Exploring Data Science Applications for Construction Project Performance in the South African Construction Industry

Wanda Buhle Mpingana[1], Opeoluwa Akinradewo[1], Clinton Aigbavboa[1] and Percy Bongokuhle Msibi[2]

[1] cidb Centre of Excellence, Faculty of Engineering and the Built Environment, University of Johannesburg, South Africa
[2] Department of Construction Management and Quantity Surveying, Faculty of Engineering and the Built Environment, University of Johannesburg, South Africa
wanda101buhle@gmail.com

Abstract
Construction projects create vast amounts of data which is produced throughout the project's lifecycle in the form of building plans, schedules, cost estimates, bills of quantities and variation orders, etc. Data science has the potential to enhance processes like project delivery, decision-making, equipment and personnel management, transparency, and communication. Hence, it has been noted as an upcoming technology that could assist the construction industry in meeting the data increase. Despite the significant benefits associated with the implementation of data science, the South African construction industry (SACI) continues to have problems with budget overruns, time delays, and the project being delivered below quality standards. Thus, the aim of the paper was to investigate data science applications for construction project performance in the SACI. The research was carried out using the quantitative approach through a well-structured questionnaire that was distributed to construction professionals in South Africa (SA). A total of 55 responses were received from the participants, and the data were analysed using descriptive statistical procedures using the computer software SPSS. The study disclosed the major applications of data science for construction project performance to be in tracking construction equipment and assets, project cost management, and operations and maintenance. The research findings prove that there is still a low adoption of data science for all facets of construction project performance as construction equipment and assets, project cost management as well as operations and maintenance relate to cost performance not quality and time performance.

Keywords
Construction project performance, Data Analytics, Data Science, Predictive Analytics, and Project Delivery.

1. Introduction

In the age of disruptive digital technologies, the rate at which data is accumulated has rapidly increased (Munawar *et al.,* 2022). This increase in data has not spared the construction industry (CI), in fact, Aghimien *et al.* (2019) strongly believe that the amount of data that is generated during the delivery of a construction project and the use of traditional tools for data analysis have made it difficult to make sense of this data. Furthermore, Tang *et al.* (2017) suggest that the exponential increase in the amount of information that is exchanged on a daily basis from multiple electronic devices such as cell phones, computers, cameras, laser scanners, and laptops has contributed immensely to the inundation of data in the construction industry (CI). Xu *et al.* (2020) postulate that poor planning, budgeting, cost overruns, miscalculations, insufficient funding for the project, low return on construction assets, and poor quality control are some of the challenges that have always affected the CI.

The nature of the CI is such that large amounts of construction information is generated throughout the lifespan of a project (Bilal *et al.,* 2019). Typically, this information is conveyed verbally or through written materials on paper, including 2D drawings. Cheng and Teizer (2010) note that traditional methods of delivering information have not always been able to make speedy and accurate decisions due to inadequate or poorly written data. The need for data science in the CI resulted from various construction stakeholders and researchers trying to gain value from this data and for construction companies to improve the overall project performance and increase profits (Renukappa *et al.,* 2021). Hence, the purpose of the study is to assess the applications of data science for construction project performance to improve the adoption of data science in aiding project delivery in the SACI. The findings of this study will create awareness among construction professionals where data science can be applied to aid with the current challenge of a data-intensive industry (Aghimien *et al.,* 2019). This paper will also contribute to the body of knowledge of data science in the CI.

2. Data Science

Data Science as a concept was first introduced in the mathematics and statistics field, which at the time only focused on data analysis (Cao, 2017). According to Xu *et al.* (2020), data science was first discovered in the 1960s. During this period, the term "datalogy" was commonly used, which was defined as the science of processing data (Hwang and Chen, 2017). However, Xu *et al.* (2020) note that this concept gave data processing status in the science community and gave origin to the "data science" field as we commonly know it today. The exponential growth of data science has been accelerated in recent years by the rise of the 4IR, rendering it a universally recognized emerging field (Ram *et al.,* 2019).

