Sustainable Construction in the Era of the Fourth Industrial Revolution, Vol. 187, 2025, 76–87
https://doi.org/10.54941/ahfe1006562

Transforming South African Cities: The Impact of Big Data Analytics on Smart City Initiatives

Thabang Tlou[1], Douglas Aghimien[2], Bankole Awuzie[1], and Clinton Aigbavboa[3]

[1]Department of Construction Management and Quantity Surveying, University of Johannesburg, Johannesburg, South Africa
[2]Civil Engineering Technology, Faculty of Engineering and the Built Environment, University of Johannesburg, Johannesburg, South Africa
[3]CIDB Centre of Excellence, Faculty of Engineering and Built Environment, University of Johannesburg, Johannesburg, South Africa

ABSTRACT

Technological advancement is rapidly changing the way cities operate. These advancements have led to the development of smart cities, which are data-driven cities designed to improve people's way of life. Analyzing the numerous data generated within cities can help drive the attainment of city smartness. Based on this knowledge, this study assessed the impact of big data analytics (BDA) on the attainment of smart cities in South Africa. This was done to propose policies guided by technological advances, such as BDA, in the quest for smart cities in the country. The study adopted a post-positivist stance with a quantitative design. Data gathered from participants with knowledge of innovative city development in the country were analyzed using relevant descriptive and inferential statistics. The study found that the major sources of big data that can be analyzed for city smartness are sensor data, banking transactions and social networking sites, which can be gathered in a hybrid mode and controlled by the government and private entities. It was further noted that carefully analyzing these data using BDA would significantly impact smart transportation, grid, healthcare, governance, economy, and environment. The study contributes to the existing smart city discourse from the view of BDA adoption in a developing country. Its findings can serve as an excellent theoretical backdrop for future works in countries where such a study has not been conducted.

Keywords: Big data, Data analytics, Information management, Smart city, Data-driven

INTRODUCTION

Cities play an important role in human existence as well as the socio-economic development of countries worldwide (Aghimien et al., 2019). Studies have noted that cities are organized communities developed through advanced innovations required for the sustainability of the nation's economy (Amer, 2014; Bawa et al., 2016). However, rapid urbanization has rendered most cities unsustainable, leading to the call for more sustainable and smart cities (Oke et al., 2020). The state of cities in most developing countries, particularly in Africa, has become a topical issue over time, with high population emanating from rapid urbanization, scarcity of basic amenities, and poor living conditions becoming evident in most cities (Aghimien et al., 2019). Although South Africa is one of the top-performing countries in Africa, cities in the country are not exempted from these problems. Cities such as Johannesburg, eThekwini and Cape Town, which are major economic hubs, are also experiencing exponential growth, which is placing planning for basic services in jeopardy (National Planning Commission, 2013). These issues have brought about the quest for smart and resilient initiatives.

Smart cities are cities that incorporate information technology, software, networks, and data in their management and in the provision of services (Al Nuaimi et al., 2015). A smart city is one that allows economic and political efficiency and the growth of social and cultural features of a community using networked infrastructures (Clarke et al., 2019). Integrating smart systems in cities will lead to digital and smarter cities (Rathore et al., 2016). Hence, smart cities have been described as connected cities (Silva et al., 2018) and data-driven cities (Govender, 2018). Evidently, for a data-driven city, a large amount of data, known as 'big data', will be generated from various data sources such as computers, cell phones, social networks, sensors, etc. These big data are characterized as data of a high volume, high velocity, and wide variety, and they require technologies and techniques to capture, store, and analyze the data (Watson, 2014). Analyzing big data using big data analytics (BDA) to make informed decisions and provide tailor-made services to the citizens becomes important.

BDA has been described as a technological advancement that helps in gathering and processing large amounts of data in real-time and allows the use of the information gleaned from the analyses of these data to make informed and beneficial decisions (Cabrera-Sánchez and Villarejo-Ramos, 2019). With the arrays of possible data sources available for attaining a 'data-driven' city, it is evident that the use of BDA is essential for the successful use of the data generated (Yaqoob et al., 2016). Based on this knowledge, studies have continued to emanate on the role of BDA on city development in countries worldwide (Al Nuaimi et al., 2015;). However, there is an absence of studies exploring the impact of BDA on city smartness in a developing country like South Africa, where technological advancement is still low. Thus, this study aimed to unearth the sources of big data required for proper analyses and to determine the potential impact of BDA in achieving smart cities in South Africa. This was done to arm policymakers with the information needed to formulate policies that are guided by technological advances in the quest for smart cities in the country.

