International Journal of Service Science, Management, Engineering, and Technology
Volume 15 • Issue 1

Big Data Analytics Artefact for Outcome-Based Funding Prediction in South African Public Universities

Anna M. Segooa, Tshwane University of Technology, South Africa

Billy M. Kalema, University of Mpumalanga, South Africa*

https://orcid.org/0000-0002-2405-9088

ABSTRACT

This study designed a big data analytics artefact for the prediction of outcome-based funding (OBF) in South African public universities. Universities in South Africa (SA) are subsidized based on their performance known as OBF that is measured depending on the outputs from teaching, research, and engagements. OBF metrics are well documented; however, public universities fail to achieve the targets for higher scores. These failures are attributed to poor decision-making resulting from limited analysis of the voluminous data generated. This study used design science methodology to develop a big data analytics artefact for prediction of OBF outcomes. The artefact was evaluated for prediction using machine learning training and tested with data collected from South African universities. Findings indicated that for better prediction using big data analytics, system characteristics, size, structure, top management support, market, infrastructure, and government regulations factors play a significant role.

KEYWORDS

Big Data Analytics, Data Processing, Decision Making in Public Universities, Machine Learning, Outcome-Based Funding, Teaching Development Grant

INTRODUCTION

Universities, globally, receive their financial income from various sources to successfully run their operations (Hearn et al., 2016). In South Africa, universities raise their finances through government funding, student tuition fees, and private income (USAF, 2016). According to the Department of Higher Education and Training DHET (2015), government funding includes the subsidies the ministry provides to the universities. Tuition fees are the costs that the students pay to the university; while private funding consists of funds the universities raise through research contracts, donations, investments, and the renting out of facilities. The DHET (2015) indicates that government funding accounts for 40% of total income, while tuition fees and private subsidies each account for 30%,

DOI: 10.4018/IJSSMET.334220

*Corresponding Author

respectively. The government funding is further categorized into earmarked and block grants. The earmarked grant is the subsidy that the government provides to the universities to improve student success, whereas the block grant amount is determined by the performance of the institution. The block grant is linked to the university performance and is also referred to as outcome-based funding (OBF) (Temoso & Myeki, 2023).

This funding covers approximately 70% of the government funding, which implies that failing to raise sufficient OBF is a challenge that may lead universities to experience financial shortfalls (Dlamini, 2016).

The South African landscape highlights that funding is linked to national policy goals and the performance of the universities. OBF is referred to as block grant, which is a funding that is incentivized through academic activities and is categorised into four, namely: teaching input grant, teaching output grant, research output, as well as the institution factor grant, which relates to a large enrolment of full-time disadvantaged students (De Klerk et al., 2017). The principle behind OBF is that of encouraging and motivating universities to improve the student success rate and research output (Temoso & Myeki, 2023). In the South African perspective, OBF focuses on four priority areas, namely, teaching input, which is linked to enrolments the university makes each year; teaching output, which is linked to the graduation of undergraduates and masters' structured research qualifications; research output, which is linked to publications and doctoral qualifications; as well as the institutional factor, which is linked to full-time enrolments of disadvantaged students, such as black and coloured South African citizens. Hafzan et al. (2019) observe that several measures have been proposed to enable universities to improve their OBF. These include improving on the universities' curricula to make such attractive to learners and the industry, provision of better pedagogical approaches, retention of good experienced academic staff, as well as proper analysis of the generated data by the university. Such actions could draw better insights, thus enabling real-time decision-making (Moyo & McKenna, 2021).

Hawthrone and Grzybowksi (2019) note that South Africa (SA) has one of the highest economic inequalities worldwide with a Gini coefficient of 0.63 in 2015. These economic inequalities are mainly attributed to the apartheid-era that introduced racial discrimination policies, which left many communities financially and socially excluded. Due to its historical nature and the increasing globalization, the country has been suffering insufficient human resources, especially in the academic domain. In the South African perspective, the demand for highly skilled academic employees in institutions of higher education and training has remained high, and this trend is likely to continue (Hoque & Tshutsha, 2022). As a result many universities fail to attract and/or retain highly qualified academic staff, which causes them to fall short of attaining the required scores for OBF (Segooa et al., 2019).

Attainment of a good OBF score is mainly dependent on good performance of an academic institution, which accrues from various factors among them is the attraction and retainment of highly experienced and skilled academic staff (Galli, 2020). Because of the poor economic stand of many citizens whose children seek a higher education, the South African government decided to significantly increase its contribution to higher education (Dlamini, 2016). The increment of the government's financial contribution, on the positive side, led to the establishment of more universities in the country. Whereas, on the negative side, it created more internal competition for skilled academic staff (De Klerk et al., 2017). Successful academic institutions depend on having the right number of skilled and capable academic staff who should be well motivated in order to contribute to the institution. However, in order to motivate these staff, factors including satisfactory salary, job design, internal communication, as well as better and real-time decision-making need to be put into consideration (Galli, 2020; Hoque & Tshutsha, 2022). It is imperative that academic institutions attract, motivate, and retain knowledgeable academic staff in order to remain abreast with OBF requirements (Arman, 2023). Thus, in addition to their obvious desirability, such institutions must also have effective strategies and techniques for analysing all the generated data within their settings to speed track decision-making (Segooa et al., 2019).

Universities generate voluminous data through digital internal and external activities in which some of the data is structured, other remaining unstructured. Because universities have different stakeholders with different socio-technical backgrounds, this data generally comes from a variety of sources, at high speed, and in a variety of formats (Gandomi & Haider, 2015; Temoso & Myeki, 2023). In that this voluminous data comes from various sources, it is accompanied by many uncertainties that complicate its analysis (Segooa et al., 2019). Academic institutions generate unstructured and semi-structured data, such as students' digital credentials, badge systems, data about virtual aspects of students' skills, such as their complex abilities, assessment of higher-level skills, including feedback based on those assessments, data from virtual learning systems, including Massive Online Open Courses (MOOCs), as well as complaints and suggestions from students and staff (Wishon & Rome, 2016). Analysing this data enables an understanding of complex issues and the ability to make informed decisions. Specific areas include academic analytics, business intelligence, and learning analytics. In spite of this, extracting knowledge from such highly rich data sets generated from diverse sources remains a great challenge (Selowa et al., 2022). The successful analysis of this data could be integrated into a coherent picture of the university and its stakeholders for better decision-making.

By leveraging BDA, analytics capabilities go beyond the conventional focus of current and historical data. They include modelling and forecasting, conducting what-if analyses to determine what business strategies will mean, and visualization, optimization, simulation, and location intelligence, in addition to traditional data warehouse and business intelligence applications. In turn, this will improve institutional performance and provide better services to stakeholders (Selowa et al., 2022). Consequently, South African universities with the capability of analysing all the available data have the potential to benefit maximally from OBF (Hafzan et al., 2019). More so, Thomson (2017) suggests that analysing all the data generated within an organisation is essential for identifying the challenges faced by its stakeholders, thus becoming easier to tailor-make better interventions. The inability by universities to analyse all available data prevents them from uncovering the hidden patterns and meaningful insight needed to support better decision-making (Agrawal, 2015; Daniel, 2015).

The voluminous data, also known as big data, requires advanced analytical skills, techniques, and tools. However, these skills are limited within universities, leading to poor utilization of such analytical tools (Klempin et al., 2018; Segooa et al., 2019). This limitation has been core for many universities' inability to achieve targets necessary for the obtaining of OBF, this is despite the government's effort to incentivize and motivate universities to achieve these targets (De Klerk et al., 2017). This study, therefore, sought to develop a big data analytics (BDA) artefact for the prediction of outcomes-based funding in South African public universities. The remainder of the paper details the literature on BDA, the research methodology, analysis and presentation of results, development of the artefact, discussion, and conclusion.

LITERATURE REVIEW

Big data analytics refers to the large-scale data analysis aimed at discovering hidden patterns, which aid prompt and proactive decisions. BDA could be categorized as descriptive, predictive, prescriptive, and diagnostic analytics (Bahrynovska, 2022; Segooa et al., 2019). Many organisations, universities included, are still using traditional methods such as relation database management systems for data analysis, and are becoming overwhelmed by big data because of the various formats in which it is generated (Bahrynovska, 2022). The lack of proper and better BDA results reduces efficiency through delayed decision-making; such delay hinders the real-time or near real-time needed to attend to stakeholders' queries (Segooa et al., 2019).

According to Chinsook et al. (2022), big data on its own does not deliver value; however, its proper analytics enables organisations to draw better insights that lead to the realization of business value and return on investment. This implies that better BDA provides ideas that can be used in the management and instruction of higher education in a variety of situations. This includes improving services that meet the needs of lecturers or learners, as well as obtaining all of the services necessary to efficiently adapt and deploy in the institutions. Tasmin et al. (2020) observe that BDA enables organisations to make evidence-based decisions rather than using intuition. The researchers maintain that despite the potential of BDA to improve decision-making within organisations, there are challenges that may hinder organisations from gaining better insights from the received data. Such challenges may include lack of required skills and proper tools for analysis, inappropriate methods of interpretation of the analysed data, as well as the failure to discover the hidden patterns and relationships from the data.

Data Protection and Related Risks in Big Data Analytics

The more data an organization collects, the more difficult it is to store it securely and such has seen data breaches of billions of records of data (Islam et al., 2023). This highlights the importance of data security, but also the challenges organizations face in protecting data. Maintaining data security becomes more difficult and costly as an organization holds more data. Big data analytics systems like any other technological innovation dealing with data, such as cloud environment, can be crashed by attacks such as information theft, distributed denial of services (DDoS) attacks, ransomware, or other malicious activities that originate offline or online (Wang, 2019). This implies that a protected security system for big data is necessary in order to address these challenges. Though a number of technological options are available to organisations to advance privacy while pursuing data analytics under the privacy by design (PbD) framework, big data security issues are still challenges that need attention during analytics. According to Tasmin et al. (2020), various approaches have been suggested to minimize these challenges. These include, but are not limited to, data minimisation that reduces the collection of personal information unless a compelling purpose is outlined, de-identification that is used to remove all information that could identify an individual from a dataset, either directly or through linkages to other datasets, as well as enforcement of user access controls, which grant or deny specific requests to obtain information. Wang (2019) adds that such processes, when coupled with other security policies during data analytics, can enable organizations to use data to move their business forward while protecting the personal information they hold through careful planning and application of privacy techniques and principles during big data analytics.

