Investigation Into the Use of Learning Analytics in Online Learning at South Africa's Higher Education Institutions

By

PALESA MARALITLE MOLOKENG

MLKPAL002

SUBMITTED TO THE UNIVERSITY OF CAPE TOWN

In partial fulfilment of the requirements for the degree

Master of Commerce Degree in Information Systems

Faculty of Commerce

UNIVERSITY OF CAPE TOWN

University of Cape Town

Date: 08 June 2021

Supervisor: Prof. Jean-Paul Van Belle

Department of Information Systems

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

University of Cape Town

PLAGIARISM DECLARATION

1. I know that plagiarism is wrong. Plagiarism is to use another's work and pretend that it is one's own.
2. I have used the APA convention for citation and referencing. Each contribution to, and quotation in, this dissertation "Investigation into the use of learning analytics at South Africa's higher education institutions", from the work(s) of other people has been attributed, and has been cited and referenced.
3. This dissertation is my own work.
4. I have not allowed, and will not allow, anyone to copy my work with the intention of passing it off as his or her own work.

Signature:

Signed by candidate

Date: 08 June 2021

Full Name of Student: Palesa Maralitle Molokeng

Student Number: MLKPAL002

ACKNOWLEDGEMENTS

My master's journey would not have been complete without my supervisor, Professor Jean-Paul Van Belle. Thank you, Prof JP for your unwavering support, guidance, constructive feedback and encouragement throughout this journey. I definitely would not have been able to achieve this without you.

My deepest gratitude to my friends who encouraged me to enrol and kept me on course. Thank you for your love, support and patience throughout. I would like to thank my IS Master's part-time classmates for being a sounding board when I needed one and for persevering through the journey with me.

I would like to thank all the participants who agreed to take part in this study. Thank you for setting aside time in your busy schedules during a pandemic to help in shaping this study.

I would like to thank the late Professor Suellen Shay for sharing so much wisdom at the start of this study, may your soul rest in peace.

Abstract

Teaching and learning have evolved over time as students desire to engage with technology as part of their learning process. Technological advances are introducing new possibilities for higher education institutions (HEIs). Developments in educational technologies as well as the need by HEIs to improve both the teaching and learning environment are contributing to the growth of learning analytics (LA). Research in LA has predominantly been exploratory in nature and a shift is needed for evaluative research that explores the impact and outcomes of LA on students, educators and the institution as a whole. Moreover, there is limited research on the use of LA at HEI in a South African context with most research focusing on the Global North and Australia.

This study aimed to understand the state of LA at HEIs in South Africa (SA) by identifying barriers of use as well as future potential of the use of LA in informing decision making, predicting learner outcomes and improving the overall learning and experience. The study draws on the Technology-Organisation-Environment (TOE) framework and DeLone and McLean IS success model to develop a theoretical integrated model to understand the state of LA at HEI in order to gain insights on the barriers and drivers of use, as well as potential of more advanced LA at institutions. To meet the objective of the study and to test the model, data was collected by conducting semi-structured interviews with participants at four prominent HEIs in SA. It was important to have a diverse representaion of roles to gain a balanced view on the use of LA across institutions. A qualitative thematic analysis approach was used following a hermeanutic cycle in the analysis process. The outcome of the study showed that a big variety of data are currently being collected at HEIs within SA and the data is used for LA to varying degrees.

These findings show that data is used for various decision-making purposes in order to improve learning and teaching. The data is also used to understand student behaviours in order to predict student outcomes and to meet student support needs. Barriers and challenges exist that have impacted the adoption of LA; these include technological, organisational and environmental challenges such as quality of systems, culture to use data and ethical concerns. Opportunities for more advanced use of LA also exist which are econcouraging institutions to prioritise LA initiatives in their drive of being more data-driven.

The main contributions that this study makes is to bring an SA context to a growing field of research and implementation of LA. This study demonstrates the challenges that are unique to SA in the adoption of LA within an online learning perspective as well as highligting challenges that are similar to other global institutions.

LIST OF TABLES

TABLE OF CONTENTS

1 Introduction

1.1 Overview

This chapter introduces the research by giving the background of the study, stating the rationale for the study, presenting the research questions and objective of the study, and detailing the process that will be followed for the research.

1.2 Research Background

Online learning is growing in South Africa and this is evident in the number of universities offering online courses as well as private online learning providers offering tuition from primary school, through high school and beyond with most offering advanced courses in various subjects (Mtebe, 2015). Students' online activities generate an enormous amount of unused data with educational systems being increasingly engineered to capture and store data on users' interactions with a system (Sin & Muthu, 2015). Researchers and designers are continously exploring creative ways to improve learning systems to better support the varying needs of learners (Poitras, Lajoie, Doleck, & Jarrell, 2016). In addition, researchers have found that penetration in the use of learner data continues to be low and opportunities exist in adaptive learning for students (Sin & Muthu, 2015), as well as curriculum construction driven by data which benefits teachers and educators (Alom & Courtney, 2018).

Learning analytics (LA) has emerged as a powerful tool for addressing a range of educational challenges and issues, including concerns over institutional retention and continuous improvement of the student learning experience through personalised learning (Bakharia et al., 2016). LA is a growing multidisciplinary field with several definitions (Dawson, Joksimovic, Poquet, & Siemens, 2019). Some scholars define LA from the perspective of the student where generated data can be used as a tool to predict educational outcomes, while others define it as a means for educators to better understand student study behaviours in order to provide support and change student learning experiences (Viberg, Hatakka, Bälter, & Mavroudi, 2018). This paper adopts the definition of LA as the "*measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimising learning and the environments in which it occurs*" (Viberg et al., 2018, p.99).

There is limited literature on the use of LA in online learning at higher learning institutions (HEIs) and even less literature of its use at South African institutions (Prinsloo & Slade, 2017). It is important to understand the extent to which LA is being used or could be used in online learning at HEIs in order to realise its potential benefits and/or effectiveness. This study aims to address this gap by providing insights from educators, administrators and support staff on the level at which they are adopting LA to optimise the learning environment, both from a curriculum and student improvement perspective, and uncovering barriers that might exist in the use of LA. The growing number of published work in LA in the last decade is indicative of the importance and growth of LA as a research field (Gasevic, Dawson, Mirriahi, & Long, 2015). Theoretical frameworks are starting to emerge in the field of LA (Lester, Klein, Rangwala, & Johri, 2017) and for this reason, an integrated theoretical framework will be applied for this study adopting the theory of the Technology-Organisation-Environment framework (TOE) and the updated DeLone and McLean's IS success model.

1.2.1 Shift From Commercial to Educational Environment

Commercial enterprises have been at the forefront of innovations in data analytics and have collected large amounts of consumer data (Delen & Ram, 2018; Rubel & Jones, 2016). The era of Big Data has introduced tremendous opportunities for organisations to innovate, leading to great advances in analytics in business, often referred to as Business Analytics (BA) (Duan, Cao, & Edwards, 2018). The shift from intuition-led decision making to data or evidence-based decision making has propelled the growth of BA with some scholars highlighting that a data-oriented organisational culture is key to successful adoption and use of BA (Delen & Ram, 2018; Duan et al., 2018; Vidgen, Shaw, & Grant, 2017). BA has evolved in the past few years and continues to gain popularity (Delen & Ram, 2018). It is considered by many to be a major part of decision making in many organisations (Bayrak, 2015). As a key business driver, BA methods are being used in many different ways to empower decision makers in the business world with data that supports them to make strategic and operational decisions (Delen & Ram, 2018; Duan et al., 2018). BA is considered a game changer for organisations enabling efficiencies and effectiveness (Wamba et al., 2017). BA gives organisations the power to integrate multiple data sources and discover insights that help them develop new or improved products, respond to changing customer demands, react swiftly to changing market conditions and gain competitive advantage (Basole, 2014; Duan et al., 2018; Wamba et al., 2017). New technologies are enabling BA to provide descriptive, predictive and prescriptive analysis to aid decision makers (Bayrak, 2015; Duan et al., 2018; Holsapple, Lee-Post, & Pakath, 2014) with advances in data processing technologies playing a role (Delen & Ram, 2018). While BA introduces many opportunities for organisations, it has been met with many challenges in its successful implementation. Researchers have identified the following challenges for BA adoption; company culture, analytics talent, technology, leadership and return on investment (Delen & Ram, 2018; Vidgen et al., 2017).

Higher learning institutions are viewed as complex environments that are often slow to adopt new technologies, tend to lag behind other industries (Lester et al., 2017; Macfadyen, Dawson, Pardo, & Gaševic, 2014) and they have different cultural, social and economic perspectives in the adoption of new technologies (Tsai & Gasevic, 2017). Higher learning institutions have long since used assessments as a way to monitor student performance (Macfadyen et al., 2014). With the evolution of LA alongside the explosion of big data, institutions are moving more towards data-driven decision making when viewing the learning environment (Lester et al., 2017). A study by Tsai and Gasevic (2017) identified some challenges affecting the adoption of LA in higher learning institutions which include; shortage of leadership at institutions to drive adoption, unequal representation of students as benefactors of LA solutions, a gap in the relationship between pedagogy and learning behaviours, a shortage of required training for educators in the use and understanding of the LA tools and a lack of emperical evidence on the benefits of adopting LA. Lester et al. (2017) identified multiple barriers that affect the use of LA in higher learning institutions which include; "*a lack of interest or awareness, time, training, resources, incentives, institutional readiness, and institutional commitment*" (Lester et al., 2017, p. 32). Other studies have shown that while the goal of LA is to improve the learner outcomes it has also been found to also improve learner support and teaching (Viberg et al., 2018).

Similarities can exist between HEIs and commercial organisations in their adoption of analytics. In both business and education, analytics are used to derive insights in order to make more informed decisions (Bayrak, 2015; Duan et al., 2018; Lester et al., 2017). Decision makers in industry use predictive analytics in BA to identify future trends, and prescriptive analytics to determine the most optimal route to take in order to gain competitive advantage and meet constantly changing customer demands (Bayrak, 2015; Delen & Ram, 2018). Educators on the other hand are able to, through the use of LA, improve learning for at-risk students by predicting student outcomes and identifying intervention measures early, transforming teaching methods and data-driven experimentation of course structure (Rubel & Jones, 2016; Sin & Muthu, 2015).

Notable differences such as motivation behind the adoption of analytics and drivers for adoption can be seen between commercial industry and the educational industry. While commercial industry is motivated by improving their bottom line and increasing capabilities (Bayrak, 2015; Duan et al., 2018) educators' primary motivator is improving the learning environment through better understanding of the educational settings and students (Avella, Kebritchi, Nunn, & Kanai, 2016).

The educational world is predominantly deemed to be complex, however, many similarities can be drawn in the use of analytics when making comparison with commercial industries, as well as the challenges in adopting analytics. Adoption of analytics in learning presents a different manner in which data is analysed and interpreted. Lessons can be identified from commercial industry in terms of some of the pitfalls to avoid, at the same time opportunities exist in employing some of the more advanced BA tools and adopting them to fit the needs of educators with the goal of optimising the learning environment.

1.2.2 South African Context

The prevalence of information technology in education is driving researchers and governments to change their view of education in HEIs (Oyo & Kalema, 2014). Online learning is changing the face of both teaching and learning with different instructional methods being introduced as well as the use of new communication channels, such as discussion boards and chat rooms, used by both educators and students (Palvia et al., 2018). Africa is generally viewed, by the rest of the world, as lagging behind in terms of key indicators of the information society (Bornman, 2016). South Africa is regarded as being amongst the most developed countries in Africa with an evolved digital infrastructure (Bornman, 2016; Palvia et al., 2018). It is also ranked as one of the most unequal societies, yet it is striving to secure a place as a hub of academic excellence on the continent (Swartz, Ivancheva, Czerniewicz, & Morris, 2019). Some African countries such as Ghana have made e-learning a priority in their higher learning institutions as benefits of online learning are realised (Wright, Cillers, Van Niekerk, Seekoe, & International Association for Development of the Information, 2017). Access to internet outside the institutional domain remains a barrier in some African countries owing to network infrastructure, and in some cases, basic infrastructure such as electricity (Bornman, 2016; Wright et al., 2017).

Africa has become the world's fastest-growing market for internet and mobile communication; however, the continent has some of the world's lowest internet penetration rates (Kotoua, Ilkan, & Kilic, 2015). The challenge of access to internet is an issue in South Africa where broadband is not accessible to the whole population due to high broadband and internet costs, as well as deteriorating and insufficient infrastructure (Bornman, 2016).

Wright et al. (2017) identified some of the challenges of e-learning adoption in Africa to include a lack of technological infrastructure, curriculum and contextual issues as well as social factors linked to motivation and confidence in the use of technology. The brain drain in Africa has been highlighted as another challenge to adoption of e-learning as more experienced and knowledgeable educators explore better opportunities at institutions with better facilities to support both teaching and learning (Kotoua et al., 2015). The Ghanaian government is pioneering online learning through various initiatives and measures to incentivise both higher learning institutions and the private sector with the goal of making online learning a success in the country (Kotoua et al., 2015).

Government funding for public universities in South Africa has decreased over time making access to higher education financially unrealisable for most low income earning households (Swartz et al., 2019).

In addition, the diversity of the student body has increases substantially due to growing inequity in the secondary school system which has resulted in many more academically unprepared students (Scott, 2018). The diversity of the student body has put more emphasis on LA in the need for early detection of academically unprepared students to better support students and ensure their academic success through early interventions.

The current transition in higher learning institutions to provide affordable online learning using either blended learning or massive open online courses (MOOCs) presents opportunities for financially disadvantaged students in the form of access to learning without the heavy financial burden (Oyo & Kalema, 2014). A challenge, however, is access to internet at home and the effect it can have on a student's progress in online learning as they are unable to make full use of the resources available and explore related topics in this way expanding their knowledge (Bornman, 2016). This study will take a South African perspective focus when exploring the use of LA at HEI by examining institutions that deliver teaching and learning through different mediums namely distance learning, blended learning as well as purely online learning.

1.3 Problem Statement

Developments in technology and its widespread adoption have greatly influenced public acceptance of online learning to deliver education on a large scale (Gašević, Dawson, & Siemens, 2015). This has fostered debate on the future of education moving towards a more digital approach (Prinsloo & Slade, 2017). Student interactions with online systems leave digital breadcrumbs which are captured and stored (Gašević et al., 2015). LA relies on these digital traces for analysis and interpretation of student data in order to provide insights on student learning practices and behaviours (Rubel & Jones, 2016). In addition, educators and policy makers have an expectation for LA to provide insights that will help improve teaching, learning and decision making (Viberg et al., 2018).

The problem identified that this research aims to uncover from a South African context is the extent to which data is used in teaching and learning, and that there is a lack in understanding of the elements affecting learner success in online learning, and a gap in knowledge how data is used at institutions to inform decision making, support students and improve teaching and learning outcomes.

1.4 Research Questions and Objectives

The objective of the study is to investigate the current status of LA in online learning at higher education institutions (HEIs) in South Africa, with the aim of identifying barriers of use as well as future potential of the use of LA in informing decision making, predicting learner outcomes and improving the overall learning experience for students.

The overarching research question is to understand what is the current status and potential for LA in online learning at South Africa's HEI. This will be addressed by the following detailed questions:

1. What type of student data can be and is currently being collected?
2. Is the data used for LA and if so, to what extent?
3. Is LA used to inform decision making?
4. What opportunities exist, if any, for more advanced LA?

1.5 Overview of Research Methodology and Process

The goal of LA is to understand and optimise learning, key to this is having a view of how learning occurs, is supported and how student characteristics influence learning (Lester et al., 2017). The phenomenon this paper is looking at is that of LA and to answer questions of the extent of use of LA to aid in decision making, support student teaching and learning, and to improve the learning experience.

The research design is influenced by philosophy, methodology and methods of the study (Creswell, 2007) It is important that researchers understand their deeply held beliefs and views as these guide choices of theory and analysis of literature (Creswell, Hanson, Clark Plano, & Morales, 2007).This section will outline the philosophical consideration that will be applied to this study and the methodology to be used to conduct the research.

The objectivists view of knowledge is that it exists and through study and research it can be uncovered and that things are what they are (Padilla-Díaz, 2015). This study will follow an objectivist view as it aims to determine the use of LA in HEI without the researcher influencing the current state of the phenomenon. The positivist view is that authentic knowledge is scientific knowledge and the universe is an ordered system that can be investigated (Walliman, 2011). Positivists view a single reality that can be measured, and this study seeks to objectively determine the current use of LA (Walliman, 2011). This study therefore adopts a positivist stance.

Qualitative research captures purposes associated with meaning and interpretation and involves close personal contact that uses the researcher as the research instrument (Creswell, 2007). In qualitative research, the researcher aims to establish meaning of a phenomenon from observing and understanding the participants views (Creswell et al., 2007). This research adopted a qualitative approach to answer the research question. A purposive sampling approach was followed in identifying participants for the study (Padilla-Díaz, 2015). This approach was used to ensure that the targeted participants were using or intended to use LA in order to get a view of the extent of use and the reasons for having adopted LA. Educators, system managers and institutional management at various higher education institutions in South Africa were approached for participation in this study. Interviews were conducted and documentation at the institutions was reviewed to ensure the credibility and validity of the study. A total of 33 interviews were conducted at three public universities and one online learning provider with participants in varying designations within the institutions. Thematic analysis was used to analyse the data and identify common themes and patterns from the data. "*Thematic analysis is a method for identifying, analysing and reporting patterns(themes) within data."* (Braun & Clarke, 2006, p. 79).

1.6 Research Contribution

Research in LA has predominantly been done in countries outside of the African continent where advancements in technology and infrastructure have propelled the growth of LA. There is limited literature regarding use of LA in Africa, and specifically in South Africa. This study aims to make a contribution by presenting a South African view to the phenomenon.

Contribution from this study will add to the body of knowledge regarding the level of adoption of LA, an understanding of the drivers for use of LA, perceived benefits of use, challenges in adoption and extent to which LA are used to improve learning in South Africa. This study will provide participants with insight into the current LA adoption landscape in South Africa which may assist them in future LA decision-making. Understanding their institution as well as other institution's level of adoption will create opportunities of collaboration and learnings across the industry.

1.7 Structure of the Thesis

This thesis consists of five chapters and is structured based on the steps followed in the research process of: stating research objectives, literature review, research methodology illustrating the framework, analysis of findings from the collected data, discussion of findings and study conclusion as illustrated in Figure 1 below.

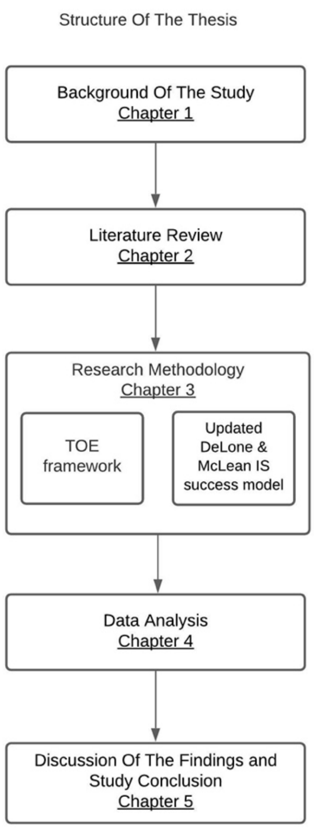

Figure 1: Structure of the Thesis

The thesis is presented according to the following layout:

Chapter 1: The first chapter gives an overview of the study motivation and background of the study. It presents the research questions and objectives for the study, outlines the research methodology used and the data research process and outlines the structure of the study.

Chapter 2: The second chapter reviews available literature in the topics of BA and LA in HEI and draws parallels between the phenomena. The following topics are covered in the review: background of BA and LA, tools used at HEIs, challenges of LA. The chapter also details the framework that will guide the study with justification of how the framework aligns with LA research in exploring the level of adoption of LA at HEI institutions.

Chapter 3: This chapter presents the methodology that was used for this study. The philosophical considerations underpinning the methodology are discussed in detail and a background of the philosophical stance of the study is given. In addition, the research design is discussed along with the data collection method and data analysis strategy followed in this study. The sampling strategy and unit of analysis are also discussed. The analysis methods to ensure validity and credibility of the study protocol are also discussed. Lastly, the ethical considerations and limitations relevant to the study are outlined.

Chapter 4: This chapter presents the case study results from the four institutions. A step-by-step explanation of how the collected data was analysed using thematic analysis in NVivo and the process followed for coding, identifying themes and subthemes linked to the study is provided.

Chapter 5: The final chapter discusses the findings of the study in relation to the literature. It examines how the findings answered the research questions, discusses implications of the study to both theory and practice and limitations of the study. It concludes by recommending topics to consider for future research.

2 Literature Review

2.1 Introduction

Learning Analytics (LA) is a growing field and advancements in technology are propelling this growth. Higher education institutions (HEIs) are embracing LA tools and techniques as they seek to improve retention rates, improve student performance and offer learning that is aligned with global technological advances. This chapter provides a narrative literature review that aims to give a comprehensive understanding of the LA field, the benefits and current challenges experienced within the field. The benefits and challenges include examples of HEIs that have adopted LA. This chapter begins with a brief background of the application of analytics in commercial enterprises. It highlights similarities and differences between the application of analytics in business and educational fields. The current uses of LA are explored with examples of existing tools highlighted. The chapter also studies online learning at South Africa's higher learning institutions and explores the impact the country's history has had on the adoption of technologies. Finally, the Technology-Organisation-Environment (TOE) framework integrated with the updated DeLone and McLean IS success model is proposed as a framework to use in conducting research into the use of LA at South Africa's institutions.

2.2 Introduction to Learning Analytics

The amount of data published has grown exponentially in the last few years driven largely by advancements in technologies, the emergence of applications and lowering costs of storage (Saggi & Jain, 2018). Technology advancements are changing how teaching happens and introducing new ways of tracking and storing student data (Avella et al., 2016). HEIs are striving to empower educators to make data-driven decisions to support teaching improvements and student achievement (Marsh & Farrell, 2015). The data infrastructure created by online learning and the availability of data collection technologies have, to a large extent, led to the rise of LA, along with the need to use data to improve the learning experience (Rubel & Jones, 2016).

The field of LA has grown over the years as institutions aim to provide flexibility in their offerings through online learning, and adopt new and robust data collection platforms (Avella et al., 2016). Concerns have also been raised in respect of challenges faced in the adoption and implementation of LA (Wamba et al., 2017). Growth in the field has raised questions regarding the effectiveness and usefulness of LA initiatives (Tsai & Gasevic, 2017).

2.3 Literature Review Approach

A narrative literature review was undertaken to gain an in-depth view of the body of knowledge available in the LA field. *"A narrative literature review critiques and summarises literature on a topic in order to generate new perspectives about the topic"* (Cronin, Ryan, & Coughlan, 2008; Torraco, 2016). The search in literature on LA began by using keywords such as "learning analytics", "analytics", "learning analytics in higher education" and "online learning analytics". Google scholar was the primary database used when searching for articles with other databases (Sage Journal Online, JSTOR and Taylor & Francis Online) used to complement the search. A simple matrix listing the databases and results was used to keep a record of keywords that yielded the most results. The inclusion/exclusion criteria illustrated in Table 1 below were used:

Include	Exclude
Full-text articles published in international conferences or workshop proceedings	Articles that do not present empirical data
Peer-reviewed journal publications	Articles that are not peer reviewed
Date from 2014 to present	Books

Table 1: Literature Review Inclusion/Exclusion Criteria

2.4 Background: Business Analytics (BA)

We live in a world where most of what we do, say or buy leaves a digital trace, and data is collected about every aspect of human life (Vidgen et al., 2017). The widespread adoption of mobile devices, applications, social media platforms, such as Twitter, and the proliferation of online shopping and other activities have led to the daily creation, collection and storage of massive amounts of data (Wamba, Akter, Edwards, Chopin, & Gnanzou, 2015). The scale of growth in data generation sources has inspired developments in various industries to better understand the economic and social benefits or challenges introduced by these new data sources (Xiang, Schwartz, Gerdes Jr, & Uysal, 2015). Easy access to computational software has enabled many researchers and organisations to exploit the opportunities presented by big data (Agarwal & Dhar, 2014; Vidgen et al., 2017). Researchers believe that the growing interest to gain new insights from data will enable societies to do things which may have seemed unthinkable or unimaginable before (Saggi & Jain, 2018). In addition, new questions and opportunities are being created as machines become smarter through the development of new and better algorithms (Vidgen et al., 2017). There is both excitement and concern around big data analytics and many organisations are seeking ways to harvest value from the data (Holsapple et al., 2014; Vidgen et al., 2017).

