Using Big Data analytics tool to influence decision-making in higher education: A case of South African Technical and Vocational Education and Training colleges

Background: Big data analytics in education is a new concept that has the potential to change the decision-making landscape in South African Colleges. Higher institutions of learning, including Technical and Vocational Education Training (TVET) colleges like all other organisations, rely on data for their decision-making. These decisions affect the way pedagogy and student management is administered. Colleges collect huge quantities of data in different formats from students, staff and stakeholders for different reasons and occasions.

Objectives: The goal of this study was to investigate how Big Data analytics and their tools may improve decision making in TVET colleges in South Africa through the lens of actor-network theory (ANT).

Method: A qualitative, interpretive inquiry was undertaken. A case study using focus group was conducted. The data collected through interviews were arranged into themes and a thematic approach was employed to analyse these themes using QDA Miner Lite software.

Results: The results from focus group interviews revealed that TVET colleges collect an enormous amount of data. These data are extracted for different reasons, yet there are no Analytics used for decision-making. Decisions are made by the highest-paid individuals (HiPPO) in colleges.

Conclusion: This dissertation recommends that the TVET colleges invest in data science skills for their staff, and Big Data infrastructure. Big Data technologies such as Mongo DB and Hadoop are recommended as the most commonly and advanced tools that can be used for Big Data analytics.

Keywords: decision-making; Big Data; Big Data analytics; TVET colleges; HiPPO; Hadoop; higher education.

Introduction

Organisations need data to inform their strategic and operational decisions. The use of data in organisational decision-making is old as mankind. Organisations have been storing and analysing large volumes of data since the advent of data warehouse systems (Daniel & Butson 2013). With petabytes of data that are generated daily, it would be imprudent to make decisions without attempting to draw some meaningful inferences from the data (Nanneti 2012). The decision makers in higher learning institutions need to be able to use analytics to enhance their decision-making. The data available for decision-making today come in huge volumes, at consistently high speed and in a variety of formats. This is what is termed as Big Data.

Big Data refers to a framework that allows the analysis and management of a larger amount of data (Moreno et al. 2016). Furthermore, Big Data is less about data that is big, but more of a capacity to search, aggregate, and cross-reference large data sets (Boyd & Crawford 2012).

Actor network theory (ANT) is used as a lens to assess and propose how the use of Big Data Analytics (BDA) in Technical and Vocational Education and Training (TVET) environment can be used to improve decision-making. The rest of the article is organised as follows: we start with the background followed by brief review of the literature of BDA, then the discussion of the four translations of ANT, research methods and then the results and conclusions.
Background

Higher learning institutions collect information about students, pedagogy and related data about staff (Daniel & Butson 2013). In post-apartheid South Africa, the former Department of Education was split into two different departments: Department of Basic Education and Department of Higher Education and Training (DHET) was then established in 2018. Higher Institutions of Learning in South Africa are classified into four divisions as illustrated in Figure 1.

These institutions also include TVET colleges. Technical and Vocational Education and Training colleges collect data from students, staff and stakeholders for different reasons and occasions. The information collected are usually from various sources and in different formats, such as from social media (Facebook, Twitter, and Instagram), CCTV cameras, fleet management sensors in college vehicles and emails as well as other forms including students’ registration data. The collected data are not being used in decision-making. The current decision-making in management is still based on the traditional way of using past experience and that of the highest paid person’s opinion (HiPPO) (Appelgren & Nygren 2019).

An example of not using available data is the stampede which took place when the college opened their gates during application and registration in one of the colleges. The stampede could have been avoided had the college analysed their social media data because the issue was raised there before the unfortunate incident happened. People were hospitalised due to the stampede after sleeping outside the college gate overnight in order to be in front of the queue when the college open their gates for registration in the morning (Tshungu 2018). All this happened because the head of college instructed the registration staff to wait for him before registration could begin.

Another example is that students who intended to study at the college sometimes make inquiries through social media platforms used by the college. During their enquiry, they indicate courses of their interest. In most cases, the students are not placed in the courses they initially wanted to register for and this leads to the students dropping out or failing the course several times. The presence of student data should help college in deciding where to correctly place students.