In recent years, data science has been gaining traction and attention and is being widely adopted in various industries such as banking, automobile, manufacturing, healthcare, etc. (Xu *et al.*, 2020). The field has gone beyond specific areas such as data mining and machine learning, which is attributed mainly to technological advancements and research efforts in developing advanced analytics, cultivating new algorithms, and improving data models (Cao, 2017). According to Cao (2017), data science is a field that consists of research efforts from different disciplines that create innovations to move beyond the discipline, which basically draws and recreates a variety of pertinent disciplines and sources of knowledge, such as mathematics, communication, statistics, computing, informatics, sociology, and management. A definition by Martinez *et al.* (2021) suggests that "data science is the intersection of computer science, business, engineering, statistics, data mining, machine learning, operations research, six sigma, automation, and domain expertise".

While a study by Hwang and Chen (2017) notes that data science is a sector that necessitates managing the massive volume of data (text, audio, images, and files) and uses it for creating prescriptive, prescribed and expected rational models. Therefore, based on the aforementioned definitions, there is an overall agreement that data science seeks to effectively look into the complex nature of data to extract value for the business. The professionals who are tasked with analyzing and processing this data to solve problems and to extract value are known as data scientists. Due to the high demand of this skill, there has been a growing interest in defining the work carried out by data scientists. According to Martinez *et al. (*2021), a data scientist manages everything from gathering data to processing it in detail and visually and presenting the data in an understandable manner. In contrast, Warden (2011) describes a data scientist as someone who excels at software engineering and statistics more than any other engineer. Data scientists seek to transform raw data to information, information to a source of knowledge, and from knowledge to decision-making (Xu *et al.*, 2020). According to Hwang and Chen (2017), the field of data science serves as the cornerstone for big data, providing the essential tools for its analysis and processing as well as a practical strategy for maximizing its potential. Furthermore, Hwang and Chen (2017) note that data science skills are needed to solve problems affecting the CI for the longest time and make building easier at all levels. In addition, Liu (2017) and Xu *et al.* (2020) note that construction organisations can use the value of data science and data scientists to manage and enhance construction sites and for problem-solving processes and systems via data-driven techniques.

3. Applications of Data Science in the Construction Industry

In SA, data science tools are employed in various construction project phases (Aghimien *et al.*, 2018a). Data science uses data analytics software to systematically evaluate patterns and to extract meaningful insight from data. The outcome of the results gathered from the data analytics is used by organisations to improve their operations, improve project delivery, increase profits, and gain a competitive advantage over their market rivals (Renukappa *et al.,* 2021). Munawar *et al.* (2022) note that for a construction company to be successful in the age of data science, companies need to embrace data science to combat some of the challenges that have been hindering the growth of the CI. Tjebane *et al.* (2022) posit the view that the SACI is one of the least digitized industries compared to other industries. Bilal *et al.* (2019) attributed this low adoption to a number of factors. Amongst others, the degree of complexity, the size of the project and the number of stakeholders involved were identified as some of the major factors hindering the adoption of technology applications in the CI.

However, the advent of the 4IR has also positively affected the CI and in adopting data science in its operations (Aghimien *et al.,* 2018b). Data science is needed to solve these issues and make building easier at various phases. Construction companies use the value of data science to manage and enhance construction sites. Furthermore, Cote (2021) notes that data science can be utilized to comprehend behaviors and procedures, develop algorithms for processing massive volumes of data quickly and efficiently, enhance the safety and confidentiality of sensitive data, and offer recommendations for data-driven decision-making. This paper will dissect the areas where data science is applied in the CI.