OVERVIEW OF SMART CITY AND BIG DATA ANALYTICS

The concept of a smart city focuses on a city as an integrated system composed of subsystems that behave intelligently and coherently, with the main aim of ensuring city services function as a single unit. To formulate common smart city features, the integration of multiple systems, such as transportation, energy, education, health, and physical infrastructure, is essential (Ghosal and Halder, 2018). These smart city features include the use of digital and electronic technology to transform lives and works and the use of information and communication technology (ICT) to develop city infrastructure and integrate these ICTs with people to enhance innovation and knowledge. For cities to be classified as 'smart', they should exhibit features such as smart living, economy, people, governance, mobility, and environment (Giffinger et al., 2007).

The term 'big data' has been used in describing data that is characterized by high volume and high velocity and which has a wide variety, requiring new techniques and technologies for the capturing, storing and analysis of the data to enhance decision making, ensure the provision of insight and discovery for the support and optimization of the processes (Mills et al., 2012). Analyzing these data to draw logical conclusions is necessary. This data analysis is a process of ensuring data is inspected, cleaned, transformed, and modelled to gain insights that provide suggestions and assist in the decision-making process (Vassakis et al., 2018).

In an ideal smart city, the interaction of people with computers, cell phones, online platforms such as e-commerce sites and social platforms generates potentially useful data. Moreover, people's movement is constantly being tracked by sensors to generate travel data that can be used to improve transport services. These digital data are large, fast and varied (Che et al., 2013). Blazquez and Domenech (2018) posited that sources of big data used in effective smart cities are transactions, public data, sensors, social media, and enterprise data. According to Askitas and Zimmermann (2015), the use of the internet leaves a digital footprint which can be tracked, and with the help of BDA, a determination on the behaviour, decision-making, and intentions of people that use the internet can be made. The internet, as a platform through BDA, provides crucial data that assists in forecasting, modelling, and explaining social behaviours.

Internet data sources include information searches, social networking sites, blogs, websites, and applications. Blazquez and Domenech (2018) stated that online search engines can capture the variation of information over some time, thus providing data on trends and societal interests and concerns. Platforms such as Google Trends can explore search categories in the Google search engine to predict critical city issues such as tourist inflows and house sales and forecast political inquiries' results, among others (Mavragani et al., 2016). Moreover, a significant number of social networking sites encourage users to express their feelings and opinions on any topic of interest. Information contained in social networking sites largely reflects the situation in society, and this has given rise to 'social big data' (Bello-Orgaz et al., 2016). For instance, a microblogging service with millions of active users who send millions of tweets per day generates large amounts of 'user-generate' data, which can assist in the prediction of current and future social and economic events or help in monitoring public perception of new policies (Arrigo et al., 2016).

In addition to social networks generating a series of big data, companies in the digital era are establishing websites to inform people about their services and products. Their corporate websites provide different functionalities, from spreading information about companies to conducting online transactions and facilitating opinions through the website. Big data approaches applied to corporate websites have allowed for sales growth and improvement of business activity, allowing innovation and development (Curme et al., 2014). Moreover, the extensive usage of Applications (apps) daily has made them a major source of data about individuals and organizations to help forecast social and economic aspects. Data logs from app usage have been proven to forecast users' intentions and automatically detect aspects such as mobility patterns (Yang et al., 2024).