The use of analytical techniques to gain insights dates back to the 18^{th} century (Holsapple et al., 2014; Wang, Kung, & Byrd, 2018). One of the early examples being the 1854 Cholera Epidemic in London where data was used to convince officials of the cause of the epidemic (Wilder & Ozgur, 2015). Processes that involve examining, calculating and making inferences of results date as far back as the dawn of the computer age with efforts such as the first weather forecast (1950) examples of how heuristic methods were used to infer solutions for different types of problems (Holsapple et al., 2014; Wang et al., 2018). Since then, the use of computers to operate on data to aid decision making has been a central practice in business and research (Holsapple et al., 2014). Advances in analytical techniques to keep abreast with the rate at which data is constantly generated through the movement of economic and social transactions online have grown tremendously (Agarwal & Dhar, 2014).

Throughout history, businesses have continuously found innovative ways to remain competitive through the use of new technologies (Wilder & Ozgur, 2015). The emergence of the BA field in recent years has been driven by the need to leverage the massive amounts of data collected and stored, and the desire to draw actionable insights (Bayrak, 2015). In recent years many business owners have viewed investments in BA as a top priority, with most viewing it as both a source for competitive advantage and an asset to their business (Holsapple et al., 2014; Wilder & Ozgur, 2015).

BA is defined as "*the application of processes and techniques that transform raw data into meaningful information to improve decision making*" (Wilder & Ozgur, 2015, p. 180). It is also defined as "*the use of data to make sounder, more evidence-based business decisions*" (Seddon, Constantinidis, Tamm,

& Dod, 2017, p. 237) as a way to leverage value (Acito & Khatri, 2014). BA is concerned with decision making and providing actionable insights for sustainable value delivery, better organisational performance and improved competitive advantage (Power, Heavin, McDermott, & Daly, 2018; Vidgen et al., 2017; Wamba et al., 2015). Definitions of BA have a common thread about them, and that is the notion of fact-based decision making (Holsapple et al., 2014; Saggi & Jain, 2018).

The use of large volumes of data in applying statistical techniques and algorithms for decision making, and understanding business performance is a phenomenon that has long been studied (Acito & Khatri, 2014; Holsapple et al., 2014). What has changed in recent years is the breadth of available opportunities realised (Acito & Khatri, 2014). "*Modern day BA is rooted in the ongoing advances of systems to support decision making*" (Holsapple et al., 2014, p. 131). BA plays an important role in today's relentlessly and rapidly changing world where businesses are under immense pressure to make the right decisions to complex and not well understood problems (Holsapple et al., 2014; Saggi & Jain, 2018). BA methods are being used for a myriad of issues; these include prediction of customer behaviour, prediction of the likelihood of a medical condition and improvements in production and delivery of goods and services (Sharma, Mithas, & Kankanhalli, 2014; Vidgen et al., 2017).

Analytics practices and technologies can be applied in various disciplines to transform evidence into insights and decisions (Vidgen et al., 2017). Analytics research and advancements in commercial enterprises have been ahead of the industry (Krishnamoorthi & Mathew, 2018). Businesses are constantly innovating in the analytics field and deploying sophisticated tools the enable analysis of rich data sets, the identification of patterns within the data and more data-driven outcomes (Papamitsiou & Economides, 2014; Saggi & Jain, 2018).

A business culture is key to a successful analytics initiative and more organisations are changing their mindsets and investing in BA to be more competitive and to achieve their goals (Seddon et al., 2017). An organisation that values data-driven decision making is more likely to succeed in the adoption of BA as it allows for enhanced visibility of operations and better performance measurement-mechanisms (Wamba et al., 2015). Data used in BA originates from a variety of sources and is often in real-time (Riggins & Wamba, 2015). The challenge businesses are faced with is how to deal with the massive amounts of data and leverage value from it (Schoenherr & Speier-Pero, 2015). Another challenge is ensuring high-quality data to ensure correct interpretations and conclusions are drawn to facilitate decision making (Wamba et al., 2015). The education industry can leverage the tools used in commercial industries in their application of analytics and apply some of the lessons learnt from these implementations.

Businesses use analytics to predict consumer behaviour to improve products and services and to gain an advantage over competitors (Reyes, 2015). Descriptive analytics in businesses are used to understand business performance and identify the root causes of failures (Saggi & Jain, 2018). While the use of analytics in both the commercial and education industry is geared towards gaining insights, predicting future outcomes and making data-driven decisions, the drivers for the output of analytics, are different between the industries (Bayrak, 2015). Analytics in business is primarily used to gain an edge over competitors and increase profits (Cao, Duan, & Li, 2015), whereas, in education, the institutions seek to improve the learning experience for both educators and students (Ifenthaler, 2017).

2.5 Background: Learning Analytics

LA is defined as the "*measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimising learning and the environments in which it occurs*" (Lester et al., 2017; Siemens & Long, 2011, p. 34; Viberg et al., 2018, p. 99). It "*involves the* *gathering, analysing and reporting of data related to learners and their environments with the purpose of optimising the learning experience*" (Reyes, 2015, p. 75).

The changing educational landscape, the need for an educated workforce and increase in the cost of accessing higher education have led educators, policy makers and business people to explore and embrace emerging approaches to education (Veletsianos, 2016). Advancements in educational technology and the changes in administration of teaching and learning have prompted the occurrence of LA (Gašević et al., 2015; J. Zhang, Zhang, Jiang, Ordóñez de Pablos, & Sun, 2018). Adoption of technology at educational institutions has historically been slow and often lags behind other industries (Klein, Lester, Rangwala, & Johri, 2019). The rate at which technology is advancing necessitates the need for educational institutions to develop new theories and employ new teaching approaches to prepare students for the new world (Veletsianos, 2016).

Online and distance learning have become an important part of higher learning institutions with many students opting to study online due to the economic benefits it provides (R. S. Baker & Inventado, 2016). The adoption of online learning and the drive by institutions to integrate the various data collection systems have contributed to the emergence of LA as institutions seek to better support students and improve performance in education (Ferguson et al., 2016; Gašević et al., 2015). LA is a field that continues to develop as is evident in the growth of research in the field, an indication of the importance of LA as a field of research (Ferguson et al., 2016; Hwang, Spikol, & Li, 2018). Most of the research focus has been on LA tools, data models and prototypes with less research focused on the behavioural elements of how LA complements education, and the improvements needed from both educators and students in order for the tools to better support their everyday learning experience (Ferguson et al., 2016).

The developing field of LA has, through the collection of large amounts of data, presented opportunities for educators and students alike to gain insights into how people learn (R. S. Baker & Inventado, 2016; Wise & Shaffer, 2015; Wong, 2017). Student data that is collected and analysed is used to improve both learning and teaching through predicting learner performance, highlighting areas of improvements in courses thus advising educators, in effect discovering social and information connections (Ferguson et al., 2016; J. Zhang et al., 2018). One such opportunity is the ability to visualise engagement and activity in tools that provide early alerts (Anderson, 2016). Visualisation provides a representation of the analysis in an easily understandable manner to support decision-making (Reyes, 2015).

In addition, LA helps tailor educational opportunities to the learner's needs, creates an environment of constant feedback between educators and students and improves the learning experience for students (Avella et al., 2016; Wong, 2017). LA methods and analysis results are observed to affect strategy and assist educators and policy makers in HEIs to make more informed decisions (J. Zhang et al., 2018). Using LA, educators are able to identify at risk students and apply the right intervention measures to enhance student's levels of achievement and retention (Wong, 2017). The access that learning institutions have to multiple data platforms enables them to consolidate, analyse and interpret the data in a cost-effective manner (Tshabalala, Ndeya-Ndereya, & van der Merwe, 2014). Educators are able to use the data to understand students' utilisation of resources, how they interact with one another and with the course content, their learning behaviours and learning outcomes (Avella et al., 2016; Wong, 2017).

2.6 Technology and Learning in South Africa

Higher learning plays an important role in the development of a society and provides advanced skills that help a society to grow and become sustainable (Kaliisa & Picard, 2017). Using technology for education improves the country's skills needs and transforms the society (Bagarukayo & Kalema, 2015). Demand for an educated workforce has grown, so too has the cost for higher education as more students seek to acquire the necessary qualifications to advance in their careers (Deming, Goldin, Katz, & Yuchtman, 2015; Kruss, McGrath, Petersen, & Gastrow, 2015). HEIs are striving to address the changing needs of both learners and industry by providing flexible, cost effective, convenient and accessible learning experiences through online learning (Tshabalala et al., 2014). The complex history of South Africa creates challenges for most South African HEIs as they seek to create an inclusive and equal educational society that meets the changing needs of learners and is on par with global HEIs (Ng'ambi, Brown, Bozalek, Gachago, & Wood, 2016). This complex history also means that South Africa has a different growth path when compared to other countries in using technology to improve the learning process (Ng'ambi et al., 2016).

The educational landscape in South Africa has changed in the last 20 years mainly influenced by global trends, advances in technology and the national development plan (Ng'ambi et al., 2016). The new generation of students are encouraging institutions to embrace technology and incorporate it in the learning and teaching environment (Tshabalala et al., 2014). This need is however, a challenge with some students starting their higher education with insufficient basic computer literacy skills (C. Brown, Czerniewicz, & Noakes, 2016). The new generation of students views technology as part of (Tshabalala et al., 2014) their everyday lives and seeks to use technology as part of their learning, in and out of the classroom (Tshabalala et al., 2014). Technology adoption in education has evolved from technology being used purely as an assessment tool in the 1990s to it playing a pivotal role in the teaching and learning process through digital learnings, flexible learning, social media and professional development (Ng'ambi et al., 2016).

Online learning is "*essential for the improvement of learners' performance, engagement, motivation, flexibility and self-regulation*" (Bagarukayo & Kalema, 2015, p. 169). Online learning is growing in South Africa with the proliferation of smartphones and advances in social media adding to this growth (Kaliisa & Picard, 2017). Educators and learners are using social networking tools and social media, such as Facebook, for learning purposes (Ng'ambi et al., 2016). These tools are user-owned and not integrated with the student learning management system (LMS), thus offering flexibility and encouraging collaboration, communication and creativity (Kaliisa & Picard, 2017; Ng'ambi et al., 2016). Students have embraced various technologies and are using resources available to them to access educational information, share information and create social learning communities (C. Brown et al., 2016). Online learning encourages availability of information and social learning communities which lead to the sharing of that information thus fostering creativity and independence of work amongst students (Ng'ambi et al., 2016).

One of the challenges facing HEIs in South Africa in the use of online learning is inequality of access for students from low socioeconomic groups and continued disparity of both access and use (Ng'ambi et al., 2016). The shift of the digital divide moving from 'have' and 'have-not' to 'can' and 'can-not' is also observed in developed contexts with students who lack digital capital, at risk of being marginalised (C. Brown et al., 2016).

2.6.1 Use of Learning Analytics

LA constitutes an ecosystem that successfully collects, analyses, interprets and acts on machine generated data on an ongoing basis to improve learning (Papamitsiou & Economides, 2014). LA encompasses different tools and techniques to analyse large amounts of student data (Rubel & Jones, 2016). New technologies have allowed educators and institutions access to data introducing opportunities to collaborate in designing single platforms that allow for the sharing of multiple data sets

(Reyes, 2015). Student online activities leave digital traces that can be analysed to identify patterns and better understand student learning behaviours (Gašević et al., 2015; Reyes, 2015). Experts in online learning at American HEIs that predict LA will be widely used in the next few years as a means to better understand student learning patterns and as a way to improve student retention (Avella et al., 2016).

Procedures in LA focus on using data analysis to inform and empower students and educators in decision making, producing actionable knowledge that can be used to improve the teaching and learning environment (Papamitsiou & Economides, 2014; Pardo, Jovanovic, Dawson, Gašević, & Mirriahi, 2019). The objective of these procedures is to have data-led student interactions with educators, course content and other students within the educational social network (Avella et al., 2016; Dawson & Siemens, 2014). The availability of educational information online is encouraging students to engage with peers and educators from anywhere at any time (Dawson & Siemens, 2014). This has allowed for the holistic analysis of a students' behaviour by using LA to get a view of their online learning and behaviours and associating it with their overall performance (J.-H. Zhang, Zhang, Zou, & Huang, 2018). The online interactions of the student within a student community create a social learning network that connects the student with their peers (these can be outside of their institution) and educators (Goldie, 2016).

The use of data in education has increased the demand for evidence-based student learning and focused the need for more data-driven decision making to improve student outcomes and the learning process (Marsh & Farrell, 2015). Data-driven decision making is viewed as the major strategy to support educators and learners in improving student achievements and the success of the institution (Marsh & Farrell, 2015). Research shows that although educators have access to vast amounts of student data, they do not always know how to use the data to make changes that will lead to improvements (Datnow & Hubbard, 2016). The basic data-driven decision-making model assumes that an "*information system is available to support the decision process, internal and external factors not available through the information system are considered, and a course or courses of action are determined*" (Picciano, 2012, p. 11). The model suggests a process for analysing student data while considering external and internal factors in order to make informed decisions that improve student outcomes (Marsh & Farrell, 2015).

Informed and actionable feedback specified in an understandable manner has been identified as one of the key factors influencing academic achievement (Pardo et al., 2019). Predicting student learning success and providing proactive feedback are key elements of LA (Dawson, Gašević, Siemens, & Joksimovic, 2014). Course Signals is an analytics application developed by Purdue University that uses learner data collected from various institutional platforms to identify students exhibiting a risk of academic failure or dropout (Dawson & Siemens, 2014; Gašević et al., 2015; Scheffel, Drachsler, Stoyanov, & Specht, 2014). According to Gašević et al. (2015) LA tools are generally not developed from theoretically informed strategies and this impacts adoption of tools that aid effective instructional and intervention practices. In a case study conducted on the use of Course Signals, the researchers concluded that while the tool offered a simple way to prompt action, it lacked alignment with empirical research on effective instructional practice and did not provide sustainable and reproduceable insights into the learning process (Gašević et al., 2015).

In a study on the Khan Academy's visual analytics module named ALAS-KA, researchers concluded that the module empowered educators to make informed decisions on teaching through analysis of the individual student's overall behaviour as well as the class tendencies with the intention of improved course content and learning environment (Ruipérez-Valiente, Muñoz-Merino, Leony, & Kloos, 2015). A learner's awareness of their progress promotes self-reflection which leads to changes in learning behaviour that will improve their performance (Scheffel et al., 2014). In this way LA helps students develop crucial skills of reflection, awareness, collaboration and linking of ideas. These skills are however difficult to measure using LA tools (Scheffel et al., 2014).

Other examples of institutions that are embracing LA are The University of Phoenix which used and merged data from its various platforms to predict the likelihood of student failure and prioritise at-risk students (Rubel & Jones, 2016). The University of Texas has started building its own adaptive learning tools to create a personalised learning experience for its students (Rubel & Jones, 2016). Table 2 below shows a snapshot of some LA tools currently used by HEI.

Tool	Description	Role of analytics
Desire2Learn	Integrated platform that addresses challenges with engagement, retention and learning outcomes.	statistical inference prediction modelling recommendation
Knewton	Adaptive learning platform that uses visualisation to support student learning, encourage engagement and visualise engagement levels.	adaptation visualisation summary and description
Loop	Learning environment tool that integrates with learning management systems (LMS) to provide visualisation of student behaviours, online interactions and course structure to support teachers and educators.	visualisation
Open Essayist	Formative, developmental tool that provides automated reflective feedback to learners on draft essays	summary and description visualisation
OU Analyse	Predicts students at risk using demographic and virtual learning environment data	alerting summary and description visualisation prediction modelling recommendation
Student Success Plan (SSP)	Case management system to enhance and manage student support services with the aim of improving retention and learning outcomes	summary and description
Tribal's Student Insights	Integrated platform that identifies students at risk using predictive modelling techniques and presents information so that educators can apply relevant intervention measures	alerting summary and description visualisation prediction modelling

Tool	Description	Role of analytics
X-ray Analytics	Predictive modelling tool that uses past and present student performance to provides visualisation of at-risk students and alert educators.	visualisation summary and description alerting
Meerkat-ED	Analyses student online activities and provides visualisation of engagement and interactions	visualisation

Table 2 Examples of LA Tools (Ferguson et al., 2016; Reyes, 2015)

2.6.2 Learning Analytics Challenges

While LA presents opportunities for educators and students, education stakeholders need to be aware of the issues related to the use of LA in higher education (Avella et al., 2016). Researchers have highlighted that the results and predictors from LA tools do not provide guidance on the course of action to be taken to improve learning and teaching (Gašević et al., 2015). There is no overwhelming evidence from the research of the efficiencies and improvements on the learning process resulting from the adoption of LA, instead, the institutions indicate an expectation of future benefits (Ferguson et al., 2016). There is limited evidence of successful LA implementations (Bakharia et al., 2016; Ferguson et al., 2016). The nature of LA introduces concerns around the quality of the collected data and notes a risk that students may adjust their behaviours to trick the system, this will result in incorrect analysis and interpretations (Knight & Shum, 2017).

LA emphasises an in-depth understanding of data analysis and interpretation on how to use the information to employ actionable steps to optimise the learning process (Avella et al., 2016). Given the importance of providing valuable actionable feedback, researchers have identified challenges in the development of informative indicators of student learning progression that can be used by educators to provide real-time feedback (Pardo et al., 2019). Many educators are hesitant to trust the interpretation outputs of LA about learning and educational effects as valid, yet they hope to gain insights from analytics results (Scheffel et al., 2014). Educators have also highlighted that LA cannot provide an objective assessment of a student's state of knowledge (Scheffel et al., 2014).

Ethical and privacy concerns have been highlighted by both educators and students in the use of LA (Arnold & Sclater, 2017). Loss of autonomy has been raised as a privacy concern due to LA tools directing the course of action to be taken thus impeding students and educators decision-making abilities (Rubel & Jones, 2016). Concerns have also been raised on the extensive scope institutions have in their data-collection activities, raising questions on the use of the data and who stands to benefit from its collection (Arnold & Sclater, 2017; Rubel & Jones, 2016). Other existing challenges are in the design of LA tools to ensure they fulfil the intended objective without them being too rigid or restricting and imposing on the student (Knight & Shum, 2017; Knight, Shum, & Littleton, 2014). Educators have highlighted concerns that LA tools do not answer questions about the qualitative analysis of the educational setting with researchers recommending more involvement of educators in the design of LA tools (Scheffel et al., 2014).

2.7 Analytical Framework

The LA field is growing and a widely accepted theoretical foundation for LA is a multidisciplinary field that draws on various theories (Rogers, Gašević, & Dawson, 2016). Humans, through active interaction with the world, make fragments of the world objects of their activity (Kozulin, Ageyev, Gindis, & Miller, 2003). To produce new knowledge and design theory, researchers need contexts of other researchers' use of theory (Clemmensen, Kaptelinin, & Nardi, 2016).

There are many theories used in Information Systems (IS) research, the most used theories for technology adoption are technology acceptance model (TAM), theory of planned behaviour (TPB), unified theory of acceptance and use of technology (UTAUT), diffusion of innovation (DOI) and the Technology-Organisation-Environment (TOE) framework (Oliveira & Martins, 2011). TAM, TPB and UTAUT focus on individual adoption of technology instead of organisational adoption and could therefore not be used for this study (Oliveira & Martins, 2011). The DOI model was considered in the context of organisational use of LA, however, the TOE framework was observed to be more fitting for the study as it incorporated the barriers and opportunities for technological adoption from an environmental context.

2.7.1 Technology–Organisation–Environment Framework

The TOE framework, illustrated in Figure 2, represents one segment of the process of innovation with a focus on how the firm context influences the adoption and implementation of innovations (J. Baker, 2012). The TOE framework suggests that the implementation and use of an information system in an organisation is influenced by the technological, organisational, and environmental context (Depietro, Wiarda, & Fleischer, 1990). The technological context includes internal and external technologies relevant to the organisation (J. Baker, 2012; Oliveira & Martins, 2011). These are technologies that are already in use or technologies accessible to the organisation externally and may include processes or tools (J. Baker, 2012; Depietro et al., 1990). Technologies are seen to contribute to the organisational performance and researchers have employed diverse approaches and empirical methods to determine the level of improvements such investments deliver (Melville, Kraemer, & Gurbaxani, 2004). Organisational context refers to the characteristics and descriptive measures about the organisation such as managerial structures, culture, and size of the organisation (Baker, 2012; Oliveira & Martins, 2011). Various organisational characteristics such as growth, size and the control and management mechanisms when viewed together instead of in isolation impact the level of technology adoption (Dewan, Michael, & Min, 1998). Depietro et al. (1990) state that the firm's scope and size are important, and Dewan et al. (1998) found that firms with a large scope have a big request for information technology (IT) investment and that firm size is related to IT investment. Environmental context refers to the industry in which the organisation conducts its business and the regulatory environment that impacts the level of adoption (Depietro et al., 1990).

According to J. Baker (2012), the TOE framework has been used in different industries to explain adoption of innovations and the three elements have been shown to influence the way an organisation identifies a need for a new technology and the resultant adoption of it. Inclusion of technological, organisational and environmental factors make the TOE framework relevant when studying technology adoption, technology use and value creation from technology innovation (Gangwar, Date, & Ramaswamy, 2015). The environmental context presents both limitations and opportunities for innovation which influence the level of adoption within the organisation (Oliveira & Martins, 2011).

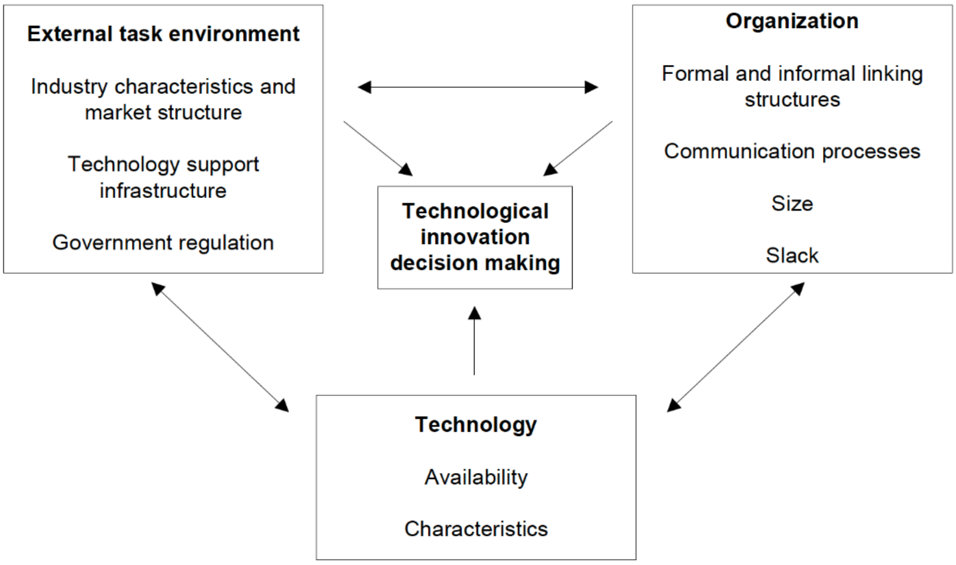

Figure 2: Technology-Organisation-Environment Framework (Depietro et al., 1990)

While the TOE theory has been used to investigate adoption of innovations, it does not provide a model explaining the organisational-adoption decisions but rather a taxonomy classifying adoption factors (Bose & Luo, 2011). TOE encourages the researcher to look at the broader context in which innovation occurs (Bose & Luo, 2011). The TOE framework is, therefore, appropriate for investigating LA adoption at HEIs in SA.

