally limited to the trained and experienced construction professionals within Nairobi City County. Views, insights, and perception on BDT were not sought from players like the developers, investors, consultants of - big data, artificial intelligence, open data, internet of things. More so, the study never touched on in-depth into what each big data skill and role really entails. This study did not deep into detailed understanding of various BD technologies, such as machine learning (ML) and artificial intelligence (AI) fields where BD statistical procedures expand punctually finding trends in the captured information which is used in producing knowledge.

6.4 Contribution to knowledge.

The results from this study provides information about the most critical gap that needs to be filled towards realisation of BDT awareness and adoption in Kenya construction industry. Formal training needs to be considered in schools, colleges and universities, for it all starts with hearing. Practising construction professionals can be sensitized via seminars and workshops. This way the huge skill gap will be filled.

The findings are important to various stakeholders and players in the construction industry in helping them to realise that BDT is here to assist them execute and manage their projects. The research undertaken is significant in that it is an eye opener towards enhancing the productivity of the contracts in Kenya through the adoption of an integrated single most source of any construction information of the previously executed and ongoing projects.

Stakeholders and players will need to expand their mind towards accepting and not resisting change from the existing traditional systems to the new integrated BDT systems on new ways of handling projects data. The general practice of the many diverse industry players therefore

should adopt and implement integration in their operations from the convectional segmented approach. This is emphasized throughout this study.

In this study costs and expenses to BDT have been well discussed. Costs which are accrued by internet connectivity, machines, cloud services, software, and expert people. It is therefore important to note that the daily input of a data professional shall be a necessity in the construction firms and companies. BDT benefits and limitations were well outlined in the literature review and subjected to field study for ranking. Legal risks to digitizing processes, which can be expertly managed were also outlined in this study in section 2.9 thereafter subjected to the field study for ranking in section 4.3.2.

6.5 Recommendations.

The following recommendations are hereby made with the focus of enhancing the extent of knowhow and uptake of BDT in the Kenya construction sector. The study established that the extent of use of BDT in the construction sector in Nairobi County and thus Kenya is yet to find its roots.

In light to these findings, the study recommends that more efforts should be directed towards learning institutions educating students of the built environment by coming up with an education programme or units on BDT concepts. Practising professionals in the construction industry also need to be made aware, sensitised and educated of the BDT through seminars and workshops so that they are able to appreciate a national centralised single most digital way of handling projects data.

The study established the following parameters: 27 possible benefits of BDT use, 19 possible limiting factors, 22 integrating principles and 15 legal risks that come with BDT. Thus, the study recommends these parameters to be considered by the stakeholders and construction industry players in adopting BDT. It is all about weighing the expenses (costs there to) and the disadvantages against the benefits thereto.

A government policy is recommended. Views of the cabinet secretary and permanent secretaries of the MoTIHUD & PW should be sought and further sensitized towards the

actualization of BDT in the construction industry. Proper lobbying would deal with possible opposition (resistance to change).

From all the frameworks in use, this study recommends that there is a lot of relevant information that needs to be professionally and expertly handled from the social media platforms and building information management projects. Establishment of a national central repository centre is recommended by this study. This would not be well thought of without proposing a common template that would guide in channeling data to the central platform.

The foregoing is crowned by this study recommending that the many diverse industry players to consider integrating their operations towards curbing the segmentation nature of the industry. This is because digital data integration is the new norm and it is here to live with us and shall eventually evolve to BDT. Consequently, the potential data sharing risks can be dealt with legally and expertly, which will help in yielding free will data sharing.

Further, future analysis should be conducted to establish the variation of the non-significance level of low BDT application on projects between government and private employers after a greater adoption of the BDT by several entities in the construction industry sector is realized.

6.6 Areas of further research.

In Kenya's construction industry BDT is a new field of thinking, which need to be researched on at reasonable intervals in reference to this study objectives for updating the findings on the BDT application status. It is however, a wide discipline that future researchers can explore beyond the objectives of this study. For instance, study on the data type that should be wired to the big data platform such as [project management data, feasibility studies, challenges and lessons learnt in every project, monitoring and evaluation reports]

Views, insights, and perception on BDT application should be sought from players like the developers, investors, consultants on 'big data, artificial intelligence, internet of things', policy makers. An in-depth study also into what each BD skill and role really entails needs to be pursued. Also, a detailed study into understanding the various BDTs, such as machine learning

(ML) and artificial intelligence (AI) fields where BD statistical procedures expand punctually in finding trends from the captured information and use them to produce knowledge.

Research on other possible integrating platforms in use such as various construction management software, enterprise resource planning (ERP) systems. Further study can be done on analytics skill, towards identifying the big data analytics software that are the most ideal ones for the construction industry.

REFERENCES

Andeso, A. (2019). *Functions of the NCA Kenya.* [Retrieved May 25, 2020] from https://www.constructionkenya.com/2619/what-role-nca/

Akbar, S. (2018). B*ig Data in construction* [Retrieved May 10, 2019] from https://geniebelt.com/blog/big-data-in-construction

Aye, K. N., & Thein, T. (2015). *A Platform for Big Data Analytics on Distributed Scale-Out Storage System.* International Journal of Big Data Intelligence, 2(2), 127-141.

Baesens, B. (2014). *Analytics in a Big Data world – the essential guide to data science and its applications.* John Wiley & Sons, Inc., Hoboken, New Jersey. [Retrieved November 9, 2018] from https://www.researchgate.net/publication/320388773_Is_BIM_Big_Enough_to_Take_Advantage_of_Big_Data_Analytics

Bello-Orgaz, G., Jung, J. J., & Camacho, D. (2016). *Social Big Data: Recent Achievements and New Challenges.* Information Fusion, 28, 45-59.