Based on the background above and the literature, BDA can help in tracing the students before, during and after the study (Oracle 2015). Making decisions using the traditional method of relying on the HiPPO leads or just the results of the administered tests to misplacement of students in courses, requesting of staff data every 6 months, long queues during registrations, and other things going wrong that could be solved through data-centric decision-making. This study therefore investigates the use of BDA in decision-making in the TVET college sector in South Africa.

In order to achieve this goal, the following research questions were addressed:

1. What role does BDA play in improving organisational decision-making?
2. What are the decision-making processes used in TVET colleges presently?
3. In what way can BDA support and enhance decision-making in TVET colleges?

Literature review

The definition of Big Data, according to Lewis (2014) is data sources whose very size and complexity creates problems for standard data management and analysis tools. Big Data is data sets characterised by their volume, velocity of change, variety of type that can add value to organisations (see Figure 2).

Furthermore, Kakhani et al. (2013) have added veracity in their definition and explain the V’s of Big Data as follows:
1. Volume: The amount of data is at very large scale. The amount of information being collected is so huge that
modern database management tools are unable to handle it and therefore become obsolete.

2. Velocity: Data are produced at an exponential rate. It is growing continuously in terabytes and petabytes.

3. Variety: We are creating data in all forms — unstructured, semi-structured and structured data. This data is heterogeneous in nature. Most of our existing tools work over homogenous data. Thus, now we require new tools and techniques which can handle such a large-scale heterogeneous data.

4. Veracity: The data we are generating is uncertain in nature. It is hard to know which information is accurate and which is out of date.

5. Value: The data we are working with is valuable for society or not.

Analytics is viewed as ‘the use of data, statistical analysis and explanatory and predictive models to gain insights and act on complex issues’ (Bichsel 2012). According to Ong (2015), analytics is the use of data, statistical analysis, explanatory and predictive models to gain insights and to present data through various forms of visualisation. These visualisations can be in the form of graphic presentations like pie charts, tree maps and others. The Figure 3, as adapted from Ong clearly shows the taxonomy of BDA.

The picture in Figure 4 shows different kinds of BDA and their applications.

According to Riahi and Riahi (2018), Big Data is referred to as the evolution and use of technologies that offer the right user at the right time with the right information from a mass of data that has been rising rapidly for a long time in our society. Similarly, the kind of visualisation Big Data can provide to enhance teaching and learning and provide a decision support system for educational establishments to attain excellence in education is unprecedented (Bhat & Ahmed 2016). In comparison to education, Nannetti (2012) states that ‘firms that emphasize decision-making based on data and analytics have performed 5% – 6% better – as measured by output and performance – than firms that rely on intuition and experience for decision-making’.

Tulasi and Suchithra (2016) indicated that BDA can be used in higher education to enhance e-learning. They mentioned that Big Data paradigms are needed in current world to add value to the processes of educational institutions. There are different databases that support Big Data, namely:

- NoSQL databases – These databases support parallel processing of the large amount of data.
- MapReduce – Provides platform to access data in distributed file systems with intermediate data being stored on local disks.
- MongoDB and Apache Hadoop – Are few platforms which have emerged to store large chunks of data.

Tulasi and Suchithra further state that data-driven decisions would help the teaching and learning process to evolve and also indulge in creation of new pedagogy. There are a few identified analytical methods that can be used to deal with educational data and they recommend Association Rule Mining (ARM). Association Rule Mining consists of two parts: the antecedent which is an element available in the data and the consequent which is the element obtained along with the antecedent. The analytics recommended in this study are however not for decision making but for teaching and learning.

Daniel (2014) in his writing stated three stages required to unlock the value of Big Data in any organisation as collection, analysis and visualisation which he represented as shown in Figure 5.