Related Studies

Advanced analytics can help leaders in most higher education institutions significantly improve their productivity and research by enabling new ways of engaging current and prospective learners, increasing enrolment, and improving learners' retention, as well as their success and completion rates. This section discusses related work that has been conducted to inform higher learning institutions, such as universities, to leverage BDA in order to achieve intended outcomes, improving their operations.

Sin and Muthu (2015) conducted a systematic literature review with a focus on educational data mining and learning analytics, and its implications for higher education. These researchers established that potential outcome from BDA in higher education is the augmenting of student success. The authors indicate that students' grades and trends in behaviour that may be discovered from Learning Management Software (LMS) and Integrated Tertiary Software (ITS) can predict academic success and student dropout rates. Sin and Muthu (2015) indicated that the leveraging of BDA influences the institution's performance by uncovering hidden areas that require management's intervention. This study highlighted areas such as performance prediction, attrition risk detection, data visualization, intelligent feedback, course recommendation, student skill estimation, behaviour detection, grouping and collaboration of students, social network analysis, developing of concept maps, constructing of courseware, and planning and scheduling could be revealed. Their study emphasized that BDA provides greater insights that help decision-making in higher education. Sin and Muthu (2015) recommended further studies to examine the attributes of BDA in education for better diagnostics, prediction, and forecasting that could help performance in higher education.

Murumba and Micheni (2017) acknowledge that many universities are moving to cloud architectures; therefore the data they generate is increasing. Such calls for the need to use BDA for enhanced decision-making. The researchers indicate that with increased globalization, public universities are faced with pressure to improve the quality of learning outcomes, while paying attention to cutting their operational costs. Murumba and Micheni (2017) emphasized that proper analysis of the huge amounts of data they collect during the enrolment stages, finance sourcing, and instructional process, is the better way to achieve this goal. Their study further indicates that BDA will also assist them to address significant numbers of pressing issues, including increasing educators' effectiveness, harnessing insights from learning experiences, equipping students with relevant skills for their future well-being and employability, improving students' pass and success rates, reducing non-productive credits, as well as optimizing non-core services and operations. Murumba and Micheni's (2017) study recommended that educational institutions, especially those in developing countries, be encouraged to invest in analytics programmes, and to develop expertise to help them take advantage of big data. Furthermore, these researchers indicated that by so doing, the big data approach to data management would help these institutions reduce difficulties associated with traditional data analysis. This would result in new learning methods and better decision-making by policymakers.

The study of Tarmizi et al. (2019) investigated factors contributing to student attrition in higher education. This study indicated that student attrition translates to students withdrawing from courses by either cancelling their programmes or simply failing to continue in the next phase of study. Tarmizi et al. (2019) indicated that the factors contributing to student attrition could be seen as complex social components; these factors negatively impacting on the academic, economic, and social well being of a student. These factors include individual characteristics such as age, race, gender, and family background, as well as their socio-economic status, technological literacy, culture, and support from the institutions. Tarmizi et al.'s (2019) study noted that among the many factors, the most contributing is the institution's lack of proper analysis of data collected from students. Such leads to failure to support the students' efforts in solving the challenges they face. Florea and Florea (2020) also observed that better analysis of students' data is essential in helping higher institutions to plan and provide necessary skills to students. Not only will this help students to complete their studies, but also to make their studies employable. These authors concur that because institutions of higher education are dealing with increasing student information in their daily operations, better analytics are needed in order to draw meaningful insights from the voluminous data generated.

Soliudeen et al. (2021) conducted a systematic review on higher education big data governance. The study explored ways in which BDA contributes to higher education. This study concurred with the findings of Florea and Florea (2020), observing that BDA has the potential to assist universities to monitor student performance and proactively provide real time feedback. More so, Soliudeen et al. (2021) indicated that trends in student habits and performance may be identified, and diagnostic measures may be implemented to ultimately provide interventions that may influence student performance, consequently fostering their progression. The study indicated the need for the contextualization of BDA studies to the developing countries' perspective, especially on the African continent. Soliudeen et al. (2021) hence recommended undertaking more studies on higher education's big data governance, focusing on the higher institutions' stakeholders.

Islam et al. (2023) conducted a swot analysis on how BDA could influence the universities' operational processes. These researchers observed that large and complicated data volumes are challenging to handle. Better analytical skills are needed in order to draw better insights and to realize business value. This study indicated that if universities were to discover the information they want, BDA should be their major goal. BDA could also be utilized to enhance the precision and effectiveness of cataloguing and classifying systems that will enable them to achieve their desired targets. Islam et al.'s (2023) study recommended the need for contextualized models to enable higher education institutions to acquire the needed skills and expertise to analyse big data. This recommendation was in agreement with that of Gourikeremath et al. (2021), who also suggested that an institution with better analytical skills will achieve increased revenue from research outputs and other side-income activities.

The findings from the related work are summarized in Table 1, which shows the objectives, similarities, differences, and research methodology of the reviewed related literature and this study.

Table 1. Summary of related literature

Author and Year	Objectives of the Study	Research Methodology Used	Similarities With Current Research	Differences With the Current Research
Sin & Muthu (2015)	To determine the applications of Big Data technologies in education through a systematic literature review	Systematic literature review to find studies relating to Educational Data Mining and Learning Analytics using content analysis.	Explanation of Big Data concepts and related applications of Big Data in education	The study lacked theoretical background even though some themes influencing data analytics were identified.
Tarmizi et al. (2019)	To review students' attrition in higher education (HE) and the contributing factors, as well as to find the existing computational model to analyse and predict student attrition in HE.	Comparative analysis of the categories of Big Data analytics.	Detailed analysis of applications and benefits of big data analytics	The study based on previous on literature without empirical data. More still, the study also lacked theoretical background that could inform big data analytics
Florea & Florea (2020)	To determine the Big Data and the ethical implications of data privacy in Higher Education research.	Theoretical examination of factors that call for a reconsideration of data privacy and access to private information in HE research.	Role of Big Data in academic institution and factors influencing its use were detailed.	The study lacked analysis of empirical data or implementation of an artefact. More still, it relied more on literature without a supporting theoretical background.
Soliudeen et al. (2021)	To carry out a systematic literature review on higher education governance of big data.	Kitchenham methodology to carry out a systematic literature review to identify themes.	Establishment of characteristics of Big Data in higher education, Big Data contribution to higher education, higher educational governance models, the roles of government in managing higher education and the Big Data initiative in the developing nations.	Implementation of artefact to enable government to better govern higher education was recommended though not carried out.
Islam et al. (2023)	To explore the potential of Big Data analytics in improving library management in Indonesia and identify the challenges, opportunities, and best practices for libraries when applying Big Data analytics.	SWOT analysis of Big Data in Indonesian library management.	Big Data analytics used to enhance the administration and operations within organizations. Identifying the challenges, opportunities, and best practices for putting Big Data analytics into reality.	Practical experience of Big Data analytics recommended from the lens of SWOT analysis where design features could not be recommended.

continued on following page

Table 1. Continued

Author and Year	Objectives of the Study	Research Methodology Used	Similarities With Current Research	Differences With the Current Research
Gourikeremath et al. (2021)	To carry out a comparative assessment of scientific productivity of science faculties of the University of Mysore and the Karnatak University using different bibliometric indicators, as reflected in publications covered in Web of Science database, during 2002-2016.	Quantitative approach to do a comparative analysis.	Highlighting factors needed for universities for increasing the research output and the impact of research leading to OBF.	The study concentrated more on research and publications as the predominant source of OBF. Also paid little attention to Big Data analytics, yet with Web of Science database by Clarivate Analytics is already dealing with Big Data.

THEORETICAL FOUNDATION

BDA for decision-making works similarly to any new technological innovation. Various information systems' theoretical frameworks and models have been suggested to inform users' acceptance, adoption, and the use of new technological innovations. Commonly used ones include the technology, organisation, and environment (TOE) theory (Depietro et al., 1990), the diffusion of innovations (DOI) theory (Rogers, 1995, 2003), the theory of the unification of the amalgamated models of information systems (Kwon & Zmud, 1987), the unified theory of acceptance use of technology (UTAUT) (Venkatesh et al., 2003), and the technology acceptance model (TAM) (Davis, 1989). These models and theories have been used singly or in composite form to inform studies of acceptance and use of technology. Researchers such as (Momani, 2020; Sigama & Kalema, 2022; Taneja & Bharti, 2021) have used these models to predict user acceptance and use of new technological innovations, finding them satisfactorily predictive.

Based on the factors that influence BDA identified from the literature, this study opted to use TOE as the underpinning theory in order to embrace all aspects of technology, organisation, and environment identified in the literature as influencing BDA. The TOE model is the implementation theory that embraces the aspects of technology, organisation, and environment. With respect to BDA, the technology aspects assess the characteristics and availability of tools needed to analyse the voluminous data. The organisational aspects, on the other hand, assess those issues relating to the structure, size, communication process, and slack in adopting the technological innovation. The environmental aspects examine the technology infrastructure support, market, and government regulations governing BDA (Islam et al., 2023; Soliudeen et al., 2021; Tarmizi et al., 2019).

TOE, or its constructs, has been used in other studies to assess the readiness and use of BDA (Kalema & Mokgadi, 2017; Kasten, 2020). These studies emphasized the role of organisational support, since modern data analytics techniques that are relevant in analysing big data require advanced data-analysis techniques. The above researchers indicated that such techniques could include machine learning, artificial intelligence, and natural language processing, to extract insights from large and complex data sets. This implies that users must be equipped with new skills in order to remain abreast of the new technologies. Hence, organisational support and clear understanding of the technology characteristics play a critical role in this situation. Other researchers such as Islam et al. (2023) indicate that BDA includes self-service tools that promote democratization, natural-language query tools, and tools used to extract insights from large and complex data sets. Such tools, even when available, require support from both the vendor and the organisation. This understanding explains the relevance of the TOE model in underpinning this study, as it embraces the major aspects of organisation and technology.

The environmental aspects, on the other hand, are essential, various organisations lagging behind in adopting new technologies at the rate they planned due to poor support from government and vendors (Kalema & Mokgadi, 2017). Assessing the environmental features helps them to know the current state of modern analytics, the challenges in moving forward to adopt modern analytics, the needed technologies and tools for modern analytics, as well as the characteristics of successful organisations within the business (Agrawal, 2015; Schüll & Maslan, 2018). Understanding these characteristics, along with individual and big data characteristics, was leveraged by this study to design the conceptual model that guided the flow of the remainder of this paper.