2.7.2 Updated DeLone and McLean IS Success Model

Measuring IS success is important in understanding the value derived from such investment and actions arising from management of the IS (DeLone & McLean, 2003). IS success evaluates the creation, distribution and use of information through technology (Petter, DeLone, & McLean, 2012). The complexity of measuring success of an information system, driven by the multi-dimensional nature of IS and its continued growth led to the conception of the DeLone and McLean IS success model (DeLone & Mclean, 2004). The original model proposed by DeLone and McLean in 1992 was based on Shannon and Weaver's communication theory and adapted by Mason to measure IS impacts (DeLone & Mclean, 2004; Mason, 1978; Shannon & Weaver, 1998). The model was constructed from an extensive review of IS literature to identify the factors that contribute to IS success (DeLone & McLean, 1992, 2003). It identified six inter-related dimensions as illustrated in Figure 3: system quality, information quality, system use, user satisfaction, individual impact and organisational impact. The model can be interpreted as: certain qualities in an information system will drive its usage and satisfaction impacting both individuals and the organisation (Chen, 2010). The model illustrates both the human and technological aspects of a system (Chen, 2010).

Ten years after the publication of the model, DeLone and McLean published an updated model illustrated in Figure 3, that made re-specifications in accordance with the changes observed in the role of IS over time (Chen, 2010; DeLone & McLean, 2003, 2004; Mohammadi, 2015; Urbach & Müller, 2012; Wu & Wang, 2006). The updated model added service quality and intention of use, following Seddon's suggestion, collapsed individual impact and organisational impact into a more parsimonious net benefits construct (DeLone & McLean, 2003; Seddon, 1997). According to Chen (2010), the IS success model *"provides a theoretical basis for empirical studies that aim to clarify the link between system use and its impact"*. The model can be evaluated from the point of view that a system can be assessed in terms of its information, system and service quality, which in turn affect the intention of use or use and user satisfaction, contributing to the net benefits derived from its use (DeLone & McLean, 2003; Mohammadi, 2015). The net benefits can be both positive and negative, which also affect user satisfaction and continued use of the system (DeLone & McLean, 2003; Urbach & Müller, 2012; Wu & Wang, 2006).

When evaluating the updated model, DeLone and McLean (2003) recommend that the quality dimensions be measured separately because each has an effect on use and user satisfaction. System use continues to be a dependent variable in the updated model and DeLone and McLean (2003) have highlighted that usage precedes user satisfaction in a process sense, however, positive experiences of use will yield greater user satisfaction in a causal sense (DeLone & McLean, 2003; Stefanovic, Marjanovic, Delić, Culibrk, & Lalic, 2016).

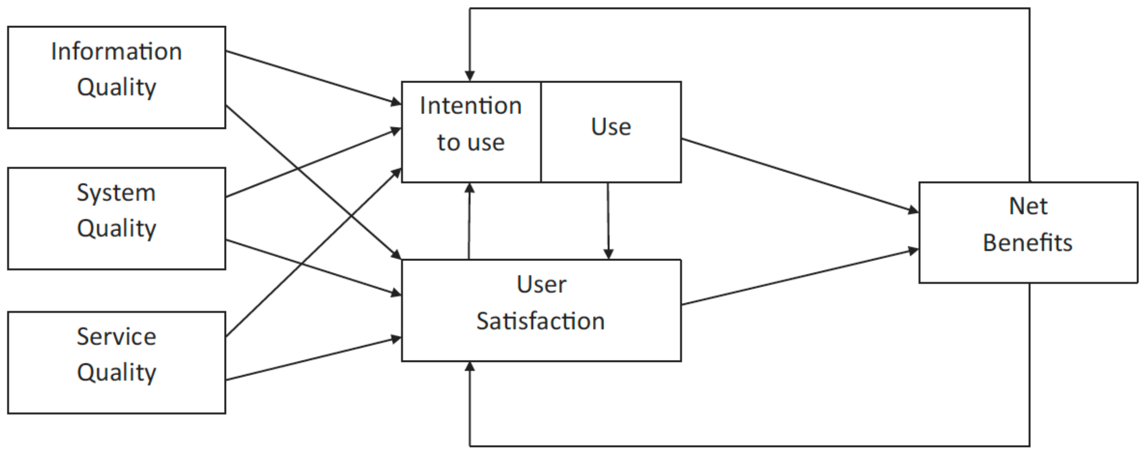

Figure 3: DeLone and McLean's Updated IS Success Model (DeLone & McLean, 2003)

The updated DeLone and McLean IS success model has been frequently adopted in evaluating success of e-learning systems (Mohammadi, 2015) which has encouraged that further studies be carried out using the model (DeLone & McLean, 2003). The information and system quality impact both use and user satisfaction (Halawi & McCarthy, 2006). It is therefore appropriate that the model be used in analysing the use of LA.

2.7.3 Research Model

LA adoption can be viewed from a university level and from a course level. At the university level LA takes a more holistic view of high-level aggregate course level cohorts, departmental performance and overall university performance. At the course level, focus is on course-level decisions, student progress and educational approaches where the course is viewed at a high level and at a lower individual student level. Applying the TOE framework and updated DeLone and McLean IS success model to these different levels helps understand initial adoption of LA as well as the level of use at the institutions and the various factors that either drive or hinder usage. Figure 4 illustrates the conceptual model to be used in this study integrating the TOE framework and the updated DeLone and McLean IS success model. This research aims to use the conceptual model as a framework to better understand how LA is currently used in online learning at South Africa's higher learning institutions.

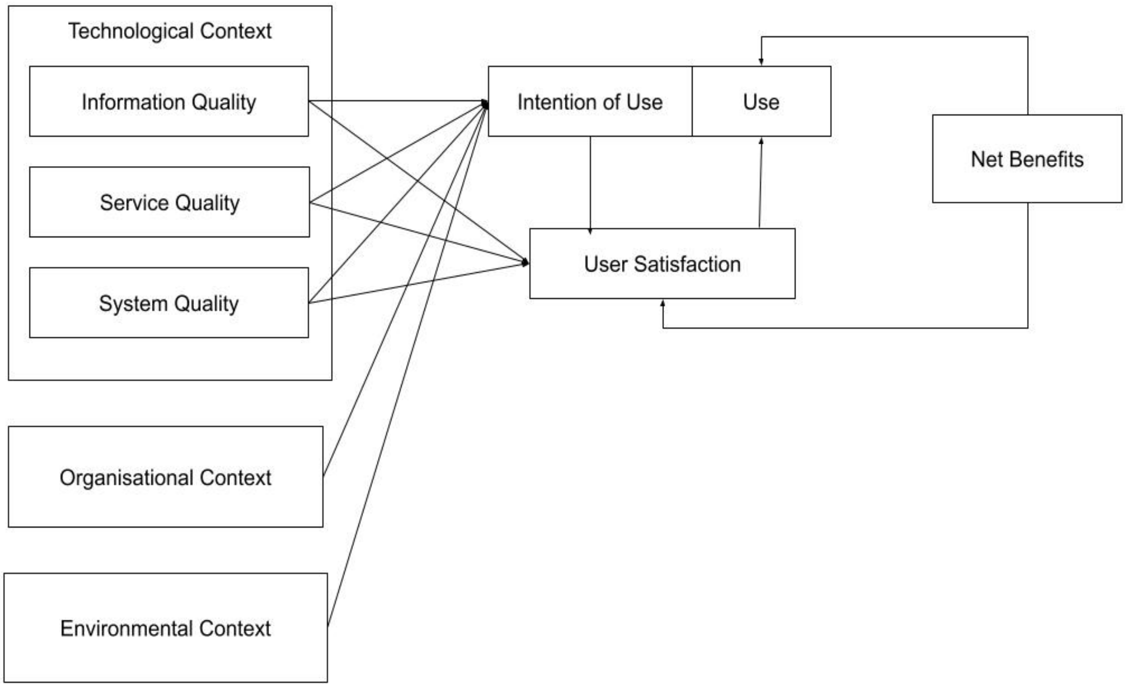

Figure 4: Learning Analytics Conceptual Model

2.8 Conclusion

This review sought to gain a more in-depth understanding of the state of use of LA in HEIs in South Africa. The review shows that research in use of LA in HEIs is slowly growing as more institutions are starting to incorporate and develop analytical tools and techniques. Literature in the field of LA has grown demonstrating a developing field. Most of the research, however, has not proposed research frameworks in the study of LA with researchers recommending conceptual models for studies. The analysis on the uses of LA shed light on the opportunities that are available when institutions incorporate LA techniques. It also emphasises the need for more informed decision making to gain better insights, an approach that is strongly supported in the commercial industry. The review also highlighted challenges in the adoption of LA. Some of the challenges, such as privacy, are in line with generally raised concerns within the commercial industry on the types of data collected as well as the use of that data. The review highlights a gap in literature of the use of LA tools and techniques in South Africa, with most of the cases under review related to institutions mainly in the United States of America (USA), Europe or Australia. The limited research in the South African context on the use of LA presents an opportunity for future research in this field. The complexities that define the country will potentially introduce different approaches in adoption of these new technologies as the country aims to be a global player in the education industry.

3 Research Design

3.1 Introduction

This chapter details the research strategy in the investigation on the use of learning analytics (LA) in online learning at South Africa's higher education institutions (HEIs). The research design illustrates the approach the researcher will follow in answering the research questions and meeting the research objectives. It is viewed as the plan for conducting the study (Saunders, Lewis, & Thornhill, 2016).

The Technology-Organisation-Environment (TOE) framework and the DeLone and McLean IS success model were integrated in designing the research instrument and defining the methodology to be followed to address the research question. The paper starts with a brief background of LA to draw attention to the need for the study. Different research philosophies and methods are discussed, and motivation is given for the research method to follow. A detailed view of the research strategy, data collection and analysis approach, timeframe and high-level timeline is defined. The chapter ends with a discussion on ethical considerations and the contribution that the study seeks to make to research.

3.2 Research Background

3.2.1 Learning Analytics Background

Every part of human life leaves a digital trace as the world becomes more connected and technology driven (Vidgen et al., 2017). Insights have become a new form of evidence owing to the growing advances in technology and improved capabilities of computer systems (Dede, 2016). Analysis of data from user interaction with information technology (IT) has attracted a lot of attention as a new way of understanding the learning process (Gašević et al., 2015). These advancements in technology and the need to gain insights have contributed to the growth of LA (Gašević et al., 2015). LA is an emerging field defined as "*measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimising learning and the environments in which it occurs*" (Siemens & Long, 2011, p. 34).

HEIs are aiming to use data to inform their decision making, having been inspired by emerging technologies (Klein et al., 2019) and educational technologies have created new opportunities to gain learner insights (Gašević et al., 2015). The digital traces left by students can be examined to discover knowledge about learning behaviour that improve the educational practice (Gašević et al., 2015). Educational institutions are under pressure to provide evidence of the effectiveness of learning interventions, improvements from adjustments to teaching practices and student-focused learning models (Klein et al., 2019). Researchers view the benefits of LA tools as their ability to inform different levels of stakeholders within the institution by creating environments that foster data-driven decision making (Klein et al., 2019).

LA is expected to provide technological capabilities to support students through interventions that will improve the learning process (Ifenthaler, 2017). Students benefit from these interventions through optimised learning pathways and regular feedback, while educators benefit through detailed monitoring capabilities of student progress on a student as well as at a class level (Ifenthaler, 2017). LA creates awareness of the learner's and educator's current learning situation which promotes informed data-driven decisions (Avella et al., 2016). It also helps instructional designers better understand performance of the designed material so adjustments can be made to better support learning and teaching (Avella et al., 2016; Ifenthaler, 2017). At an institutional level, LA helps provide a view on attrition rates thus helping institutions to devise measures to ensure students achieve the best outcomes and benefit from learning (Ifenthaler, 2017).

Viewing use and adoption from the African context, thought needs to be applied on the complex past of most African countries, and their development and adoption of technologies, to fully embrace LA in their educational institutions (Kotoua et al., 2015). The expansion of broadband and investments in infrastructure are helping expand technology access in many developing countries (Mtebe, 2015; Tshabalala et al., 2014). Online learning is changing both teaching and learning and encouraging new mediums and instructional methods to be introduced (Palvia et al., 2018). Africa has often been viewed as lagging behind in innovations and adoption of technologies, however, the improvements in infrastructure and investments in technologies are challenging that perception (Bornman, 2016; Mtebe, 2015). South Africa, while regarded as one of the growing economies on the continent, is also viewed as having an unequal society from an economic perspective (Bornman, 2016; Swartz et al., 2019). Improvements in infrastructure are encouraging online learning because of its appeal of providing flexibility and education at a lower cost (Tshabalala et al., 2014).

With all the benefits presented for using LA, barriers to LA adoption and use have slowed down growth which has resulted in institutions lagging behind other industries in their use of analytics (Klein et al., 2019). The success of the adoption of LA is driven by the institution's capacity, readiness and willingness to promote the use of the tools and encourage a data-driven decision-making culture (Klein et al., 2019). One of the questions asked by researchers stems around the readiness of institutions to adopt LA given the massive amounts of student data available, the growing number of tools to choose from, and the need for interactive data visualisation to support both lecturers and students (Ifenthaler, 2017). HEIs are aiming to create a culture where decisions are informed by data and are encouraging educators to embrace this changing landscape in order to keep abreast with the changes that are being introduced by technology (Marsh & Farrell, 2015).

3.2.2 Research Questions and Objective

The objective of the study is to investigate the current status of LA in online learning at HEIs in South Africa, to identify the drivers, barriers and potential use of LA in informing decision making, predicting learner outcomes and improving the overall learning experience for students.

The research aims to meet the objective by addressing the following detailed questions:

1. What type of student data can be and is currently being collected?
2. Is the data used for LA and, if so, to what extent?
3. Is LA used by administrators, educators and lecturers to inform decision making?
4. What opportunities and barriers exist for more advanced LA?

3.3 Integrated Theoretical Model

LA research has emphasised multiple areas when investigating adoption of LA in a HEI yet there is limited theory available when assessing LA adoption (Dawson et al., 2019). The TOE framework and the updated DeLone and McLean IS success model each provides a limited view when applied in isolation in understanding LA adoption (Prinsloo & Slade, 2017). Integrating the theories and models introduces a more holistic view in assessing the adoption of LA and continued use and incorporating these theories with the literature creates a good model that can be tested.

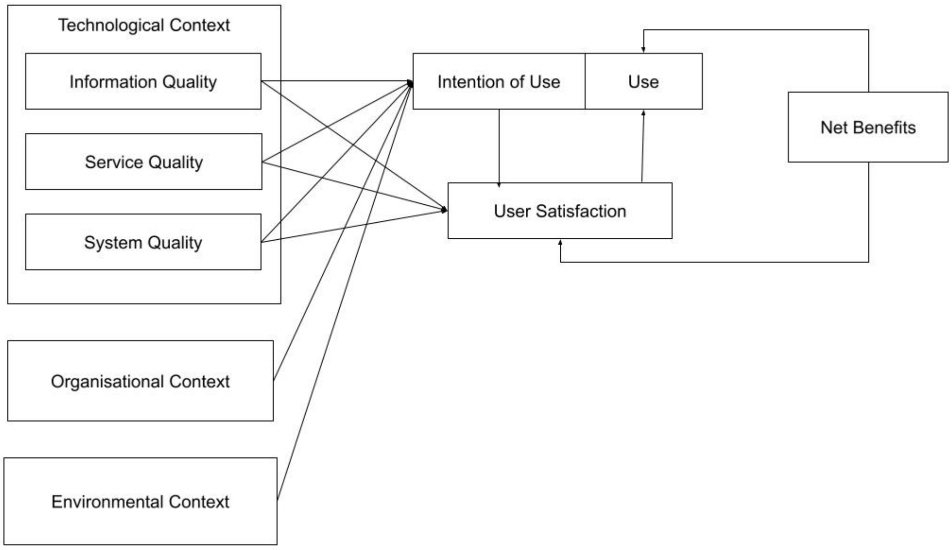

Figure 5: Learning Analytics Conceptual Model: Integrated TOE and Updated DeLone and McLean's IS Success Model

A conceptual model was developed, illustrated in Figure 5, to capture the salient factors in assessing LA adoption and use through integrating the TOE framework and the updated DeLone and McLean IS success model. Environmental context addresses social factors impacting the institution (C. Zhang & Dhaliwal, 2009, p. 255). From an LA-adoption perspective, it aims to establish factors relating to the use of LA by institutions in relation to other institutions and reasons behind adoption or non-adoption (Ferguson et al., 2016). It also takes into account regulatory factors that may impact adoption (Prinsloo, Slade, & Khalil, 2018). These factors are also relevant to the type of institution and how dependent it is on its peers. Organisational context looks at the organisation as a whole and the focus from an LA-adoption perspective would be from the HEI level. Using TOE's institutional and environmental context can potentially give insights into factors driving institutional adoption. The DeLone and McLean IS success model is a level lower and views LA use on an ongoing basis, such as at a course level. The DeLone and McLean model is concerned with continual use of a technology and provides a perspective of the benefits derived from the continual use of LA to support educational approaches and monitoring of student progress (Tsai & Gasevic, 2017). The TOE model is concerned with the initial use of an innovation and focuses on the initial use of LA without evaluating activities beyond initial adoption.

Incorporating the updated DeLone and McLean IS success model by integrating the technology factors; system quality, service quality and information quality, to TOE's technological context helps evaluate or understand use in relation to external technology factors as well as internal system-specific factors. The updated DeLone and McLean's IS success model also looks at the quality of the information produced by the system which potentially drives or influences whether or not the LA tools will be used. This gives a different perspective when analysing initial technology use. There is an overlap as the model highlights the crucial elements of the system-specific technology factors which TOE does not accentuate. In addition, TOE does not take into account benefits or barriers and impact on continued use and satisfaction which the DeLone and McLean model emphasises. The rest of the DeLone and McLean's model looks beyond technology and explores continual use and benefits derived. This is an important element that can relate to the effectiveness of LA by giving input into the value derived from its adoption, opportunities that it presents and also highlighting the negative benefits derived and challenges observed.

3.4 Research Philosophy

The research-design process starts with an assertion of the philosophical assumptions the researcher makes, along with their worldviews and sets of beliefs (Creswell et al., 2007). According to Saunders et al. (2016), research philosophy relates to "*the development of knowledge and the nature of that knowledge"*. A paradigm represents the researcher's worldview and basic beliefs about the nature of social reality (Shanks, 2002). Enquiry paradigms define the scope of legitimate enquiry and what falls within the limits of the enquiry (Guba & Lincoln, 1994). The choice of philosophical assumptions has implications for designing and conducting the study (Creswell, 2007). The basic beliefs that define enquiry paradigm can be summarised in the ontology, epistemology and methodology (Bhattacherjee, 2012).

3.4.1 Ontology

Ontology indicates the stance towards the nature of reality and its characteristics (Creswell, 2007). The two main assumptions in research are subjectivism and objectivism. The subjectivist perspective is that social actors create the social phenomenon through their opinions and activities (Saunders et al., 2016). In subjectivism, the researcher is the instrument and findings are created as the study progresses (Guba & Lincoln, 1994). Subjectivism is based on the concept that knowledge exists with the knower, reality is a product of the mind and to know and understand it, the researcher needs to experience it, therefore, knowledge is socially constructed (Saunders et al., 2016).

Objectivism assumes that social phenomenon is independent of social actors (Creswell, 2007) and it adopts scientific methods that are replicable (Saunders et al., 2016). Objectivism holds that knowledge about the way things are exists outside of social actors (Guba & Lincoln, 1994). The ontological lens chosen by the researcher guides the conclusions derived from the research, this is evident as different researchers conducting the same study, using the same methods but taking different ontological stances can come to different conclusions (Aliyu, Bello, Kasim, & Martin, 2014). This study assumed an objectivist stance because it aimed to study the extent to which LA was used at HEIs without the researchers' influence. All of these concepts are considered to be fairly objective.

3.4.2 Epistemology

"*Epistemology is the theory of knowledge, especially about its validation and the methods used*" (Walliman, 2011, p. 16). Epistemology refers to assumptions about what constitutes acceptable knowledge in a discipline and the world (Bernard, 2017). The three main approaches are interpretivism, critical theory and positivism. The choice of paradigm for researchers depends on the paradigmatic considerations of the phenomenon and the best way to conduct the study (Bhattacherjee, 2012).

The positivist view states that authentic knowledge is scientific knowledge and the universe is an ordered system that can be investigated (Walliman, 2011). Truth and reality are free and independent of the viewer and observer (Aliyu et al., 2014). Positivist work entails an observable social reality where outcomes can be generalised, similar to those produced in natural science enquiry (Amaratunga, Baldry, Sarshar, & Newton, 2002). The enquiry is concerned with facts not impressions and the observable phenomenon can lead to credible outcomes (Saunders et al., 2016). Such studies serve to test a theory in an attempt to increase predictive understanding of the study (Orlikowski & Baroudi, 1991). Positivist research uses theories to guide the data-collection strategy (Saunders et al., 2016).

Interpretivists view the world as a creation of the mind, what we see is what the mind has conjured (Walliman, 2011). Interpretivist paradigms often seek experiences, understandings and perceptions of individuals in their role as social actors (Thanh & Thanh, 2015). It holds that people create meaning through their interactions with the world (Orlikowski & Baroudi, 1991). An interpretivist research is embedded in the belief that there is no universal truth or a single world view and often holds subjective and biased impression about the world (Aliyu et al., 2014). Researchers seek to understand the phenomenon by immersing themselves in the setting of the inquiry and accessing the meanings participants give to the phenomenon (Orlikowski & Baroudi, 1991). This dictates that the researcher must adopt an empathetic stance (Saunders et al., 2016). The point of research using this paradigm is to gain a deeper understanding of phenomenon which can be used to inform other settings (Bhattacherjee, 2012). In interpretive research, reality is interpreted through sense-making and the goal is to construct theory instead of testing a theory (Bhattacherjee, 2012).

Objectivity was key for studying this phenomenon and it was also important to have a scientific method to ensure repeatability of the study (Walliman, 2011). Positivists conducting interviews aim to discover the reality of the phenomenon as it is (Aliyu et al., 2014). This study followed the positivist paradigm as this would help maintain an objective stance while investigating the phenomenon.

3.5 Research Purpose and Approach

Research can follow two main approaches: inductive or deductive (Bhattacherjee, 2012). Inductive research aims at building or constructing a theory where theoretical inferences are made from the observed data (Morgan, 2013). With deductive research, existing theories are rigorously tested to gain new insights and improve, refine or extend the theory (Saunders et al., 2016). A deductive approach was followed in designing the study protocol using the Learning Analytics Conceptual Model. While an inductive approach was followed in the initial data analysis, iterations of the analysis indicated a more deductive approach as the themes emerging were aligned to the conceptual model.

Research purpose can be descriptive, exploratory and explanatory (Bhattacherjee, 2012). Exploratory research is often conducted in a new area of enquiry or to seek new insights and assess a phenomenon in a new light (Bhattacherjee, 2012). Descriptive research seeks to portray an accurate view of the phenomenon through observations, guided by a scientific method and detailed documentation (Bhattacherjee, 2012; Walliman, 2011). Explanatory studies establish causal relationships by seeking explanations of observed phenomenon (Bhattacherjee, 2012). The descriptive approach of observing a phenomenon suited the study and was, therefore, followed.

3.6 Research Method

Development of a discipline is achieved through research; a means of developing a unique body of knowledge (Amaratunga et al., 2002). Research is about advancing knowledge and finding out about things that are not known (Walliman, 2011). The body of knowledge of the IS discipline can be advanced only through the use of appropriate methodologies and the rigorous application of methods of research (Amaratunga et al., 2002). Research in IS draws on a variety of disciplines, such as social science, anthropology, computer science and economics, which are then applied to the IS context (Peffers, Tuunanen, Rothenberger, & Chatterjee, 2007). Research methods are the techniques used to conduct research; they represent tools to use to collect and analyse data so conclusions can be drawn (Amaratunga et al., 2002).They also provide a good measure for determining validity of the study and credibility for the new knowledge (Walliman, 2011). A choice of method refers to the approach for collecting and analysing data (Bhattacherjee, 2012). Qualitative work has often been associated with interpretivist philosophy (Creswell, 2007; Shanks, 2002), however, it can also be used as an approach in positivism (Lin, 1998). The main difference between qualitative and quantitative methods is in the way they are used and the purpose (Morgan, 2013). The choice of method should, therefore, be guided by the questions asked in the enquiry and the conclusions drawn (Lin, 1998). The combination of diverse epistemologies and methodologies must be applicable to the situation being investigated (Aliyu et al., 2014).