Bilal, M., Lukumon O. O., Pasha, M., Qadir, J., Munir, K., Ayaji, S. O., Akinade, O. O., Owolabi, H. A., & Alaka, H. A. (2016). *Big Data in the construction industry: A review of present status, opportunities, and future trends.* Advanced Engineering Informatics 30(3):500-521. doi:10.1016/j.aei.2016.07.001 [Retrieved March 25, 2020] from https://www.researchgate.net/publication/305641258_Big_Data_in_the_construction_industry_A_review_of_present_status_opportunities_and_future_trends

Black, K. (2010). *Business Statistics: Contemporary Decision Making.* 6th Edition, John Wiley & Sons. [Retrieved May 20, 2020] from https://research-methodology.net/sampling-in-primary-data-collection/purposive-sampling/

Bloomberg, L. D., & Volpe, M. (2019). *Completing Your Qualitative Dissertation: A Road Map from Beginning to End.* Thousand Oaks, CA: Sage. [Retrieved June 6, 2020] from https://books.google.co.ke/books?id=bJGbDwAAQBAJ&pg=PA84&lpg=PA84&dq=A+case+study+is+a+comprehensive+analysis+of+an+event,+program,+group,+or+other+defined+topi

c+of+interest+(McMillan+2012)&source=bl&ots=JfmOqG_a_H&sig=ACfU3U367Qklyz9vwGmqwjpZ1ipbGUZYkg&hl=en&sa=X&ved=2ahUKEwjSoMK6pYXqAhWj6uAKHZtfCYsQ6AEwAHoECAwQAQ#v=onepage&q&f=false

Board of Registration of Architects and Quantity Surveyors of Kenya (BORAQS)

Bosch, E. (n.d). *LibGuides: Library Basics: Primary and Secondary Sources.* libguides.bgsu.edu. [Retrieved May 17, 2020] from https://en.wikipedia.org/wiki/Secondary_research

Bughin, J., Chui, M., & Manyika, J. (2010). *Clouds, Big Data, and Smart Assets: Ten Tech-Enabled Business Trends to Watch.* McKinsey Quarterly, 56(1), 75-86.

Building SMART alliance. (2017). *Transforming the Building Supply Chain Through Open and Interoperable Information Exchanges* [Internet]. National BIM Standard - United States (R) Version 3. Washington, USA. [Retrieved April 11, 2020] from https://www.researchgate.net/publication/322980261_Big_Data_in_Building_Information_Modeling_Research_Survey_and_Exploratory_Text_Mining

Burger, R. (2018). *How the Construction Industry is Using the Big Data*. [Retrieved November 15, 2018] from https://www.thebalancesmb.com/how-the-construction-industry-is-using-big-data-845322

Burger, R. (2019). https://www.thebalancesmb.com/how-the-construction-industry-is-using-big-data-845322 [Retrieved April 25, 2019]

Business dictionary, 2010.

Cao, G. (2010). *A four-dimensional view of IT business value. Systems Research and Behavioral Science,* 27(3), 267-284.

Chen, H., Chiang, R. H., & Storey, V. C. (2012). *Business Intelligence and Analytics: From Big Data to Big Impact.* MIS quarterly, 36(4).

Church, R. M. (n.d). *The Effective Use of Secondary Data* (Issue 1, Vol. Volume 33). [Retrieved May 25, 2020] from https://en.wikipedia.org/wiki/Secondary_research

CRC Construction Innovation. (2007). *Adopting BIM for facilities management: solutions for managing the Sydney Opera House*, Cooperative Research Center for Construction Innovation, Brisbane, Australia. [Retrieved November 9, 2018] from https://www.researchgate.net/publication/320388773_Is_BIM_Big_Enough_to_Take_Advantage_of_Big_Data_Analytics

Data Mining. (2017). *Data Mining*. [Retrieved April 6, 2020] from https://simple.wikipedia.org/wiki/Data_mining

Definition of Interoperability. (n.d). *Dedicated Website for A Definition of Interoperability at Interoperability - Definition Info.* Copyright AFUL under CC BY-SA. [Retrieved April 6, 2020] from https://en.wikipedia.org/wiki/Interoperability

Data in the Construction Industry. (2019). *Data in the Construction Industry* [Retrieved March 6, 2020] from https://www.designingbuildings.co.uk/wiki/Data_in_the_construction_industry Last edited 15 Mar 2019

Deutsch, R. (2015a). *Data Driven Design and Construction: 25 Strategies for Capturing, Analyzing, and Applying Big Data*. New Jersey: John Wiley & Sons. [Retrieved March 3, 2020] from http://www.tojdac.org/tojdac/VOLUME8-SPTMSPCL_files/tojdac_v080SSE197.pdf

Digital Technologies Hub. (2019). General-purpose programming. [Retrieved May 7, 2020] from https://www.digitaltechnologieshub.edu.au/teachers/topics/general-purpose-programming-coding

Dodge Data & Analytics. (n.d). *Find the right projects. Build the winning bids.* [Retrieved April 30, 2020] from https://www.construction.com/company/about

Dong, L., Neufeld, D., & Higgins, C. (2009*). 'Top Management Support of Enterprise Systems Implementations',* Journal of Information Technology, 24(1), 55-80.

Ferkoun, M. (2014). *Cloud Computing and Big Data: An Ideal Combination.* [Retrieved June 4, 2020] from https://www.ibm.com/blogs/cloud-computing/2014/02/cloud-computing-and-big-data-an-ideal-combination/

Fichman, R. G., & Kemerer, C. E. (1993*). Adoption of Software Engineering Process Innovations: The Case of Object Orientation.* Sloan Management Review. 34(2), 7-22.