The Conceptual Model

The composite BDA conceptual model was derived from the five constructs, namely, technological, organisational, environmental, and individual characteristics, as well as data-quality management. As demonstrated by researchers such as Islam et al. (2023) and Schüll and Maslan (2018), proper BDA leads to informed decision-making after which universities will be in a better position to identify those elements that demand attention to improve their OBF. The conceptual model embraces those attributes needed for BDA in order to draw better insights relevant to educational institutions. As the demand for better analytics and prediction increases, organisations must be guided by better models to remain competitive in the big data era. Holistically, organisations must address those aspects that will help them realise trends, opportunities, and adoption strategies, including support, analytics culture, and the technical foundations of data (Galli et al., 2021; Radwan et al., 2023). Such would include data management, data preparation, data visualization, as well as formal infrastructure to establish procedures for the exchange and flow of data and information, such as information technology (IT) that collaborates business units to address dynamic stakeholders' requirements (Galli et al., 2021). Additionally, the BDA infrastructure requires the identification and implementation of strategies in relation to adaption of informal and formal procedures, organization aspects, including internal analysis, influence of the external environment, as well as the individuals' perceptions. Hence, the conceptual model needs to address theses aspects in order to create a better artefact design (Galli et al., 2021; Segooa et al., 2019). Figure 1 represents the conceptual model.

Explanation of the Constructs and Hypothesis Development

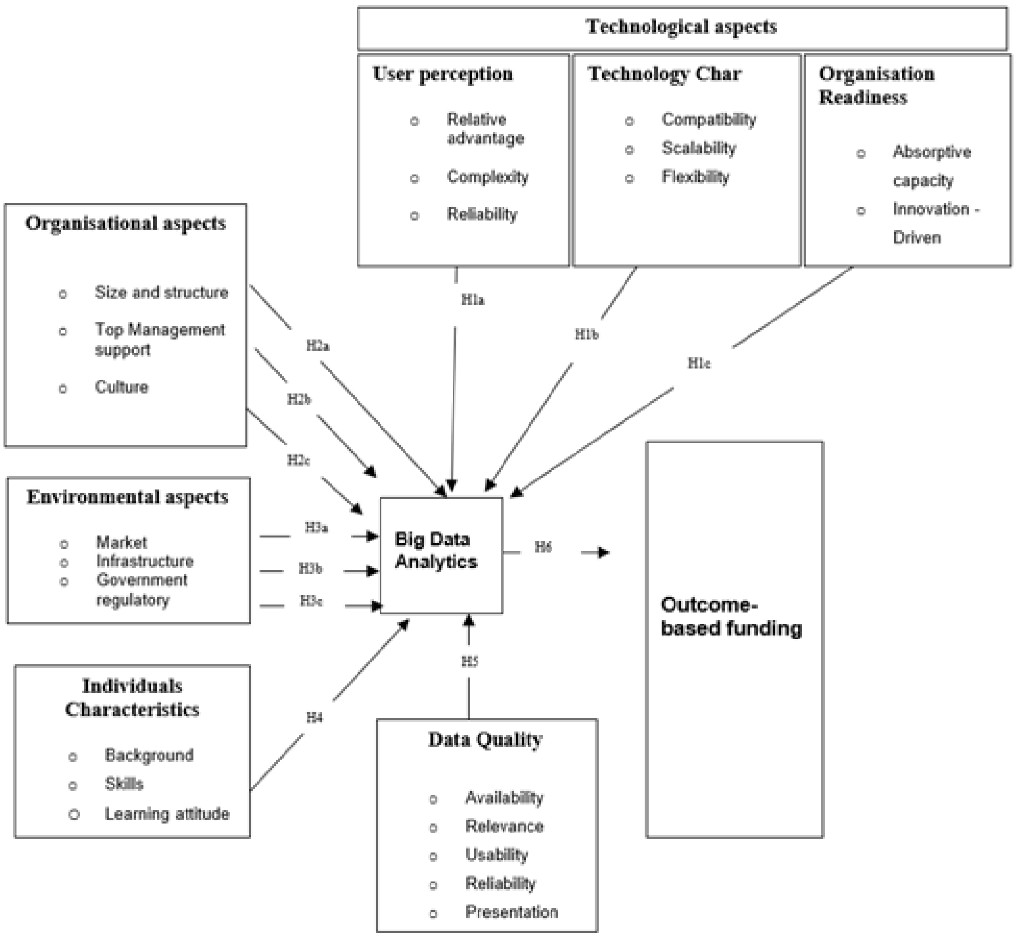
As demonstrated in Figure 1, five constructs informed the development of the conceptual model. From the understanding of these constructs, the hypotheses for this study were derived. These constructs are discussed in the following paragraphs.

Technological Aspects

As data volume and diversity expand, managing these becomes increasingly difficult. Automation and augmentation of tasks associated with the data and its analytics are becoming more common. These include tools for identifying sensitive data, classifying and tagging it, building data warehouses, surfacing insights, and building models automatically (Bahrynovska, 2022). Technology aspects thus play a significant role in BDA. This study classified technological aspects into three categories, namely: technology characteristics, user perception of the technology, and organisational readiness for technological innovations. From these subconstructs, hypotheses H1a to H1c were derived.

H1a: User perception of the technology influences BDA to achieve OBF.
H1b: Technology characteristics influence BDA to achieve OBF.
H1c: Organisational readiness for technological innovations influences BDA to achieve OBF.

Figure 1. The conceptual model



Organisational Aspects

The power of big data analytics lies in the ability to use massive amounts of data from diverse sources to identify opportunities and risks, providing organisations with a faster and more cost-effective way of making decisions and improving their business outcomes (Okour et al., 2018). However, to achieve this, the size and structure, as well as the support that management gives to sustain BDA, plays a significant role. The larger the organisation, the greater the expected amount of data to be generated. Additionally, modern analytics requires using sophisticated tools and technologies to process, transform, and analyse data in real-time or near-real-time (Pedro et al., 2019). To achieve this objective, the information technology team needs support from the top management. This led to the development of hypotheses H2a to H2c.

H2a: The size and structure of an organisation influences BDA to achieve OBF.
H2b: Top-management support influences BDA to achieve OBF.
H2c: Organisational culture influences BDA to achieve OBF.

Environmental Aspects

Big data analytics, as with any technology, is influenced by the availability of tools to use. Government, as well as bodies, set policies and standards, as well as support from vendors. In business, environmental factors refer to the surroundings in which an organisation operates. Environmental context is important because it directly affects decision-making processes in organisations, including the decision to adopt new technologies such as BDA (Mezghani et al., 2022; Pedro et al., 2019). In relation to BDA, environmental aspects are incorporated in order to fully understand the decision-making process associated with the way in which the environment impacts the availability of tools and support needed for analytics. These factors include market or competitive pressure, availability of tools, infrastructure, as well as government policies and standards. From this construct, hypotheses H3a—H3c were derived.

H3a: The environmental aspects due to market availability influence BDA to achieve OBF.
H3b: The availability of infrastructure influences BDA to achieve OBF.
H3c: Government regulatory standards influence BDA to achieve OBF.

Individual Characteristics

As the volume, complexity, and diversity of data continues to increase, various challenges emerge as a result of rapidly evolving volumes of data. These challenges may include the need to understand the core of big-data-management practices, the ability to apply such practices to complex issues, and the continuously drawing of insights from the voluminous data (Mezghani et al., 2022). The IT teams on hand need to empower users with skills and techniques for use of the needed tools for BDA, as well as for building new solutions from the analysed big data (Bahrynovska, 2022). In order to achieve this, individual characteristics play a serious role, users' attributes, such as attitude, educational background, and skills, come to centre stage. This construct has led to the development of the fourth hypothesis, H4.

H4: Individual characteristics influence BDA to achieve OBF.

Data Quality

Data quality is as diverse as the data it represents. This is one aspect that plays a critical role in any data-driven applications, including end-to-end data management. The quality of data is essential for organisational growth as it is leveraged to make meaningful decisions. In relation to BDA, an organisation must govern the procedures, and applications must analyse the generated data. It should be noted that high-quality data could be processed and analysed quickly, resulting in better insights and bigger data-analytics efforts (Pedro et al., 2019). Data quality is essential. With quality data, organisations are in a position to give customers the best experience when decisions are made, because accurate data has been applied. Good customer experience is paramount for their satisfaction, leading to brand loyalty and higher revenue (Arman, 2023; Islam et al., 2023). This construct formed the basis for the development of the fifth hypothesis, H5.

H5: Data quality influences BDA to achieve OBF.

Big Data Analytics

While there are many opportunities for businesses to use big data in the digital age, especially with the advent of the Fourth Industrial Revolution (4IR), IT and data-management teams struggle to deliver insights from the generated data within the time frame and with the quality standards that businesses require (Soliudeen et al., 2021). This implies that BDA within any organisation, higher education institutions inclusive, plays an important role in holistically understanding the stakeholders' challenges, thus forging solutions to those challenges (Kasten, 2020). This implies for universities seeking to achieve OBF, better analysis of the generated big data remains the best option. Based on this understanding, the sixth hypothesis was derived.

H6: Big data analytics directly influences the achievement of OBF.

RESEARCH METHODOLOGY

This study was conducted in two phases. The first phase involved the collection of data following a quantitative approach, whose findings informed the architecture of the artefact. The second phase focused on designing of the artefact by following the Hevner (2007) approach.

Data for the quantitative phase was collected from two public universities in South Africa selected based on proximity, willingness to assist with data collection, as well as their number of campuses spread across various provinces of South Africa. The intended respondents were lecturers, heads of departments, research assistants, librarians, and registrars' personnel. These respondents formed the population of the study. Participants were presumed to have sufficient exposure and interaction with students' data and reports that inform decision-making towards OBF. Based on the statistics of post-secondary school education and training in South Africa during 2016, the population of both academic and non-academic staff combined for one university was 5,331. That of the other university was 3,089. Being finite populations, this study leveraged the Krejcie and Morgan's (1970) tool for computation of a sample size from a finite population. The tool was used to determine the sample size of a population greater than 8,000 but less than or equal to 9,000. The sample size of the study was approximately 368. This was based on distributing the measuring instrument—a questionnaire with close-ended questions.