The Daniel and Butson (2014), is composed of the following constructs:

- Institutional analytics: These are a variety of operational data that can be analysed to help with effective decisions about making improvements at the institutional level.
- Information technology (IT) analytics: The purpose of IT analytics according to Daniel is to integrate data from different systems such as student information, learning management and alumni systems, and also systems that manage learning experiences outside the classroom.
- Academic/programme analytics: These provide data that administrators can use in order to support strategic decision-making processes and provide ways for benchmarking for comparison with other institutions.
- The final construct of the UO-TEA by Daniel is Learning analytics: Its main concern is measuring, collecting, analysing and reporting of data about learners and their contexts, for purposes of understanding and optimising learning and the environments in which it occurs.
Theoretical foundation

Underpinned by, and using ANT, this study embraces both human and non-human actors in the network and considers them equally valuable and having the potential to be equally influential in the network, during decision-making processes. Through the review of the literature, the researchers (Alexander & Silvis 2014; Cresswell et al. 2010; Paledi & Alexander 2017; Rivera 2013; Tatnall 2005) noted that ANT has been successfully used to carry out interpretivist studies making it relevant to this study.

Latour (1996) indicated that ANT has been developed by students of science and technology, and its claim is that in order to understand what holds society together, we have to reintject the facts made by natural and social sciences and the artefact designed by engineers. Actor network theory does not accept any form of separation between the participants; it emphasises that human and non-human actors should be analysed in the same way (Paledi & Alexander 2017).

The four moments of translation in the theoretical framework are problematisation, interessement, enrolment and mobilisation.

Problematisation

During this moment, the primary actor/s identifies, defines and proposes a solution to a situation. The primary actors also:

[D]etermine a set of actors and defined their identities in such a way as to establish themselves an obligatory passage point in the network of relationships they were building, making them indispensable to the network. (Callon 1986)

In this moment of translation, we look at what decisions can be problematised. Is there any issue that the organisation is not deciding on and which might be causing or solving a problem that the college is or might be facing?

Interessement

During this moment, the primary actors use different processes and negotiations to get the other actors interested and buy into their proposed solution. The result is that the other actors either affirms to the proposed solution or become the primary actor’s allies or not. Callon and Latour (2002) explain that:

[7]The model of interessement sets out all of the actors who seize the object or turn away from it and it highlights the points of articulation between the object and the more or less organised interests which it gives rise to. (pp. 196–233)

This is the stage where the college might look at who are the custodians of information or data that may help in the reaching of the decisions that are to be made. We further see how these custodians of information will provide the right solution or better decision on the problems that the college is facing, thus putting a recruitment of members of a network (actants) that will be involved in solution.

Enrolment

Enrolment is described as various types of negotiations and hurdles that go with interessement that will help them to succeed which results if interessement is successful (Callon 1986) ’and a network of aligned interest starts when actors accepts the roles defined for them’ (Iyamu 2011) but it does not necessary mean interessement always leads to enrolment.

Now that we have had a successful building of interest, members of the network or actants are now moving towards the solution using whatever tool is available in the network.

Mobilisation

During mobilisation, all the actors have reached consensus and working towards the same goal to implement the targeted solution.

The actants will now be able to reach a common understanding of who is to do what and when given the availability of all members. These members (actants) are both human and non-human. The human actants in this stage have accepted the usefulness of their no-human counterparts and are using them because they know the important role each actant plays in the decision-making process through BDA.

Only two actor network concepts were adopted for this study, namely, translation and inscription. These concepts illustrate the research model for this study. In order to have the moments of translation, actors or actants will need to work together. They both recruit each other by recognising the problems in the organisation or college (problematisation); provide how each one will play a role in the solution of the identified problem (interessement and mobilisation) and finally start using all resources they have and acknowledgement if any recommendations made. It is only after a balanced network that inscription can occur. All of that will lead to a harmonious BDA. Figure 6 is a representation for the research model in this study.
Methodology
Given the fact that this study aimed to investigate how BDA can influence decision-making in TVET colleges in South Africa, it, therefore, calls for data to be collected within the participants’ natural setting and from the participants who are involved in the organisational decision-making in the TVET colleges. Semi-structured interviews were employed for this study. The researcher interviewed decision-makers (management teams) in the TVET colleges of Limpopo. The teams interviewed are either the campus management team or college management team depending on their level of participation in the decision-making in the college. This study employed interviews as the data collection method because interviews are a more personalised form of data collection method than questionnaires (Bhattacherjee et al. 2017).