Frost, L. A., & Sullivan, D. L. (2018). *Application of Big Data Analytics in Construction Industry 2018-2022: With Big Data Coming into Picture Construction Industry is Expected to Witnesses Transformational Growth*. [Retrieved November 9, 2018] from https://www.reportlinker.com/p05567610/Application-of-Big-Data-Analytics-in-Construction-Industry-Forecast-to.html

Fujikawa, K. (n.d). *The 6 Challenges of Big Data Integration.* [Retrieved October 25, 2018] from https://www.flydata.com/the-6-challenges-of-big-data-integration/

Gambatese, J. A., & Hallowell, M. (2011). *'Enabling and Measuring Innovation in the Construction Industry',* Construction Management and Economics, 29(6), 553-567.

Gharajeh, S. M. (2018). *Big Data Technology: A Deep Dive into NoSQL Databases: The Use Cases and Applications.* [Retrieved November 9, 2018] from https://www.sciencedirect.com/topics/computer-science/big-data-technology

Goes, P. B. (2014). *Editor's Comments - Big Data and IS Research.* MIS Quarterly, 38(3).

Heger, C. (2014). *Perfection at the connection: BIM in the field bringing an assembly line approach to modern construction. In BIM Forum*. Dallas, USA. [Retrieved November 9, 2018] from https://www.researchgate.net/publication/320388773_Is_BIM_Big_Enough_to_Take_Advantage_of_Big_Data Analytics

Hendrickson, C. (1998). *Project Management for Construction, Fundamental Concepts for Owners, Engineers, Architects and Builders,* 2nd Edition. Prepared for world wide web publication in 2000: Version 2.2 prepared Summer, 2008. [Retrieved March 5, 2020] from https://www.cmu.edu/cee/projects/PMbook/index.html

Hosey, T. (2018). *Reinventing Construction Industry with Big Data.* [Retrieved February 25, 2020] from https://www.geospatialworld.net/article/reinventing-construction-industry-with-big-data/

Institution of Construction Project Managers of Kenya – ICPMK. (2020).

Interesting Engineering. (2018). *Connected Vehicles in Smart Cities: The Future of Transportation* Published by interestingengineering.com on 16 November 2018, [Retrieved April 2, 2020] from https://en.wikipedia.org/wiki/Smart_city#cite_note-1

Jadeja, B., & Issa, T. (2017). *Big Data - A New Technology Trend and Factors Affecting the Implementation of Big Data in Australian Industries.* [Retrieved January 7, 2020] from https://link.springer.com/chapter/10.1007/978-3-319-67925-9_11

Jiao, Y., Wang, Y., Zhang, S., Li, Y., Yang, B., & Yuan, L. (2013). *A Cloud Approach to Unified Lifecycle Data Management in Architecture, Engineering, Construction and Facilities Management: Integrating BIMs and SNS.* Advanced Engineering Informatics, 27(2), 173-188.

Jreport. (n.d). *What are Business Intelligence Tools?* [Retrieved on April 6, 2020] from https://www.jinfonet.com/resources/bi-defined/business-intelligence-tools/

Karola, K. (2015). 16 Essential Project KPIs That Benefit the Entire Team. [Retrieved June 6, 2020] from https://www.scoro.com/blog/16-essential-project-kpis/

Kenya National Bureau of Statistics - KNBS. (2019).

Kiganda, A. (2017). *How to register with Institution of Construction Project Managers of Kenya (ICPMK).* [Retrieved May 25, 2020] from https://constructionreviewonline.com/2017/05/register-institution-construction-project-managers-kenyaicpmk/ Modified date: May 2, 2017

Kitchin, R., & McArdle, G. (2016). "What makes Big Data, Big Data? Exploring the ontological characteristics of 26 datasets". Big Data & Society. 3 (1): 205395171663113. doi:10.1177/2053951716631130. [Retrieved October 30, 2019] from https://en.wikipedia.org/wiki/Big_data

Kiran, R. (2019). *What is Big Data Technology?* [Retrieved May 10, 2020] from https://www.edureka.co/blog/top-big-data-technologies/

Kothari, C. K. (2004). *Research Methodology: Methods and Techniques.* New Delhi. New Age International Publishers Ltd.

Koutsogiannis, A. (2016). *What are the Main Stages of Construction Projects?* [Retrieved March 4, 2020 from https://www.letsbuild.com/blog/6-stages-of-a-project-in-construction

Kraemer, H. C., Mintz, J., Noda, A., Tinklenberg, J., Yesavage, J. A. (2006). *Caution regarding the use of pilot studies to guide power calculations for study proposals.* Arch Gen Psychiatry. 2006; 63((5)):484 – 9. [PubMed] [Google Scholar] [Retrieved June 6, 2020] fromhttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC4453116/#:~:text=A%20pilot%20study%20can%20be,to%20identify%20potential%20problem%20areas

Kruskal, W. H., & Wallis, W. A. (1952). "Use of ranks in one-criterion variance analysis". Journal of the American Statistical Association. 47 (260): 583–621. doi:10.1080/01621459.1952.10483441 [Retrieved June 4, 2020] from https://en.wikipedia.org/wiki/Kruskal%E2%80%93Wallis_one-way_analysis_of_variance

LaMorte, W.W. (2017). *Non-parametric tests. Mann Whitney U Test (Wilcoxon Rank Sum Test).* [Retrieved June 2, 2020] from http://sphweb.bumc.bu.edu/otlt/MPH-Modules/BS/BS704_Nonparametric/BS704_Nonparametric4.html

Lancaster, G. A., Dodd, S., Williamson, P. R., (2004). *Design and analysis of pilot studies: recommendations for good practice.* J Eval Clin Pract. 2004; 10((2)): 307 – 12. [PubMed] [Google Scholar] [Retrieved June 6, 2020] from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4453116/#:~:text=A%20pilot%20study%20can%20be,to%20identify%20potential%20problem%20areas

Laney, D. (2001). *"3D data management: Controlling data volume, velocity and variety".* META Group Research Note. 6 (70). [Retrieved September 27, 2019] from https://en.wikipedia.org/wiki/Big data

Lee, J. (1994). *"Odds Ratio or Relative Risk for Cross-Sectional Data?"*. International Journal of Epidemiology. 23 (1): 201–3. doi:10.1093/ije/23.1.201. PMID 8194918. [Retrieved June 7, 2020] from https://en.wikipedia.org/wiki/Cross-sectional_study#cite_ref-1

Liu, J., Byna, S., & Chen, Y. (2013). *"Segmented Analysis for Reducing Data Movement", In Proceedings of the 2013 IEEE International Conference on Big Data* (IEEE Big Data 2013), October 6-9, 2013, Santa Clara, CA, USA, pp. 344-349.