Qualitative research emphasis is on generating meaning and understanding through rich description and explores the depth and complexity of a phenomenon (Morgan, 2013; Saunders et al., 2016). It attempts to describe people in natural situations and involves prolonged periods of contact with the life situation (Ormston, Spencer, Barnard, & Snape, 2014). Qualitative research is concerned with understanding the meaning social actors ascribe to a social phenomenon (Amaratunga et al., 2002). Data is typically collected from participants and analysed inductively to discover general themes (Creswell, 2007). The qualitative method has been accused of bias and lacking objectivity, and some researchers have questioned its generalisability (Aliyu et al., 2014; Amaratunga et al., 2002). The debate draws on the opinion that the manner in which the research is carried out can remove the risk of bias and subjectivity (Aliyu et al., 2014; Lin, 1998). The researcher's systematic approach to data analysis can ensure rigour (Bhattacherjee, 2012).

Quantitative research focuses on using numbers to represent opinions or concepts and is characterised with the view that human behaviour can be explained by social facts (Morgan, 2013). Quantitative research tests objective theories by examining relationships between variables to identify causal relationships (Amaratunga et al., 2002). Enquiry in quantitative research is external to the phenomenon studied (Amaratunga et al., 2002; Morgan, 2013).

3.6.1 Study Method

This study adopted a qualitative approach to answer the research question. The study aimed to investigate the phenomenon at different levels; at the individual level focusing on monitoring student progress, at the course level focusing on course design and, lastly, at the institutional level focusing broadly on educational management and strategy (Sclater, Peasgood, & Mullan, 2016).

3.7 Research Strategy

The research situation determines the research strategy which should have its own characteristics for collecting and analysing data (Amaratunga et al., 2002). The choice of a strategy should be based on its suitability to help the researcher answer the research question and meet the research objectives (Creswell, 2003; Saunders et al., 2016). The case study approach has been followed in many IS studies (Shanks, 2002). A case study is an empirical analysis of a phenomenon within a natural setting over an extended period (Benbasat, Goldstein, & Mead, 1987; Bhattacherjee, 2012). It is also a formal technique that uses scientific methods to describe a phenomenon where focus is on the richness and depth of understanding the phenomenon (Bhattacherjee, 2012). Although case research investigates a predefined phenomenon, it does not control or manipulate the variables (Benbasat et al., 1987; Shanks, 2002). Case studies award the researcher to adjust the instrument to better capture the phenomenon of interest once data collection has started (Bhattacherjee, 2012) and typically combines techniques such as interviews and documentation analysis (Saunders et al., 2016). Case study research adopts qualitative or quantitative methods, and can be used to provide descriptions of a phenomenon, test or develop a theory (Benbasat et al., 1987; Shanks, 2002). It can also be used to build a theory, applying an interpretivist approach, or to test a theory, applying a positivist approach (Bhattacherjee, 2012; Shanks, 2002). In positivist case research, enquiry is designed and evaluated according to the natural science model of research (Shanks, 2002).

The case research strategy was followed to gain insights into how participants are using LA. Multiple cases were considered to establish whether the findings are common across the various cases, determine the extent of generalisability and reliability through inter-case comparison (Paré, 2004; Saunders et al., 2016).

3.7.1 Timeframe

Time horizon is determined by the research question and the main time horizons are cross-sectional and longitudinal (Saunders et al., 2016). Longitudinal studies refer to data collected over a long period of time where the same respondents can be surveyed at different times (Bhattacherjee, 2012; Walliman, 2011). Cross-sectional design analyses a particular phenomenon at a particular time (Saunders et al., 2016). A cross-sectional design provides a snapshot of reality at a point and does not cater for future changes in the area of analysis. The cross-sectional timeframe was chosen as the strategy to collect data for this study.

3.7.2 Sample

Sampling is the method of selecting a unit of analysis that is representative of the population so that outcomes of the study can be related back to the population and generalisations can be made (Bhattacherjee, 2012). Sampling methods provide tools and techniques to assist researchers to focus their data collection by considering data from the representative sample (Marshall, 1996). There are different approaches used for sampling a population, this study adopted the purposeful sampling method. Purposeful sampling is a non-probability technique based on the principle of selecting information rich cases which are knowledgeable about the focus of the study (Hennink, Kaiser, & Marconi, 2016; Patton, 2005). LA is a relatively new area and the cases that were selected for the study were leading HEI and commercial online learning providers that were considered to have the most value-add from LA. These are top rated institutions as well as one institution with the largest student body in South Africa.

In non-probability sampling, the appropriate sample size is viewed as one that answers the research question (Marshall, 1996; Saunders et al., 2016). In qualitative research, there is no agreed number of respondents that deem a study credible, however, a suggested 25 to 30 respondents is suggested to be sufficient (Saunders et al., 2016). The point of saturation is a point where no new insights or themes emerge from continued data collection and all categories have been identified and explored (Boddy, 2016; Hennink et al., 2016). Following these principles, the point of saturation method guided the number of interviews conducted.

3.7.3 Unit of Analysis

Various institutions make use of a learning management systems (LMS) for capturing student learning data (Mtebe, 2015). The institutions use different systems for student interaction. All these leave digital traces which can provide insights into student behaviour and give a holistic view of the student. Students do not form the target audience for this study because of the focus of the enquiry. The focus will be towards those interested in the data collected from a student's digital footprint with the institution with stakeholders with varying levels of interest.

Specific courses or modules that have a high number of enrolment (more than 250 students), blended courses as well as the main Massive Open Online Courses (MOOCs) formed the basis of the study. The profile of institutions that participated in the study were public universities as well as an online learning provider. Below are the profiles of participants that were targeted for the study based on their goals for using LA:

- Education administrators: determine which data is included in the data collection and make decisions on whether a tool is used or not
- Educators and lecturers: interested in insights that monitor student performance and success of a course. They make decisions on adjusting course content and intervention strategies for at risk students.
- Management at Massive Open Online Courses (MOOCs): insights on what can be achieved with learning analytics, decisions on the type of data collected and analysed and understand how they are using the data.

It is common practice to conduct a pilot study with a small number of people before data collection can commence (Walliman, 2011). A subset of the sample population formed part of the pilot study to examine the comprehensibility and validity instrument.

3.7.4 Data Collection Method

Collecting data using different types of methods produces a wider scope of coverage and presents a fuller picture of the phenomenon (Paré, 2004). Key to data collection is ensuring access to participants in order to conduct the study (Creswell, 2007). The following were conducted: voice recorded structured interviews were carried out with participants and publicly accessible documentation was analysed. The duration of the interviews was 60 minutes.

Interviews are purposeful discussions that help researchers gather valid and reliable data relevant to the study (Saunders et al., 2016). Interviews are the main data collection method for case studies with their primary goal of eliciting participant's experiences and views in their own terms (Paré, 2004). Interviews serve as a way to objectively collect data by gathering information from the interviewee about a reality that exists independently from them (Saunders et al., 2016). The duration of each interview was one hour. The first few interviews were conducted at the institution's premises with subsequent interviews conducted online using different online conferencing tools. The decision to conduct interviews online was due to the lockdown imposed in South Africa due to the COVID-19 pandemic where travel and in-person engagements were restricted. Documents were analysed to determine the extent at which LA are used, the challenges faced in its use and establish the various ways it is used, if at all. Prior to conducting the interview for data collection, the researcher clearly outlined the type of data that would be gathered. This ensured proper use of the time with participants and good coverage of the research question (Benbasat et al., 1987). The research instrument used was created based on existing literature and the conceptual model illustrated in Figure 5.

3.7.5 Validity and Reliability

The goal of reliability in positivist case research is to minimise errors and bias in a study (Paré, 2004). Reliability is generally achieved when a different researcher can conduct the same study in the same manner and arrive at the same conclusions (Paré, 2004). Recording how the study is conducted provides detail to allow replicability and therefore test for reliability.

Validity gives the research a high level of confidence that the method chosen was appropriate and useful in meeting the research objectives (Straub, Boudreau, & Gefen, 2004). Concerns of construct validity can be addressed by using multiple sources of information because multiple sources provide multiple perceptions to clarify meaning (Paré, 2004). This method of confirming the reliability, credibility and authenticity of a study is referred to as triangulation and also serves as a way to eliminate bias (Saunders et al., 2016).

3.8 Data Analysis

In any research, the researcher must provide a detailed account of the approach followed in analysing the collected data to establish credibility of the study (Darke, Shanks, & Broadbent, 1998). In addition, the researcher must present a well-argued and summarised view of the enquiry (Darke et al., 1998). Analysing text and other forms of collected data in qualitative research can present challenges (Creswell, 2007; Darke et al., 1998). Data analysis in qualitative research involves preparing the data, identifying themes and patterns in the data through a process of coding, and representing the output in the form of graphs, tables or discussion (Creswell, 2007). The aim of data analysis is to summarise in a concise and truthful way, what has been learned in the study (Gillham, 2000)

This study aimed to investigate adoption of LA at different HEIs in SA and was conducted by interviewing participants and reviewing publicly accessible documentation at the institutions. The conceptual model guided analysis of the collected data. Soon after conducting the audio recorded interviews, the recordings were transcribed using the computer software NVivo. Each interview was allocated a unique identifier to allow for easy recognition of the participant interviewed, topics covered and to identify common statements (Saunders et al., 2016). Transcribing is reproducing the audio recording verbatim into words and noting important aspects, such as the tone in which things were said, so they can be linked to contextual information (Saunders et al., 2016). Transcribing is a time-consuming process (Saunders et al., 2016) and was best done soon after the interview while the interview was well remembered.

The transcribed content was reviewed again, once all the interviews had been transcribed, to check if any substantive statements had been missed and if nonessential statements were highlighted.

This thesis adopted the thematic analysis method to analyse the data. The process followed a hermeneutic cycle for identifying codes that were categorised into the main themes of the study. It is important to note that while the analysis was inductive, the research protocol was designed according to a theoretical framework that framed the analysis and matched what the research aimed to uncover (Braun & Clarke, 2006). It is important to clearly indicate the theoretical lens used when using thematic analysis and this study was guided by the conceptual model illustrated in Figure 5.

The process itself involved constant review of the content, highlighting statements, and identifying categories. At the end of this exercise, common categories were combined and others those that were incorrectly combined were either removed or split. The identified categories and substantive statements were cross checked against the transcribed content to check that they were classified correctly. The categories, once confirmed, were captured and substantive statements linked to a category. The analysed documentation followed the same process and was coded into categories in the same way as the interviews.

3.9 Access and Ethics

In every research, the researcher has to consider how they will gain access to the data they need and how they will explain to participants why the data is needed (Saunders et al., 2016). Gaining access to a source to obtain data pertinent for the research is a key element of any study (Saunders et al., 2016).

Ethics are concerned with the researchers behaviour (Saunders et al., 2016). Ethical dilemmas and concerns are part of doing research (Guillemin & Gillam, 2004) and are important aspects for the success of the study (Bhattacherjee, 2012; Saunders et al., 2016). Research ethics questions whether a researcher upholds the expected moral code and are honest and responsible when conducting their research (Saunders et al., 2016). Abiding by ethical standards ensures that the study is free from manipulation to advance a specific agenda (Bhattacherjee, 2012). Ethics in qualitative research is an area that has been explored since the 1960s (Guillemin & Gillam, 2004). Guillemin and Gillam (2004) highlight two types of ethics in qualitative research; procedural ethics where approval is sought from an ethics committee to conduct a study involving human participants, and ethics in practice which refers to everyday issues that arise during research. It is important to consider ethical factors such as informed consent, ethics committee approval to conduct the study, confidentiality and handling of sensitive information.

Prior to data collection, ethics approval to embark on the study was acquired from the University of Cape Town's ethics committee. This involved submission of the relevant ethics approval information pack which included ethics application form, research proposal, research design including the instrument, consent form and supervisors' signature. In addition, approval to conduct the study at the targeted HEIs was obtained prior to commencing data collection to get permission to conduct the study and engage with the participants. This was important due to the sensitivity associated with obtaining strategic information and other institutional information. Before each interview, consent was obtained from the participant to confirm their willingness to participate in the study and to highlight that participation was voluntary and they could withdraw should they wish to do so. Participants were given a background of the study, so they understood why the study was being conducted and any risks that existed regarding the study. Participant contribution to the study remained anonymous throughout the data collection, analysis and discussion process.

3.10 Conclusion

This chapter outlines the intended approach and strategy to be followed for the research. The research question and research model to be followed link with the research instrument to illustrate the process of answering the research questions. The case study method of investigation will be followed applying interviews and documentation analysis. The paper also details how the data will be collected and analysed to ensure objectivity, credibility and reliability.

4 Data Analysis and Findings

4.1 Introduction

This chapter presents the results of the data analysis for the study. The thematic analysis approach was followed in analysing the collected data and this chapter discusses the results of the analysis performed and how this answers the research questions. The research aimed to investigate the current status of learning analytics (LA) in online learning at higher education institutions (HEIs) in South Africa (SA) with the objective of identifying the drivers, barriers and potential use of LA in informing decision making, predicting learner outcomes and improving the overall learning and teaching experience for learners.

This chapter is structured into the following sections; Section 4.2 details the process followed in analysing the data, section 4.3 provides a description of the case studies, sections 4.4 to 4.9 present the thematic analysis findings of the study, and section 4.10 concludes the chapter.

4.2 Data Analysis Process

Thematic analysis was used to analyse the data and identify common themes and patterns from the data. "*Thematic analysis is a method for identifying, analysing and reporting patterns (themes) within data"* (Braun & Clarke, 2006, p. 79). It is important to understand how data was analysed in order to determine validity and reliability of the research and to compare findings with other similar research (Braun & Clarke, 2006).

The coding of the interview transcripts and documents was undertaken with a specific objective: to establish whether or not LA was being used at SA's HEIs. The analysis process began with a review of the transcribed interviews for each case and identifying items of interest in the data. The research questions and objective served as a lens for analysing the data and guiding the surfacing of the themes.

The hermeneutic cycle was followed for the data-analysis process in identifying the codes. This entailed reading through the text, reflecting on the text through conducting detailed analysis and gaining understanding and interpreting the text as a result. This followed a back-and-forth recursive approach and required a few iterations before the final codes were identified. The coding process began with extracting text from the interviews where items of interest and patterns of meaning were identified and using open coding to get an idea of the patterns that could be grouped and coded. Through each iteration of the process, more codes emerged while others were combined with codes that had similar meanings. It was important to keep an objective view of the data.

The next step in the process involved grouping the codes into overarching categories of potential themes. Similar codes were clustered into categories and the data was again reviewed to identify potential themes. Part of this process was starting to consider the relationships between the themes and identifying the overall story emerging from them.

The potential themes were then reviewed to establish whether the themes had a significant organising concept, whether there was sufficient meaning in the data to support the theme, whether the theme was related to the extracted text as well as questioning the quality of the themes. A critical view of the data was needed to establish the relevance of the identified themes and, through this process, following some iterations, related themes started to fall into place.

The themes started to take shape and some of the themes could be classified into the Technology-Organisation-Environment (TOE) framework and the updated DeLone and McLean IS success model. The use of LA and impact of LA category proved to be complex as they incorporated a considerable number of various concepts and were coded as unique themes.

After multiple iterations of reviewing the data and the potential themes, five main themes were defined and classified. While the generation of the themes were guided by the research questions, the process followed remained inductive. From this point on the researcher could analyse each theme and subsequent subthemes.

4.3 Description of the Case Studies

This section summarises the profiles of the four HEIs interviewed to set the context. The HEI will be referred to by the codes illustrated in Table 3 and referred to as institutions as opposed to universities because InstA is not a university. Table 3: Participant Institution Profile

4.3.1 Organisational Profiles

Case	Code	Type of Institution
Case 1	InstA	Private online learning provider
Case 2	InstB	Public university
Case 3	InstC	Public university
Case 4	InstD	Public university

Table 3: Participant Institution Profile

Case 1: InstA is a private online teaching and learning provider partnering with institutions across many countries with several hundreds of employees. Their goal is to make education accessible to a wide range of students at different points in their lives and careers through partnering with the leading universities. As a purely online organisation, data forms an essential part of its day-to-day operations. The organisation prides itself on being data driven in its decision making and general operations and thus seeks to use data in improving the programmes that it offers as well as enhancing the end-to-end student experience. As an online provider its focus is on using LA to understand challenges experienced by students at any point of their studies spanning from enrolment through to completion and are deliberate about continuously improving their offering and ensuring that they are meeting student learning objectives.

Case 2: InstB is a public university and one of the largest HEIs in SA. The university has a global reach and supports students from various countries. It provides both vocational and academic programmes and has received international accreditation on its qualifications. As one of the leading research universities on the African continent, it seeks to harness technology as it grows into a digital future. In line with this goal, the institution seeks to use the extensive data available to it to make data-driven decisions, support students, improve teaching and create an environment of growth for both learners and academics. The university has invested extensively in data-related projects and continues to use the lessons from these investments to improve.

Case 3: InstC is a public research university and one of the oldest universities in SA. It is one of Africa's top teaching and research universities and is globally recognised and ranked highly alongside other reputable international universities in the university world rankings. As a leading research and learning university, the use of technology and technological tools has become widespread within the university and so has the amount of data available within the university. The university has observed challenges resulting from the inequalities within the country and has sought to play its part in creating sustainable solutions and enhancing its responsiveness to developmental challenges facing the country. As part of this goal, the university has sought a more holistic approach that incorporates data-driven decision making thus emphasising the importance of having access to and using data that is collected through the various student touchpoints.

Case 4: InstD is a public research university organised into nine faculties. It is one of the leading teaching and learning universities and has global accreditation and recognition for teaching and research. The university is one of a few that has incorporated a mixed teaching model of both online and face-to-face instruction in its programmes and has successfully applied this model for more than two decades. This hybrid model and the need to incorporate data in improving their offering has led the university to be purposeful in its goal of being a data-driven university. The tools that the university has been using in its teaching and learning approach has set it apart from its peers and often seen as exemplary in the use of LA as well as academic analytics to bring about change and growth.

4.3.2 Profile of Interviewees

Table 4 illustrates the profiles of the interviewees who took part in the study and the total number of interviewees. It also categorises participation by role to demonstrate the diversity of interviewees with the aim of getting multiple viewpoints on the state of LA in the sampled institutions. It was important to have a diverse representation of roles in the study as people's perceptions in different roles and their experiences have an impact on the level of adoption, if any, of LA and their views on the use of analytics. Having a balanced view across the roles (e.g., executive vs course convenor) helps identify differences and commonalities in views of the phenomenon.

Role	Description	Institutional Code	Interviewee Code	Number of Interviewees
Coordinator: Learning Technologies	Manages the portfolio of teams supporting teaching and learning processes including online learning environments and other software within the institution.	InstC	CI	1
Director/Head of Teaching and Learning	Reports to the executive and different stakeholders on the use of LA within the institution and also serves as the LA practitioner within the institution.	InstA InstB InstD	VL JU JC	3
Deputy Dean	Supports the dean of department and provides academic leadership within the faculty.	InstC	DC	1

Role	Description	Institutional Code	Interviewee Code	Number of Interviewees
Head of Department (HOD)	Leads, manages and develops the department to ensure it achieves the highest possible standards of excellence in all its activities.	InstA InstC	QE EM, HD, HH	4
Course Convenor	Sets the aims and learning outcomes for the course and ensures delivery and assessment of the course.	InstB InstC InstD	AG, BB LC, AH, CS, HI LE, FL, IU	9
Lecturer	Delivers the aims and learning outcomes of the course.	InstC	AC, HL, TS	3
Institutional Researcher	Supports academics and management with information and statistics at a student, course and faculty level. This is a support function that has standard reporting and also deals with ad hoc requirements.	InstB InstC	PJ, TR PS	3
Engineering Manager	Manages the various technology systems used within the institution to ensure optimum availability and implements changes and improvements.	InstA InstC	MK CH	2
Information Services	Collects, analyses and reports on data collected at a course level.	InstA InstB	CR DJ	3
Student Success Practitioner	Supports academic staff in their development and promotes teaching and learning best practices.	InstB InstD	EO, SL KT, NI	4
Enrolments	Manages enrolments within the institution.	InstA	ED	1
Total number of interviewees interviewed				**33**

Table 4: Profile of Study Interviewees

Document analysis was used in combination with the interviews as a form of triangulation. The two methods were used as a means to seek convergence and corroboration. Table 5 provides a list of the analysed documents and the institutions from which the documents were obtained.

Document	Description
InstB Annual Report, 2019, 254 pages	A report published annually detailing the institution's achievements in the year, highlighting challenges and detailing current and planned initiatives. The 2020 report was analysed.
InstC Data Analytics for Student Success Proposal, 2019, 7 pages	Proposal for an institution-wide strategic project aiming to place data-informed approaches to student success at the heart of all levels of the teaching and learning leading to improved success for all students.
InstC Online Education Policy, 2017, 8 pages	The online institution's policy for various levels of the student cohort published in 2017.
InstC Teaching and Learning Report, 2018, 222 pages	Annual report that examines, by faculty, the overall teaching and learning environment in terms of student enrolments and outcomes, academic staffing and the outcomes of quality assurance reviews compiled during the most recent academic year. The 2018 report was analysed.
InstD Education Innovation Review, 2018, 70 pages	Annual department report that highlights achievement in the year and lays out planned projects. The 2018 report was analysed.

Table 5: Analysed Documents

The next section of the analysis presents the detailed data analysis of the study. The sections are structured in the following order: technological factors, organisational factors, environmental factors, use of LA and impact of LA along with the relevant subthemes. The analysis is supported by quotes from the interviewees supporting each theme. To maintain the confidentiality of all interviewees, aliases have been used instead of their real initials.

4.4 Technological Factors

The technological factors are the Information Technology (IT) related factors influencing the use or lack of use of LA within the institution. These were based on the components of the systems, physical architecture in place as well as the data and integration points within the institution. The theme focuses on the type of data available, the quality of the data and accessibility of the data. This category provided insights into the challenges and barriers that IT systems can have on the use of LA and how the systems can also become a driver for use. It also looks at the tools that are available at institutions in the adoption of LA and the extent to which the tools are perceived to be usable and helpful in achieving participant goals. Table 6 summarises the factors uncovered related to the technology theme and illustrates the incorporation of the updated DeLone and McLean IS success model's dimensions of information quality and system quality. In addition, Table 6 provides a view of where the sub-theme was mentioned at the four institutions.

Subtheme	Description	InstA	InstB	InstC	InstD
Analysis of collected data	Various student interactions are collected and stored for analysis. In order to add value, the collected data has to be understood, analysed and interpreted so that insights and meaning can be derived from it.	✓	✓	✓	✓
Information quality	The quality of the information collected encourages trust and drives usage. Information quality incorporates the accuracy and consistency of the data captured and inspires reliability of the insights the data presents.	✓		✓	✓
System quality	Reliability of the tools used within the institution, the level of security, efficiency and availability of the tools to ensure that both educators and students are able to complete their tasks with minimal interruptions. The reliability of the systems also impacts the extent to which data can be or is collected.	✓			✓
Use of Analytics Tools	Extent to which the analytics tools can be accessed and used and the level of complexity of the tools.	✓	✓	✓	✓

Table 6: Technological Factors

4.4.1 Analysis of Collected Data

All institutions in this study collect many different student interactions, given that every student interaction with the institution's online systems presents an LA opportunity. The institutions collected a wide range of student data which included student registration data, data on students physically accessing the institution, students accessing the learning management system (LMS) and students utilising the library resources. Any interaction that the student does where they either log in or use their student identification to access a resource or facility leaves a digital footprint. The vast amount of captured student data required analysis and correct interpretation for the data to add value. In addition, different stakeholders within the institution sought different meanings from the data based on their needs which influenced how they used the data.

> *"The way the data is derived and described has got certain meanings … what is included and what's excluded … then you have a data dictionary to provide that kind of explanation but even data dictionaries have the tendency to be very complex in their description … therefore, you need to have interventions to explain what they are looking at. Unfortunately, that is a challenge." JC*

An element of interpreting the student data is understanding how the data is modelled and knowing what insights are being observed from the data. When student data is incorrectly interpreted, this can have negative consequences on the student. A challenge experienced at InstC and InstD centred on having a common understanding of the data models in the data systems and alignment and consistency in the classifications used. This has impacted the extent to which the data has been translated and meaningful insights observed thereby impacting users of the data. A level of familiarity with the data models has ensured ease of extracting meaning from the data. This level of familiarity is established through regular interactions with the data which can be challenging for those engaging with the data on an ad hoc basis.