After determining the sample size, simple random sampling was used to distribute the questionnaires to the respondents. Overall, 400 questionnaires were distributed to the two universities of which 270 (67.5%) were returned. Of these, only 219 were usable. The rest had incomplete data and were discarded. Thereafter, data was imported onto the statistical package of social sciences (SPSS) for analysis. After data analysis, results were used to inform the development of the artefact's architecture. Artefact design was conducted using Windows Apache MySQL and PHP (WAMP), with the Laravel framework as the backbone. The data-gathering interface and backend were designed using the concept of three-tier architecture that makes use of the model view controller (MVC).

Reliability Testing of the Questionnaire

The questionnaire's reliability was tested using Cronbach's alpha; reliability was found to be 0.879. Consequently, the reliabilities of each independent construct were also tested. None of the constructs showed a reliability that called for deleting of the construct on items. This gave a clear indication to proceed with data analysis.

RESULTS

Descriptive analysis results indicated that most respondents 63.9% (n = 140) were lecturers; followed by librarians 13.7% (n = 30). On the other hand, the fewest participating respondents 6.8% (n = 15) were from the registrar's environment, followed by heads of department (HODs) with 7.3% (n = 16), and research officers with 8.2 (n = 18).

Correlations Analysis

Correlation analysis was also conducted to determine how each construct related to another, giving the relationship between the variables. The analysis was conducted on the constructs of technology characteristics, users' perception towards technology, organisational readiness for technological innovations, organisation size and structure, top-management support, culture, availability of market, university infrastructure, and government regulation. Others were individual characteristics, BDA and OBF. Results are as illustrated in Table 2.

Correlation tests were conducted on 0.01 level (2-tailed) and 0.05 level (1-tailed), and the level of the relationship was flagged, as illustrated on Table 2. Values close to one are said to be highly correlated, whereas those below 0.5, if not significant at either 0.01 or 0.05, are said to be least correlated. Technological characteristics and organisational culture, as well as availability of market, were seen to be of least non-significant correlation. As demonstrated in Table 2, the rest of the constructs showed a good correlation with one another.

Regression Analysis

Regression analysis helps to show how each construct contributes to the overall prediction of the model. Using the bootstrap approximates to normality at a 95% confidence interval reflected t-value > = to 1.96 at 0.05 level for significance (Hair et al., 2021). Furthermore, the variance inflation factor (VIF) was also included to check the existence of multicollinearity. The multicollinearity is said to exist if VIF >5. Table 3 presents the regression results.

Results as demonstrated in Table 3 indicate that 12 constructs and nine sub-constructs made a significant contribution, with big data analytics (BDA) having the higher contribution of 35.1% (β = 0.351 and p = 0.001). On the other hand, three constructs of environmental aspects due to government

Table 2. Correlation results

UserPerc	Pearson Correlation	1												
	Sig. (2-tailed)													
	N	219												
SysChar	Pearson Correlation	.183**	1											
	Sig. (2-tailed)	.007												
	N	219	219											
OrgReadTech	Pearson Correlation	.174**	.255**	1										
	Sig. (2-tailed)	.010	.000											
	N	219	219	219										
OrgSizeStruc	Pearson Correlation	.084	.198**	.254**	1									
	Sig. (2-tailed)	.218	.003	.000										
	N	219	219	219	219									
OrgMgtSup	Pearson Correlation	.148*	.246**	.239**	.400**	1								
	Sig. (2-tailed)	.028	.000	.000	.000									
	N	219	219	219	219	219	1							
OrgCul	Pearson Correlation	.100	.132	.303**	.107	.154*								
	Sig. (2-tailed)	.140	.052	.000	.114	.023								
	N	219	219	219	219	219	219							
EnvtMkt	Pearson Correlation	.358**	.104	.333**	.327**	.196**	.330**	1						
	Sig. (2-tailed)	.000	.124	.000	.000	.004	.000							
	N	219	219	219	219	219	219	219						
EnvtInfra	Pearson Correlation	.071	.231**	.226**	.257**	.266**	.195**	.180**	1					
	Sig. (2-tailed)	.292	.001	.001	.000	.000	.004	.008						
	N	219	219	219	219	219	219	219	219					
EnvtGovReg	Pearson Correlation	.207**	.179**	.200**	.388**	.330**	.128	.318**	.226**	1				
	Sig. (2-tailed)	.002	.008	.003	.000	.000	.058	.000	.001					
	N	219	219	219	219	219	219	219	219	219				
IndChar	Pearson Correlation	.125	.101	.106	.343**	.350**	.101	.179**	.199**	.259**	1			
	Sig. (2-tailed)	.064	.135	.117	.000	.000	.135	.008	.003	.000				
	N	219	219	219	219	219	219	219	219	219	219			
DataQual	Pearson Correlation	.122	.212**	.276**	.402**	.429**	.128	.181**	.268**	.325**	.295**	1		
	Sig. (2-tailed)	.072	.002	.000	.000	.000	.059	.007	.000	.000	.000			
	N	219	219	219	219	219	219	219	219	219	219	219		
OBF	Pearson Correlation	.170*	.251**	.243**	.423**	.455**	.125	.215**	.294**	.383**	.354**	.915**	1	
	Sig. (2-tailed)	.012	.000	.000	.000	.000	.065	.001	.000	.000	.000	.000		
	N	219	219	219	219	219	219	219	219	219	219	219	219	
BDA	Pearson Correlation	.173*	.191**	.051	.238**	.261**	.052	.165*	.188**	.289**	.279**	.262**	.630**	1
	Sig. (2-tailed)	.010	.004	.453	.000	.000	.440	.014	.005	.000	.000	.000	.000	
	N	219	219	219	219	219	219	219	219	219	219	219	219	219

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

Table 3. Regression analysis

Coefficients								
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	4.287	0.582		7.362	0.000		
	UserPerc	0.383	0.103	0.308	3.718	0.004	0.795	1.259
	SysChar	0.587	0.280	0.159	2.097	0.026	0.850	1.176
	OrgReadTech	0.088	0.085	0.081	1.038	0.301	0.758	1.319
	OrgSizeStruc	0.312	0.092	0.267	3.395	0.011	0.662	1.510
	OrgMgtSup	0.152	0.076	0.101	1.994	0.034	0.683	1.465
	OrgCul	0.042	0.093	0.034	0.452	0.652	0.831	1.203
	EnvtMkt	-0.287	0.136	-0.178	-2.106	0.036	0.651	1.535
	EnvtInfr	0.301	0.103	0.251	2.918	0.018	0.833	1.201
	EnvtGovReg	0.030	0.080	0.030	0.379	0.705	0.737	1.358
	IndChar	0.111	0.056	0.095	1.988	0.043	0.778	1.286
	DataQual	0.196	0.086	0.195	2.283	0.020	0.693	1.443
	BDA	0.334	0.076	0.351	4.393	0.001	0.823	1.216
a. Dependent Variable: OBF								

regulations, organisational culture, and organisation readiness for technology, were found not to make significant contributions.

Testing the Hypotheses

By using the non-parametric t-value, the set hypotheses were tested at a 95% confidence interval. Results are as presented in Table 4.

As illustrated in Table 3, three hypotheses in relation to the constructs of organisational readiness to accept technological innovations, organisational culture, and government regulatory standards were rejected.

Table 4. Hypotheses testing

Hypothesis	Result	Action
H1a: Users' perception of technology influences BDA to achieve OBF.	$p = 0.004 < 0.05$	Accepted
H1b: Technology characteristics influence BDA to achieve OBF.	$p = 0.026 < 0.05$	Accepted
H1c: Organisational readiness for technological innovations' influences BDA to achieve OBF.	$p = 0.301 > 0.05$	Rejected
H2a: Organisation size and structure influence BDA to achieve OBF	$p = 0.011 < 0.05$	Accepted
H2b: Top-management support influences BDA to achieve OBF.	$p = 0.034 < 0.05$	Accepted
H2c: Organisational culture influences BDA to achieve OBF.	$p = 0.652 > 0.05$	Rejected
H3a: The environmental aspects due to market availability influences to BDA to achieve OBF.	$p = 0.036 < 0.05$	Accepted
H3b: The availability of infrastructure influences BDA to achieve OBF.	$p = 0.018 < 0.05$	Accepted
H3c: Government regulatory standards influence BDA to achieve OBF	$p = 0.705 > 0.05$	Rejected
H4: Individual characteristics influence BDA to achieve OBF.	$p = 0.043 < 0.05$	Accepted
H5: Data quality influences BDA to achieve OBF.	$p = 0.020 < 0.05$	Accepted
H6: Big Data analytics directly influence the achievement of OBF.	$p = 0.001 < 0.05$	Accepted

DISCUSSION OF RESULTS

As presented in Table 3, technological aspects play a significant role in analysing big data. Both hypotheses H1a relating to the influence of user perception of BDA to achieve OBF, and H1b relating to the technological characteristics, were accepted. On the one hand, user perceptions of the technology may include its ease of use, usefulness, as well as efficacy while using it. The technological characteristics embrace those aspects relating to flexibility, scalability, as well as security. Both positive beliefs and its features increase user trust and confidence to use it effectively and efficiently. On the other hand, hypothesis H1c, relating to the organisational readiness to use technology, was not accepted. Such could imply that a technology may be used whether the organisation is ready or not, as long as the need calls for its use. Kalema (2022) indicated, during the peak of Covid-19, that even those organisations not savvy in using technology were forced to automate their activities, thus enabling their employees to work from home. Additionally, researchers Moyo and McKenna (2021) also observed that using big data analytics tools requires a positive attitude and perceptions of users, which will help the organisation to draw better insights from the generated data.

Results also indicated that the organisational size and structure (H2a), as well as the support received from top management (H2b), are significant in the use of BDA to attain OBF. Organisational size and structure is proportional to the data generated within an organisation; hence the need for better analytics to gain better insights. Additionally, organisations with formalised structures have their own tailored rules and procedures by which business activities are performed. Employees are supposed to adhere to these prescriptions in order to know their stakeholders' needs. More still, the support users receive from their top management and the buy-in smooths over the implementation of new technological innovations. Support, such as training, appropriate IT budgets, upskilling, and reskilling enables users to apply the technological innovation with ease. These findings concur with those of researchers Islam et al. (2023) and Soliudeen et al. (2021), who alluded to the importance of supporting users for effective BDA. The findings of this study, however, indicate that organisational culture has no significant influence on BDA in achieving OBF (H2c). Kalema (2022) indicates that the advent of the 4IR caused pressure on organisations to digitize their services in order to have a share of the global market. Such pressure ultimately compelled organisations to adopt technological innovations with no consideration for their traditional way of doing things. Culture, thus, did not have significant influence in the adoption of new technological innovations, yet it caused market share availability to be salient (H3a) due to increased global pressure.