Lu, J., Yu, C-S., & Liu, C. (2005). *'Facilitating Conditions, Wireless Trust and Adoption Intention',* The Journal of Computer Information Systems, 46(1), 17-24.

Lynch, J. (2018). *Why Using Big Data Mitigates Risk and Helps Construction Businesses Take Flight*. [Retrieved May 10, 2019] from, https://www.autodesk.com/redshift/big-data-in-construction/

Lynch, J. (2018). *Why Using Big Data Mitigates Project Risk and Boosts Outcomes*. [Retrieved May 10, 2019] from https://www.enr.com/articles/45934-why-using-big-data-mitigates-project-risk-and-boosts-outcomes

Maaz, Z. N., Bandi, S., & Amirudin, R. (2018). *Big Data in The Construction Industry: Potential Opportunities and Way Forward.* [Retrieved January 7, 2020] from http://www.tojdac.org/tojdac/VOLUME8-SPTMSPCL_files/tojdac_v080SSE197.pdf

The Turkish Online Journal of Design, Art and Communication – TOJDAC ISSN: 2146-5193, September 2018 Special Edition, p.1470-1480

Magoulas, R., & Lorica, B. (2009). "*Introduction to Big Data*" [Retrieved March 3, 2020] from https://www.oreilly.com/data/free/release-2-issue-11.csp Release 2.0. Sebastopol CA: O'Reilly Media (11).

Mahajan, R., Parthasarathy, S., & Jain, V. (2018). *Managing Risk in Digital Transformation Risk Advisory.* [Retrieved May 19, 2020] from https://www2.deloitte.com/content/dam/Deloitte/in/Documents/risk/in-ra-managing-risk-in-digital-transformation-1-noexp.pdf

Manuel, N. (2017). *Top 7 Big Data Skills That Are High in Demand*. [Retrieved on 25th Oct 2019] from https://www.springpeople.com/blog/top-big-data-skills-that-are-high-in-demand/

Martínez, M., Marin, N., & Vila, M. A. (2015). *The Role of Information Technologies to Address Data Handling in Construction Project Management*. Journal of computing in civil engineering 30(4) July 2016. [Retrieved October 17, 2018] from https://ascelibrary.org/doi/full/10.1061/%28ASCE%29CP.1943-5487.0000538

Mbaluka, W. (2013). *Big Data Management and Business Value in the Commercial Banking Sector in Kenya:* A Research Project Submitted in Partial Fulfillment of the Requirements of the Degree in Master of Business Administration University of Nairobi (Unpuplished). [Retrieved November 9, 2018] from http://erepository.uonbi.ac.ke/handle/11295/58632

Mencarini, R. (2014). *Big Data and BIM: turning model information into insight. In BIM Forum*. Dallas, USA. [Retrieved November 9, 2018] from https://www.researchgate.net/publication/320388773_Is_BIM_Big_Enough_to_Take_Advantage_of_Big_Data_Analytics

Meola, A. (2016). *What is the Internet of Things (IoT)?* [Retrieved on 6th June 2020] from http://www.businessinsider.com/what-is-the-internet-of-things-definition-2016-8/?IR=T

Mohan, C. (2013*). History Repeats Itself: Sensible and NonsenSQL Aspects of the NoSQL Hoopla (PDF).* Proc. 16th Int'l Conf. on Extending Database Technology. [Retrieved April 6, 2020] from https://en.wikipedia.org/wiki/NoSQL#cite_note-7

Mugenda O. & Mugenda A. (1999). *Research Methods: Quantitative and Qualitative Approaches.* Nairobi: African Centre for Technology Studies.

Mugenda O. & Mugenda A. (2003). *Research Methods: Quantitative and Qualitative Approaches.* Nairobi: African Centre for Technology Studies.

My Market Research Methods. (2018). *Types of Data & Measurement Scales: Nominal, Ordinal, Interval and Ratio.* [Retrieved June 6, 2020] from https://www.mymarketresearchmethods.com/types-of-data-nominal-ordinal-interval-ratio/

National BIM Standard. (2014). *Frequently Asked Questions About the National BIM Standard* - United States - Archived from the original on 16 October 2014 [Retrieved March 17, 2020] from https://en.wikipedia.org/wiki/Building_information_modeling

National Construction Authority -NCA. (2020).

Ndambo, D. (2016). *Big Data Analytics and Ccompetitive Advantage of commercial Banks and Insurance Companies in Nairobi, Kenya.* A Research Project Submitted in Partial Fulfilment of the Requirements of Master of Business Administration Degree, School of Business, University of Nairobi (Unpuplished). [Retrieved November 9, 2018] from http://erepository.uonbi.ac.ke/handle/11295/58632

Novotny, R. (2018). *What is Real Time Data? How Real Time Information is Essential in the Field.* [Retrieved January 9, 2020] from https://esub.com/what-is-real-time-data-why-real-time-information-is-essential-in-the-field/

Nyaga, E. N. (2016). *An Assessment of the Level of Adoption of Information Communication Technology by Local Building Contractors During Project Implementation (a Case for Kenya) Jomo Kenyatta University of Agriculture and Technology (Unpublished).* [Retrieved on April 5, 2021] from http://ir.jkuat.ac.ke/handle/123456789/2449