> *"For those of us who work with the data a lot … critically understand the academic structure at InstC … don't really struggle, but often … we're providing data to people who are not frequent users of it, and that is a barrier explaining all the codes and so on to them." PS*

In addition, it is critical to have checks in place where the output of a report can be compared to other measures or validated by looking at the raw data to qualify the insights presented in a report. A key challenge at InstC has been that most educators are not expert users of the data systems and lack the expertise or capacity to validate the outputs of the reports they use. This presents the risk of working off incorrect information and taking steps that do not address the problem at hand.

> *"If you've got nothing to check it against, you can actually be working with incorrect data, definitely, so that is a problem. That, of course, is one of the problems of having a number of users who maybe aren't expert users, writing reports and possibly generating or pulling reports that aren't working properly, they've got nothing to check it against, so they just assume it's working. And that's a real problem." PS*

LA measures data about learner behaviours and the majority of the data is collected from students' online activities. While institutions have a big online presence and most student activities happen on these platforms, it is equally important to note that learning can still occur outside of the online systems where students opt to learn in an 'offline' mode and only engage with the systems once they have completed a task or have a better understanding of the material. Students may choose to download their material to give themselves space to learn, while others may request that the institution provide the material in a printed format and only engage with the LMS or other institutional systems when required. These offline interactions, which don't leave an electronic footprint, have been observed as a major challenge in understanding student learning behaviours and in understanding the effectiveness of the teaching as many of them are unrecorded. It has also presented opportunities for educators to identify creative ways of looking beyond just the online activities in better monitoring students and understanding their engagement.

> *"And with the data that we get from the online course, it's a lot more reliable because it's like logs type of data. However, with it being an online course, there's nothing stopping someone downloading content and working on it in their own time .... So, you need to account for things like that. So, we've got a lot of rules in place to try and account for that, but you can't always account for all of those nuances." QE*

> *"We are working with a lot of gaps in the data … we don't know how many hours the student is spending on their studies … and unless they do come into the LMS we sit with fairly big gaps in our system. So, the pedagogy really shapes the analytics as well." TR*

> *"The student activity is outside the LMS so time on task is lost for us in that but … hopefully in the near future we will get that information but at this stage some data sources we've got access to and some data sources we don't have … access to." JC*

4.4.2 Information Quality

The quality of the information or data collected affects the degree to which it can be used and trusted. Various elements can affect the quality of the information and these span from the technology systems in place to human errors at the point of data capture to incomplete data being captured. The technological systems at the institutions must have measures in place to ensure a high level of quality in the data. In essence, the quality of the data collected determines the credibility of the insights derived from it.

> *"It relies on the whole thing of garbage in garbage out, so we must make sure that when the data is collected … majority of sources have got a high percent of accuracy because it is compulsory for these sources to be captured." JC*

The data-collection mechanism used has a big impact on the level of trust placed on the data. There is a high level of trust in data collected via the LMS at the institutions because of the online nature of the data which has capabilities, such as providing an audit history of the interactions with the data creating a low likelihood of information manipulation. Educators at the various institutions found the LMS data to be more trustworthy than data captured on Microsoft Excel spreadsheets where information can be changed without requiring any authorisation or view of the history of the changes. They also found the LMS data trustworthy and more accurate when compared to data captured on other systems.

> *"With regard to more access statistics than usage statistics, there would be no reason for me to doubt the statistics, and also because we use a rather expensive learning management system … I can trust its statistics." LE*

> *"Well, it's changing because well, now in an online environment so everything is there. So, this is ideal situation where I will trust the data 200% because everybody's online everybody's using it." KT*

As part of improving the quality of the information at the institution, one of the departments at InstC ran a project over time to try and get a better understanding of their students and their learning behaviours. In its first year of the project, they experienced immense challenges with the data and found lots of gaps. In the second year of the project, they saw improvements in the quality of the data, an indicator that the institution had already started taking steps to improve the quality of their information across the faculties.

> *"The first year … of the project there was a big challenge getting the data, they didn't really have the [warehouse] data set up properly yet so the data came from different sources, had to be joined together they still found lots of missing variables, lots of missing data, challenges of international students not being comparable to local students in terms of background, errors in the data so the same school captured in a number of different spellings and so on. The second year we ran it and in fact it was smoother. So, I do think we're getting better at collecting data." CS*

Data accuracy, completeness, consistency and ensuring that all student interactions with the institution are recorded are important elements of having a high level of data quality. Data accuracy and completeness refer to whether the information that is stored is correct to ensure the quality of the data when accessed. At InstA, they have implemented procedures in their data warehousing systems to check that the data stored aligns with the data from the source systems. This, in turn, qualifies the validity of the various reports and dashboards that they have in place that are used for LA. These procedures, however, are complex and the institution has highlighted that these are areas with which they struggle.

> *"We are moving towards building out a proper data warehouse design so that just makes it a bit easier to track different metrics and dimensions and making sure that they then match sources properly." MK*

Data inaccuracies have been an area of concern at InstB where data is either duplicated, captured incorrectly on the system or in a different field impacting the data interpretation and analysis. Another issue at the institution is related to the accuracy of the data at the time of access. This is impacted by how data is transferred to the data warehouse and the timing of the transfer from the different source systems. Timings of the data movement can result in inaccuracies when using the data as the data may not be up to date.

> *"My problem is am I getting the right data … I can get the data but I'm not sure it is right, and it actually happens to me that sometimes I get data which I ask my colleagues to look at from the different point of view, from different systems they say … here is a discrepancy." PJ*

Similarly, at InstC data accuracy has been a major pain point where information gaps have been identified with certain information missing, incorrectly captured or inconsistencies with the fields data is captured against. Another challenge identified at the institution is in the definition of certain classifications and how that information is obtained from students. If the definition of the classification is incorrect, the data obtained within that classification will also be inaccurate.

> *"I trust the integrity of the academic data – I'm not so sure about the admissions data because I keep getting different figures and I think different people calculate different things." DC*

Another challenge observed at InstB and InstC is getting students to provide data that the institution needs for legislative, operational or support purposes. When students do not provide this information or partially provide it, this creates gaps and inaccuracies when analysing and interpreting it.

> *"Another big challenge with our data is that we have to do a lot of data by equity in terms of the SA legislation, but more and more students are refusing to declare their status, so we have to infer it which is dangerous." DC*

> *"From a data perspective, yes, there's challenges in terms of the quality of data and sometimes we rely on students to provide us with information when they register." DJ*

The inconsistency of capturing student performance information on the student information system (SIS) has also been highlighted as a key challenge at InstC. This is closely related to the institution's culture where faculties function independently of each other and there are no enforced rules on how performance information should be captured and the timings of having that information captured. This is a challenge within the institution and affects the accuracy and quality of the data when observing student performance holistically.

> *"So, a lot of departments don't put all their tests and scores up into SIS so, when you look at something, the data is not there … and if you're not capturing the mark onto the [InstC] system where you've missed out on that too and if they're making decisions based on that, then the information is not available." AC*

A constraint at InstB is ensuring that the data is up to date, and this is impacted by data movements between systems and the frequency at which that happens. These delays in the data becoming available have a big impact on decision making and could also result in misinterpretation of the information.

> *"Then is the issue, for example, of the data movement. I cannot, unfortunately, get the data from the source. I have to wait until the data moves to the warehouse, and this sometimes takes time and sometimes this is just a partial movement …. So, I see that the data is not always very much clean … because of this complexity of the system." PJ*

Having accurate data captured and having data captured timeously are important in ensuring the usability of that data. Consistency of the information is key in having a high quality of data. Having the same naming conventions and definitions for fields within the system ensures that the same rules are applied when the data is captured and when input fields are defined. Where data is captured manually, the same rules apply to ensure consistency across the board.

The majority of the data captured at InstA is done directly on their systems, however, there are departments that manually capture data, and this has been identified as a challenge in ensuring data consistency. The institution is, however, aiming to move away from manual data capture to limit errors and inconsistencies in the data.

> *"Generally, the student data is very good because that's all recorded online but, for example, there are certain departments that sometimes still capture data in Excel spreadsheets or in Google Sheets, and often courses would have different spellings or it's just the consistency that's not necessarily there. So, we are trying to move as far away as possible from manual data collection, or manual inputs of data and rather try and pull it from the source systems where changes are once off and easy and it's clean." CR*

Data inconsistencies at InstB are a major challenge and impact decision making. The institution has disparate systems where data is captured, and all these are fed into the institution's warehouse. The challenge is where data is manually captured and inconsistencies are observed when extracting reports from the data, that the information either doesn't exist or partially exists. The major challenge is that data formats are not consistent, and errors are not rectified when the information is captured. Another cause of data inconsistencies at the institution is a lack of communication when new data fields are added by the areas capturing the data. This creates problems downstream where the data is mined and interpreted and the areas responsible for extracting the data and reporting on it do not have any knowledge of the definitions of the new fields, their purpose and reasons behind their creation.

> *"So, in terms of collecting the data, one of the major challenges we had as a department that processes it, we don't capture any of the data ourselves …. And you'll find that they will make a change to the system that they don't log anywhere, and then suddenly your analyses are off." TR*

4.4.3 System Quality

The quality of the technological systems in place at an institution can be a barrier or driver in enabling the effectiveness of LA. The interviewees viewed the quality of the technology systems as the reliability of the tools they use in the institution, the level of security of the tools, efficiency and availability of the tools to ensure that both educators and students can complete their tasks with minimal interruptions.

InstA is primarily online and has to ensure that their systems are available 24/7 to meet the needs of both students and educators who are dispersed around the globe. The institution has built a robust system that is able to support a large number of students accessing resources at the same time with minimal interruptions. In contrast, system availability has been identified as an issue at InstC where students are sometimes unable to access the LMS which impacts them meeting deadlines or submitting their work.

> *"Something like [an LMS] that's down every now and then when the capacity gets too high or when there's maintenance … I've actually had to shift submission dates because of it." AH*

System capacity is also an issue at InstB where the large volumes of submissions by students can sometimes result in system failures, highlighting a problem in the performance of some of the systems.

> *"You are talking about a university that says it wants to be online but every time when the due date for submission of assignments is, then the online system crashes. So, that cannot be something that can come back to the module, but it's a university-wide challenge" BB*

Common across the various institutions is the type of data that is collected with interviewees all mentioning registration, financial, and LMS data to name a few. While all this information is common, there are differences in how it is stored and accessed and the institution's ability to link the various data inputs to a specific student. The system architecture plays an important role in ensuring the quality of the systems. At InstA, the key to ensuring the success of their online offering is creating an infrastructure that ensures a seamless flow of data across the institution and ensuring accuracy, usability and accessibility. At InstB, the data is stored in various fragmented systems in different databases, but there isn't a seamless integration between all the data sources. In addition, the institution has legacy systems that have been modified over time without improving the system, thus causing even more challenges. The system infrastructure at the institution is a major obstacle because of the lack of integration between the various systems. InstD, like InstB, has a dispersed system infrastructure, however, the data between the various systems is integrated and links exist to ensure that the data flows seamlessly. This, however, is not a perfect system and there are situations where the data is not easily accessible or there are system failures.

> *"There's no central warehouse, so the data are stored in different databases that often don't talk to one another … if I understand the field of data management in higher education a bit then I think there's two options to either have a central warehouse where all the data goes, often a black hole or there is different data sources that are seamlessly integrated and talk to one another and InstB have the second part but they don't talk to one another. The data are not in the same format, not same quality, not the same governance, not the same." EO*

InstC has a data warehouse that stores various types of information and feeds different systems used across the institution. Within the institution, different types of information are stored to meet specific needs and, while the information exists, it can be a challenge to draw the information because, while there is a central data warehouse, it is made up of different data marts which have different owners who define data according to their own requirements. The institution has invested heavily in working towards integrating these various data marts to enable users of the data to get a holistic view of the student cohort. With most educators within the institution overloaded with teaching responsibilities, most do not have capacity to gain knowledge on how to retrieve a picture of their students from the warehouse and, therefore, the capability to draw on different sources of data to provide a meaningful picture is not a priority. Educators are left to draw out information from the different systems (e.g., student information system) and correlate information across these systems to get the view they need, and this approach has its limitations because not all the data is available. One of the challenges relates to the institution's culture where different faculties function independently of one another, therefore, resulting in a lack of alignment in how data is captured and stored and alignment of the systems across the institution.

"At InstC there is a system called [name of system] which is a data warehouse basically and lots of other systems feed into that …. So, if you look at how the university works as a whole, like different people in the university use different bits of information for different purposes and it's very difficult to find, like, a complete list … it is composed of … separate individual databases collections of data and those all come from different functional areas or business owners in the university." CI

4.4.4 Use of Analytics Tools

Accessibility and the extent to which the analytics tools within the institution are easy to use drive their usability and their use. Interviewees at InstB referred to dashboards that they use both on the LMS and looking broader on other systems. Interviewees found that the user interface to access data, looking at specific variables to answer certain questions or to get a bigger picture for novice users, proved challenging. However, regular users who have acquainted themselves with the rules and system variables find it easy to use. The complexity of the dashboard discouraged those unfamiliar with it to use it.

"There's training that we provide. If you only extract statistics once in a year, you don't need to be on it constantly then you might not find it as easy … What happens then is that they send a request to our department, and we conduct the extract on their part." DJ

The user interface of the warehousing system at InstC was built so that novice users of the system can easily find their way around it and draw the data they need. A recent upgrade to promote this objective has, according to interviewees, made the system more complex with an interface that is not user-friendly equating to a low desire to use the system. The challenge is that the system requires extensive system knowledge to understand the data structures and rules on retrieving specific types of information. Some reports have been created by the specialists of the system, but the challenge comes in where an understanding of a specific use case looking at specific variables is needed. These challenges have demotivated most users from enquiring the data themselves and led most to refer to administrators of the system to provide them with the data or reports they need.

"The hope was that we would actually have a suite of reports that people could run so just putting in the parameters, as you've described. It's actually, because of the way the warehouse is structured, all the data stamps, or the date stamps shall I say. The date-stamp nature of the records actually makes it very difficult to write reports." PS

"If you have a new question that you want to ask the data, there is absolutely no way to find out whether a report exists or a version of a report exists, that will enable you … to answer the question of the data that you're asking." EM

At both InstB and InstC users have found workarounds to meet their needs by retrieving data from different sources without having to use the complicated warehousing tools. They have identified data sources and downloaded data into tools such as Microsoft Excel, married it all together and analysed it.

"The data that I use will be stuff which I'll go and dredge out. So, I'll go and figure out all of the student records, and I'll pull them into a spreadsheet and then I'll fiddle with it there. You know, so, whatever that means, you know, rank them to figure out who's doing what, create pivot tables." AG

The LMS data has been extensively used by interviewees across the various institutions and they have found the dashboards particularly helpful in giving a view of their student cohort. Educators were able to look at student engagement in the course, view student activity and link various activities to better understand a student's performance.

4.5 Organisational Factors

Organisational considerations will be explored and the impact these have on adopting LA will be considered. This theme explores non-technological aspects and relates factors such as the culture of the institution, influencers and or issues that may either drive the use of LA or hinder it. These tend to differ between organisations and are likely to be unique as each institution has its own identity and way of doing things. The manner in which an institution is run or operates gives valuable insight into its attitudes towards LA and how it has been adopted, if at all. It also speaks to the issues that drive adoption and the extent to which those within the institution are motivated to adopt LA. Table 7 gives a summarised description of the subthemes and an illustration of where the theme was mentioned for each institution.

Theme	Description	InstA	InstB	InstC	InstD
Organisational culture	An organisation's attitude towards learning, using data to aid decision making and the level of sharing insights can be a driver or barrier in its adoption of LA. In addition, the level of support from the organisation's executive in investing in analytics-related projects encourages a culture of data-driven decision making.	✓	✓	✓	✓
Data-focused initiatives	The extent to which the organisation prioritises projects that focus on data to improve teaching and learning.		✓	✓	✓
Educator capacity	Educator capacity plays a key role in improving learner outcomes through the use of LA. Analysis and interpretation of student performance, engagement and other activities requires time and, if the tools available within the institution do not readily and easily provide this view, there may an impact on the amount of time available for educators to invest in such an analysis. Educators who have capacity to gain insights from analytics and intervene where needed will be open to doing so.		✓	✓	✓

Theme	Description	InstA	InstB	InstC	InstD
Ethical use of student data	Organisations are obliged to safeguard student data and ensure access to this data will not impair or prejudice any students.	✓	✓	✓	✓

Table 7: Organisational Factors

4.5.1 Organisational Culture

The culture of an organisation can contribute positively towards adoption of an LA or be a hindrance. There are various relevant components of the organisational culture, and these will be analysed in detail across the various institutions.

4.5.1.1 Access to student data

Access to data as well as having the right level of expertise to interpret the data are important in the adoption of LA. At InstC and InstD, access to a wide range of student data was highlighted as a limitation with some participants never having sight of certain types of student data, such as demographic data, or not having the means or understanding of the processes to obtain the data. Other participants indicated that having access to individual student information, such as each student's demographic information, would not provide value in better understanding student behaviours or improving course design. An aggregated view at a class level would influence how they deliver content with the hope of having a bigger impact.

> *"But the one thing is, of course, you don't always have access to these systems …. I don't know how to get onto the student information system so, I would then ask our admin team to help get that information." HD*

While educators had limited access to student data, the various institutions had administrative bodies dedicated to providing support, where needed, in the form of responding to data requests and providing dashboards where needed. From the interviewees' perspective, the need for access is exclusively to gain a better picture of the student, get a view of their support needs, understand their performance relative to peers on the course and improve teaching.

> *"It's not open for all to use; I think that's a challenge. It's usually one or two people who are administrators on the tools and if you want something or need something, you have to email them with your request or your question and then they go into that system and they draw the reports for you." KT*

In addition, interpretation of the data was deemed to be an important component, however, a key challenge shared across the institutions was knowing what data to collect, what is being achieved from collecting the data, what are the motivators for collecting it and understanding reasoning for its use. It can be very unclear what indicators to measure and why those are important to measure and what benefits their outputs will provide.

> *"I think the biggest challenges are just knowing why, what are you trying to achieve, and usually those questions aren't very clear so, it becomes a bit murky as to, like, well, what do you actually need to record and/or produce data artefacts from … I think by and large the main problem we have is we don't really know what it is that we want to know." MK*

4.5.1.2 Collaboration and information sharing

The level of autonomy or collaboration between the various departments within the institution posed as a barrier or enabler to the use of LA. The more collaboration that existed, the more they could learn from one another. The various faculties at InstC work independently of one another and each faculty has considerable autonomy on how it functions. Various projects have been initiated to use data to drive decision making, improve student performance and improve learning, however, these have been initiated within faculty and sometimes within a department. There has been little or no intra-faculty collaboration or sharing of knowledge and resources.

> *"There are too many individual initiatives deployed at the university without a roadmap that shows where we are going and how we are to get there together." InstC Teaching and Learning Report*

In contrast, InstA, InstB and InstD have a more collaborative and centralised approach to decision making and implementing changes. This has resulted in a less siloed organisation where ideas are shared, and information flows easily contributing to a more open culture amongst departments. While there are benefits to a top-down approach of decision making and implementing changes, a challenge identified at InstB is that this can create a misalignment in understanding what is needed and what is happening on the ground as management is removed from the day-to-day challenges experienced by educators.

> *"But the goal of us having a centralised data team is for us to have things talk to each other … we had so much information that we were just missing out if we weren't sharing it across departments." QE*

> *"And then some of the usual stuff where top management think the other way and the lower management and the people who are operational are not involved." JU*

Other challenges raised stem from the bureaucratic structures within InstB which impact how changes are brought forward and implemented. This was, however, not mentioned in the other institutions.

> *"And then there are structures, you know the bureaucratic structures … put us in particular positions where you can't change things at will or without going through the proper channels." HL*

4.5.1.3 Executive-level support

The level of senior-level support in creating a data-oriented organisation where evidence drives decision making influences the vigour, motivation and interest with which this can be achieved. Strong leadership has played a pivotal role in the adoption of LA across the institutions where funding is provided and the needed time for experimenting and learning is allocated. The teaching and learning executives at InstC have provided great support for using evidence from the data collected at the institution to make decisions and answer critical questions that aim to improve the students' learning experiences. In addition, evidence-based decision making using data is a strategic goal for InstD within its teaching and learning strategy, further illustrating the importance of LA within the institution.

> *"Institutional leadership is critical in this area and there's no doubt that she's interested in answering the right questions and she believes that data will enable her to do that." EM*

> *"I don't think you can manage these days in higher education institutions without having access to data and more and more universities are realising that … we know, in a budget cycle, we based our acquisition of software on impacting on student success. It must have measurable impact on student success." JC*

At InstB, however, a lack of executive-level support and a lack of prioritisation of data-driven decision making has been a limitation in broadly implementing LA across the institution. An important issue that was noted at InstB was that the collapse of the working LA system that the institution had in place resulted in a loss of trust and support from the executive due to a lack of alternatives and the length of time it took before a workable alternative was created. In addition to having executive-level support, the outcomes of the LA must be usable and accurate and must demonstrate benefits on student learning and success.

> *"In certain quarters of the management there is support. But what I also should mention is that with the LA system collapsing and us taking very long to put a workable replacement in place, we lost a lot of credibility and a lot of support from management and from academic staff. So, we first kind of have to now prove ourselves again and show them, here's the data, it's accurate, we've cleaned it this time, we know it's going to work … we've put in place measures so that it won't fall over again … we need a system that can be trusted to be used for decision making, before we can ask them to support it and use it for decision making." TR*

Along with executive-level support, in order to reap the benefits of an LA system, the various stakeholders within the institution need to play their part in adopting the system and the technological tools in place need to support this.

> *"If you don't have a strategic drive or strategic alignment, then you will not be able to reach any outcome. So, what we have is a Data Coordinating Committee focusing on student success." JC*

4.5.1.4 Openness to change

A common thread amongst the interviewees has been that for anything new to be adopted at their institutions there needs to be a shift in attitude to one that is open to change at varying degrees. There were similarities within the institutions regarding how well feedback is accepted in order for changes to be suggested and implemented. A culture that is resistant to change, whether at an individual or organisational level, has been associated with the level of expertise those in academia have in their areas of specialisation as well as their tenure at their relevant institutions.

> *"Academics – they're not open for change. In this world, I think we need to be agents of change and do things differently. Because when you come up with suggestions other people will talk about years of experience, you cannot tell me to do this." JU*

> *"It's very difficult for lecturers to do that because if I need to tell them my colleagues look at this feedback you've been getting this feedback, can you do something to change it? It can become very tricky because there's nothing wrong with their style of teaching, it's just not working for the majority of students. It used to work, it doesn't anymore. To get people to change that is really difficult." AH*

Another noted challenge is related to the insights observed from using the data, as the data has revealed realities that institutions were either not ready to engage with or equipped to deal with, and data analytics has focused on those realities in the past. The political environment within the institution and the sensitivity with which findings and insights are presented have an impact on its change appetite and its motivation to use data.

> *"The results of the analytics scared management and it scared academics because, for the first time, they were being confronted with longitudinal data of how students interact with the system and the impact that it's having on their performance." TR*

The institution's view on the opportunities and value derived from using analytics and its approach to the use of student data influences the degree of adoption of LA. Institutions where student data has been used for institutional-reporting purposes only, instead of being used to improve teaching and learning, have derived limited or no benefits from the data available. At InstA and InstD, LA has been used to improve teaching and learning as well as to understand the impact that programmes have in a broader societal context engaging with students post-completion to get feedback on the benefits derived from their choice of study. Role-players within the institution must be empowered in their use of student data to meet their various objectives. While educators have expressed a willingness to explore the various tools and systems available within the institutions, they find challenges in the ease with which to use the systems as well as in understanding how to derive what they need from the data. Some have continued to utilise the various tools available to them to incorporate data in their day-to-day teaching and decision making, despite the challenges.