Data science can play an essential role in construction health and safety by detecting risky patterns or behavior, and safety hazards with real-time data analytics before an accident can occur (Liu, 2017). Moreover, Liu (2017) goes on to say that data science in the CI can also be used to reduce accidents on-site. Through predictive analytics, construction professionals can forecast common accidents before they occur. As a result, worker injuries and related costs are reduced. To make precise predictions of incidents, predictive analytics depends on data from previous safety occurrences and a deep understanding of the business and its operations (Zhang, 2020). These estimations cover both workers and equipment as well as the environment. A study conducted by Oracle (2021) revealed that most AEC companies are looking into predictive analytics to enhance project delivery and reduce risks. According to Sivarajah *et al.* (2017), the ability to foresee future events using both current and previous data is at the core of predictive analytics. Predictive analytics use methods like statistical modelling and machine learning to generate predictions about the future, this has been game-changing for the CI mainly because of its ability to track raw data and transform it into meaningful information which can be used for prediction (Oracle, 2021). According to Sivarajah *et al.* (2017), the construction team can use predictive analytics to swiftly identify potential issues and rank them according to likelihood before selecting when to handle them. According to Xu *et al.* (2020), predictive analysis allows the contractor to make smarter decisions when bidding and budgeting a project. Contractors can avoid losses due to on-site conditions or design errors by analysing past project data prior to establishing an estimate. By using multiple scenarios insights can be applied to make estimations and avoid failures in the future. By acting proactively, the construction stakeholders can avoid additional costs, delays and accidents caused by dealing with changes to the design and schedule (Sivarajah *et al.,* 2017). Furthermore, Zhang (2020) notes that predictive analytics provide organisations a chance to investigate the future which allows construction professionals to mitigate risk unlike any other tools available on the market.

The process of cost optimization is a procedure that must be done throughout the lifespan of a construction project. According to Lu (2019), project cost management involves estimating, budgeting, planning, managing, financing, controlling, and funding so that the project's success parameters can be met, which are the project being finished on budget and on time. The use of AI systems in conjunction with data science in construction projects can allow the construction team to plan their activities and estimate appropriate costs on the project (Phaladi *et al*., 2022). Intelligent AI-powered tools and trackers assist in centralizing financial data, receipts, criteria, restrictions, and documentation while also analyzing expenditures in real time (Motau and Kalema, 2016). Data science tools will then automatically break down the construction process and schedule the construction activities (Phaladi *et al*., 2022). In addition, data science is becoming an essential part of workforce management (Ismail *et al*., 2018). The increase of mobile technology has made it easy to track worker and equipment movements on site (Bello *et al.,* 2021) and contractors can estimate when maintenance or updates are required utilizing data analytics tools now that mobile technology is starting to have predictive capabilities. Real-time adjustments can be made and applied across the job site thanks to this mobility, which boosts productivity and optimizes time management for the best possible product output (Gotthardt *et al.,* 2019). According to Motau and Kalema (2016), by forecasting future requirements based on past project experiences, data science can increase efficiency and make important resources and information easily accessible when needed. Mobile technology can help supervise technicians on construction sites, update and analyze crucial paperwork in real time, and make real-time modifications to schedules or project scope (Bello *et al.,* 2021).

According to Mathenjwa (2020), there can be information and data gaps when working with subcontractors, such as hindered visibility into their activities on the clock. Data Science software ensures that information collected on subcontractor activities is analysed to monitor subcontractors and ensure that they are compliant and perform their duties according to the contract (Ismail *et al.,* 2018). Through data analytics, design and schedule details are automatically recorded to avoid disputes, according to Bello *et al.* (2021), this is done to ensure accountability among various subcontractors. Secondly, according to Ismail *et al.* (2018), data science encompasses predictive analytics, which is a software development responsible for BIM simulation. BIM simulation helps the project team to view an entire project prior to the actual commencement of the works, taking collaboration between all the stakeholders involved to greater heights. Design clashes can be easily identified through prediction, helping the contractor avoid possible problems in the erection, operation, and maintenance. Furthermore, Sibiya *et al. (*2015) note that infrastructure projects need close monitoring of all activities, and the contractor must complete the project on schedule, within the allocated budget, and to the highest standards of quality and safety. According to the Luthuli (2019), data science technologies can also be used to retain vital information about the building site and to monitor actions taking place there, such as workers who are not productive during working hours or who are not wearing protective gear.