Olive, R. (2019). *Key Performance Indicators in construction - Why it matters.* [Retrieved on June 5, 2020] from https://wollamconstruction.com/key-performance-indicators-in-construction-why-it-matters/

Omran, B.A. (2016). *Application of Data Mining and Big Data Analytics in the Construction Industry.* [Retrieved February 25, 2020] from https://etd.ohiolink.edu/pg_10?::NO:10:P10_ETD_SUBID:119954

Ormsby, T. (n.d). *Search Smart: Types of Research: Secondary Research.* flinders.libguides.com. [Retrieved May 9, 2020] from https://en.wikipedia.org/wiki/Secondary_research

Otali, M. & Ujene, A. O. (2015). *Comparative Evaluation Of Managers And Construction Workers Productivity Factors In Nigerian Built Environment - Used Mann-Whitney U Test.* [Retrieved June 6, 2020] from https://www.researchgate.net/publication/328873970_comparative_evaluation_of_managers%27and_construction_workers%27_productivity_factors_in_nigerian_built_environment/figures?lo=1

Pal, K. (2015). *The Impact of the Internet of Things on Big Data.* [Retrieved June 5, 2020] from http://data-informed.com/the-impact-of-internet-of-things-on-big-data/

PAT Research. (2019). *Top 53 Bigdata Platforms and Bigdata Analytic Software*. [Retrieved September 25, 2019] from https://www.predictiveanalyticstoday.com/bigdata-platforms-bigdata-analytics-software/

Qubole. (2017). *Big Data Cloud Database & Computing.* [Retrieved June 17, 2020] from https://www.qubole.com/resources/big-data-cloud-database-and-computing/

Rahman, N. (2016). *Factors Affecting Big Data Technology Adoption.* [Retrieved May 20, 2020] from https://pdxscholar.library.pdx.edu/cgi/viewcontent.cgi?article=1099&context=studentsymposium

Regoniel, P. (2015). *Conceptual Framework: A Step by Step Guide on How to Make One*. [Retrieved May 10, 2019] from https://simplyeducate.me/2015/01/05/conceptual-framework-guide/

Richards, N. M., & King, J. H. (2014*). Big Data Ethics.* Wake Forest Law Review, 49, 393-432.

Rouse, M. (2018). *Data Warehouse: Big data management* [Retrieved May 18, 2019] from https://searchdatamanagement.techtarget.com/definition/big-data

Rouse, M. (2019). *Internet of Things (IoT) Agenda*. [Retrieved March 11, 2020] from https://en.wikipedia.org/wiki/Internet_of_things

Rukwaro, R. (2016). *Research Writing in Research.* Nairobi: Applied Research & Training Services.

Schmidt, C. O., & Kohlmann, T. (2008). *"When to use the odds ratio or the relative risk?".* International Journal of Public Health. 53 (3): 165–167. doi:10.1007/s00038-008-7068-3. PMID 19127890. [Retrieved June 7, 2020] from https://en.wikipedia.org/wiki/Cross-sectional_study#cite_ref-1

Sharron, S., Bosman, E., Do, K., & Lyon, C. (n.d.) *Five Key Legal Risks for Data-Centric Technology M&A.* The M&A Journal, 18(7). [Retrieved May 25, 2020] from https://media2.mofo.com/documents/180711-five-legal-risks-technology-ma.pdf

Sibiya, M., Aigbavboa, C., & Thwala, W. (2015). *Construction Projects' Key Performance Indicators: A Case of the South African Construction Industry.* [Retrieved June 6, 2020] from https://www.researchgate.net/publication/301376459_Construction_Projects'_Key_Performance_Indicators_A_Case_of_the_South_African_Construction_Industry

Stewart, P.W. (n.d). *Small or pilot study, GCRC protocols which propose "pilot studies".* Cincinnati Children's Hospital Medical Center. [Retrieved June 6, 2020] fromhttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC4453116/#:~:text=A%20pilot%20study%20can%20be,to%20identify%20potential%20problem%20areas

SurveyAnyPlace website. (n.d). *Closed Ended Question*. [Retrieved April 6, 2020] from https://help.surveyanyplace.com/en/support/solutions/articles/35000042308-closed-ended-question#:~:text=Closed%2Dended%20question%20advantages%3A,clarify%20question%20meaning%20for%20respondents

Syagga, P. (2018). *Research Methodology Class Notes*. University of Nairobi, Nairobi.

The Economist. (2010). *Data, Data Everywhere*. [Retrieved November 9, 2018] from https://www.economist.com/special-report/2010/02/27/data-data-everywhere Authored on 25 February 2010.

Thomas, L. (2020). An Introduction to Simple Random Sampling. [Retrieved June 6, 2021] from https://www.scribbr.com/methodology/simple-random-sampling/

Verma, A. (2018). *Top 10 Big Data Skills to Get Big Data Jobs*. [Retrieved April 6, 2020] from https://www.whizlabs.com/blog/big-data-skills/ Amit Verma

Volk, R., Stengel, J., & Schultmann, F. (2014). *"Building Information Modelling (BIM) for Existing Buildings - Literature Review and Future Needs," Automation in Construction,* vol. 38, pp. 109 – 127. [Retrieved April 11, 2020] from

https://www.hindawi.com/journals/complexity/2018/9691868/ and from https://www.sciencedirect.com/science/article/abs/pii/S092658051300191X?via%3Dihub

Wang, D., Fan, J., Fu, H., & Zhang, B. (2018). *Research on Optimization of Big Data Construction Engineering Quality Management Based on RNN-LSTM*. [Retrieved May 10, 2019] from https://www.researchgate.net/publication/326229619_Research_on_Optimization_of_Big_Data_Construction_Engineering_Quality_Management_Based_on_RNN-LSTM

Wasonga, M. (2019). *Leverage on technology to meet Kenya's construction demand.*

[Retrieved October 25, 2019] from https://www.cio.co.ke/leverage-on-technology-to-meet-kenyas-construction-demand/

We Build Value Digital Magazine. (2018). *Megatrend and Innovations on Big data: revolutionizing construction*: [Retrieved October 11, 2018] from https://www.webuildvalue.com/en/megatrends/big-data-revolutionizing-construction.html

Wilkinson, R. W. (1984). "*Computerized Specifications on a Small Project*," *ASCE Journal of Construction Engineering and Management,* Vol. 110, No. CO3, pp. 337-345.