> *"Previously when we were looking at learning analytics, it was purely from the education standpoint, whereas now, it can also be from a market-research side of it and bring in other elements into that." QE*

> *"It's very academic dependent so depending on who the academic is or who the convener of the course is that you're talking to, they'll use different approaches to understanding how the students are doing, for some there's no [external department] involved in that engagement whatsoever." EM*

At InstB there is a culture of accounting for everything that happens within the institution and reporting on it. The shortcomings highlighted were that being too focused on only reporting can result in omitting important aspects of student experiences and leaving certain variables out. While the institution does focus heavily on academic reporting, it recognises the value that LA can add when it is driven by pedagogy and not by the tools or reports generated. Another key element in how data is viewed is ensuring that the meaning of the data is clearly understood, and various factors and aspects are considered. Reporting at an academic level, while important, only provides an aggregated high-level view and more detailed lower-level analysis through the use of LA gives a more detailed insight.

> *"I think there are a lot of times where the university uses analytics is about the numbers only, and not about what lies behind the numbers, the struggles … so, I think that that for me, in terms of the issues of trust, for example, that might be one of the things that I worry about, just to almost use the numbers to obscure true transformation of what is happening at an institution because I think that it masks some of the deepest stuff that are happening that you can't just get from numbers on a system." FL*

A common question that has arisen from some interviewees is regarding what can be done or if enough is being done to improve student learning and resolve the failures and problems that have consistently been observed over the years, some of which are environmental and not directly related to any institution.

> *"One of the big challenges I've sort of found is yes, we can find out more and more, but what are we going to do about it … can we do anything about it? ... I think the answer has to be collaborative across departments." CS*

InstC has acknowledged that it lags behind some of the institutions within SA in its use of analytics for teaching learning and is committed to closing this gap by initiating an institution-wide project that seeks to help the institution become more data-driven in its decision making.

> *"InstC lags behind the initial cohort of five Siyaphumelela universities in terms of its use of data to inform and improve student support. However, there are currently many separate conversations taking place across the university about the need for more data-informed decision making about the student experience." InstC Teaching and Learning Report*

4.5.2 Data-Focused Initiatives

All the institutions have various initiatives as part of their strategy to be more evidence-driven in their decision making. While the institutions have some initiatives in common, there are some differences in the approaches followed in their implementation. One initiative that is common across the institutions has been using data to identify courses that have high failure rates, thus keeping students from graduating and completing their programmes. This is a priority project at InstD as they aim to improve student performance and success. The institution established a department dedicated to working on identifying these courses across the board and assigned a representative within each faculty to delve deeper into exploring and interpreting the data and identifying faculty-specific causes for the high failure rates. Part of the process followed at the institution entails findings being presented to convenors and heads of department (HOD) whose courses are defined as having high failure rates. The responsibility for devising plans to improve the performance of the courses and to continually monitor whether the changes are creating any improvements will be that of the convenor and HOD.

> *"There is this high-impact modules project, and it is run by someone within the [department name] department. What he does is he heads how the improvement of these modules should happen … he generates the data and then from there conducts a meeting […] not the whole faculty [but only] the departments which are involved, departments in which those modules are in." NI*

Similarly, InstB also had a dedicated department that worked primarily with the LA; however, these initiatives were externally funded with a limited lifespan meaning that should the institution wish to continue with the initiative at the end of the funding period, it would then need to prioritise funding for it. Given the university's many priorities, however, the department was scaled down at the end of the period and a lot of the work had to stop. A gap has been noticed since the department was scaled down and the institution has observed a need for more dedicated analytics teams. The initiatives that were prioritised and findings made available as part of tackling the issue of low pass rates in modules and using data to identify common issues were used as an early warning system. It was an important tool in identifying students who are either at risk or who might be at risk based on defined characteristics. This early warning system aimed to use student data to provide better support for students through interventions aligned with the need to improve student success.

"The recommendation was basically given that an early alert system, and you know, this is something that is happening all around the world, so we are nothing very original we are original in the approach … and then the referral mechanism, which links it together." PJ

In addition, at InstA and InstB, in collaboration with identifying students at risk, initiatives have been undertaken to better identify student profiles and behavioural trends in learning to ensure students receive the right level of support early on. A key project at InstB, following the collapse of their former LA system, has been to procure an LA system that is aligned with the institution's goal to become more evidence-driven while providing the right level of support to students to improve their learning outcomes. The intent is to consult various stakeholders to ensure that the tool is used by various role players and that it supports their goals.

"We're in the process of procuring the system to rebuild the learning analytics functionality at the institution. We've bought the warehouse and we are busy populating that and then we need to start the process of consulting the colleges, consulting the regions to build the indicator set and, of course, consulting the student body because we can't exclude them out of the process …. And we're in the process of setting up the student tracking system and we should have that up in about x months' time." TR

Over the years, various initiatives have been started at InstC, however, these have always been at a faculty or department level with little or no knowledge sharing across the institution. The initiatives, while not started explicitly under an LA banner, have focused on using data to understand the cause of the high failure rates in courses, the reasons for a low completion rate by students based on specific variables as well as finding ways to support students who drop out while in good academic standing, to name a few. As the institution has moved towards a more critical awareness in understanding what is happening in the courses using LA to influence and make changes, some of these initiatives have been incorporated into the institution-wide programme where all faculties play a role.

"That work has been absorbed into the … group where there's now the institutional-wide project which is looking more broadly across a variety of different courses and different resources, etc. … to have a far more institutional focus on collecting data, collecting the right data, putting it in the right place, looking at it through the right lens and making decisions based on being informed by a series of metrics, which I suppose we'll have to define what those are at some stage." EM

Understanding student engagement and being able to measure the level of engagement has been identified as one of the key components of LA across the various institutions. InstC has implemented a robot system for students on a specific programme to better support them by having metrics in place to alert both educators and students when performance was declining. Unfortunately, given the costs involved with running such an initiative, it was limited to only a small number of students, and the tools used were not available across the institution.

"Have the robot system for each individual student, so they get the red, green or orange. We've done that with x foundation students, but the cost involved for resources, and you have to say, for the few students we are doing could we not use that money to better help everybody?" DC

The technological challenges at InstC have impacted its ability to be more data-focused, as this limitation has led to inconsistencies in the use of data between faculties and made it difficult to assess the effectiveness of data-driven initiatives.

> *"There is no systematic use of evidence to support teaching and learning interventions and to assess their effectiveness at central, faculty or departmental level, or programme, major or course." InstC Student Success Proposal*

4.5.3 Educator Capacity

Academic staff at institutions have, given their heavy teaching and other responsibilities, little room and time for additional tasks. When using LA in day-to-day teaching, it is important to ensure that course data is correctly structured and captured to get real value from LA. In addition, educators must familiarise themselves with the tools that the institution has in place to know how the data is structured to enable them to extract the relevant information to meet their teaching objectives and monitor the progress of their students. A barrier observed at InstC and InstD has been the lack of capacity to delve deeper into their use of LA and a shortage of time to familiarise themselves with the complexities of the tools in order to get meaningful outputs and explore the various opportunities that the data presents.

> *"And then also the academics are overworked …. There's so many things in our staff meetings that we discuss that we want to do, but we never get to it because someone needs to do it, then they're really overloaded in other capacities." AH*

> *"So, for lecturers to try and keep up with the expectations of the institution considering the success rates and they then must also do the research, they don't necessarily have time to sit and look at the numbers, and then say, if this is what is happening, how can I use this data to improve my offering." NI*

A view shared equally at InstD, where they believe the true value of student data is in the actioning of it to improve teaching and learning, the actioning of which requires time and capacity from educators.

Similarly, at InstB, where the classes are much bigger and the use of the LMS is not compulsory for students, and a majority opting not to use it for a variety of reasons, educators have had to define different means of using LA outside the LMS and defining variables that provide consistency in their measurements in an offline environment.

Educators with big classes have found that using LA at an aggregated level and identifying groups of students who need support and implementing interventions at a group level based on the students' learning behaviours have added value. With the challenges in educator capacity, it is equally important that the data must be presented in a manner that will be easy to use to gain real value. Where there have been complexities in understanding the data, there has been little motivation to dedicate time to gain an understanding of student learning behaviours in addition to using performance data.

> *"For the majority of our courses, 'online' is not a compulsory element, so, not all students use the LMS …. So, if the data doesn't come to the lecturer in a dashboard format that is easy for them to have an idea of what is happening in their courses, and even if a dashboard was there, they are just overwhelmed about how should they respond, they just don't have time to look at the 20 students that haven't logged on for this week to send them an email." EO*

While capacity is a major constraint, personal interest and drive have motivated educators and administrators to allocate time to review the student data from their courses and apply the needed interventions to support students. They have also created metrics that allow them the ability to identify the various students' personas in their courses to help improve the support given.

> *"What I can say is that as an individual, it's sometimes your interest, your curiosity, you're doing, things according to the facts, it really empowers you because now I can see, I've now got wider*

lenses to say, I can even zoom into other departments and see that there's something wrong in this department. Then I can take that, share with the school directors … based on the information that comes to my table, then I can do something." JU

4.5.4 Ethical Use of Student Data

Ethics is an important aspect of LA with big responsibilities being put on institutions to ensure that student data is well protected, students are not prejudiced based on the outputs of the data and ensuring that educators act responsibly in accordance with the insights derived from the data.

All the institutions have very strict protocols in place in the use of student data as well as access to the data. Access to sensitive student information across the different institutions is permission and role-based to safeguard the information and to keep a history of access should there be a breach of any type. A major challenge at InstB has been the unauthorised distribution of student data that has led to imposters contacting students under the pretence of the institution. The institution has prioritised identifying the source of these leaks and also reviewing its systems to prohibit unauthorised distribution of student data.

"The leakage of student personal data that is then irregularly used to market services of external parties … is an ongoing problem. The [department name], working together with the [department names], investigated the source of the leaks and remedial action is being put in place to manage this vulnerability." InstB Annual Report

Institutions have limited the level of access and the type of information that can be viewed to ensure protection of the privacy of their students and in accordance with the Protection of Personal Information Act (POPIA). The data also has to be anonymised before being provided to the researcher to further protect the student's privacy.

"The policy is they need to get ethical approval; we give them the data and we also anonymise the data …. So, a lot of the data that we provide is all anonymised, it doesn't relate back to anything that can identify the student, like the name, the ID number, or even the campus ID, so we have to comply with the POPI Act." CH

When asked about their LA policy, most institutions didn't have an LA-specific policy, however, they indicated that one is needed and cited other policies within the institution as providing the guidelines. InstD, however, has invested a considerable amount of work in developing their policy around the use of student data and ensuring that it is aligned with all legislative requirements while keeping students at the heart of it.

"InstB doesn't have a policy on the ethical use of student data." EO

"We went through a project just now centring on a gap analysis of our data privacy where we used legal experts who contributed to the POPI Act … to write changes in our policies. It's being embedded in all the policies. I contributed to learning analytics ethics and guidelines. We do not yet have this kind of institutional learning analytics ethics, but it's embedded under the bigger umbrella." JC

"And then my biggest thing is there needing to be a policy. There needs to be an assessment policy, curriculum policy, e-learning policy, all those policies need to be in place." KT

4.6 Environmental Factors

Environmental factors relate to the external environment that the institutions operate in and the issues that may either drive or limit the adoption of LA. The cases in this study are all within SA and operate in the same physical environment. The regulatory reporting requirements are the same for HEIs and public institutions have to use the same systems to meet these requirements. The cases in this study have similar systems and tend to use the same vendors even though they might be on different versions of the systems or have access to different modules. Table 8 demonstrates where the theme was mentioned for each institution regarding the environmental factors that influence the adoption of LA.

Theme	Description	InstA	InstB	InstC	InstD
External/Regulatory Reporting Requirements	Adherence to legislative rules that govern the organisation in terms of teaching and learning as well as the country and having policies in place to adhere to them.	✓	✓	✓	✓
Vendor Support/Maturity of Tools	Maturity of the tools in place and vendors providing support for the tools in place.	✓	✓	✓	✓
COVID-19 Impact	The COVID-19 pandemic forced institutions to change their teaching formats and also highlighted socioeconomic challenges that impact students engaging and performing in their studies. The use of data aided institutions in identifying some of the gaps that existed.		✓		✓

Table 8: Environmental Factors

South Africa has a complex history, and the historical inequalities of the country are constantly manifesting themselves in various ways. Some of the barriers to online learning within South Africa have been equated to access to technological resources in the form of hardware and devices as well as access to affordable and functioning internet, regardless of location. These socioeconomic challenges were made more prevalent with the COVID-19 pandemic and institutions adapting their teaching to be fully online where students without the necessary devices or bandwidth on their phones could not access resources. When students are unable to connect to the LMS, it makes it difficult to track their learning and when their reasons are socioeconomic in nature, the data can help the institution provide needed support. This is a countrywide challenge that has impacted all institutions.

4.6.1 External/Regulatory Reporting Requirements

Institutions in South Africa have a statutory reporting requirement to submit statistics about student enrolments and other statistics to the department of higher education. Submissions are made three times a year to the department of a snapshot of student enrolments, the institution's teaching input and output units and number of graduates, and. It is a regulatory requirement and all institutions report on similar variables. This reporting also affects the funding provided by the department to the institution. All institutions, except for InstA use the Higher Education Management Information System (HEMIS) to meet this reporting requirement.

> *"HEMIS is our higher education management information system service … the data that we provide there is part of statutory requirements … they need to submit statistics about their student's enrolment, staff and space." DJ*

There are legislative requirements that all institutions must adhere to in line with both South African and international regulatory laws. In South Africa, all institutions have to adhere to the Protection of Personal Information Act (POPIA) which aims primarily to safeguard the information collected by the institutions. This refers to all data points from both students and staff members, and assurance that the data is being used for its intended purpose. In addition, the international profiles of the students at institutions that deliver online courses and Massive Open Online Courses (MOOCs) globally, those institutions must also adhere to the European Union (EU) General Data Protection Regulation (GDPR); regulation that, in addition to safeguarding personal information, addresses the transfer of said information. The legislation ensures protection of the rights of students empowering them to request evidence from the institution of how the data is being used. It also allows students to request that their data be removed at any time should they request for it.

> *"And so, everybody at the university is quite acutely aware of the value of data, the implications of sharing data, we are evaluating software acquisitions based on the POPI [Act]." JC*

> *"So … we are POPIA compliant, we have a data governance structure in place to say we will protect student privacy." EO*

A major challenge at InstC has been its readiness to adhere to the rules of the acts. Adherence to legislation requires that processes within the institution be changed and mechanisms in which data is stored and shared need to be re-examined.

> *"The difficulty of our data and how we collect data and how we secure data to be able to be compliant means that we're going to have to change [a] significant [number] of our business processes." EM*

Institutions collect vast amounts of student data and, with the various compliance requirements, it is imperative that they maintain a high level of security and privacy. Safeguarding student data was an important part of collecting data with all interviewees stressing the sensitivity of the data that their institutions collect and the need to keep it secure and protect the privacy of the students.

> *"In terms of the legal data or in terms of data in general, obviously we do work with quite sensitive data in terms of, we've got email addresses, we've got phone numbers, we've got ID numbers in certain instances. So, we go through quite a number of steps in order to ensure that the identities of students are protected at all times." CR*

Where data is shared across the institution, institutions have guidelines in place that must be followed due to the sensitivity of the data and to align with the regulatory requirements. Interviewees from the various institutions indicated that there are strict protocols in place to protect the data that the institution has, in turn protecting the privacy and confidentiality of students and educators.

> *"We're very strict on our personally identifiable information and so we have very strict protocols in place of how we manage that … so, there are a lot of restrictions that we've put in place to prevent any sort of personally identifiable information being shared."* QE

For the institution to execute its mandate as a place of learning and to provide support to its students to ensure their success, it needs to use the data available to it. The empowerment of students in how this data is utilised and adherence to the legislation presents challenges in fulfilling this mandate as questions around student consent are raised on whether students should provide consent for their data to be used and whether they should have a say in how their data is used.

> *"It's actually a bit of a double-edged sword because on the one hand you need the level of detail to work effectively with the data but on the other hand, you don't want to compromise confidentiality. So, it's very much, I think, a growing issue at this point more and more so for the last couple of years with the whole Protection of Personal Information Act and so on. So, yeah, there are protocols and procedures and I think people are very frustrated by them."* PS

> *"There's obviously data about students being accessed without their permission, according to the POPI Act, for example and so, we have to be very clear about why we are accessing that data."* FL

4.6.2 Vendor Support/Maturity Tools

What is common across all institutions is their use of the LMS and the SIS. There are various tools used within the various institutions. While vendors are different, some tools serve the same function and are used in similar ways while other tools are specific to an institution. Table 9 lists the various systems used in the institutions and the extent to which the tool is supported by the vendor.

Systems/Tools	Description	Vendor Support
Visual analytics platform	Interactive data visualisation software used to help represent data in a user-friendly and easy-to-understand manner.	The tool is used by three of the institutions in the study and they have access to different modules and instances of the tool. The interviewers who used the platform found that the vendors provided support and regular improvements on the platform to improve ease of use. InstA, for example, uses the tool to visually represent insights from the data.
Learning management system (LMS)	The institutions use different LMSs to teach, track student activities, and report on various metrices.	All institutions use the LMS for various teaching activities, such as delivering lectures to students, online quizzes, tests and obtaining feedback from students. The LMS at InstA is an open-source tool with different contributors who may have specific requirements and is not customised to InstA's needs. InstD uses advanced modules in their LMS to better support students. The LMS used by each

Systems/Tools	Description	Vendor Support
		institution has seen many improvements over the years to adapt to the evolving teaching landscape.
Student information system (SIS)	The institutions use similar SISs to store student-related information which includes enrolment data, student demographic data, student performance data, etc.	The public institutions use a similar SIS and have access to very similar modules. The vendor that provides this system is well established across institutions and interviewers indicated that they can retrieve different types of data from the system in accordance with their user permissions.
Data warehouse dashboard	Visual representation of the raw data from the various source systems. The tool is permission based and users can view data and reports according to their roles.	Institutions all have bespoke data-warehouse solutions to support their different needs. At InstC the tool is used as a report repository for users based on their roles and, while there are challenges in retrieving data, these are more operational and not vendor related.

Table 9: Vendor Support of Institutional Tools

All institutions have LMSs in place to run their courses and serves as a space for students to engage. A limitation highlighted with the LMS used at InstA is that it is one developed by an open-source community and the direction that the tool takes is based on the majority of the community use cases which, at times, do not align with the needs of InstA.

> *"The direction that it takes is based on what the greatest number of people are asking for and so, if we need something specific in our use cases, we'll probably have to do that ourselves." MK*

A constraint issue of the LMS at InstB and InstC is technology-based where the tool does not always support the capacity and, when capacity is high, it can fall over making it inaccessible and unusable. In addition, maintenance on the tool causes disruptions. Educators have had to change schedules and work around these limitations to limit the number of disruptions on students as far as possible.

> *"Something like [the LMS] that's down every now and then when the capacity gets too high or when there's maintenance." AH*

The limitation at InstD is related to how the LMS is structured when retrieving data from it. Interviewees indicated the complexity of retrieving student-activity data for courses with a big class complement where the process is manual and time-consuming and the information is not structured in a consolidated manner for specific needs.

> *"For instance, my first years are over 300 students, and then if you ask [the LMS] to generate a report on access, or on usage or on performance, you have to actually sift through a big load of data to get to what you want." LE*

The rules defined in an LA system are a critical element in any LA system for it to have a real impact on learning and teaching. This means that the course design and the learning outcomes of the course are key in defining the indicators that need to be in place and in defining items of measure. These can vary between courses and between the type of medium that the course is delivered in. A purely online course with a high level of online engagement and participation will have different measures to a hybrid of online and face to face, which will, in turn, have different measures to a face-to-face course. It is crucial that these indicators are understood and measured correctly to ensure that the LA delivers outputs that support educators in their pursuit of meeting student and course outcomes.

> *"The assumptions behind your learning analytic system are crucial, because they dictate your design and one of the assumptions we always worked off of, with the system that we were building was that the indicators will be derived from the pedagogy of the institution" TR*

4.6.3 COVID-19 Impact

The COVID-19 pandemic and the lockdown restrictions that were imposed in South Africa from March 2020 led to face-to-face teaching institutions having to shut down their campuses and adopt a fully online mode of teaching. Some institutions were already running a fully online teaching model, while others had a hybrid teaching model where some activities were online and others face to face, and others had adopted a blended teaching model. Interviewees interviewed during the COVID-19 pandemic expressed both challenges and new lessons learnt since adopting a purely online mode of teaching. One of the major advantages that interviewees highlighted as a result of going fully online was a more deliberate focus in terms of student tracking and monitoring in that students had to access the institution's LMS in order to attend lectures, tutorials, and other activities which might have taken place in a classroom setting. This assisted the institutions to better track student engagement and quickly identify students who were not accessing resources.

> *"At the moment it's quite clear if a student doesn't log on to [the LMS] there's a problem … Whereas when they're on campus, if they don't log on it's not necessarily a problem because they could have got this stuff from their mate, but at the moment there's no mate to get it from. So, at the moment it's quite clear. If you don't log on, if you don't engage with the stuff and you can't check the stuff, it's a problem." EM*

This ability to closely monitor engagement has assisted the institutions in better understanding challenges students were experiencing due to their socioeconomic circumstances. Access to resources, such as laptops, was something that was taken for granted when students were on campus because the majority made use of the computer labs provided by the institutions, however, during the pandemic the institutions were able to identify students who did not have computers to access material, participate in lectures and class activities. In addition, students who did not have access to data which is critical in order to access any resources online and students in remote areas where bandwidth was a constant challenge were also identified. This data assisted institutions in supporting these students by providing them with the resources they needed so they could continue with their education. Other challenges were also unearthed through constant interrogation of the data and flagging students who appeared to need other interventions.

> *"Now we are dealing with COVID as well, so most of the challenges we only realised them during this time because now you have students in the remote areas that were struggling with access, we've students that only relied on libraries, you have students that can never work from home you know, some lecturers can't even work from home so imagine if you are a student you are now suddenly expected to study from home full time. InstB had an agreement with MTN to give students data, still, some faced challenges in terms of connection." BB*

From a teaching perspective, educators and administrators devised innovative ways to teach and assess performance. Some of these innovations were initiated years prior as a result of the various protest actions that took place at HEIs across the country in 2016 that led to campuses having to close. These innovations have now been adopted in a more concerted manner and incorporated in the teaching with the LMS playing a pivotal role in their delivery.

"We are definitely changing the way we work – not necessarily because of COVID but we are using [the LMS] more …. It's been very easy to pick up copying and things …. So, we were forced before COVID actually to explore some of the functionality of [the LMS]. The only thing maybe that I can add now is, we have to make sure that we are very clear and consistent in the way that we post content so that students don't get confused." UI

4.7 Purpose of Collecting Data and Learning Analytics Use Cases

This theme examines why data is used and what opportunities exist with its use. It explores various avenues on the intentions behind collecting, analysing and interpreting data within the different institutions and the impact these intentions have on the current state of LA. It is important to note that the reasons behind collecting data are only those that surfaced from this research. By understanding some of the motivators of collecting data, it may facilitate the discovery of other areas in which LA can be used within institutions and possible opportunities. Table 10 illustrates where the theme was mentioned for each institution.

Theme	Description	InstA	InstB	InstC	InstD
Decision making	Student data has been used as an input in decision making along with other available information.	✓	✓	✓	✓
Holistic view of student cohort	Use of student data provides a rounded view of students and course activities equipping educators and administrators to identify students in need of support as well as courses that are deemed to be problematic.	✓	✓	✓	✓
Student support	Using student data has helped identify students with different types of support needs where interventions have been identified and implemented.	✓	✓	✓	✓

Table 10: Purpose of Collecting Data

Most institutions collect data as part of their general operations and as a consequence of using technology in their day-to-day activities. While institutions have historically collected data to meet their academic reporting needs, to establish performance metrics and identify problem areas at a micro level, the type of data collected and its use have equipped educators, administrators and the executive to make more informed decisions, shape the direction of the courses and better understand student learning behaviours and needs. With the vast amounts of data collected by the institution, a question on whether the data is sufficient is often posed and interviewees mostly agree that the data collected is sufficient, it is more a question of access to the data and the type of data collected.