Construction asset management and tracking are critical for successful construction project performance (Itemit, 2021). For construction fleets, reporting on essential preventative maintenance reduces the possibility of breakdowns or equipment problems (Aghimien *et al.,* 2018a). Installing a Global Positioning System (GPS) in construction equipment which is used to track the exact location of the organisation's equipment is essential in monitoring and controlling equipment on site (Itemit, 2021). The tracking device uses GPS to communicate with a satellite to triangulate its current position. These technologies aid in managing and allocating costs as well as tracking the equipment inventory in real-time. Equipment and personnel tracking may be beneficial in preventing theft and equipment loss (Phaladi *et al.*, 2022). On the other hand, general contractors can identify areas for improvement and produce more accurate estimations by monitoring the performance of subcontractors during the project's execution phase (Ismail *et al.*, 2018). According to Barker (2020), the client benefits from optimized contractor performance in gaining better value through quicker turnaround times; higher-quality interactions; safer work environments; effective utilization of resources; less wastage of time and money; and significantly efficient use of project funding in the project. However, Rhumbix (2021) notes that you can jointly measure contractor performance and maximize efforts by utilizing data analytics. Therefore, data analytics can assist in determining the degree of performance of the contractor and highlight areas that could want improvement (Oncioiu *et al.,* 2019).

Risk analysis and risk management are top of the list when it comes to successful project delivery and ongoing operations. Planning, identification, monitoring, classification, response analysis, and numerous other factors go into managing processes and risks (Rhumbix, 2021). Analysis of risks determines the potential pitfalls in future resources, results, and impacts. Meng *et al.* (2022) claim that by creating a probability density function and construction cost through a comparison of data for a proposed project and a previous project that is similar to the one being planned, the risk analysis framework can calculate the likelihood-cost interval for a planned project's final estimate. Furthermore, Rhumbix (2021) notes that by examining factors affecting project risks, the data analytics model develops probabilistic estimates of project costs. In this model, data from the relevant region are gathered and analyzed in comparison to the data for a prior similar project before the planned project is implemented. Lastly, operations and maintenance are the final stage of the construction process. After construction is complete, construction data is gathered and saved for later use throughout the building's life cycle. According to Tejjy (2022), the building information can be used by the maintenance crew to carry out repairs and maintenance on the building. Munawar *et al.* (2022) suggest that construction-related data can be used to guide future renovation projects and give maintenance teams structural information. Furthermore, Wilcock (2022) asserts that data analytics are important post construction operations and maintenance. Ismail *et al.* (2018) argue that you won't be able to choose and track KPIs, set realistic targets, and model a different and more lucrative future without the data necessary to analyze past performance, identify issues and forecast probable results. The areas for data science applications for construction project performance in the SACI are summarised in table 1. The data science tools that are most commonly used in the industry are Geographic Information Systems (GIS), Building Information Modelling (BIM), and the Internet of Things (Rhumbix, 2021).

Table 1: Applications of Data Science in the Construction Industry

S/N	Application Areas of Data Science in the Construction Industry	Authors
1	Construction Health and Safety	Zhang (2020) and Liu (2017)
2	Predictive Analytics	Oracle (2021); Xu *et al.* (2020); Zhang (2020) and Sivarajah *et al.* (2017)
3	Project Cost Management	Phaladi *et al*. (2022); Lu (2019) and Motau and Kalema (2016).
4	Workforce Management	Gotthardt *et al.* (2019); Bello *et al.* (2021) and Ismail *et al*. (2018).
5	Subcontractor Performance Analytics	Mathenjwa (2020); Bello *et al.* (2021) and Ismail *et al.* (2018)
6	Construction Planning and Modelling	Luthuli (2019); Ismail *et al.* (2018) and Sibiya *et al. (*2015)
7	Tracking Construction Equipment and Assets	Phaladi *et al.* (2022); Aghimien *et al.* (2018a); and Itemit (2021)