Hypothesis H3b theorized the influence of availability of infrastructure in the organisation's environment. This hypothesis was accepted. BDA requires tools that work beyond the traditional transaction processing systems in order to analyse both structured and non-structured data. Hence, infrastructure, such as reliable internet connections and network connections, play an important role. The findings of this study are in agreement with those of [REMOVED HYPERLINK FIELD] who noted that reliable infrastructure increases BDA capability, assists in the conforming to data-processing standards, and improves data quality. Such elements are essential for effective decision-making. However, hypothesis H3c, relating to government regulations and standards, was not accepted. The implication of this could be that much as there is an OBF framework to be used, this can only be achieved should the available data be properly analysed, better insights, therefore, being drawn. Hence, for universities to conform to the OBF government framework, they need first to find better ways of analysing the data within their environment.

Results further indicated that individual characteristics' influence on BDA in achieving OBF (H4) is significant. BDA, as with any technology, requires an individual conducting it to have a positive attitude and beliefs, the required skills, willingness to learn new skills, as well as awareness of what needs to be done. This allows individual characteristics to play a significant role in BDA. The findings of this study concur with those of Walls and Barnard (2020) and Mezghani et al. (2022), who note that an organisation's competitiveness in the big data era largely depends on individuals. These individuals should be willing to go the extra mile to learn new skills needed for better analytics.

The fifth hypothesis, H5, predicted the significance of the influence of data quality on BDA to improve OBF. This hypothesis was accepted. The implication of this finding is that high-quality data provides accurate insights that enable better decision-making. This also implies that organisations must ensure that generated data be made available to all users in a better format, clean, reliable, and relevant to the context. Data should be disseminated in such a manner that it is readable and visible (Soliudeen et al., 2021). The findings of this study are in support of those of Kasten (2020), who noted that poor data quality leads to poor insights and poor decision-making. Conversely, good data quality enables an organisation to perceive external data benefits and to increase internal data use and facilitation of resources needed to draw actionable insights.

Lastly, results indicated that the sixth hypothesis, H6, relates to the direct influence of BDA on OBF. This hypothesis was accepted. The implication of this finding is that the potential to draw better insights informed by BDA can reflect accurate reporting and deliver timeous information for the stakeholder. This also suggests that better BDA reveals all those challenges that may hinder progress or improvement of performance that could reduce the attainment of OBF. These findings are in agreement with those of researchers such as Janssen et al. (2017) and Kasten (2020), who indicated that BDA has the potential to improve quality and trust of data needed for decision-making. The findings are also in agreement with Murumba and Micheni (2017), who highlighted that should universities leverage BDA, they will gain business value. Thereafter, decision-making can be improved.

ARTEFACT DEVELOPMENT

The quantitative findings, coupled with the literature that was reviewed relating to OBF, informed the development of the artefact, including the identification of both functional and non-functional requirements. The interface was designed using Windows Apache MySQL and PHP (WAMP), with the Laravel framework as the backbone. Designing a friendly and suitable user interface is essential in allowing users flexible interaction with the system so as to provide meaningful operation and production of good results (Akkaya & Ovatman, 2022). The data-gathering interface and backend were designed using the concept of three-tier architecture that makes use of the model view controller (MVC). From the interface, a user selects the factors that influence the teaching output (TO), research output (RO), and enrolment from students with disadvantaged backgrounds, which is regarded as an institutional factor (I). The captured data are saved in the backend. The system exports the data from there to a comma-separated value (CSV) file, which is used to run the prediction of OBF.

Implementation of the Artefact

South African landscape highlights that funding is linked to national policy goals and the performance of the universities. The OBF is measured by scores of academic activities categorised into four aspects that include: teaching input grant, teaching output grants, research outputs, and institution factor grant. These aspects forms the South Africa block grant/OBF Universities grant model (De Klerk et al., 2017). The model works on generated funding framework submitted to DHET by universities to compute the OBF scores. From the literature review, teaching input category 23, variables were identified as: teaching output 8 variables, research output 30 variables, and institutional factors, 7 variables. Based on machine learning training, data could be generated automatically if this approach was used with the identified 78 variables to generate data for system training and the data was saved in the backend, from where the system exported it a comma-separated value (CSV) file. Of the generated data, 80% was used to train the system, whereas 20% was used for system testing. Figure 2 presents the datasets that were used to train the system.

The coding of the process steps of the system training and the output of the result from the prediction are as demonstrated in Figure 3.

Figure 2. Sample of training datasets


```
10,210,210,211,209,210,211,209,210,210,210,211,211,210,209,211,210,211,209,209,211,58,60,58,58,58,58,58,59,58,58,58,58,60,58,58,58,59,58,30,30,31,30,31,30,30,
11,210,211,210,211,209,210,209,210,209,209,209,210,210,211,210,210,209,211,209,211,58,59,58,58,58,58,58,59,58,57,58,58,60,58,58,58,60,58,30,30,31,30,31,31,31,
11,210,209,209,211,211,209,210,211,210,211,209,209,210,211,210,211,211,209,211,210,58,60,58,58,58,58,58,59,58,56,58,58,60,58,58,58,59,58,30,31,31,30,30,31,31,
11,210,211,211,210,210,209,211,209,210,211,209,210,209,211,210,210,209,209,209,210,58,59,58,58,58,58,58,59,58,56,58,58,60,58,58,58,60,58,30,30,31,30,31,31,31,
10,210,210,210,211,210,211,211,211,210,211,211,210,211,211,210,211,210,210,210,209,58,59,58,58,58,58,58,59,58,56,58,58,60,58,58,58,60,58,30,30,30,31,30,30,30,
10,210,209,210,209,210,211,210,211,210,211,211,209,210,209,210,209,210,210,210,211,58,60,58,58,58,58,58,60,58,57,58,58,60,58,58,58,60,58,31,31,31,30,30,30,30,
10,209,209,211,210,211,211,209,210,210,209,210,210,210,209,210,210,209,209,209,210,58,59,58,58,58,58,58,59,58,56,58,58,59,58,58,58,60,58,30,31,30,31,30,30,31,
10,210,209,210,209,209,210,210,209,211,209,209,209,211,209,210,211,210,211,211,211,58,59,58,58,58,58,58,59,58,58,58,58,60,58,58,58,60,59,30,31,30,31,30,31,30,
11,211,210,210,210,211,210,211,209,210,210,211,209,210,211,210,209,209,210,209,210,58,59,58,58,58,58,58,59,58,57,58,58,59,58,58,58,60,59,30,30,30,31,30,31,30,
10,209,209,210,211,210,209,209,209,211,209,210,211,211,211,209,211,209,211,211,211,58,59,58,58,58,58,58,59,58,56,58,58,59,58,58,58,59,58,31,31,31,30,30,31,30,
10,209,210,209,210,210,211,209,211,211,210,210,211,211,211,209,211,209,211,209,209,58,60,58,58,58,58,58,60,58,57,58,58,60,58,58,58,59,58,31,30,31,30,31,31,31,
10,209,211,209,209,210,209,209,209,211,211,209,211,211,210,211,209,210,211,211,210,58,60,58,58,58,58,58,59,58,58,58,58,59,58,58,58,60,58,31,30,30,30,30,31,31,
10,210,209,209,210,211,211,209,210,210,209,209,211,211,209,211,210,211,209,209,211,58,59,58,58,58,58,58,59,58,56,58,58,59,58,58,58,59,58,31,30,31,31,31,30,31,
10,209,209,209,210,209,211,209,209,211,209,209,210,211,211,209,211,210,211,211,210,58,60,58,58,58,58,58,60,58,56,58,58,59,58,58,58,59,58,31,30,30,30,30,30,30,
09,211,211,209,211,210,211,209,210,209,211,210,211,211,209,210,211,210,211,209,211,58,59,58,58,58,58,58,59,58,56,58,58,59,58,58,58,60,59,31,30,31,31,30,31,30,
11,210,210,211,211,211,210,211,210,210,211,209,211,209,211,209,210,209,209,210,210,58,59,58,58,58,58,58,60,58,56,58,58,59,58,58,58,60,58,31,30,30,31,31,30,30,
09,211,211,210,210,209,211,210,210,211,209,209,210,209,210,211,211,211,209,209,211,58,59,58,58,58,58,58,59,58,58,58,58,60,58,58,58,60,58,30,30,31,31,31,30,31,
09,210,210,211,209,209,209,210,211,210,209,209,209,209,210,209,209,209,209,209,209,58,59,58,58,58,58,58,60,58,56,58,58,59,58,58,58,59,58,31,31,30,31,31,31,31,
09,211,210,211,209,209,210,209,211,209,211,210,211,209,211,210,211,211,211,211,209,58,60,58,58,58,58,58,59,58,56,58,58,60,58,58,58,60,59,31,31,31,30,31,30,30,
11,209,211,210,211,209,210,211,209,211,211,211,210,209,211,209,209,210,209,210,211,58,59,58,58,58,58,58,59,58,57,58,58,59,58,58,58,60,59,30,30,31,30,30,30,30,
11,211,209,210,211,210,211,211,209,209,211,210,210,209,210,209,210,209,210,210,211,58,60,58,58,58,58,58,60,58,57,58,58,59,58,58,58,59,59,30,30,31,30,31,31,31,
10,209,210,211,211,210,209,210,211,209,211,209,209,209,210,209,209,210,209,211,209,58,59,58,58,58,58,58,59,58,57,58,58,60,58,58,58,59,58,31,30,30,31,31,30,30,
09,209,209,211,209,209,209,209,210,209,210,209,211,209,209,210,211,210,209,210,210,58,60,58,58,58,58,58,60,58,57,58,58,59,58,58,58,60,58,31,30,31,31,31,31,31,
```