Wood, L. (2019). *Singapore Data Centre Market Report 2019-2024 - Data Centre Construction is Being Led by Large Global Multimedia Companies - ResearchAndMarkets.com.* [Retrieved April 30, 2020] from https://www.businesswire.com/news/home/20191016005349/en/Singapore-Data-Centre-Market-Report-2019-2024--

Xie, I., Krystyna, K., & Matusiak, PhD. (2016). *Digital Preservations*, *in Discover Digital Libraries, Digital Repositories*. [Retrieved June 6, 2020] from https://www.sciencedirect.com/topics/computerscience/digitalrepository#:~:text=Digital%20repositories%20are%20information%20systems,provide%20access%20to%20digital%20content.&text=Institutional%20digital%20repositories%20serve%20multiple,%2C%20dissertations%2C%20and%20research%20data.

APPENDICES

Appendix 1: Field Study Introduction Letter

UNIVERSITY OF NAIROBI
DEPARTMENT OF CONSTRUCTION MANAGEMENT & QUANTITY SURVEYING
P.O. Box 30197, 00100 Nairobi, KENYA. *Tel: No. +254-020-491 3531*
E-mail: *dept-cmqs@uonbi.ac.ke*

Ref: B53/12221/2018 **Date:** 18th June, 2020

To Whom It May Concern

Dear Sir/Madam,

RE: MUCHOKI JEDIDA WAGUTHII

This is to certify that the above named is a student in the Department of Construction Management and Quantity Surveying, pursuing a course leading to Master of Arts in Construction Management degree. She is in her second year of study.

She is carrying out a research entitled ***"Application of Big Data Technology in the Construction Industry in Kenya: A Case Study of Nairobi County"*** in partial fulfillment of the requirements of the degree programme.

The purpose of this letter is to request you to allow her access to any kind of material she may require to complete her research. The information will be used for research purposes only.

CHAIRMAN
DEPARTMENT OF CONSTRUCTION
MANAGEMENT AND QUANTITY SURVEYING
UNIVERSITY OF NAIROBI

Isabella N. Wachira-Towey, (PhD)
Chair & Senior Lecturer,
Dept. Construction Management & Quantity Surveying

Appendix 2: Questionnaire

INTRODUCTION

Dear Respondent,

My name is Muchoki Jedida Waguthii and I am a postgraduate student at UoN studying Master of Art in Construction Management. This survey forms part of my thesis whose research topic is APPLICATION OF BIG DATA TECHNOLOGY IN THE CONSTRUCTION INDUSTRY IN KENYA: A Case Study of Nairobi County.

This questionnaire aims to collect information related to application and status of big data technology (BDT) in the construction industry. Perception of users on a central digital centre for data storage, analysing and mining of information is what this study is seeking to know. Further, collecting information on the factors that may determine the application of BDT and information related to extent of use of various integrating frameworks available in application of BDT such as building information management (BIMa), real time data (RTD) models, cloud computing, Internet of Things, and Social Media Platforms' in Nairobi County.

To sum up, it is all about the extent of adoption, factors/parameters to be considered in acceptance of BDT, perception of users on BDT, and strategies that can be put in place to increase the uptake of BDT and related platforms.

Declaration: You have been selected to take part in this study, your participation is of importance and is greatly appreciated. The study is governed by the ethics of Social Science Research. The information given is for academic purpose only and will be treated in anonymity and with utmost confidentiality and it will be analysed together with that of other respondents. Kindly tick (√) the box that matches your answer to the questions and give the answers in the spaces provided as appropriate.

SECTION A: BACKGROUND INFORMATION

1. What is your Profession

☐ Construction Project Manager ☐ Architect ☐ Quantity Surveyor

☐ Engineer _____ (SPECIFY) ☐ Facility Manager

☐ Other (Please specify)

2. Do you also double up as

☐ an Official in any professional anybody ☐ Lecturer ☐ Contractor ☐ **NONE n/a**

☐ Other (Please specify)

3. Education: What is the highest degree or level of school you have completed?

☐ Certificate ☐ Diploma ☐ Bachelor's degree ☐ Master's degree ☐ Doctorate

☐ Other (Please specify)

4. Who is your Employer?

☐ Government Agency / Department ☐ Private Entity ☐ Non-Governmental

☐ Other (Please specify)

4. What is your Gender?

☐ Female ☐ Male

5. Number of years of experience?

☐ 0-5 ☐ 6-10 ☐ 11-15 ☐ Over 15

SECTION B: BDT EXTENT OF ADOPTION

1. Please select your most preferred choice for the following BD definition:

Big Data is a large volume of structured, semi-structured and unstructured data that has the potential to be mined for information and used in machine learning projects and other

advanced analytics applications.

☐ Strongly agree ☐ Agree ☐ Uncertain ☐ Disagree ☐ Strongly disagree

2. Have you adopted BDT in your current organization?

☐ Yes ☐ No ☐ Planning to adopt ☐ Don't plan to adopt

☐ Other (please elaborate)

3. For how long have you used big data technology (BDT)?

☐ 0 ☐ **1-5** ☐ 6-10 ☐ 11-15

☐ Other (Please specify)

4. Have you ever (in previous organisations) used BDT in any construction projects?

☐ Yes ☐ No ☐ Unsure

☐ Other **similar technologies** (please elaborate)

If NO, why? / reason

.....