8	Optimizing Contractor Performance	Barker (2020); Rhumbix (2021); (Oncioiu *et al.*, 2019) and Ismail *et al.* (2018)
9	Risk Assessment Optimization	Meng *et al.* (2022) and Rhumbix (2021)
10	Operations and Maintenance	Munawar *et al.* (2022); Tejjy (2022); Wilcock (2022) and Ismail *et al.* (2018)

5. Research Methodology

For this study, the quantitative approach was considered appropriate. As Queirós *et al.* (2017:371) define a quantitative research approach as a well-structured theoretical framework that collects data which is a comprehensive view of the population where the analysis of the numerical data will be done through a statistical procedure. Through statistical processes, data can be validated and tested. It is for this reason that a quantitative approach was utilized to gather empirical data for this study. This study was done in Johannesburg, South Africa, Gauteng Province. This is due to its central location and since it has a large concentration of construction companies. The target population in this study consisted of construction professionals registered with various professional bodies and actively practicing in the construction and property industries in South Africa.

A well-structured questionnaire was distributed to the respondents in the target population, this is important for the study as it will reflect the true reflection of data science adoption in the SACI. The respondents consisted of architects, data scientists, facility managers, construction managers, building managers, project managers, engineers, and quantity surveyors from the public and private sectors. A total of 55 responses were received, and they were all usable. Statistical Package for Social Science (SPSS) and a Microsoft spreadsheet were used to analyze and interpret the results.

The SPSS program allowed for the recording of original data to determine the mean item scores, standard deviations, ranks, and Cronbach's alpha, both SPSS and Microsoft Excel were used to create the graphs and charts. The reliability of the study of the study was tested via internal consistency. Where Pallant (2020) defines internal consistency as the degree where variables of the same scale correlate with each other which is often assessed by Cronbach's Coefficient Alpha. The coefficient is effective in research studies where the Likert scale was used for responses. The range is from 0 to 1, and a Cronbach's Alpha score of 0.7 is considered to be an acceptable reliability coefficient as the higher the value, the greater the reliability (Pallant, 2020). The Cronbach's Alpha value for the application of data science for construction project performance in the SACI was 0.946, which proves the reliability and validity of the questionnaire.

6. Research Findings

6.1. Demographic Characteristics of the Respondents

The majority of the respondents to the questionnaire were females sitting at 60%, while 40% were male, with more than 50% of the participants older than 26 years old. Figure 1 shows the various professionals who participated within the sample; architects, data scientists, and building managers specified others contributing 34% of the respondents.