Figure 3. The coding steps of the model in the spider environment testing the artefact

allnOneEvaluationAlgorithm.py | finalPrediction.py | finalPhDPredictionModel.py | PhDAllInOneTestModelAnalytics.py

```python
# -*- coding: utf-8 -*-
"""
Created on Thu Feb 13 01:16:56 2019

@author: Anna Segooa
"""

from pandas import read_csv
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LogisticRegression
from pickle import dump

#LOAD THE DATASETS
filename = 'finaloutcomebasedprediction.csv'

#CREATE THE COLUMN NAMES FOR HEADING FOR THE DATASETS
colnames = ['High workload', 'challenging course content','student demographics','family income','high school Grade Point Average',
            'remedialcourse for firstyear students','Reputation of institution','Socioeconomic opportunities of the university',
            'cultural opportunities of the university','Lecturer respect towards students', 'Academic Intervention',
            'Housing and Residence Life','Financial Aid and Scholarships', 'Library access', 'Enrolment on student success course',
            'costs on university fees', 'Location From Students Home', 'Students Academic Goals support','peers relationship and
            'Equipped classroom', 'English language', 'Teaching method', 'warning system', 'classroom environment', 'student assessmen
            'student performance monitoring', 'feedback from lecturer', 'subject coordination', 'Tuition fees', 'student support facil
            'Experience with staffs', 'Analysis of attendance register', 'intervention academic support and social counsellors',
            'Learning with peers', 'Student tertiary preparedness', 'financial condition', 'availability of tutors and mentors', 'soci
            'Class participation', 'Speciality choice on passion','Self-regulation skill', 'Assessment criteria to enroll doctoral stu
            'Inadequate information resources', 'Poor research project management', 'Poor communication on seminars', 'workshops and c
            'Poor support to workshops and conferences','lack of sponsorship and scholarships','One on one meetings with supervisors',
            'Group meetings with peers', 'Writing retreats','Mentoring approach','Supervision methods','Research office culture', 'Pub
            'Libraries support','Research environment','Subject specialist requirements for core or main supervisor','Researcher skill
            'Supervisors training and workshops','Lack of students commitment','Inadequate computer literacy','Personal issues such as
            'lack of support from spouse and family members','Poor time management','Inadequate writing skills','Lack of time to consu
            'Mutual trust and respect between supervisor and Leaner','Clear  expectations','Teaching and professional behavior','Fee I
            'Deep inequality','Socio-economic deprivation','Economic political and cultural spheres','Social arrangements respect all
            'National Student Financial Aid Sscheme Funding Strategies', 'class']

reader = read_csv(filename, names=colnames)

#PASS ALL THE VALUES IN THE DATASETS INTO AN ARRAY VARIABLE
array = reader.values


#NOW SPLIT THE DATASET INTO INPUT AND OUTPUT
X= array[:,0:78]
Y=array[:,78]

#The next thing is to then split our data into test data and train data
test_data = 0.25
seed = 8

X_train,X_test,Y_train,Y_test = train_test_split(X,Y,  test_size=test_data, random_state=seed);
model = LogisticRegression(solver='liblinear');
model.fit(X_train, Y_train);

#Now, we need to save our Model to disk
modelname='universityoutcomebasedprediction.sav'
dump(model, open(modelname, 'wb'))
```

The testing of the artefact was conducted from the inception, using unit testing methodology. Thereafter, a full system testing was conducted. During the testing process, it was important to note that the artefact is capable of performing both descriptive and diagnostic analytics that focus on what has happened and why, during the prediction of OBF. This is allowed to determine the ratio of the contribution of each factor towards predicting OBF, and to forecast what happens should the university fail to meet the metrics of such a factor. This prediction helped to provide insight into how to change the future planning of the university. Moreover, the artefact was also tested for compatibility in the various business environments to verify that it could work with the existing infrastructure and interoperability.

Prediction Output

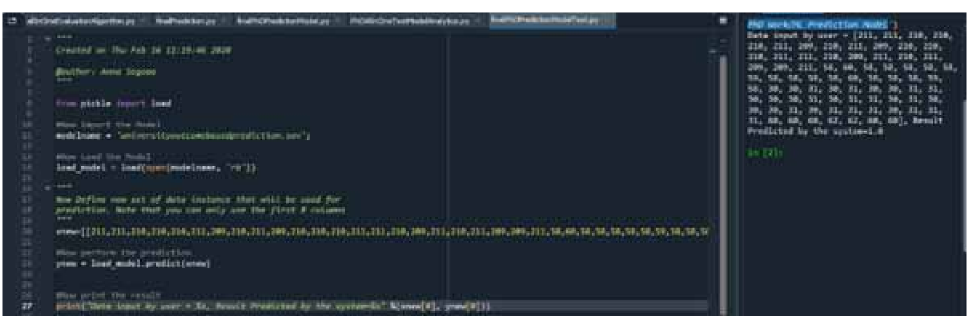
Some background mathematical calculations were embedded in the CSV file generated using the Excel formula. These calculations were used to determine if the funding application was successful or not. Based on the result of the calculations, a new column was then added to the CSV files named "result." The CSV file was then used by the Python modelling framework through the Anaconda Platform. The Spyder data science and data analytics environment was used to perform data modelling. The artefact was implemented in machine learning by using the CSV file generated from the backend server. The data was loaded into the modelling system and used to generate columns indicating the output, demonstrating whether funding was successful or not. The output was represented with 1 or 0, where 1 = successful and 0 = unsuccessful. All the values in the columns were then passed into an array that was split into input and output. The output became the result column that was generated, and other columns became the input data. The "result" column formed the basis of the data modelling. The next step was to use the dataset to train the system to recognize the patterns; this performed the modelling of predicting OBF.

Figure 4 illustrates the prediction results obtained after the machine learning training of the system. The output can either be "1" for successful or "0" for unsuccessful.

CONCLUSION, LIMITATIONS, AND DIRECTIONS FOR FUTURE RESEARCH

Recent technological advancements are poised to significantly impact the way in which organisations work. The vibrant development of technological evolution and innovative applications are compelling organisations to collect voluminous, diverse, and disparate data, whose analytics are increasingly becoming too complex for the traditional systems to handle (Bahrynovska, 2022; Chinsook et al., 2022). For better analytics and realization of value, trustworthy data is paramount for any form of data architecture; such makes big data analytics a top priority. As a result, organisations must modernize their data-analytics approaches in order to gain competencies needed for BDA. Furthermore, literacy is

Figure 4. The output



essential for analysing big data as the new pervasive technologies evolve in the day-to-day operations of organisations.

Data remains an asset of many organisations including universities; however, its value can only be realised when organisations are in the position of analysing it and drawing better insights therefrom. This study revealed that in order to effectively analyse the increasing volumes of data, the "big data" organisations must pay attention to a variety of factors including technology, organisation, characteristics of individuals consuming the data, as well as the environment in which they operate. Universities, likewise, should note that it is through successful BDA that they can attain sufficient competencies to gain a competitive advantage and better funding.

LIMITATIONS AND FUTURE WORK

This study was informed by data collected from only two public universities in South Africa that were selected based on their willingness to participate in the study. Much as the factors that were used to model the OBF artefact are generic to all universities in South Africa, the system was trained based on the data collected from these two universities. This study acknowledges that individual universities may have different processes that could influence the data used to train the system for prediction of OBF. Therefore, the findings of this study could be limited to generalisation to all South African universities. This study recommends that future research consider widening the population of the study so as to have a large representative sample data on which to train the system for predicting OBF. More still, for the South African universities to be competitive, the government, through its DHET, is supporting universities basing on their productivity, hence, the essence of OBF. However, some Sub-Saharan Africa countries do not apply these constraints, as is the case with South Africa. Therefore, the findings of this study are limited in generalization to all Sub-Saharan African universities. Even while this is so, other universities could effectively use the findings of this study for competitiveness. Similarly, the artefact could be applied in any university that needs to improve on its competitive advantage and service delivery.

Normally, organisations would want to work seamlessly by increasing their stakeholders' satisfaction through improved flexibility, innovativeness, effective collaboration, and risk mitigation (Arman, 2023; Radwan et al., 2023). However, organisations are usually challenged by the advent of new technological trends (Kalema, 2022). Such new trends hinder organisations' agility. Agility calls for the need to modernize through upskilling on issues of data virtualization and visualization, data fabrics and democratization, automation of data catalogues, as well as on data silo consolidation into unified cloud data management. Much as this study included training within its organisation, factors that influence the upskilling and reskilling were not detailed. Nevertheless, BDA requires modernization and use of advanced tools and software. This study recommends that future research emphasize the need for modernizing organisations' capabilities for BDA.

CONTRIBUTIONS OF THE STUDY

Theoretically, this study is based on the literature and primary data to develop a model using BDA to predict OBF for South African universities. Both the model and the artefact make significant contributions to research. First, they inform how big data can be effectively analysed, and second, they predict how universities receiving funding via OBF fare in the face of crises, such as that of the Covid-19 pandemic, which paralysed most activities. This successful analysis could help to inform how OBF financed universities could be more resilient to future shocks. Additionally, this study demonstrates to management the importance of leveraging the power of machine learning for prediction through BDA. However, to use such tools for BDA, organisations must equip their stakeholders with the skills needed to use modern technological tools and techniques. The artefact will serve as a good reminder to organisations' management of the need to train its employees, reskilling when the need arises.

AUTHOR NOTE

Correspondence concerning this article should be addressed to Billy M Kalema, School of Computing and Mathematical Sciences, University of Mpumalanga.

Email: billy.kalema@ump.ac.za

REFERENCES

Agrawal, K. P. (2015). *Investigating the determinants of Big Data Analytics (BDA) Implementation in emerging economies*. Academic Press.

Akkaya, K., & Ovatman, T. (2022). A comparative study of meta-data-based microservice extraction tools. *International Journal of Service Science, Management, Engineering, and Technology*, *13*(1), 1–26. doi:10.4018/IJSSMET.298677

Arman, A. (2023). How perceived inclusion impacts minority employees' organizational satisfaction in the UAE banking industry. *International Journal of Service Science, Management, Engineering, and Technology*, *14*(1), 1–14. doi:10.4018/IJSSMET.326551

Bahrynovska, T. (2022, September). *Big Data in Education. How It Transforms the Industry?* Forbytes. https://forbytes.com/blog/use-of-big-data-in-education/

Chinsook, K., Khajonmote, W., Klintawon, S., Sakulthai, C., Leamsakul, W., & Jantakoon, T. (2022). Big data in higher education for student behavior analytics (Big Data-HE-SBA System Architecture). *Higher Education Studies, 12*(1), 105–114. doi:p10510.5539/hes.v12n1

Daniel, B. (2015). Big data and analytics in higher education: Opportunities and challenges. *British Journal of Educational Technology*, *46*(5), 904–920. doi:10.1111/bjet.12230

Davis, F. D. (1989). Perceived usefulness, perceived ease-of-use, and user acceptance of information technology. *Management Information Systems Quarterly*, *13*(3), 319–340. doi:10.2307/249008

De Klerk, D., Spark, L., Jones, A., & Maleswena, T. (2017). Paving the road to success: Reflecting critically on year one of an undergraduate student support programme at a large South African university. *Journal of Student Affairs in Africa*, *5*(1), 1–13. doi:10.24085/jsaa.v5i1.2478

Department of Higher Education and Training (DHET). (2015). *Ministerial Statement on University Funding: 2015/16 and 2016/17*. https://www.dhet.gov.za/Financial%20and%20Physical%20Planning/Ministerial%20Statement%20on%20University%20Funding%202016-2017%20and%202017-2018,%20November%202015.pdf

Depietro, R., Wiarda, E., & Fleischer, M. (1990). The context for change: Organization, technology, and environment. *The Processes of Technological Innovation, 199*(0), 151–175.