In this study, data collected during interviews were arranged in themes and patterns according to the constructs of the framework and relation to the research questions. A thematic approach to analyse these themes was used. The coding resulted in several themes that were revisited and refined. QDA Miner Lite was used as an analysis tool in this study. The programme was designed to assist researchers in managing, coding, and analysing qualitative data. The emerging themes from the interviews were generated by QDA Miner Lite. They are presented in Figure 7 and Figure 8.

After identifying the themes, we reviewed the categories and themes that were identified in order to find if there are any new insight that may emerge from the data. We repeated this action until we reached what is called saturation. Saturation is achieved when no new themes can be revealed after any further observations and analysis (Lowe et al. 2018).

Ethical considerations
A presentation for full ethical approval was made to the ethics committee and ethics consent was received on December, 13, 2017. The ethics approval number is FCRE/ICT/2017/11/015(2).

Results
It was illustrated through literature that organisations that use BDA in their decision-making gain competitive advantage and can solve problems proactively. Decision-making is a process that involves problem-solving, ‘it is assumed that the decision is the basic activity of management staff and the information, which is the enabler of the decision’ (Kościelniak & Puto 2015). In an environment where Big Data is processed, the staff in the organisation are using data to guide their decisions to a much higher extent than previously (Björkman 2017). Furthermore, by applying analytics to Big Data, organisations are able to extract and exploit valuable information to enhance decision-making and support informed decisions (Elgendy & Elragal 2016). Insight-Driven Organisations, an increasing number of organisations are utilising BDA across their enterprise for improved decision-making and not limiting them to a single function or process (Tembhek 2016).

In the interviews of this study, respondents mentioned that they become aware of the problems through different personnel (staff, students). IT and marketing become aware of problems through newsfeeds from college website, social media accounts. Education Management Information Systems (EMIS) become aware of issues through Coltech SQL queries. Coltech is the most common information system that is used to store student data). Admin Managers mentioned that it is the duty of marketing and IT to ensure that the colleges are marketed or advertised well as they communicate with students through notice board, suggestion box but the staff through emails. The EMIS Managers confirmed that they use online application only and the information will be captured by admin of the college using Coltech. IT Managers mentioned that they use college website, Facebook, Instagram and WhatsApp for marketing. The respondents also indicated that the college principal and senior management as well as council are the ultimate decision makers when it comes to the college strategic plan and their job is to implement the strategic plan.

In the interview results, the participants emphasised that even though they do value team efforts in their planning, reporting and decision-making, the colleges’ decisions are made by the highest person in authority. These results also show that the decision-making is also influenced by the traditional way of deciding. These colleges adhere to the HiPPO framework in their decision-making.

Furthermore, there is a clear indication that there is a huge volume of data from different sources and in different formats that are being processed in their environment. However, the data is not analysed for decision-making. Big Data tools are not available in colleges and therefore that skill, investment and knowledge in BDA is not there.

Recommendations
In this network, decisions have to be made by management, guided by data (information) from college websites, social
FIGURE 7: QDA Miner Lite – screenshot of the themes from interviews.

FIGURE 8: QDA Miner Lite – screenshot of themes from literature.
In light of the findings that show the lack of analytics and lack of Big Data tools, the study recommends that colleges should invest in analytics skills for personnel and invest in Big Data frameworks/tools.