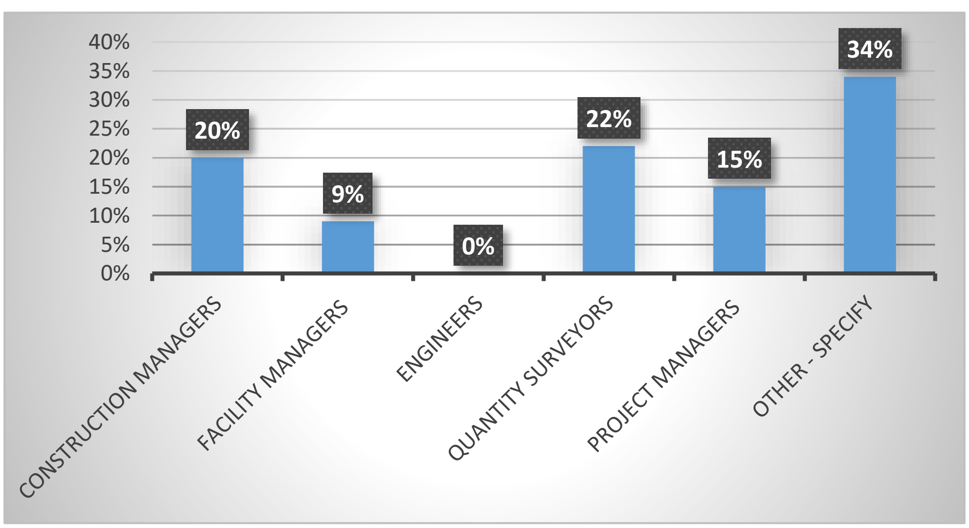

Figure 1 Respondents' occupational status

6.2. Descriptive Analysis on the Applications of Data Science in the Construction Industry

The participants were asked to indicate where data science is applied mostly in the SACI using a 5-point Likert scale of 'Strongly disagree' to 'Strongly agree' on the questionnaire. Table 1 shows the results relating to the applications of data science uptake in the SACI. Based on the rankings, the top five applications of data science in the construction industry are in tracking construction equipment and assets, project cost management, operations and maintenance, risk assessment optimization, and construction planning and modelling.

Table 2: Applications of data science in construction

Applications of data science in SA construction industry	**MIS**	**SD**	**R**
Tracking construction equipment and assets	3.98	1.209	1
Project Cost Management	3.95	1.177	2
Operations and Maintenance	3.93	1.215	3
Risk Assessment Optimization	3.91	1.159	4
Construction planning and modelling	3.89	1.242	5
Construction health and safety	3.87	1.171	6
Optimizing Contractor Performance	3.87	1.233	6
Subcontractor performance analytics	3.80	1.078	8
Workforce Management	3.78	1.182	9
Predictive Analytics	3.75	1.040	10

MIS = Mean Item Score, SD = Standard Deviation, and R = Rank

7. Discussion

Based on the respondents' average rating of the variables, the findings revealed a low level of adoption of data science in construction project performance in South Africa. The results are in agreement with studies conducted by Renukappa *et al.,* (2021); Aghimien *et al.,* (2018b), and Oracle (2021), which indicated the low-level adoption of data science. Despite the significant benefits associated with the adoption of data science, the CI is still lagging behind compared to other industries such as the automobile, finance, and healthcare industries (Ram *et al.,* 2019). The top three data science applications used for construction project performance in the SACI were in the tracking of construction equipment and assets which was ranked first, project cost management which was ranked second and operations and maintenance ranked third. These results do not agree with studies conducted by Bilal *et al.* (2019) and Ram *et al.* (2019), which ranked predictive analytics, Augmented Big Data and BIM integration as the most used big data applications in the CI.

Nevertheless, studies conducted by Bello *et al.* (2021) suggest that the most usage of data science is in the GPS tracking system aiding in monitoring and controlling equipment on construction sites. This is in agreement with the participant's response ranking tracking of construction equipment and assets first. Project cost management, the second-ranked application of data science, supports the findings of Phaladi et al. (2022) who point out that the utilization of AI technologies in collaboration with data science can bring out realistic cost estimates and well-structured construction programme. The respondents ranking third operations and maintenance is in agreement with Wilcock (2022), who asserts that data analytics are essential to post-construction operations and maintenance.

8. Conclusion

The findings in relation to the literature revealed that the research participants believe that tracking of construction equipment and assets, project cost management and operations and maintenance to be the top three applications of data science in the construction industry than subcontractor performance analytics, workforce management and predictive analytics which were the least ranked applications. There is a lack of awareness of data science applications in the industry, but this study focused only on the application areas in construction companies in Gauteng, SA. The government can increase this awareness by introducing stringent regulations promoting data science in public-sector construction projects.

The reviewed literature revealed that there are numerous applications of data science in the CI. It was noted from the literature the various data science that can be used in the CI. These included facilities management, risk assessment optimization, social media analytics, project cost management, data visualization, energy management analytics, operations and maintenance, performance prediction, generative design and tracking construction equipment and assets. The areas for further studies might look at the use of predictive analytics in improving project delivery in the construction industry.