Data is collected to understand performance in a course and to look at specifically what went on in the course. What are the things that went well and what are the things that didn't, for students who performed well, what are their profiles and, for those who didn't perform well, what are the possible reasons? To understand descriptive data, the LMS and other data sources are used as input to get a broad perspective. In addition, the collected data also provides insights into student performance to incorporate all student activities to better understand where the student is in relation to their peers in the course. As more institutions move towards using technology, educators are also incorporating the use of the LMS data more in their courses and are introducing more online engagement with students. In addition, institutions collect data from students in surveys and incorporate the LMS data to understand drivers of student successes and/or failures and also incorporate predictive modelling techniques to determine the probability of a student passing given certain variables. The range of roles that the interviewers in this study perform provided diverse viewpoints on the uses or need for LA in the different institutions and motivators of use, if any. Table 11 provides a summary of the different uses or needs of LA mentioned by the different roles within the four cases.

Uses or Needs of LA	Role
Institution-wide decision to support students during the pandemic by providing resources such as laptops and bandwidth to those in need so they can continue with their studies. The data collected revealed that students were experiencing challenges accessing resources and some students did not have the correct tools to access the tools.	Deputy Dean and HOD
Making recommendations to the executive on the LA initiatives that the institution needs to embark on based on insights observed to better improve teaching and learning at the institution.	Director/Head of Teaching and Learning and Coordinator: Learning Technologies
Identifying students needing financial support and provide support by guiding them to the relevant resources and by motivating those who are at risk of dropping out due to financial reasons to receive assistance from the institution or partner donors.	HOD
Making changes to the design of the course in light of the insights observed from the data to improve teaching and learning. Frequent use of the LMS data has assisted convenors to better understand their students and identify areas of improvement in some of their courses.	Course Convenor
Identifying students (individual or aggregated) who are at risk of failing the course and implementing intervention measures to help mitigate the risk through support initiatives.	Course Convenor and Lecturer
Identifying students (individual or aggregated) where there is low engagement and implementing measures to improve engagement. Some interviewers introduced nudges where they send students emails to remind them of submissions or send them short quizzes to get an indication of where a student might be struggling.	Course Convenor and Lecturer
Tracking student performance and engagement levels to gain insights on learning behaviours and effectiveness of the course.	Course Convenor
Focusing on student success and retention to help them meet their learning objectives. Where lecturer, convenor or student advisor-specific interventions are needed, the student success practitioners alert the relevant roles to ensure students are supported early enough so improvements can be made.	Student Success Practitioner

Uses or Needs of LA	Role
Creating reports for the department or convenors to show performance and student engagement. In addition, designing the look and feel of the dashboards used at a course level and defining the data that needs to be retrieved and what views are needed.	Information Services and Student Success Practitioner
Technical design of dashboards and tools that can be used to view student and institution-wide performance. Relying on academics and administrators to understand what the requirements are for what needs to be designed and built.	Engineering Manager
Reporting at a macro level on the overall performance of the institution, identifying trends in different departments and analysing and interpreting new insights.	Institution Researcher

Table 11: Uses or Needs of LA

There are many benefits that have been observed from the use of LA at the institutions. As a purely online learning institution, InstA has emphasised LA as a key driver towards ensuring that students meet their learning outcomes, while at InstB, LA is gaining more focus as it is seen to be a critical area. An important aspect that institutions have been trying to understand is where learning actually happens; whether it happens based on the amount of time students spend working on the material online or whether they find it easier to learn by downloading the material and spending time working on it offline. This is an area that InstA and InstB have been working on getting a better understanding of. This has also shed light on whether the student needs are being addressed through the courses and programmes for which they enrol.

> *"It should also be noted that learning analytics and the various ramifications of using student and other institutional data (in terms of privacy, ethics and profits, for example) is a growing conversation to the extent that it would appear to constitute a quasi-education discipline." InstB Annual Report*

> *"We're very much an analytical and data-driven company. And I think it's part and parcel of being an online company is you can gather so much data." QE*

An important aspect of using LA for improving teaching and learning is ensuring that the insights derived from the data are actioned upon and not ignored. When a student is identified to be at risk or in need of additional support and that information is ignored, the educators or administrators viewing the information are perceived to be acting in an unethical manner and are thus not addressing the issues the student might be facing in order to help improve their performance. Other risks that could introduce unethical behaviour stem from definitions used in identifying such students without either putting them at more risk or prejudicing them in any way.

> *"You can't trust it blindly, I suppose, that's the lesson involved. And you should be abou,t I suppose, the consequences of what you use the data for … again, for risk scores you could give someone a risk score that was based on dodgy data and then you've inadvertently caused harm because you've demotivated someone who really shouldn't have had any concerns." CI*

4.7.1 Decision Making

The data collected has been used to meet specific objectives and analysed to answer specific questions. Most found value in the data when informing decision making related to the course and curriculum design, introducing changes on how a course was structured to support student active learning and management of a course.

> *"We use the descriptive data to analyse what devices the students are using to access the LMS, from that we derived the lack of access to devices and that are, therefore, impacting our support." JC*

The COVID-19 pandemic has also played a role in motivating educators to find new ways of assessing student progress and engagement through frequent use of the LMS data. Due to the socioeconomic challenges related to bandwidth that some students have in accessing online learning tools and material, the pandemic has only heightened these issues and through the use of data assisted institutions to identify students in need of support. At a course level, the LMS data has equipped educators with the information needed to improve their courses.

> *"Now in COVID we would, for instance, have a check-in, a little short test that we forced them to do to just see who's got active connection and who's actively participating in the course. So, we'd use that data to then inform our decision, and we'd have a mock test and see how many people has activated that and accessed that to then inform our decision how our tests will be set up." LE*

As part of their goals to improve the design of their courses and to continuously improve teaching and learning and ensure students are meeting their learning goals, interviewees agreed that more use of data assisted them in meeting these goals. At InstC, predictive models have been used in some courses to allocate students to tutorial groups that harness their diversity in both abilities and backgrounds so they can learn to work together as a collective to meet the course objectives. The model helps educators identify a blend of student skills, traits and learning behaviours so they can pair them in groups that will help them flourish by exploiting their strengths and weaknesses.

> *"The data we get, we go through it and then we use it to see how we can improve the quality of the course, but we also look at if the data we collect is actually telling us what we want to be able to improve upon the course." HI*

4.7.2 Holistic View of Student Cohort

Having a complete view of the student cohort assists in understanding the profile of students in a course, department or faculty. In addition, having the right information at hand allows those in academia to gauge student performance, identify issues with their learning to better support them and identify bottleneck courses to add support mechanisms or redesign them.

A holistic view of the cohort also gives insight into how the students are experiencing the course, an important element in ensuring that their learning objectives are being met. Interviewees stressed the value of the student experience because it can either motivate them to do better or be a demotivating factor that may negatively affect the students' performance. In addition, it also assists educators to observe their teaching critically and identify areas that may be lacking and needing to improve as well as things that are working and helpful to students. In addition, having an all-round, integrated view of a student's performance across courses has helped educators in their decision making as they can determine whether a student will be successful early on in the programme, given their performance in courses that are key to the programme. While useful, at InstC there is an inconsistency in capturing student marks on the SIS and thus, getting a rounded view of their performance in relation to success measures was challenging.

"You can then look at the suite of courses that the student has taken and the level of success in those particular assessments and make some form of rational decision about how they're doing and then be able to advise them accordingly. The problem with that is that it's been very poorly implemented across the institution and deadlines by which these grades need to be captured have not been met." EM

Another identified challenge has been with the MOOCs, where course convenors and lecturers found that students were not forthcoming with their feedback in terms of how they are finding the coursework and often have to rely on the data to get insights into the performance. Another noted challenge with the MOOCs is that due to the high drop-off rate, insights on the reasons for dropping off are not readily available as students do not provide that feedback.

"On the MOOC you don't know, just don't know how people are doing ..., We go on about 10%, which is good for MOOCs, it's really high. Like, a lot of MOOCs get, like, 1% or 2% and we don't get an indication along the way, but we've run it enough times now that we sort of know what to expect and the numbers of the fall-off rate is fairly similar each time." HH

4.7.3 Student Support

Student support comes in many forms and one of the drivers of using LA at the institutions has been to identify students that are seen to be at risk, either of failure or of discontinuing their studies and to offer the appropriate intervention. Early intervention and the correct type of intervention is said to help students achieve their learning objectives and improve retention rates. InstA has a dedicated team whose primary focus is to support students to help them meet their learning objectives. The institution looks at different variables and is guided by pedagogy in the definition of these variables to ensure that students who are at risk are identified early and given the right level of support. InstB has a unit dedicated to ensuring the retention of students. While the unit started with a focus on their first-year cohort, it has grown to include all years of study and different faculties within the institution have a student retention subunit to address student needs. At a faculty level, this unit has seen improvements in student performance and retention, indicating that the interventions that are being applied have had a positive effect. InstC and InstD have dedicated student advisors in each department, and they play a critical role in supporting students and helping them through the various challenges they may have. At both institutions, where students have been identified as high risk because, either they have not been engaging with the material or attending lectures or submitting assignments and quizzes, to name a few, and where lecturers in the course have observed consistent concerning behaviour, the student advisors would then intervene by contacting the students to establish the needed intervention. Access to the student's activity data in some cases may be limited where sensitive student information is required and, in those cases, the advisors will not have all the information available to them to successfully assist a student. Some of the information that has been cited as confidential, is student-counselling data, to which the student needs to give consent first before the advisor can be privy to it.

"We have one particular member of our academic staff who is tasked with providing curriculum advice for particularly vulnerable students and, in this instance, I'm talking about vulnerable students who are signalled as such by having been excluded by the university and have been allowed to return on appeal. And so, they have a particular set of restrictions on them when they return, and we have one academic who is involved with trying to ensure that they have the best possible trajectory on their continuation in their programme." EM

An area of support that is consistent amongst the institutions has been on the student's financial aid. Students face financial challenges that often result in their performance declining or them dropping off without completing their studies. The use of data has assisted interviewees to be more proactive in identifying these students and, where possible, supporting them in obtaining financial assistance from the institution. While performance data indicates a problem, where interviewees have access to a broader spectrum of information, they can identify the cause of the problem, in the absence of this they can obtain detailed information from the student, so they know which intervention is needed.

> *"With the student financial aid department, it's just sometimes the information is shared … and then when we look into our enrolment planning, and then we look into the students who drop out … it will be financial constraints. Then we look into what is available, we check as a college management committee what other funds are available." JU*

Institutions have devised interventions tailored to the situation that the student was in as opposed to applying a blanket solution in supporting students and helping them improve their performance. An early-alert mechanism, which uses of algorithms to identify students who may be at risk early on, has been a helpful way of supporting students and giving them the aid they need. InstB is in the process of building an early-alert system with proposed interventions. The key to any alert system is ensuring that once students at risk have been identified, the correct steps and intervention measures are taken to help them improve. While an early-warning system is useful, the point at which the warning is made is vital. The earlier a student is identified as at-risk and the correct interventions are taken, the better, otherwise the warning can come too late, and the institution cannot help the student.

> *"The university has developed an appropriate ris- prediction strategy that provides a framework for identifying at-risk and high-risk students within the institution based on current and historic data." InstB Annual Report*

Since the COVID-19 pandemic, institutions have increased their reliance on the data collected and have used it to make decisions to better support students. One of the key decisions made by most of the institutions was to better support students who were identified as needing resources by providing them with laptops so they could continue with their studies while at home as campuses have been closed. In addition, institutions were able to gain an understanding of student needs and sought to assist students who required bandwidth by getting into agreements with mobile network operators in South Africa to provide students with free mobile data.

> *"If it was a device problem, InstD actually had the option where students were given laptops or tablets and based on those statistics, I could then motivate that that student is a student that deserves to have access to that device because prior to COVID she would have accessed [the LMS] 99.9% of times and now during COVID being at home, she cannot access it anymore." LE*

4.7.4 Lack of Adoption and Missed Opportunities

Institutions have invested time and resources in better understanding and supporting students through the use of data. While institutional strategy is geared towards a broader use of data for LA, the barriers and limitations of LA have led to a lack of or slow level of adoption and limited use of LA. The consequence of which is missed opportunities within institutions that may have resulted in improvements in both teaching and learning.

A concern was raised by a participant at InstB, questioning the use of data within the institution to drive curriculum design and changes. The interviewee indicated that while data is used, there isn't a concerted effort at the institution to use LA to inform course design and to make improvements. It was only in some courses where the participant noted that this might be happening.

"The fully online modules – I do think they may have a better use of the data to inform curricula or pedagogy but, again, not in a committed, concerted, focused way to say that after the semester let's look at all the data; how many students participated, at what times, so I don't think there is, it's culture." EO

A question raised on the use of LA within the organisation is related to whether or not the data has been useful and if changes and benefits have been derived from its use. Socioeconomic conditions continue to be observed as having an impact on student performance and an important aspect of this is the correct classification of issues. The inequalities and history of SA often result in misclassification of variables that contribute to the high failure rates observed, however, interviewees viewed the socioeconomic realities of the students as a factor in their performance challenges.

"What we found is generally in most of our courses it is racially skewed …. In this country, race is often linked to socioeconomic issues. So, I think calling something race is not correct, it should be more socioeconomic …. Consistently, unfortunately, poor socioeconomic backgrounds underperform compared to privileged backgrounds." DC

The collection of data at InstC and InstD has grown significantly over the years, so too have the opportunities that come from insights derived from the data. The opportunities that the LMS has presented in the institutions have been immense and the ability to track various student activities and associate LMS data with other source systems has proven beneficial. A limitation has been that not all educators have made use of the LMS in a consistent manner, thus not deriving the same benefits across the institution.

"Provided those particular courses have activities in the LMS … any activity that you do in the LMS provided the course has LMS data associated with it. We can, we can link it back to the student." CH

"But I think the possibilities are endless with data science and the amount that you could do with the data that these learning management systems are actually spitting out at us, we must just have better training on how to interpret it." LE

The need for more specialised skills in the interpretation and mining of data has been highlighted as a major drawback as these skills are not necessarily available at the institution and the specialisation of the skill also requires dedicated time and attention. While this is a missing skillset, the institutions do still mine and interpret the data with the limited skills available.

"There's a lot of different attributes in the data warehouse, but I think what we miss, almost what we're missing is a data scientist type role. They can take all that data, analyse it and pick up markers for improving success of students." CH

4.8 Impact of Learning Analytics

Interviewees have highlighted the impact that LA has had in their teaching to improve learner outcomes. One such impact has been observed when applying interventions to support students where the use of data has highlighted risks for students. Courses that keep students from graduating have consistently been observed as an impediment across the various institutions and the data has given a more detailed view of the problem as well as assisting in the identification of the most suitable interventions. The common definitions of these courses are those that have a large number of registered students have had a consistently high failure rate.

"We know that there are certain courses that are impediments to throughput and to a point, if you fix one of them, often the trouble is you're actually just pushing the problem further down. It's about tracking it in that regard." AC

One of the benefits of using data observed at InstB was identifying inconsistencies in the admissions process for certain courses and re-examining that to ensure that students are given admission for programmes that they will be successful in without setting them up for failure. The institution was able to, through the use of data, recognise that different rules were being applied at the detriment of students and the changes implemented were enough to ensure that students were admitted to courses where they could be successful. The use of LA at InstB allowed the institution to identify students who were in the system for a long time without completing any qualification. The institution was able to identify these students and give them the support they needed in order to complete their qualifications.

"We also picked up when we started doing an analysis of students at the individual level, and we started picking up massive inconsistencies regarding the time that students have been registered at the institution … that was also something that we would have been unaware of otherwise." TR

A key highlight at InstC has been the introduction of MOOCs and the impact these have had on the course design of the modules. This has been helpful in improving the approach to course design both on the MOOCs and on other face-to-face courses and applying the lessons to improve learning outcomes. The introduction of the MOOCs has also resulted in students enrolling for a programme at the institution after completing the MOOC. The institution has used this data to measure the effectiveness of the MOOC in promoting learning and to learn from the feedback and insights to improve and better structure their courses. Another lesson has been the uptake of the MOOCs at InstC and the high completion rate when compared to the average completion rate on MOOCs in general.

"So, because of the postgraduate diploma, the blended online module, we've now been able to launch six MOOCs … and the benefit has been that some of the students now we're getting for the postgraduate diploma are people who have done the MOOC." HI

While data serves as a useful tool in providing insights, interviewees noted that data in isolation does not lead to improvements in teaching and learning, but rather this is a combination of gaining insights from the data, applying various interventions and other measures. A combination of these measures is what has been observed to lead to improvements in meeting student outcomes.

"In no way can we isolate any improvement to a single variable, but based on the initiatives and projects, we attribute that the improvement may relate to our interventions or up to our student success projects." JC

4.9 Barriers and Constraints to Learning Analytics Use or Impact

While positive impacts have been highlighted, drawbacks exist in the use of LA which have either slowed adoption or resulted in no adoption in certain areas within the institution. Algorithms play a significant role in the design of LA systems and are used to predict learning behaviours and identify high-risk students early on. Algorithms may also cause harm where variables used perpetuate biases or incorrectly classify students. Another constraint is in the technologies used where there are system failures, inaccuracies in data due to a lack of system integration as well as complexities in the tools built where the users' needs are not understood or catered for. A shortage of skills in the interpretation of complex and large amounts of data have also been highlighted as a constraint at some institutions. The culture of an institution and the level of autonomy in different departments were raised as constraints in the adoption of LA. Table 12 summarises the factors negatively impacting the use of LA across the various institutions.

Factors Negatively Impacting LA	Reference
Algorithms used for predictive analysis may be designed to perpetuate biases by including variables and making assumptions that may reinforce existing inequalities.	*"Threat of algorithmic bias … very often we're trying to use learning analytics to predict success, but we have more data that shows us what failure is, so instead of predicting success, we're predicting failure." TR*
Disruptions in the availability of the information system impacting its usage.	*"System failures and the systems' inability to support a high-user access at the same time is a barrier in the use of LA and interactions are not catalogued as a result.* *" … other challenges can be ICT challenges … sometimes our system crashes because it was not made for more than a million submissions per day … the system is overloaded, then it crashes." JU*
Integration between the various source systems in the institutions to ensure that data flows seamlessly and is accurate and available. A lack of integration affects the interpretation of the data due to gaps in the data.	*"There's no central warehouse, so the data are stored in different databases that often don't talk to one another …. The data are not in the same format, not same quality, not the same governance, not the same." EO*
A shortage of skills and understanding of how raw data can be analysed and translated into meaningful insights.	*"The main constraint is actually some kind of mediation between the data and the person that needs it … there's a big gap between the data that's collected and what you can actually use at the end." HD*
Complexity of system and tools discourage willingness to learn skills for using the analytics systems or tools with most users finding alternative means to gather the information they need.	*" … there's a big barrier there … it's the interface of [system name] that people find very, very difficult to navigate." PS*
Level of autonomy and a lack of collaboration across the institution.	*"InstC has a lot of distributed expertise, it's a very distributed institution in how it operates culturally, organisationally … at InstA if someone thinks this is important, it is a more complicated process because faculties have strong autonomy and there are lots of different pockets of expertise. It's quite hard to just align everybody." CI*
Lack of access to data is a barrier. An inability to obtain certain types of quantitative data has made it challenging to measure the outcomes of this change and instead, they have relied more on qualitative data.	*"… that's an area that we've been trying to do but we haven't really been able to keep track of their progress because then we can say that it's fair, this process is fair. So, we've been trying to do that but we haven't done it very well because of this registration data that's such a struggle to get." HD*

Table 12: Factors Negatively Impacting Use of LA

4.10 Conclusion

This chapter presented a detailed analysis of the study based on the data collected from the four institutions. The findings focused on the five themes that emerged from the analysis, namely technological factors, organisational factors, environmental factors, use of LA and impact of LA. Further discussion relating the findings to the research to the literature will be done in the next chapter to complete the analysis.

A summary of the themes and subthemes arising from the analysis has been illustrated in Figure 6 and the linkages between the model and emergent themes has been shown in below

Framework Component	Emergent Theme
Technological context	Section 4.4
Information quality	Section 4.4.2
System quality	Section 4.4.3
Service quality	-
Organisational context	Section 4.5
Environmental context	Section 4.6
Intention of use	Section 4.7.1 Section 4.7.2
Use	Section 4.4.4 Section 4.7.3
User satisfaction	Section 4.8
Net Benefits	Section 4.8 Section 4.9

Table 13 Linking Emerging Theories to Conceptual Model

.

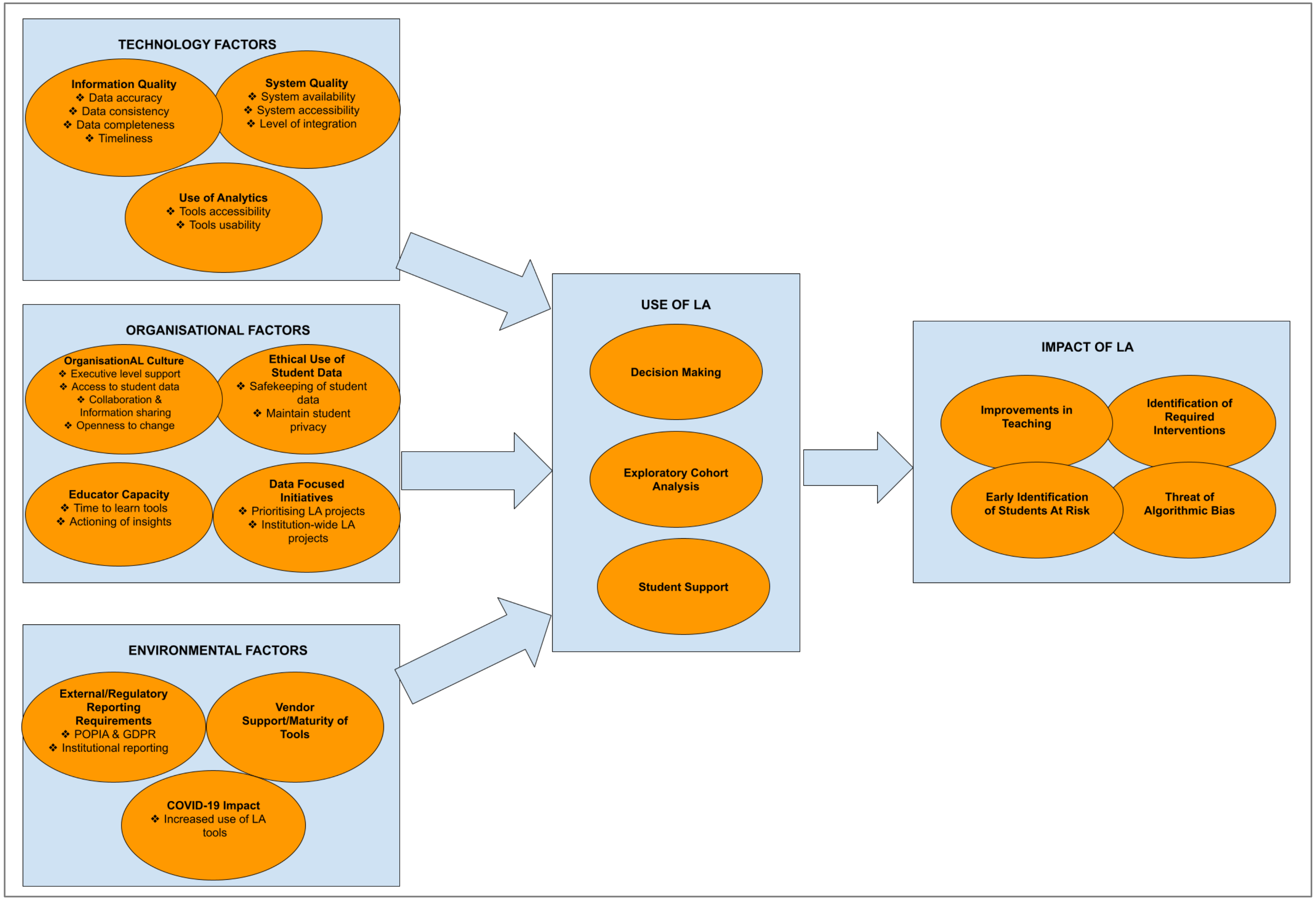

Figure 6: Summary of Data Analysis Themes

5 Discussion of Findings

5.1 Introduction

This chapter presents the findings based on the research questions posed. The overarching research question of this study was to investigate the current status of learning analytics (LA) in online learning at higher education institutions (HEIs) in South Africa (SA). The following detailed questions were posed in meeting the research objective:

RQ1: What type of student data can be and is currently being collected?