**Invest in personnel skills**

Technical and Vocational Education and Training Colleges need to invest in information architecture and BDA. There is a skills gap in analytics. Big Data Analytics skills have the ability to improve operational effectiveness and efficiencies of generating great revenues in business (Alshaie 2017). There are some institutions in South Africa like Nelson Mandela University, ATTI, and CTU that are able to train people for Big Data and Analytics through the Cisco academy. Furthermore, there are online organisations such as Eudemy, Coursera, Dataquest and Cisco networking academy that offer the same training for free or a small fee for certification. Mokwena (2011) suggested that institutions can improve the usage of information systems by providing adequate training offered by experienced trainers, providing enough tools, improving the support offered to personnel as well as making sure there are incentives for personnel who take up such information systems. Furthermore, ‘the onus is then on the management to ensure that they provide adequate support needed by implementers and users of the proposed technological intervention’ (Ilorah et al. 2017).

**Invest in Big Data frameworks/tools**

In order for TVET colleges to gain more insight and improve their work in learning, preventive, predictive, prescriptive and institutional analytics through Big Data, they must invest in Big Data tools. There are many tools used for BDA such as MapReduce, MongoDB, Azure, Hadoop, Flink, Storm, and Samza.

The most recommended tool that is recommended is Hadoop. The advantage about Hadoop is that it is a scalable open-source computation framework that works across many host servers which may not necessarily be high-performance computers (da Silva et al. 2018). Furthermore, Hadoop is an open source project of Apache. Some enterprises launched their own Hadoop distributions with tools to manage and administer the cluster and also with a free/premium policy (Lavanya & Murali 2016).

Big Data Analytics that can be used in TVET colleges.

According to Riahi and Riahi (2018), BDA can be categorised into four types, namely:

- **Descriptive analytics**: It is the initial stage of data processing wherein a set of historical data is created. Descriptive analytics produces future probabilities and trends and provides an idea about what is likely to occur in the future.

- **Diagnostic analytics**: Diagnostic analytics investigates and focuses on the root cause of a problem. It is used to determine reasons why things happened. This kind of analytics attempts to identify and understand the causes of events and behaviours.

- **Predictive analytics**: These analytics use past data to predict the future. It is mainly focused on forecasting or predicting what is likely to happen. Predictive analytics uses many techniques such as data mining and artificial intelligence (AI) to analyse current data and make scenarios of what is likely to occur.

- **Prescriptive analytics**: It tries to answer the question: What should be done? Its dedication is focused on finding the right action to be taken. Descriptive analytics produces historical data while prescriptive analytics uses these parameters to find the best solution.

**Conclusion**

This study uses ANT as a lens to investigate how decision-making in institutions of higher learning, particularly TVET colleges in South Africa can be integrated with BDA. We propose that the use of BDA can improve decision-making and propose different BDA tools available in the market. Big data analytical tools produce instant alerts and afford feedbacks to instructors and scholars on academic performance by analysing fundamental complex data patterns (Riffai et al. 2016). This approach will help in predicting a dropout student, student who needs additional help or even a student who needs more challenging assignments ((Riffai et al. 2016). TVET colleges already are processing Big Data even though some may not even be aware of it. Furthermore, this study will propose/demonstrate how a balanced actor-network can ensure smooth and enhanced decision-making through BDA.

**Future research**

This study intends to promote knowledge exploitation where the research contributions address problems relevant to the use of BDA for decision-making in the higher education sector in South Africa, especially in the TVET college sector. However, there are fewer studies about Big Data, BDA, or BI in the South African TVET context. Future research could investigate BI or Big Data use or readiness in this sector.

**Acknowledgements**

**Competing interests**

The authors declare that they have no financial or personal relationships that may have inappropriately influenced them in writing this article.
Authors’ contributions

K.T.S. declared that the work in the article is a study that he undertook for his Master’s Degree at Tshwane University of Technology under the supervision of S.N.M. and A.I.I.

Funding information

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Data availability

Some of the themes/codes from the interviews are in Figure 8.

Disclaimer

The views and opinions expressed in this article are those of the authors and do not necessarily reflect the official policy or position of any affiliated agency of the authors.

References


Selowa, K.T., 2022, Using big data analytics to enhance decision-making in higher education: A case of South African TVET colleges, Master of computing dissertation Tshwane University of Technology, Pretoria.


http://www.sajim.co.za