9. References

Aghimien, D.O., Aigbavboa, C.O. and Oke, A.E. (2018a). Digitalisation for effective construction project delivery in South Africa. *Contemporary Construction Conference: Dynamic and Innovative Built Environment (CCC2018).*

Aghimien, D.O., Aigbavboa, C.O. and Thwala, W.D. (2019). Microscoping the challenges of sustainable construction in developing countries. *Journal of Engineering, Design and Technology*. 176: 1110-28.

Aghimiena, D.O., Aigbavboa, C.O. and Oke, A.E. (2018b). A Review of Data Mining for Sustainable Construction in Nigeria. *10th International Conference on Applied Energy (ICAE2018), China*

Bello, S.A., Oyedele, L.O., Akinade, O.O., Bilal, M., Delgado, J.M.D., Akanbi, L.A., Ajayi, A.O. and Owolabi, H.A., (2021). Cloud computing in construction industry: Use cases, benefits and challenges. *Automation in Construction*, *122*, p.103441.

Bilal, M., Oyedele, L.O., Kusimo, H.O., Owolabi, H.A., Akanbi, L.A., Ajayi, A.O., Akinade, O.O. and Delgado, J.M.D. (2019). Investigating profitability performance of construction projects using big data: A project analytics approach. *Journal of building engineering*. 26: 100850.

Cao, L. (2017). Data science: a comprehensive overview. *ACM Computing Surveys (CSUR)*. *50*(3): 1-42.

Cheng, T. and Teizer, J. (2010). Real-time data collection and visualization technology in construction. In *Construction Research Congress 2010: Innovation for Reshaping Construction Practice*. 339-348.

Cote, C. (2021). *What is Data Science? 5 Applications in business*. Available from: https://online.hbs.edu/blog/post/what-is-data-science [Accessed 25 August 2022].

Gotthardt, M., Koivulaakso, D., Paksoy, O., Saramo, C., Martikainen, M. and Lehner, O.M. (2019). Current state and challenges in the implementation of robotic process automation and artificial intelligence in accounting and auditing. *ACRN Oxford Journal of Finance & Risk Perspectives*. 8: 31-46. Accessed from: http://www.acrn-journals.eu/resources/SI08_2019c.pdf [Accessed 08 July 2022].

Hwang, K. and Chen, M., 2017. *Big-data analytics for cloud, IoT and cognitive computing*. New York: John Wiley & Sons.

Ismail, S.A., Bandi, S. and Maaz, Z.N. (2018). An appraisal into the potential application of big data in the construction industry. *International Journal of Built Environment and Sustainability*, *5*(2).

Itemit. (2021). Tracking Construction Equipment. Available from: https://itemit.com/tracking-construction-equipment-why-and-how/ [Accessed 25 September 2022].

Liu, H (2017). *The Impact of Owner Practices and Procedures on Construction Project Safety Performance*. Theses and Dissertations--Civil Engineering. 51. Available from: https://uknowledge.uky.edu/ce_etds/51 [Accessed 04 August 2022].

Lu, W. (2019). Big data analytics to identify illegal construction waste dumping: A Hong Kong study. *Resources, conservation and recycling*. *141*: 264-272.

Luthuli, C.A. (2019). White paper on Science, Technology and Innovation. Available from: https://www.dst.gov.za/images/2019/White_paper_web_copyv1.pdf [Accessed 25 August 2022].

Martinez, G., Renukappa, S. and Suresh, S. (2021). Business model innovation in small enterprises from developing countries during COVID-19 outbreak: exploring drivers and BMI outcomes. *International Journal of Business Environment*. *12*(4): 364-388.

Mathenjwa, M.M. (2020). *Challenges faced by main contractors when working with inexperienced subcontractors on South African construction projects* (Doctoral dissertation).

Meng, X., Zhu, J., Chen, G., Shi, J., Li, T. and Song, G. (2022). Dynamic and quantitative risk assessment under uncertainty during deepwater managed pressure drilling. *Journal of Cleaner Production*. *334*: 130249.