According to Aghimien et al. (2021), ubiquitous computing has developed exponentially in the digital era, resulting in the generation of inconspicuous sensors that are inexpensive to gather data on people's daily activities. The urban and embedded sensors integrated into smart city infrastructure are highly potential social and economic data generators. Xiong et al. (2013) noted that the credit card reader is one of the most widely used sensors. Transactions recorded provide data for detecting and predicting activities in online stores. Besides, retail scanner data can be used to forecast consumer behaviour, sales, and prices to model market trends and suggest a competitive strategy in alignment with the data obtained. Sensors embedded in mobile phones are a source of social data that generates data about user locations to study social parameters like behaviours and preferences. The data obtained from sensors are useful in detecting places of interest and personality traits that companies can use to personalize services. Mobile phone data have been used to recreate and draw mobility maps showing population distribution and can detect behavioural patterns associated with emergencies such as earthquakes (Deville et al., 2014).

The value of the data extracted and processed is critical in the integration of people, processes, and technology in a smart city (Ullah et al., 2024). According to Al Nuaimi et al. (2015), many cities compete to become smart cities with the hope of reaping economic benefits, and the vast majority are looking at using BDA in their smart city applications. This can be in the form of predictive analytics, security analytics, edge analytics or semantics analytics. These diverse analytics can be conducted on data gathered in the different aspects of smart cities, vis: smart transportation, smart grids, smart weather forecasting, smart agriculture, smart healthcare, smart education and smart governance (Oke et al., 2020; Ullah et al., 2024).

RESEARCH METHODOLOGY

The study adopted a post-positivist stance with a quantitative design, which entails gathering data through a structured questionnaire. A purposive and snowball sampling approach was adopted to gather data from experts involved in construction and urban development in Ekurhuleni, the City of Johannesburg, and the City of Tshwane in Gauteng province. Purposive sampling was first used to identify some set of experts with at least 5 years of active involvement in construction and urban development. However, because it was difficult to determine the total number of experts within the defined target criteria at the time of conducting the research, a snowball approach was then employed to reach out to more experts through a referral from those initially identified. Based on the approach adopted, 43 usable samples were derived and considered fit for data analysis in the study.

The questionnaire was designed in sections, where the first section sought answers to specific demographic questions. The second section assessed some of the sources of big data, the platforms and control measures for handling generated data on a five-point agreement scale, with one being strongly disagreed and five being strongly agreed. The third section of the questionnaire explored the potential impact of BDA on diverse aspects of smart cities using a five-point impact scale, with one being very low and five being very high. Data analysis was done using percentage (%) for the data on the demographic questions. The Cronbach alpha (α) test was used to determine the reliability of the questions asked in sections 2 and 3 with a threshold of > 0.60 for an acceptable level. Mean item score ($\overline{X}$) was used to rank the respondent's rating of the variables in both sections. Since the respondents were drawn from diverse professions, the Kruskal-Wallis H-test (*K-W*) was used to determine the significant difference in their rating of the different variables. This test gives a chi-square (χ^2) and p-value. For a significant difference to exist in the respondents' rating, a p-value of below 0.05 must be achieved (Pallant, 2020).

FINDINGS AND DISCUSSION

Background Information

The analysis of the background information shows that the participants for the study cut across diverse related fields such as IT (16%), engineering (35%), quantity surveying (5%), construction and project management (37%), as well as urban planning (7%). These respondents were drawn from both public (53%) and private (47%) organizations in Gauteng province. The majority of these respondents (88%) possess a bachelor's degree, with the remaining 12% having a master's degree. In terms of years of experience, 68% have between 6 to 10 years of experience in construction and urban development, while 28% have between 11 to 15 years, and 4% have above 20 years of experience. The average years of experience of these respondents is calculated as 9 years. This background information revealed that the respondents for the study are knowledgeable in terms of their academic qualifications and experience to answer the questions given in this research.

Sources of Big Data for Smart City Attainment

Table 1 shows the result of the respondents' ratings of the different variables in relation to the data sources, platforms, and control systems. The α test revealed that internal reliability was achieved as α-value of 0.834, 0.817, and 0.835 were derived. Four major sources of big data were assessed: sensor data, transactions, social networking sites, and enterprise data, as indicated in past literature reviews. The result in Table 1 revealed that using urban and mobile sensors through GPS and location data is a major potential source of big data that can be used for the delivery of smart cities in South Africa.