Dlamini, W. (2016). Factors affecting effective management of limited resources in South African comprehensive universities. *Research Journal of Business and Management*, *3*(3), 219–233. doi:10.17261/Pressacademia.2016321994

Florea, D., & Florea, S. (2020). Big data and the ethical implications of data privacy in higher education research. *Sustainability (Basel)*, *12*(20), 8744. doi:10.3390/su12208744

Galli, B. J. (2020). Impact and role of motivation theories in continuous improvement environments a reflection of literature. *International Journal of Service Science, Management, Engineering, and Technology*, *11*(1), 1–13. doi:10.4018/IJSSMET.2020010101

Galli, B. J., Bongo, M. F., Yamagishi, K. D., & Ocampo, L. A. (2021). A critical literature analysis of the relationships of marketing and strategic planning under project environments. *International Journal of Service Science, Management, Engineering, and Technology*, *12*(2), 1–24. doi:10.4018/IJSSMET.2021030101

Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, *35*(2), 137–144. doi:10.1016/j.ijinfomgt.2014.10.007

Gourikeremath, G., Kumbar, B. D., Hadagali, G. S., & Hiremath, R. S. (2021). A comparative study of University of Mysore and Karnatak University in science: Research output and citation impact during 2002-16. *Journal of Indian Library Association*, *56*(2), 75–83.

Hafzan, M. Y. N. N., Safaai, D., Asiah, M., Saber, I. M. M., & Syuhaida, S. S. (2019). Review on predictive modelling techniques for identifying students at risk in university environment. In *MATEC Web of Conferences* (Vol. 255, p. 03002). EDP Sciences.

Hair, J. Jr, Hair, J. F. Jr, Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2021). *A primer on partial least squares structural equation modeling (PLS-SEM)*. Sage Publications. doi:10.1007/978-3-030-80519-7

Hawthrone, R., & Grzybowksi, L. (2019). Benefits of regulation vs competition where inequality is high: The case of mobile telephony in South Africa. *Economic Research Southern Africa Working Paper,* (791), 1–43.

Hearn, J. C., Warshaw, J. B., & Ciarimboli, E. B. (2016). Privatization and accountability trends and policies in U.S. public higher education. *Education in Science*, *41*(184), 1–26. doi:10.15390/EB.2016.6270

Hevner, A. R. (2007). A three cycle view of design science research. *Scandinavian Journal of Information Systems*, *19*(2), 87–92.

Hoque, M., & Tshutsha, N. (2022). Exploring factors affecting the attraction and retention of academic employees at the Durban University of Technology. *International Journal of Service Science, Management, Engineering, and Technology*, *13*(1), 1–12. doi:10.4018/IJSSMET.290332

Islam, N., Islam, K., & Islam, M. (2023). Exploring the potential of big data analytics in improving library management in Indonesia: Challenges, opportunities, and best practice. *Internet Reference Services Quarterly*, *27*(2), 111–120. doi:10.1080/10875301.2023.2184900

Janssen, M., Van Der Voort, H., & Wahyudi, A. (2017). Factors influencing big data decision making quality. *Journal of Business Research*, *70*, 338–345. doi:10.1016/j.jbusres.2016.08.007

Kalema, B. M. (2022). Developing countries' continuance usage of e-services after Covid-19 in the 4IR era. *Proceedings of NEMISA Summit and Colloquium*, 4, 1-11. doi:10.29007/2cdm

Kalema, B. M., & Mokgadi, M. (2017). Developing countries organizations' readiness for big data analytics. *Problems and Perspectives in Management, 15*(1), 260–270. 10.21511/ppm.15(1-1).2017.13

Kasten, J. E. (2020). Big data applications in healthcare administration. *International Journal of Big Data and Analytics in Healthcare*, *5*(2), 12–37. doi:10.4018/IJBDAH.2020070102

Klempin, S., Grant, M., & Ramos, M. (2018). *Practitioner perspectives on the use of predictive analytics in targeted advising for college students* (CCRC Working Paper No. 103). Columbia University, Teachers College, Community College Research Center.

Krejcie, R. V., & Morgan, D. W. (1970). Determining sample size for research activities. *Educational and Psychological Measurement*, *30*(3), 607–610. doi:10.1177/001316447003000308

Kwon, T. H., & Zmud, R. W. (1987). Unifying the fragmented models of information systems implementation. In *Critical issues in Information Systems Research (pp. 227–251)*. John Wiley & Sons.

Mezghani, K., Alsadi, A. K., & Alaskar, T. H. (2022). Study of the environmental factors' effects on big data analytics adoption in supply chain management. *International Journal of E-Business Research*, *18*(1), 1–20. doi:10.4018/IJEBR.309395

Momani, A. M. (2020). The unified theory of acceptance and use of technology: A new approach in technology acceptance. *International Journal of Sociotechnology and Knowledge Development*, *12*(3), 79–98. doi:10.4018/IJSKD.2020070105

Moyo, T., & McKenna, S. (2021). Constraints on improving higher education teaching and learning through funding. *South African Journal of Science*, *117*(1–2), 1–7. doi:10.17159/sajs.2021/7807

Murumba, J., & Micheni, E. (2017). Big data analytics in higher education: A review. *The International Journal of Engineering and Science*, *6*(06), 14–21. doi:10.9790/1813-0606021421

Okour, M., Chong, C. W., Asmawi, A., & Akour, M. (2018, July). Knowledge management systems usage from the user's perspective: The influence of organisational factors in Jordanian banking sector. In *2018 8th International Conference on Computer Science and Information Technology (CSIT)* (pp. 207–212). IEEE.

Pedro, J., Brown, I., & Hart, M. (2019). Capabilities and readiness for big data analytics. *Procedia Computer Science*, *164*, 3–10. doi:10.1016/j.procs.2019.12.147

Radwan, A. F., Snoussi, T., Mousa, S. A., & Abdulzaher, M. (2023). Using smart applications in delivering government services in the UAE: Factors of adoption and satisfaction. *International Journal of Service Science, Management, Engineering, and Technology*, *14*(1), 1–15. doi:10.4018/IJSSMET.326520

Rogers, E. M. (1995). Diffusion of innovations. *The Joint Commission Journal on Quality Improvement*, *21*(7), 324–328. doi:10.1016/S1070-3241(16)30155-9 PMID:7581733

Rogers, E. M. (2003). *Diffusion of innovations.* Academic Press. doi:10.4102/sajim.v24i1.1489

Selowa, K. T., Ilorah, A. I., & Mokwena, S. N. (2022). Using big data analytics tool to influence decision-making in higher education: A case of South African technical and vocational education and training colleges. *South African Journal of Information Management*, *24*(1), 1–8. doi:10.4102/sajim.v24i1.1489

Schüll, A., & Maslan, N. (2018, March). On the adoption of big data analytics: Interdependencies of contextual factors. ICEIS, (1), 425–431.

Segooa, M. A., Kalema, B. M., & Zolait, A. H. (2019). Predictive analytics to improve outcome-based funding for the public universities in South Africa through big data. In *2019 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT)* (pp. 1–6). IEEE. doi:10.1109/3ICT.2019.8910285

Sigama, K., & Kalema, B. M. (2022). A Model for MOOC Implementation in Areas of Low Bandwidth in Developing Countries. *International Journal of Distance Education Technologies*, *20*(1), 1–17. doi:10.4018/IJDET.312182

Sin, K., & Muthu, L. (2015). Application of big data in education data mining and learning analytic-A lterature review. *ICTACT Journal on Soft Computing*, *5*(4), 1035–1049. doi:10.21917/ijsc.2015.0145

Soliudeen, M. J., Adenuga, K. I., & Sadiq, F. I., (2021). Higher education governance of big data: A systematic literature review. *Digital Solutions and the Case for Africa's Sustainable Development*, 152–172. 10.4018/978-1-7998-2967-6.ch010

Taneja, B., & Bharti, K. (2021). Mapping unified theory of acceptance and use of technology (UTAUT) 2: A taxonomical study using bibliometric visualisation. *Foresight*, *24*(2), 210–247. doi:10.1108/FS-08-2020-0079

Tarmizi, S. S. A., Mutalib, S., Hamid, N. H. A., & Rahman, S. A. (2019). A review on student attrition in higher education using big data analytics and data mining techniques. *International Journal of Modern Education and Computer Science*, *11*(8), 1–14. doi:10.5815/ijmecs.2019.08.01

Tasmin, R., Muhammad, R. N., & Nor Aziati1 A. H. (2020). Big data analytics applicability in higher learning educational system. In IOP Conference Series: Materials Science and Engineering (Vol. 917, No. 1, p. 012064). IOP Publishing. doi:10.1088/1757-899X/917/1/012064

Temoso, O., & Myeki, L. W. (2023). Estimating South African higher education productivity and its determinants using Färe-Primont Index: Are historically disadvantaged universities catching up? *Research in Higher Education*, *64*(2), 206–227. doi:10.1007/s11162-022-09699-3 PMID:35669093

Thomson, S. D. (2017). Preserving transactional data. *International Journal of Digital Curation*, *11*(2), 126–137. doi:10.2218/ijdc.v11i2.419

Universities South Africa (USAF). (2016). *Universities funding in South Africa; A fact sheet*. Universities South Africa. https://www.uct.ac.za/usr/news/downloads/2016/UniversitiesFundingSouthAfrica_FactSheet.pdf

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Towards a unified view. *Management Information Systems Quarterly*, *27*(3), 425–478. doi:10.2307/30036540

Walls, C., & Barnard, B. (2020). Success Factors of Big Data to Achieve Organisational Performance: Theoretical Perspectives. *Expert Journal of Business and Management*, *8*(1), 1–16.