Motau, M. and Kalema, B.M. (2016), August. Big data analytics readiness: a South African public sector perspective. In *2016 IEEE International Conference on Emerging Technologies and Innovative Business Practices for the Transformation of Societies (EmergiTech)* (pp. 265-271). IEEE.

Munawar, H.S., Ullah, F., Qayyum, S. and Shahzad, D. (2022). Big data in construction: current applications and future opportunities. *Big Data and Cognitive Computing*. *6*(1): 18.
Oncioiu, I., Bunget, O.C., Türkeș, M.C., Căpușneanu, S., Topor, D.I., Tamaș, A.S., Rakoș, I.S. and Hint, M.Ș. (2019). The impact of big data analytics on company performance in supply chain management. *Sustainability*. *11*(18): 4864.
Oracle. (2013). *Big Data Analytics. An Oracle White Paper*. Available from http://www.oracle.com/us/products/database/big-data-for-enterprise in April 2022
Pallant, J. (2020). *SPSS survival manual: A step by step guide to data analysis using IBM SPSS*. London: Routledge.
Phaladi, M.G., Mashwama, X.N., Thwala, W.D. and Aigbavboa, C.O. (2022). A Theoretical Assessment on the Implementation of Artificial Intelligence (AI) for an Improved Learning Curve on Construction in South Africa. In *IOP Conference Series: Materials Science and Engineering* (Vol. 1218, No. 1, p. 012003). IOP Publishing.
Queirós, A., Faria, D. and Almeida, F. (2017). Strengths and limitations of qualitative and quantitative research methods. *European journal of education studies*. 3(9): 369-387.
Ram, J., Afridi, N.K. and Khan, K.A. (2019). Adoption of Big Data analytics in construction: development of a conceptual model. *Built Environment Project and Asset Management*. *9*(4): 564-579.
Renukappa, S., Suresh, S. and Reyes-Veras, P. (2021). Awareness of big data concept in the Dominican Republic construction industry: an empirical study. *Construction Innovation: Information, Process, Management*. 1-19. DOI: 10.1108/CI-05-2021-0090
Rhumbix. (2021). 10 Examples of Data Analytics Sweeping Construction Project Management. Available from: https://www.rhumbix.com/blog/10-examples-of-data-analytics-sweeping-construction-project-management [Accessed 25 September 2022].
Sibiya, M., Aigbavboa, C. and Thwala, W., 2015. Construction projects' key performance indicators: a case of the South African construction industry. In *ICCREM 2015* (pp. 954-960).
Sivarajah, U., Kamal, M.M., Irani, Z. and Weerakkody, V. (2017). Critical analysis of Big Data challenges and analytical methods. *Journal of Business Research*. 70: 263-286.
Tang, L., Zhang, Y., Dai, F., Yoon, Y., Song, Y. and Sharma, R.S. (2017). Social media data analytics for the US construction industry: Preliminary study on Twitter. *Journal of Management in Engineering*. *33*(6): 04017038.
Tejjy. (2022). What is BIM (Building Information Modelling). Available from: https://www.tejjy.com/what-is-bim-building-information-modeling/ [27 September 2022].
Tjebane, M.M., Musonda, I. and Okoro, C. (2022). Organisational Factors of Artificial Intelligence Adoption in the South African Construction Industry. Front. *Built Environ*. 8: 823998.
Warden, P. (2011). *Why the term "data science" is flawed but useful*. Available from: http://radar.oreilly.com/2011/05/data-science-terminology.html [Accessed 25 August 2022].
Xiao, H. and Proverbs, D. (2003). Factors influencing contractor performance: an international investigation. *Engineering, construction and architectural management*. *10*(5): 322-332.
Xu, J., Lu, W., Ye, M., Xue, F., Zhang, X. and Lee, B.F.P. (2020). Is the private sector more efficient? Big data analytics of construction waste management sectoral efficiency. *Resources, conservation and recycling*. 155: 104674.
Zhang, Z. (2020). Predictive Analytics in the era of Big Data: Opportunities and challenges. *Annals of Traditional Medicine*. 8(4): 68.