However, transactional data from bank cards and e-payments, as well as social media text and communication, cannot be overlooked as they appeared to be key sources of big data with a shared $\overline{X}$-value of 4.72. The last source (enterprise data) has a $\overline{X}$-the value of well above the average of 3.0. This implies that website logs and online citizens' activities can also help generate significant data that can be analyzed for smart city actualization. The *K-W* test revealed that despite coming from diverse professional backgrounds, the respondents for this study had a unified view of the significance of these assessed sources of big data for smart cities as a *p*-value of above 0.05 was derived for all four sources.

The platforms from which these data can be harnessed were also analyzed. The result revealed that a hybrid system allowing offline and online data generation can be most beneficial. This platform revealed a $\overline{X}$-value of 4.67. Also, having either an offline or online data generation platform can prove to be useful as they have a $\overline{X}$-the value of 4.56 and 4.53 respectively. Also, the *K-W* test shows no discrepancies in the rating of these platforms by the different professionals, as a *p*-value of above 0.05 was derived for all three assessed. In terms of control of these data, the result revealed that joint control by public and private entities in managing the big data life cycle would be the most efficient approach, as this was rated first with a $\overline{X}$-value of 4.63. No significant difference was also observed among the respondents regarding the rating of this control system based on the *K-W* test.

Table 1: Big data sources, platform and control for smart city attainment.

			K-W Test	
Big Data Sources	$\overline{X}$	**Rank**	χ^2	***p*-Value**
Data sources ($\alpha = 0.834$)				
Sensor data - urban and mobile sensors	4.74	1	6.587	0.680
Transactions data	4.72	2	11.498	0.243
Social networking data	4.72	2	9.519	0.391
Enterprise data	4.63	4	13.194	0.154
Data platforms ($\alpha = 0.817$)				
Hybrid Systems	4.67	1	12.217	0.201
Offline Platforms	4.56	2	8.765	0.459
Online Platforms	4.53	3	14.080	0.120
Data control (α =0.835)				
Public-Private Partnership	4.63	1	16.180	0.063
Government/Public	4.26	2	9.077	0.430
Private Sector	4.21	3	11.437	0.247

Impact of BDA on Smart City Attainment

Careful analysis of the big data generated from different sources can significantly impact diverse aspects of a smart city. Table 2 gives the result of the analysis of the respondents' feedback on the impact of BDA on diverse smart city components. A cursory look shows that internal reliability was achieved as α-value ranging from 0.652 to 0.928 were achieved. Thus, the questionnaire used was reliable in answering the questions for which it was designed. Furthermore, the K-W test revealed that across all six smart city components explored, there is no divergence in the rating of all variables by the different groups of respondents. All the assessed variables showed a p-value of above 0.05. this implies that despite their different professional backgrounds, the respondents had a unified view regarding the impact BDA will have on these identified smart city components.

On the overall, the respondents indicated that BDA will significantly impact the identified smart city components immensely as a $\overline{X}$-the value of well above average of 3.0 was derived for all components. In terms of transportation, effective analysis of traffic movement data will impact the development of Smart traffic lights in providing traffic patterns and aid in traffic flow ($\overline{X} = 4.91$) and assist in traffic flow and elimination of congestion ($\overline{X} = 4.91$). For smart grid systems, BDA can help Improve the development of power supply ($\overline{X} = 4.65$) and monitoring of user consumption ($\overline{X} = 4.65$) while being able to forecast the production, transmission, and distribution requirements ($\overline{X} = 4.63$). In the area of smart healthcare, BDA promises easy treatment of patients through audio analytics ($\overline{X} = 4.70$) and images from sensors ($\overline{X} = 4.67$). Also, BDA offers Improved public ICT and increased e-government ($\overline{X} = 4.79$) and increased transparency and open data for the public ($\overline{X} = 4.77$) in the area of smart governance. It can also promote green urban planning ($\overline{X} = 4.84$) and smart waste management systems ($\overline{X} = 4.84$) in a smart environment while increasing productivity ($\overline{X} = 4.84$) and economic image and trademarks ($\overline{X} = 4.81$) in the smart economy dimension.

Table 2: Impact of BDA on smart city attainment.