Wang, C. (2019). The strengths, weaknesses, opportunities, and threats analysis of big data analytics in healthcare. *International Journal of Big Data and Analytics in Healthcare*, *4*(1), 1–14. doi:10.4018/IJBDAH.2019010101

Wishon, G., & Rome, J. (2016). Institutional analytics and the data tsunami. *Educause Review*. https://er.educause.edu/articles/2016/12/institutional-analytics-and-the-data-tsunami

APPENDIX A

Please Put a Cross in the Box Corresponding to an Answer That You Feel Suitable for the Question. In this study Outcome-Based Funding (OBF) is the subsidy institutions receive based on their outcomes of enrolments, graduates, and research outputs. Big data refers to data generated in high volumes, high speed, from different sources with different data formats. Big data analytics is the process of analysing data that comes in high volumes, high speed, which is generated from different sources and formats to gain more insights for better decision-making.

SECTION A: General Information

1. Which University do you work for?

University A	
University B	

2. What is your position?

Directors /DVC	
Dean/Assistant dean	
Head of Department	
Lecturer	
Registrars	
Librarian	
Research officer	

3. For how long have you been working at the university?

0 -3 years	**4- 7 years**	**8 -11 years**	**12-15 years**	**16+ years**

4. What is your overall work experience?

0 -3 years	**4- 7 years**	**8 -11 years**	**12-15 years**	**16+ years**

5. What system do you use to extract or analyse data for reporting?

Microsoft Excel	
ITS	
Microsoft Access	
Higher Education Data Analyser (HEDA)	
Cliq view, Cognos, SQL)	
Other: Please specify	

6. Have you heard of Big Data?

Yes	**No**

SECTION B: Constructs to Measure Big Data Analytics Adoption and Data Quality

By using the rating scale from 1–5 where; 1 = Strongly Disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly Agree. Indicate your level of agreement or disagreement for the following statements:

7. Technological Issues	Questionnaire Item	Rating Scale				
7.1. User perception (***Assess the perception of the user towards a technology***)		**1**	**2**	**3**	**4**	**5**
7.1.1 Relative advantage (perceived usefulness)	Using technology helps me to do my work faster.					
	Using new technologies helps me to effectively do my work.					
7.1.2 Complexity (perceived ease of use)	I find technology difficult to use when doing my work.					
	I prefer doing my work manually rather than using technology.					
7.1.3 Reliability	The current system we use to do our work is reliable enough for us to depend on it.					
7.2. System characteristics (***Assess the technology towards capability of the system itself***)		**1**	**2**	**3**	**4**	**5**
7.2.1 Scalability	The current system we are using has the potential to accommodate and analyse different types of data formats.					
7.2.2 Compatibility	The current system we are using can be accessed from different platforms (i.e. Windows, Lynx, and Mac etc.).					
7.2.3 Flexibility	The system we are currently using has the ability to extract student data from different business units where it is captured.					
	The system we are currently using has the ability to analyse student data from different business units.					
7.3. Organization readiness (***Assess the likelihoods of the university to adopt new technology)***						
7.3.1 Absorptive capacity	Our university is likely to invest funds in innovative systems.					
	We have the required knowledge and skills to analyse volumes of data as generated within our university.					
	Our university has policies in place to sensitize staff on the new technological innovations to be implemented.					
7.3.2 Innovation driven	Management involves employees in making decisions regarding the implementation of new technologies.					
7.3.3 Funding	The management allocates funds for new IT projects.					

8. Organizational Issues	Questionnaire Item	Rating Scale				
8.1 Size and Structure (***The factor assess the university setting which can potentially influence the adoption of new technologies***)		**1**	**2**	**3**	**4**	**5**
8.1.1 Size	Our university has organisational structure with different supporting units.					
	Our university has a large number of staff and students that need to be supported by our IT department.					
8.1.2 Formalisation	We have policies in place that relate to data analysis and reporting.					
	All employees are encouraged to follow the approved policies and rules of the university.					
8.1.3 Centralisation	Our data analysis and reporting are prepared at individual business unit					
	Top management makes decision based on reports that are compiled by dedicated individuals.					

continued on following page

Table. Continued

8. Organizational Issues	**Questionnaire Item**	**Rating Scale**				
8.2 Top Management Support ***(the factor assess the level of management supports towards new technologies)***		**1**	**2**	**3**	**4**	**5**
8.2.1 Training	We have training policies in place to adjust to new technologies.					
	Management allocates funds for training and developments.					
8.2.2 Awareness	Employees are made aware of the technologies introduced to the institutions to improve work performance .					
8.2.3 Financials	The university supports new Information Technology (IT) projects budgets.					
8.3 Culture*(the factor assess the culture of decision making within the university and its flexibility to leverage on technology innovation)*		**1**	**2**	**3**	**4**	**5**
8.3.1 Decision making culture	Our university relies on data analytics to make decisions.					
	Our university supports data-driven innovations to improve service delivery.					
	The university has a technological system for managements to view reports.					
8.3.2 Agility	We continuously examine innovative opportunities to better our processes.					
	Our university takes advantage of new technologies to improve our efficient and effectiveness.					
8.3.3 Norms and standards	We have set-standards of analysing generated data within our business units.					

9. Environmental Issues	**Questionnaire Item**	**Rating Scale**				
9.1 Market (The factor assess the influence of market on the university to adopt new technologies)		**1**	**2**	**3**	**4**	**5**
9.1.1 Vendor*(service providers of I.T systems i.e. HEDA and ITS)*	Service providers of our IT systems give us support whenever we need help.					
	Our service provider gives training support after installation and upgrade of the systems.					
9.1.2 Competition	Our university experiences competition intensity from other institutions.					
	Our university has pressure to deliver reports to stakeholders like department of higher education.					
9.1.3 Environment uncertainty	The pressure from stakeholders, such as students, forced us to adopt to new technological innovations.					
	The stakeholders requests prompt responses on assessments.					
9.1.4 Customer diversity	Our students comes from different locations and backgrounds.					
	Our student enrolment tries to balance according to race, gender, and disability.					
9.2 Infrastructure ***(The factor assess the state of the infrastructure to potentially adopt new technologies)***		**1**	**2**	**3**	**4**	**5**
9.2.1 Sharing of resources	It is easy for me to access resources whenever I am in need of them.					
	Our network is stable and reliable to allow access of resources.					
	The IT infrastructure of my university is flexible to support new technological innovations.					
9.2.2 Access to data	We have authentication and access procedures to access data within our business units.					
9.3 Government **(*This factor assess the government regulatory and support effects towards new technologies and information dissemination)***		**1**	**2**	**3**	**4**	**5**

continued on following page

Table. Continued

9. Environmental Issues	Questionnaire Item	Rating Scale				
9.3.1. Regulations	I am aware of the national policies of handling and dissemination of information.					
	Our universities follow national policies and procedures of handling information.					
9.3.2. Financial support	The government funds initiatives that support technology innovations.					
9.3.3 Learning platforms	The government assists and supports in the development of technological skills.					
10. Individual Characteristics ***(This factor assess the capability and skills of the users to adopt Big Data Analytics)***		**1**	**2**	**3**	**4**	**5**
10.1 Background	I have relevant IT experience to enable me analyse data within my business unit.					
10.2 Analytics skills	I have done training in data analytics.					
10.3 Learning Attitude	I am willing to train in order to acquire skills needed for data analytics.					
11. Data Quality ***(The factor assess the quality of data that is used in reporting students enrollments, graduates and publications)***		**1**	**2**	**3**	**4**	**5**
11.1 Availability	Data that is used for reporting is available to users whenever they need it.					
11.2 Usability	We do data clean-up to assure the correctness of the content before it is used for reporting.					
	We have better expertise and skills to retrieve previous data for future reference.					
11.3 Reliability	Data used for reporting is always complete and fit for reporting.					
11.4 Relevance	We use and analyse the data we generate from students and lecturers to forecast future enrolments and to improve graduation rates and publications.					
11.5 Presentation quality	The presentation of our reports is sufficiently enough for management to draw insights required for decision-making.					

SECTION C: Overall Big Data Analytics Role Towards Outcome-Based Funding

From your experience, opinion, or observation, how can you rate the role of Big Data analytics in improving decision making towards outcome-based funding (OBF)?

12. The Role of Big Data Analytics	1	2	3	4	5
12.1 By Top management in the university					
12.2 By other staff and employees not in the IT department					
12.3 By IT personnel within the university					

SECTION D: Demographic Information

13. What is your age group?

18-20 years	21-25 years	26-30 years	31-35 years	36-40 years	41+ years

14. What is your gender?

Male	**Female**

15 What is your highest qualification?

National Diploma	**BTech/Degree**	**Postgraduate/Masters**	**DTech /PHD**

Mmatshuene Anna Segooa is a Doctor of Computing (Informatics) at Tshwane University of Technology, South Africa. She has experience in teaching of undergraduate and Advanced diploma. She serve as a postgraduate supervisor for Masters and Doctoral studies. Administratively she is involved in International Collaboration and partnerships as a projects Lead. She is a reviewer of International Journals and conferences. She is also involved in the organisation of international conferences such as International Conference on Advanced Information Systems Engineering (CAISE23). She is a research consultant for the government public service sector in South Africa. Her main research interest includes Big Data Analytics, Artificial Intelligence, Design Science, and Systematic Literature Review primarily for business application and enhancement.

Billy M. Kalema holds a PhD in Computer Science and has over 20 years teaching experience at a University level. He is a researcher, and mentor and has graduated over 50 masters and Doctoral students. He is an external examiner for postgraduate students for many universities in South Africa and Internationally. He has published widely in the areas of socio-cognitive aspects of human response to information technology including acceptance, use, utilization and evaluation of technology for decision-making, ERPs, E-learning, ICT4Education, ICT4Business Enhancement, MOOCs, Big Data and the Statistical Methods for Data Analysis. He has spoken at various international conferences, Doctoral symposiums, seminars and workshops. He is a National Research Foundation (NRF) C-rated researcher, and a member of several academic bodies including; Association of Information Systems (AIS), Institute of Electrical and Electronics Engineers (IEEE); Information Society for Africa (IST-Africa), the International Association of Computer Science and Information Technology (IACSIT) and the Asian Council of Science Editors (ACSE). He serves on several technical committees as an Editorial board member and peer reviewer for both journals and conferences. His current and future research plans revolve around the practical application of research in daily life by putting IT to use especially in the economically and technologically disadvantaged developing countries.