.....

5. Has your organisation hired an 'information management specialists (IMS)'? For instance, a BDT manager, a data scientist and such.

☐ Yes ☐ No ☐ Planning to adopt ☐ Don't plan to adopt

☐ Other (please elaborate)

6. Do you capture information electronically and store it in your organisation?

☐ Yes ☐ No

7. Real time data (RTD) is data that is up to date and viewable the moment it is available. e.g. (WhatsApp groups, material and fleet tracking, determining clock ins & clock outs).

Do you use real time data exchange in your projects?

☐ Yes ☐ No

Comment further ...

8. Do you use building information modelling in executing your projects?

☐ Yes ☐ No

9. Do you use drones in executing your projects?

☐ Yes ☐ No

10. A central digital repository is information systems that ingest, store, manage, preserve, and provide access (retrieval) to digital content.

Have you ever heard of a centralized database management system (CDMS) or one stop search engine?

☐ Yes ☐ No

11. Do you normally submit data to a central digital repository platform?

☐ Yes ☐ No

Comment further …

12. In which stage of a project's lifecycle would you or your organization introduce BDT to projects?

☐ Conception ☐ Design stage ☐ Pre-Construction ☐ Construction ☐ Post-Construction ☐ Operation ☐ Maintenance Stage

13. What V's or characteristics of BDT are you conversant with? Tick as many as applicable

☐ Volume ☐ Variety ☐ Velocity ☐ Veracity ☐ Value ☐ Variability ☐ **NONE n/a**

☐ Other (please state)

BDT Technology

1. Are you conversant with these Big Data Technologies fields (BDT fields)? (Please tick on as many as it applies to your knowhow)

☐ Data Storage ☐ Data Mining ☐ Data Analytics ☐ Data Visualization ☐ **NONE n/a**

4. Are you conversant with these BD Technologies? Please tick on as many as it applies to your knowhow)

☐ Hadoop ☐ MongoDB ☐ RainStor ☐ Hunk

☐ Presto ☐ RapidMiner ☐ Elasticsearch

☐ Apache Kafka ☐ Splunk ☐ KNIME ☐ Spark ☐ R-Language ☐ BlockChain,

☐ Tableau ☐ Plotly

☐ TensorFlow ☐ Apache Beam ☐ Docker ☐ Airflow ☐ Kubernetes ☐ **NONE n/a**

5. Are you conversant with these BD skills? Please tick on as many as it applies to your knowhow)

☐ Apache Hadoop ☐ Apache Spark ☐ Data Mining and Machine Learning ☐ NoSQL

☐ Data Visualization ☐ General Purpose Programming Skills ☐ Quantitative Analysis Skills ☐ Problem Solving Skills ☐ SQL Skills ☐ **NONE n/a**

6. Please select your most preferred choice for the following statement on BDT hardware: Hardware must be as recommended by the type of software in use.

☐ Strongly agree ☐ Agree ☐ Uncertain ☐ Disagree ☐ Strongly disagree

BDT Education and Training

1. Did you get educated on BDT from University/College?

☐ Yes ☐ No

2. If no in the above, who gave you your Primary initial BDT know-how/ training?

☐ Software companies ☐ Software sellers ☐ BDT champion ☐ My employer

☐ Self-Drive

☐ Other (please specify)

3. General purpose programming is a coding language, because computers are unable to think for themselves, therefore they require users to give them sets of ordered instructions to know what to do.

As such, it is becoming necessary for all professionals in whichever industry to have knowhow of general purpose programming (GPP). Do you agree?

☐ Yes ☐ No

4. Would you advocate for inclusion of IT (GPP) in the construction related courses in the tertiary curriculum? e.g. Construction management with IT, Engineering with IT

☐ Yes ☐ No

SECTION C: IMPORTANT COMPONENTS OF BDT

1. Please rate each importance of BDT components in BDT adoption.

	1	2	3	4	5
	Very important	Important	Uncertain	Less important	Not important
Benefit based statements.					
1. Quicker works execution timelines are realized (time saving). Shortens construction time.					
2. Improved research on new proposed projects					
3. Improved project planning & scheduling.					
4. Provides accurate data.					
5. Increases efficiency in the workplace.					
6. Gives answers to complex questions.					
7. Improved resource allocations					
8. Real time capabilities.					
9. Transparency (information is accessible).					
10. Allows predictions to reduce the risks in future.					
11. Increased data collection techniques.					
12. Integration and collaboration of platforms.					
13. Managing projects efficiently. Allows better project management.					

14. Determine the most profitable projects to pursue.					
15. Informed decision making. Brings more insights, ultimately leading to more information driven decisions.					
16. Cost saving (lowering costs) curbs cost overruns.					
17. Quality improvement					
18. Improved return on investment – ROI (the profit from an activity for a particular period compared with the amount invested in it)					
19. Enhances productivity.					
20. Consistent, up to date project information, that is, (lifecycle information management)					
21. Risk management. Reduces project risks (due to access to the relevant historical information).					
22. Allows improved safety and health management.					
23. Capability to handles huge amounts of data.					
24. Quick response times is realised.					
25. Opens up for central repository storage & retrieval.					
26. Improved budget estimates and superior understanding of costs.					
27. Brings knowledge thus better planning.					

2. Please rate the limiting factors in BDT adoption (reasons for not using BDT)

	1	2	3	4	5
	Strongly agree	Agree	Uncertain	Disagree	Strongly disagree
1. Lack of top management support.					
2. Upfront cost of adoption (too expensive to set up the BD infrastructure)					
3. Technological capabilities.					
4. Pressure to remain competitive.					
5. Size of a project.					
6. Information privacy.					
7. Information security.					
8. Skill gap in BDT					
9. Formal training is lacking					
10. The complexity behind getting data in the BD structure.					
11. Data un-synchronizing risks.					