			K-W Test	
Impact of BDA	$\overline{\mathbf{X}}$	**Rank**	χ^2	***p*-Value**
Smart Transportation ($\alpha = 0.853$)				
Smart traffic lights aid in traffic flow	4.91	1	12.506	0.186
Assist in traffic flow and eliminate of congestion	4.91	1	5.891	0.751
Provide real-time speeds, direction, and location	4.88	3	7.600	0.575
Smart Grid ($\alpha = 0.926$)				
Improved and reliable power supply	4.65	1	16.808	0.052
Improved monitoring of user's consumptions	4.65	2	7.580	0.577
Forecasting the production, transmission, and distribution requirements	4.63	3	10.890	0.283
Maintenance management that predicts faults	4.56	4	11.883	0.220

Continued

Table 2: Continued

			K-W test	
Smartt Healthcare ($\alpha = 0.652$)				
Treat patients through audio analytics	4.70	1	12.037	0.211
Enhance services using sensors and camera images	4.67	2	8.161	0.518
Improvement in the diagnosis and prescription through machine learning	4.65	3	9.133	0.425
Smart Governance ($\alpha = 0.800$)				
Improved public ICT and e-governance	4.79	1	12.031	0.212
Increased transparency and open data	4.77	2	12.498	0.187
Increase public participation	4.70	3	9.957	0.354
Creation of integrated digital public services	4.60	4	17.569	0.041
Data-driven policies to align with people's needs	4.56	5	11.764	0.227
Smart Environment ($\alpha = 0.777$)				
Promote green urban planning	4.84	1	20.056	0.018
Smart waste management systems	4.84	2	9.338	0.407
Smart water, air, and land management	4.79	3	11.460	0.245
Accelerate the adoption of green buildings	4.72	4	5.355	0.802
Generation and use of green energy sources	4.58	5	10.144	0.339
Smart Economy ($\alpha = 0.928$)				
Increased productivity	4.84	1	11.388	0.250
Increased economic image and trademarks	4.81	2	14.449	0.107
Create high entrepreneurship and innovation	4.79	3	11.872	0.221
High labour market flexibility	4.79	3	5.801	0.760
Local interconnectedness of businesses	4.74	5	8.480	0.487

Discussion and Implication of Findings

The findings revealed that the major sources of big data for attaining smarties in South Africa are sensor data gathered from urban and mobile sensors through the usage of GPS and location data as well as through social networks and transactional data. This finding supports Blazquez and Domenech's (2018) earlier submissions that sensors embedded in mobile phones are a potential source of big data that show the location of users and can be used in studying the social behaviours, preferences and mobility of users. Also, the ranking of social networking sites as one of the major sources of big data aligns with the works of Askitas et al. (2015), which noted that the internet is being used for the explaining, modelling, nowcasting and forecasting of social behaviours. The study's findings in terms of data platforms indicate that urban development experts believe that a hybrid system that allows for offline and online generation of big data is a critical component of the major sources of big data. From the perspective of big data control, a public-private partnership is required to manage the big data life cycle.

The results on the impact of the usage of BDA in the attainment of smart cities in South Africa indicate that the experts in urban development expect the impact in transportation will be realized through the assistance with traffic flow and elimination of congestion and providing real-time speeds, direction, and location of transportation of modes in smart cities. In the perspective of smart grids, the impact is perceived to provide an improved and reliable power supply that is self-healing and can predict the energy consumption of citizens. This is expected due to the continued energy supply constraints that South Africa faces, which have resulted in power interruptions. Furthermore, urban development experts viewed the impact of BDA in smart healthcare as having the ability to treat patients through audio analytics and enhancing services offered by healthcare facilities. In smart governance, BDA offers to improve the ICT, increase e-governance, ensure transparency with regard to public data gathered and increase public participation. The respondents perceived the smart environment impact of BDA to promote green urban planning and create a smart waste management system which will alleviate current South African challenges such as illegal land occupation and uncontrolled waste management. Lastly, from the perspective of the smart economy, BDA will lead to high levels of productivity and increase the economic image of smart cities.

CONCLUSION

The need for smart and intelligent cities has become crucial with emerging technological advancement and the need to provide more liveable and enjoyable cities for people. Using a survey of participants with knowledge of smart city development, this study concludes that in the quest for smart cities in South Africa, big data can be gathered and analyzed from data gathered using sensors, banking transactions and social networking sites. These can be gathered in a hybrid mode and controlled by the government and private entities. Furthermore, analyzing these data using BDA would impact smart transportation, smart grid, smart healthcare, smart governance, smart economy, and smart environment. This implies that South Africa stands the chance to transform its cities through effective analyses of crucial big data, which can, in turn, significantly shape the smart operations of diverse sectors in urban areas. Furthermore, the findings of the study serve as an excellent theoretical platform for future works in countries where such studies on the impact of BDA on city smartness have not been conducted. While the findings of this study offer significant contributions to the existing smart city discourse from the South African perspective, the findings are limited by several factors. Firstly, the response was gathered from experts in a single province. Further studies can explore other provinces to get more responses. Also, the practical assessment of big data harnessing and analysis is needed to understand salient issues, including the ethical use of these data. As such, further studies are needed to explore the digital and data ethics of deploying BDA in attaining smart cities in the country.

REFERENCES

Aghimien, D. O., Ikuabe, M., Aigbavboa, C. O., Oke, A. E. and Shirinda, W. (2021) 'Unravelling the factors influencing construction organizations' intention to adopt big data analytics in a developing country', *Construction Economics and Building*, 21(3), pp. 262–281.

Aghimien, D. O., Aigbavboa, C. O., Thwala, W. D. and Ohiomah, I. (2019) 'Challenges hindering the attainment of smart cities', *Interdependence between Structural Engineering and Construction Management (ISEC-10)*, Chicago, May 20–25, pp. 1–6.

Al Nuaimi, E., Al Neyadi, H., Mohamed, N. and Al-Jaroodi, J. (2015) 'Applications of big data to smart cities', *Journal of Internet Services and Applications*, 6(25), pp. 1–15.

Al-Nuaimi, E., Al-Neyadi, H., Mohamed, H. and Al-Jaroodi, J. (2015) 'Applications of big data to smart cities', *Journal of Internet Services and Applications*, 6(25), pp. 6–25.

Amer, S. B. (2014) 'City definitions', *Paper Presentation at DTU Management Engineering*, Technical University of Denmark, May 25.

Arrigo, E., Liberati, C. and Mariani, P. (2016) 'A multivariate approach to Facebook data for marketing communication', in *Proceedings of the 1st International Conference on Advanced Research Methods and Analytics, CARMA2016* (pp. 66–74). Editorial Universitat Politècnica de València.

Askitas, N. and Zimmermann, K. F. (2015) 'The internet as a data source for advancement in social sciences', *International Journal of Manpower*, 36(1), pp. 2–12.

Bawa, M., Caganova, D., Szilva, I. and Spirkova, D. (2016) 'Importance of Internet of Things and big data in building smart city and what would be its challenges', in A. Leon-Garcia et al. (eds.) *Smart City 2015* (pp. 605–616).

Bello-Orgaz, G., Jung, J. J. and Camacho, D. (2016) 'Social big data: Recent achievements and new challenges', *Information Fusion*, 28, pp. 45–59.

Blazquez, D. and Domenech, J. (2018) 'Big data sources and methods for social and economic analyses', *Technological Forecasting and Social Change*, 130, pp. 99–113.

Cabrera-Sánchez, J. and Villarejo-Ramos, A. F. (2019) 'Factors affecting the adoption of big data analytics in companies', *Journal of Business Management*, 596, pp. 415–429.

Che, D., Safran, M. S. and Peng, Z. (2013) 'From big data to big data mining: Challenges, issues, and opportunities', in Hong, B., Meng, X., Chen, L. and Winiwarter, W. (eds.) *Database Systems for Advanced Applications*. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer.

Clarke, N. J., Kuipers, M. C. and Roos, J. (2019) 'Cultural resilience and the smart and sustainable city: Exploring changing concepts on built heritage and urban redevelopment', *Smart and Sustainable Built Environment*, 9(2), pp. 144–155.

Curme, C., Preis, T., Stanley, H. E. and Moat, H. S. (2014) 'Quantifying the semantics of search behavior before stock market moves', *Proceedings of the National Academy of Sciences*, 111, pp. 11600–11605.