12. Ability to ensure that you are getting the right data.					
13. Trusting data integrated from various sources.					
14. Uncertainty created by wide range of BD tools					
15. No client / government requirement to use it					
16. Satisfied with existing system					
17. Lack of standards and guidelines					
18. Lack of government policy					
19. Resistance to change					
20. I don't really understand BDT					

☐ Other limitations (please list here)

SECTION D: PROJECT DATA / INFORMATION

1. Effectively managing construction voluminous data is an important managerial task?

☐ Yes ☐ No

☐ Other opinion (Please specify)

2. Please select your most preferred choice for the following statements:

Poor or missing information can readily lead to project delays, **uneconomical decisions, or even** the complete failure of the proposed project.

☐ Strongly agree ☐ Agree ☐ Uncertain ☐ Disagree ☐ Strongly disagree

3. Well thought approach in actualization of a detailed digital data management system is a huge task that lies ahead of the Kenya's construction industry?

☐ Yes ☐ No

4. Data creates information, produces knowledge and it is applied to give wisdom in allowing decisions to be made. Can the 'past historical projects data' be used towards mitigating time schedule breaches throughout the project cycle?

☐ Yes ☐ No

☐ Other opinion (Please specify)

5. Availability of past or ongoing project information guides the research of a new such like similar project and sets pace?

☐ Strongly agree ☐ Agree ☐ Uncertain ☐ Disagree ☐ Strongly disagree

6. Legally and contractually, well managed data protects the many construction industry players.

☐ Yes ☐ No

7. In your opinion who are the consumers of data in the construction industry in Kenya? Select as many as applicable.

☐ Construction project managers ☐ Architects ☐ Quantity surveyors ☐ Engineers

☐ Contractors ☐ Site supervisors ☐ Policy makers ☐ The ruling government

☐ Academicians & Students

☐ Other opinion (Please specify)

8. Do you concur with that digital data mining can solve the consistent and historical plague of schedule overruns, cost overruns, ineffective construction management and poor record keeping?

☐ Yes ☐ No

9. Would you advocate for a one stop search engine / shop' platform, otherwise a "centralized database management system (CDMS)" of various BIM projects?

☐ Yes ☐ No

10. Why would you need data? (Tick all applicable to you)

☐ To make work execution go faster ☐ To make construction sites safer ☐ To know factors affecting profitability ☐ To forecast future best or worst case events ☐ For fore warning on specific occurrences ☐ Accurate financial and project information ☐ To understand behaviour ☐ To improve market competitiveness ☐ To help allocate resources.

SECTION E: BDT AND INTEGRATION PLATFORMS

1. Which integrating platforms are you conversant with? (Tick all that you have used)

☐ Building Information Management (BIMa) ☐ Internet of Things (IoT) / Sensor Networks / Telematics (GPS devices) ☐ Cloud computing ☐ Real time data (RTD) ☐ Social Media Platforms e.g. WhatsApp groups/Facebook/LinkedIn/Twitter ☐ **NONE n/a**

☐ Any other (Please specify)

2. Please select your most preferred choice for the following listed requirements and statements for the BDT integration platform to function.

	1	2	3	4	5
BDT Platform Integration Principle	Strongly agree	Agree	Uncertai	Disagree	Strongly disagree
1. Internet usage and strong connectivity is a necessity.					
2. Enhanced interoperability support integration.					
3. Integrations demands computational expertise.					
4. Information is transformed.					
5. Knowledge is discovered.					
6. Better prediction outcomes.					
7. Processes and analyses are automated.					
8. Optimizes decision making.					
9. BDT is fed by the integration platforms while integrating platforms extract information from BDT.					
10. Management information' factors (such as project schedule, safety, cost, quality) are integrated.					
11. Integration births full information model throughout full project lifecycle.					
12. Integration births transparency and a more proactive approach.					
13. Integration enhances communication.					
14. Integration reduces construction time.					
15. Integration reduces construction cost.					
16. Promotes positive return on investment (ROI).					
17. Integration improves work quality.					
18. Integration improves good working relationship.					
19. Integration aids in early problem detection.					
20. Integration helps to reduce variations.					
21. Integration helps to reduced claims.					
22. Implementation of integration demands many					

resources (machines, cloud services, software, and expert people).					

3. There are several challenges one can face during the integration, such as: data capturing challenges, data analysis challenges, content curation challenges, data sharing challenges, data search challenges, data visualization challenges, information privacy, data transfer challenges, data querying challenges, data updating and data storage challenges. Do you agree?

☐ Yes ☐ No

3. Please rate the legal risks that come with BDT platform.

	1	2	3	4	5
	Very important	Important	Uncertain	Less important	Not important
1. Intellectual property rights.					
2. Sharing of copyright data.					
3. Professional liability.					
4. Condition of contracts.					
5. Data interoperability.					
6. Processes and responsibilities.					
7. Data security.					
8. Information privacy.					
9. Lack of BDT standards.					
10. Standard of care and professional negligence.					
11. Model management.					
12. Admissibility of electronic based documents.					
13. Legislation and judicial precedence.					
14. Legal validation of design (NCC submissions).					
15. Cost compensation.					

4. BDT brings in new data roles and titles through integration to BDT in the construction industry (Select as many as applies to your knowhow)

☐ Data Analyst ☐ Data Administrator ☐ Data Scientist ☐ Data Architect ☐ Data manager ☐ Data engineer ☐ Data Consultant

☐ Other (please specify)

5. Which regulating/professional body or state department would you propose as a national central data repository centre in Kenya? (Tick only one and state the reason why to your choice)

☐ NCA ☐ PUBLIC WORKS ☐ BORAQS ☐ ERB ☐ AAK

☐ Other (please specify) _____

Explain why your choice…

6. Do you foresee emergence of a private owned central data repository centre for the construction industry in Kenya?

☐ Yes ☐ No

THANK YOU.