Deville, P., Linard, C., Martin, S., Gilbert, M., Stevens, F. R., Gaughan, A. E. and Tatem, A. J. (2014) 'Dynamic population mapping using mobile phone data', *Proceedings of the National Academy of Sciences*, 111(45), pp. 15888–15893.

Ghosal, A. and Halder, S. (2018) 'Building intelligent systems for smart cities: Issues, challenges and approaches', in Mahmood, Z. (ed.) *Smart Cities, Development and Governance Frameworks*. Derby, UK: Springer, pp. 107–125.

Giffinger, R., Fertner, C., Kramar, H. and Meijers, E. (2007) *Smart cities: Ranking of European medium-sized cities*. Centre of Regional Science (SRF), Vienna University of Technology, Vienna, Austria. Available at: http://www.smart-cities.eu/download/smart_cities_final_report.pdf.

Govender, V. (2018) 'Laying down the groundwork for smart cities', *Smart Cities Africa Summit*, Emperors Palace Hotel, Johannesburg, South Africa, July 3–4, 2018.

Mavragani, A., Nikolaou, I. E. and Tsagarakis, K. P. (2016) 'Open economy, institutional quality, and environmental performance: A macroeconomic approach', *Sustainability*, 8(7), p. 601.

Mills, S., Lucas, S., Irakliotis, L., Ruppa, M., Carlson, T. and Perlowitz, B. (2012) *Demystifying big data: A practical guide to transforming the business of government*, Washington: Tech America Foundation. Available at: http://breakinggov.com/documents/demystifying-big-data-a-practical-guide-to-transforming-the-bus/.

National Planning Commission (2013) *National development plan vision 2030*. Available at: https://www.nationalplanningcommission.org.za/National_Development_Plan.

Oke, A. E., Aghimien, D. O., Aigbavboa, C. O. and Akinradewo, O. (2020) 'Appraisal of the drivers of smart city development in South Africa', *Construction Economics and Building*, 20(2), pp. 109–126.

Pallant, J. (2020) *SPSS Survival Manual: A Step-by-step Guide to Data Analysis Using IBM SPSS*, London: McGraw-Hill Education. ISBN: 9780335249503.

Rathore, M. M., Paul, A., Hong, W. H., Seo, H., Awan, I. and Saeed, S. (2018) 'Exploiting IoT and big data analytics: Defining smart digital city using real-time urban data', *Sustainable Cities and Society*, 40, pp. 600–610.

Silva, B. N., Khan, M. and Han, K. (2018) 'Towards sustainable smart cities: A review of trends, architectures, components, and open challenges in smart cities', *Sustainable Cities and Society*, 38, pp. 697–713. doi: 10.1016/j.scs.2018.01.053.

Ullah, A., Anwar, S. M., Li, J., Nadeem, L., Mahmood, T., Rehman, A. and Saba, T. (2024) 'Smart cities: The role of internet of things and machine learning in realizing a data-centric smart environment', *Complex Intelligent Systems*, 10, pp. 1607–1637.

Vassakis, K., Petrakis, E. and Kopanakis, I. (2018) 'Big data analytics: Applications, prospects and challenges', in *Mobile Big Data* (pp. 3–20). Springer.

Watson, H. J. (2014) ‘Tutorial: Big data analytics: Concepts, technologies, and applications’, *Communications of the Association for Information Systems*, 34(1), p. 65.

Xiong, T., Wang, S., Mayers, A. and Monga, E. (2016) ‘Personal bankruptcy prediction by mining credit card data’, *Expert Systems with Applications*, 40(2), pp. 665–676.

Yang, J., Fricker, P. and Jung, A. (2024) ‘From intangible to tangible: The role of big data and machine learning in walkability studies’, *Computers, Environment and Urban Systems*, 109, p. 102087.

Yaqoob, I., Chang, V., Gani, A., Mokhtar, S., Hashem, I. A. T., Ahmed, E., Anuar, N. B. and Khan, S. U. (2016) ‘Withdrawn: Information fusion in social big data: Foundations, state-of-the-art, applications, challenges, and future research direction’.