RQ2: Is the data used for LA and if so, to what extent?

RQ3: Is LA used by administrators, educators and lecturers to inform decision making?

RQ4: What opportunities and barriers exist for more advanced LA?

5.2 What Type of Student Data Can Be and Is Currently Being Collected?

The HEIs observed in this study all collect various types of student interactions however, the manner in which the data was integrated varies across the institutions in accordance with their information systems architecture. The nature of the HEIs in this study and the need to become more technology focused has resulted in extensive use of their learning management systems (LMS) as well as their student information systems (SIS), both of which were viewed as the core of any HEI's technological ecosystem (Gasevic, Tsai, Dawson, & Pardo, 2019). The institutions were found to collect various types of student demographic information, student financial information as well as student's previous education history in order to make enrolment decisions. The institutions also had online capabilities where students access course, library and other institutional resources online which have presented opportunities for data collection. Every online interaction that a student has with the institutions in this study left a digital impression, a claim supported by many studies (Reyes, 2015).

The SIS and LMS were found to be key in the adoption of LA at the institutions due to extensive use by both students and academic staff. The findings suggested that the assumptions made in the design of the institutions LA system were key in ensuring that the LA system would be effective in meeting set out goals and that these assumptions must be embedded in pedagogy. These findings are supported by Tsai et al. (2020) who posit that the extent to which LA deployment is informed by learning science impacts its effectiveness and it is crucial to have a mechanism in place that ensures that LA aligns with learning science.

The findings claim that the level of trust that interviewees expressed to have in the data was driven by the quality of the information as well as the perceived quality of the system where the data was sourced from. It was observed that there was a high level of trust of the LMS data as it was perceived to be complete and accurate with a low possibility of manipulation or human error. Trust building has been shown to be an important aspect in the adoption of LA and is viewed as a foundational element in demystifying perceptions and encouraging the use of the technology (Klein et al., 2019).

While the LMS and SIS provided institutions with valuable insights, it was observed that student interactions outside the LMS were not recorded and insights on learning and teaching for courses which did not rely heavily on the use of the LMS were limited to student marks derived from written tests, exams or assignments. In addition, data on students' socioeconomical situations was not collected and it was observed that instances where the data was collected was for specific students, for example students undergoing a review due to appealing academic exclusion, in certain categories and not for all students. In their study, Wilson, Watson, Thompson, Drew, and Doyle (2017) other factors impact a student's final performance and a diverse approach to predict performance and learning behaviours was more beneficial for students. Studies have also highlighted that establishing engagement and future performance on online activity requires further research in designing models that are reliable and account for non-online activity (Bodily, Graham, & Bush, 2017).

The findings suggest that the degree to which the institutions collect data about students raises concerns regarding privacy of both students and teachers, and ethical issues in the collection of data and adherence to legislation aimed at protecting personal information. Added to this is the need to adhere to the Protection of Personal Information Act (POPIA) and questions on whether consent should be given in order to collect and use student data in an ethical manner. Concerns regarding ethical and privacy issues in LA have been echoed by other studies with many researchers observing these to be key considerations in the successful adoption of an LA solution (Gasevic et al., 2019). While these issues are viewed as important, there are still a few studies that have approached the issues of ethics and privacy in LA in a systematic manner and there is a hope that these will increase as more research is conducted in this area (Viberg et al., 2018).

The vast amount of collected student data requires analysis and interpretation in order for data to be of any value (Reyes, 2015). The findings claimed that having the skills to analyse and understand the meaning derived from the data was important in order to gain actionable insights from the student data. This was highlighted as a challenge that existed within institutions due to various reasons such as having the right skills to correctly analyse and interpret the data, ensuring that there are common definitions on how the data is stored and categorised. The correct understanding of student data and having the right skills to interpret the data is important to assure educational benefits through the use of LA (Avella et al., 2016).

The findings suggest that challenges exist that influence the use of data. One of the challenges observed from the findings that impacts adoption was related to the availability and quality of technological and manual systems that institutions had in place. The findings suggest that having seamless integration and flow of data across various information systems within the organisation and ensuring data is kept up to date influences the extent to which the data is used and trusted to yield reliable insights. The findings also suggest that that the type of system infrastructure in place directly impacts how data flows between systems in the institution thus impacting the quality of the data and its usability. The diversity of the technological infrastructures employed by the HEIs highlighted the benefits and challenges presented by either a centralised or fragmented infrastructure. What was important was that the systems were integrated ensuring that data flowed seamlessly across the systems, so the data is up to date, accurate and usable. Reyes (2015) claimed that when data does not flow timeously and seamlessly, concerns have been raised on the accuracy of the insights observed and the completeness of the information.

Furthermore, understanding what question needs to be answered through the use of LA was an important observation as institutions indicated that while there are large amounts of student data being collected, the objective of collecting the data can at times be unclear. Previous studies demonstrate that having clear strategic objectives for use of LA in an institution and alignment on the priorities that LA solutions will yield and having clearly defined questions to be answered through the use of analytics helps institutions effectively meet their objectives (Gasevic et al., 2019).

5.3 Is the Data Used for Learning Analytics and, if So, to What Extent?

Data collected from students in the form of surveys during and at the end of a course was found to be used to make changes in the course to improve learning outcomes and teaching by changing how the course is delivered, adjusting assessments and the types of assessments and changes to the course structure. LMS data was observed to be used to also inform decisions on curriculum structure and to measure student engagement. LMS data was found to be used by course convenors and lecturers to identify students needing support, identify modules that were either too easy or too difficult and to measure the effectiveness of some of the changes implemented in the course. In addition, LMS data was also used to highlight course performance over a period of time. The study found that library data was not used as an input when analysing student data. The findings claim that since the COVID-19 pandemic, data regarding how students accessed physical institutional resources was compared to their online activities to better understand the cause of the changes in performance and for some, their lack of engagement with online resources.

LA uses both static and dynamic information about students in order to understand their contexts as well as optimise learning and teaching to improve student outcomes (Siemens & Long, 2011; Viberg et al., 2018). The study revealed that while institutions primarily use technology for student learning activities, there were many interactions that happened outside the online environment that were unrecorded and, therefore, not measured. The participants in the study stressed the importance of ensuring that the LA system assumptions were driven by pedagogy instead of following a single approach in the measurement of learning and understanding behaviours. This finding is supported by Tsai et al. (2020) that meaningful insights can be observed when collected data and indicators in place are aligned with learning theory. The findings are further corroborated by Wilson et al. (2017) who in their study emphasised the importance of acknowledging diverse approaches of study and ensuring that these are incorporated in the design of the LA tools.

5.4 Is Learning Analytics Used to Inform Decision Making?

The findings suggest that LA has been used within institutions for different types of decisions. LA was used for routine decision-making to answer specific questions to meet goals and objectives related to a course and these included course design and performance or was used for student focused decisions in better understanding student behaviour. In the case of InstA, data was used extensively to track the student from the point of showing interest in a course through to completion of the course and determining whether the course met their learning objectives by obtaining feedback once the course was completed. In addition, data was found to be used widely by the different types of stakeholders to make non-routine decisions which were primarily driven by the different uses of LA and stakeholder needs. As corroborated by Ferguson et al. (2016) there are many different stakeholders involved in LA with different objectives and different types of decision-making needs. The importance of pedagogy driving design of LA tools was found to be a key consideration in the extent to which the insights observed were credible. The finding was consistent with other studies in LA that emphasise the importance of understanding how people learn in order for analytics to provide useful insights (Ferguson et al., 2016; Ifenthaler, 2017).

The findings claimed that Organisational culture and the level of executive level support in an institution were an important driver or barrier in an institution moving towards a data-driven outlook and being more evidence based in its approach. The findings suggest that organisations that have a centralised decision-making approach were more likely to adopt LA because of the level of collaboration within faculties and alignment on the goals of the LA strategy. This finding aligns with Klein et al. (2019)

assertion that institutional commitment to implement LA promotes use of the tools. Tsai et al. (2020) also supports the claim that key to sustainable adoption of LA is executive level support and clear direction for institutions to be evidence-driven in their decision-making. Institutions in the study that gave autonomy to faculties were found to have different LA strategies within each faculty and did not share information on lessons learnt and pitfalls to avoid. In their studies, West, Heath, and Huijser (2016) and M. Brown et al. (2020) found cross collaboration within an institution as being an important driver in the implementation and adoption of LA as it promotes information sharing, removes silos and encourages a level standardisation across the institution. It was also found to create alignment and encouraged faculties to work towards a common goal and objective (West et al., 2016).

The extent to which the analytics tools are accessible and easily understandable was found to be either a motivator or deterrent in the adoption of LA. The findings suggest that educators and lecturer capacity is limited, and the LA tools implemented at the institution ought to be intuitive so they can be easily understood. The findings suggest that perceptions regarding the ease of use of LA tools was inconsistent amongst the different roles that participants held within the institution. The diversity of participants in the study highlighted these differences. Participants who regularly interacted with the tools at the institution found them to be user friendly while those with limited tool engagement found them to be too complicated and often required extensive time investment. A large number of participants who were in a non-technology focused roles or faculties found the tools and systems complex, while those in technology or engineering related faculties designed their own workarounds to get around the complexities of the systems or tools, other participants opted to liaise with the departments responsible for data to provide the information they needed. Support from the relevant areas within the institution is essential in driving adoption (Gasevic et al., 2019). The study found that while stakeholder involvement is important in designing or selecting the LA tools, there didn't appear to be consultation with academic staff in tool selection or design. In their study, Klein et al. (2019) emphasise the importance of actively including different types of stakeholders in the process to get buy in, gain alignment and incorporate user needs and interests in the LA capabilities. Involving different stakeholders within institutions in the design and selection of LA has benefits that include understanding on how to use the tools and also serves to have alignment within institutions on the initiatives that support the use of LA (West et al., 2016).

The findings suggest that the institutions in this study have struggled with courses which inhibit students from graduating, also known as 'bottleneck courses'. LA was used by the different institutions to inform decisions regarding approaches to be taken in order to tackle the issue of students failing to complete their studies. The decisions on investing in working to resolve these issues have largely been driven by institutions using student data to better understand what the challenges were that students experienced with the courses, what the profile of the courses were, what behaviours students exhibited when registered for these courses. Using available data to understand these questions helped inform recommendations and suggested changes needed embedded in educational research and practice in order for the interventions to be effective (Gasevic et al., 2019). The findings claim that improvements that were classified as high priority incorporated insights from analytics and required support from the executive in order for funding and time to be adequately allocated to make improvements, this finding supports the claim by Tsai et al. (2020) of the importance of institutional management and leadership in the prioritising LA initiatives and allocated the needed funding in order for improvements to occur.

The findings claims that LA was being used for early identification of students at risk of failure or dropping out and decisions were being made on ways to support students early enough to avoid this eventuality. Dawson et al. (2019) highlights the benefits that LA affords educators by identifying students that may be at risk early enough to intervene and Papamitsiou and Economides (2014) demonstrate that the use of LA could accurately predict learner dropout and retention rates at early stages. Avella et al. (2016) also posits that LA offers interventions for education institutions through application of predictive models. The findings also raised concerns related to algorithmic biases that may occur which has the potential of causing more harm to students who might already be at risk. This concern has been echoed by other studies in LA which highlighted concerns when using technology to classify students as being high risk due gaps that may exist in the data, quality of the data as well as prejudices and biases that could result in misclassification of students thus causing more harm (M. Brown et al., 2020).

The findings claim that measuring student engagement and better understanding where learning takes place were found to be areas that institutions sought to gain insights on through the use of LA. Institutions capture all student interactions with the online systems and these create records which store the interaction and the duration of the interaction. These interactions can be viewed and analysed at a student level or at an aggregated level and were observed to inform decisions on the design of the course to better address student needs (Wong, 2017). The findings raised concerns related to the manner in which these interactions are measured and further questioned whether the principles applied in measuring performance and online activity were a good determinant of a student's future success or failure. These concerns are supported by Wilson et al. (2017) who argue that various factors need to be considered when using students online activities as a predictor of success or failure.

5.5 What Opportunities and Barriers Exist for More Advanced LA?

This study found that an organisation's culture as well as attitudes and perceptions within the organisation can serve as either a barrier to LA adoption or encourage adoption. The findings claim that the institutions in this study are at different stages of their LA journey, yet none are seen to be at an advanced stage of adoption. Alignment within an institution regarding the LA strategy plays an essential role in an institution's willingness to adopt LA (Tsai et al., 2020). Of the institutions in this study, two were observed to have reached the maturity stage of LA having shown evidence of using data to drive decision-making and having initiatives implemented that are informed by data while the other two were on a journey to be more data-driven. In addition, the level of collaboration or a lack of it within an institution were found to impact the extent to which LA was adopted as stated by Tsai and Gasevic (2017). The findings also observed the drive and focus by the institutions to be more collaborative in sharing lessons from the various LA initiatives undertaken and for more cross faculty knowledge sharing to better adopt LA and implement insights derived from it. This supports M. Brown et al. (2020) claim that one of the interrelated factors in the maturation of LA is the cross collaboration within the institution which has been observed to be a driver at institutions within the study.

The study found that institutions were collecting large amounts of student data and that the LMSs were extensively used and found to be easy to use. The findings claim that the quality of the LMS data was deemed to be high and trusted by users. Participants highlighted many opportunities that the LMS presented in improving their understanding of student behaviours, in empowering them to identify students experiencing challenges early and introducing targeted interventions either to an individual student or a group experiencing similar challenges. The extensive use of the LMS data has been observed in studies to help educators in better supporting students and improving teaching and learning (Duin & Tham, 2020).

For LA to be effective, relevant interventions need to be implemented based on the insights from the data. The findings claim that academic staff had multiple responsibilities and required easy to understand tools that enabled them to answer the questions that they had using student data. The issue of capacity impacted the extent to which the academic staff were able to apply the relevant intervention in order to improve student outcomes. It was observed that some institutions did not have the capacity to implement the necessary changes or to intervene due to the responsibilities and expectations laden on the academic staff. Studies in LA have suggested that the ability to generate actionable and timely feedback is a goal of LA practice and the importance of applying interventions in order to support student outcomes is a measure of how effective LA was in the institution (Dawson et al., 2019; Tsai et al., 2020).

Complexity of the tools within the institution was found to be a challenge experienced by mainly educators thus making it difficult to use the tools available at the institution. The findings were supported by Ferguson et al. (2016) who suggests that there was limited or no consultation with teachers and educators on the features needed in the analytics tools with the focus primarily on implementation of a tool as opposed to meeting the users' needs. In addition, other studies have observed that input from academics who contribute to the design of the LA tools is rare (Wilson et al., 2017). Training and support on how the tools need to be used for them to add value were found to be lacking and a suggested area in order to improve adoption and encourage use of the LA tools (Ferguson et al., 2016).

Specialised skills and capability to analyse and interpret data, advanced skills in mining data and data science were observed to be in short supply in most of the institutions as they seek to apply more advanced mining techniques and answer new questions using the available data. This finding is corroborated by Ifenthaler (2017) who observed the difficulty of finding staff within institutions with specialised skills in data science as well as learning and teaching as these types of stakeholders were in short supply. Furthermore, it is unclear what skills are required for LA adoption and the number of people that currently possess them while previous studies suggest that the right skills are required for LA adoption (Ferguson et al., 2016). The study also found that institutions did not have staff members with multidisciplinary competencies in both learning design as well as data science who could drive LA initiatives. In addition, in some institutions where the skills did exist, it was observed that application was limited to a faculty or course and not widely shared or used across the institution. This finding is consistent with Ifenthaler (2017) finding that academic staff available for LA projects are in short supply and these are important in driving adoption.

The findings suggest that there is difficulty attributing the effectiveness of changes implemented within an institution to insights derived from LA and the extent to which improvements can be credited to LA. In addition, the findings suggest that insights from LA should not be viewed in isolation and other factors outside of the data including interventions taken need to be considered. This is supported by other findings that highlight the difficulty of measuring the effectiveness of LA because more often than not, LA initiatives are usually part of a wider programme within the institution incorporating other factors (Ifenthaler, Mah, & Yau, 2019; Wong, 2017).

The findings suggest that there is limited technological capability to support LA at some institutions for use of basic LA. While some of the institutions have taken steps to update their technologies to better support their LA needs, this was viewed as an area that would require strategic focus and funding. Other institutions in the study were found to be at a more mature level from a technological infrastructure as well as capability perspective with departments created to focus on LA and better use of data within the institution. The need for having technological capability has been observed as a challenge in LA studies as many HEIs have been found to lack the necessary technologies and capabilities such as dedicated analysts, human and financial resources for LA (Ifenthaler, 2017).

Ethics and safeguarding student data were found to be very important factors in the use of LA. The Consensus amongst all participants was that student privacy must be upheld at all times and there were responsibilities and ethical considerations that must be observed in the use of student data.

While there are legislative requirements in ensuring ethical considerations are observed, participants stressed the importance of adhering to the measures that institutions had in place and also identified areas where policies were needed in giving guidelines on using LA in an ethical manner in the institution. In their studies, Avella et al. (2016) and Viberg et al. (2018) claimed that ethical and privacy issues were key to adoption of LA yet they introduced challenges and complexities in their implementation that HEIs must balance with the need to improve learning outcomes. In addressing these challenges, Gasevic et al. (2019) recommended having policies in place within the institution on the use of student data and LA as these would provide a way for institutions to have guidelines and frameworks that can be followed in ensuring that ethical and privacy issues are observed. This study found that while the institutions did not all have LA specific policies in place, they all had intentions to develop the policies. Some institutions were already in the process of actively developing policies on the use of LA which would incorporate the ethical and privacy issues.

5.6 Conclusion

This chapter presented the findings of the study by answering the research questions posed. The chapter gave conclusions made on the current state of LA at HEIs in SA through detailed discussion of the findings and relating the findings back to the literature.

6 Conclusion

6.1 Introduction

This chapter concludes the study by briefly presenting the practical and theoretical implications of the study, highlighting the limitations of the study and offering recommendations for future research.

6.2 Summary of the Findings

To answer the research questions an integrated model of the updated DeLone and McLean IS success model and Technology-Organisation-Environment (TOE) framework was used. The model evaluates adoption of Learning Analytics (LA) at institutions by understanding the technological, organisational as well as environmental factors and the impact all of these have on the use of LA and the net benefits (positive or negative) derived from its use. There is limited research in LA from a South African (SA) context and this research aimed to address the gap in knowledge by investigating the extent to which LA is used at South African HEIs. This study aimed to gain an understanding of the state of LA by establishing whether student data is being collected and if LA is used to support decision making in order to improve teaching and learning outcomes. The study collected data from three public universities and one private online learning provider. Collection of data was done through 60-minute interviews with 33 participants in different roles to gain a diverse perspective in the use of LA. This chapter summarised and discussed the findings organised by the study research questions

This study has given a comprehensive analysis on the factors promoting and/or hindering adoption of LA. These factors were identified and mapped up using a theoretical lens from integrating the TOE framework and the updated DeLone and McLean IS success model and illustrating how the various factors encourage use of LA and the implications these have on the benefits derived from its use. These factors illustrated in Table 14 influenced the level of adoption within an institution and the extent of use of LA:

Factors Influencing Adoption	Relevance to SA HEI
Technological factors:	
Quality of the data	The quality of the data in the institution impacted the level of adoption. The main issues impacting adoption were the accuracy of the data, completeness of the data and consistency. While these factors were not unique to South Africa, the study found that these challenges existed in all but one institution.

Factors Influencing Adoption	Relevance to SA HEI
Quality of the systems in place	• The level of integration of the information systems was found to be an important factor and one that impacts the quality of the data. It was found that majority of the institutions in SA have fragmented systems in place which serve different data needs. The infrastructure drives use when the systems are integrated thus giving an accurate picture of the data which was found to be a goal for the institutions. • Availability of the information systems were a key driver in the use of LA. SA has encountered rolling blackouts as the national power utility tries to manage peak power supply. These have negatively impacted the institutions as learners and teachers are unable to access the information system.
Complexity of the LA tools in place	Tool complexity is a factor that has been found in literature as a challenge for HEIs seeking to adopt LA. The main issue with the SA institutions was a lack of training on the tools and a lack of consultation and involvement in the design of the tool. Institutions highlighted that the tools should be driven by pedagogy instead of technology in order to provide accurate results. Where courses were not reliant on the LMS different measures were used to predict learner performance and better understand behaviours.
Organisational factors	
Organisational culture	A collaborative culture was found to be a driver for LA adoption. Institutions in this study had different power structures where there was either a centralised decision-making structure or where each Faculty had full autonomy in decision-making. It was found that a centralised structure promotes adherence to decisions made. Adoption was found to be lagging at institutions with high autonomy in faculties and these institutions also collaborated less with similar LA initiatives happening at the same time within the institution. Knowledge sharing was also found to be low resulting in departments experiencing the same pitfalls as they embarked on their initiatives.
Level of focus on ethical and privacy issues	Ethical and privacy issues were found to be important for the institutions in SA. It was found that the institutions did not have an LA policy, even though they are at different levels of maturity in their adoption of LA, but they all agreed having one was important. It was also found that there is no general policy that exists within SA.
Capacity of staff in various roles	The study found that educators had multiple responsibilities in addition to their teaching responsibilities which left them little time to learn how the LA tools work in order to adopt them in their teaching. The COVID-19 pandemic increased the extent to which institutions were using data and also assisted educators in accessing the training material as the training was changed from in-person training to online training. Conducting training in person was found to be unique to SA and the pandemic has helped align SA HEIs with other international institutions in how training on the tools is administered.

Factors Influencing Adoption	Relevance to SA HEI
Level of executive support for LA initiatives	The extent to which the institution's executive supports the use of LA is a key driver in ensuring resources are made available, funding is allocated, and priority is given to LA initiatives. Institutions in this study at a more advanced level of adoptions were found to have a strong level of support in driving LA initiatives.
Environmental factors	
Regulatory requirements	Institutions in SA have a regulatory requirement to adhere to the Protection of Personal Information (POPI) Act and all institutions cited the measures they already have in place to adhere to the act.

Table 14: Factors Impacting Adoption in SA

In addition, integrating the updated DeLone and McLean IS success model highlights the importance of the use and usability of the LA tools that institutions have implemented. The extent to which users trust the data, are comfortable that the data and systems are of a high quality, and their capacity to interrogate and analyse the data motivates them to use the tools and see benefits from the insights. The benefits can, at times, be negative and it is important to illustrate both.

6.3 Implications of The Study

This section presents the implications of this research with regards to contribution to theory and practical contribution towards further understanding the state of LA in SA.

6.3.1 Theoretical Implications

Many studies have been conducted in the field of LA and while it is still an area that is maturing, studies focusing on the state of LA from an SA perspective are scarce – most studies focus on Europe, USA and Australia (Tsai et al., 2020; Viberg et al., 2018). The 2020 EDUCAUSE Horizon Report was one of the first to mention LA initiatives at HEI in South Africa and acknowledges that as more understanding of how LA is being used to improve learning and teaching, knowledge on results of implementations will start crossing both national and institutional boundaries (M. Brown et al., 2020). This study has given a high-level view of the emerging maturity of LA at four institutions by establishing to what extent institutions are collecting student data, whether the data is being used for LA and highlighting the challenges and opportunities that exist with the use of LA at these institutions. The study has established that many different types of data are, in fact, being collected and different tools and systems exist in its collection, analysis and reporting.

The TOE framework and the updated DeLone and McLean's IS success model guided the research methodology and analysis of the findings. Integrating the two models makes a contribution in terms of advancing research in the area of LA. It broadens the scope of understanding outcomes of LA by viewing initial implementation factors and ongoing continual application through use of the tools, learning from the insights and intervening where needed. This then increases knowledge in the use of LA at HEI and, as this study was conducted with South African institutions, it also increases knowledge from the SA context.

6.3.2 Practical implications

This study offers a perspective of where SA higher education institutions are in their adoption of LA. Online learning is not widely embraced in SA due to challenges within the country, many of which are social in nature and the education system in SA is predominantly face to face and online education is still perceived as expensive or not well adopted (Mpungose, 2020). The findings can help executive level stakeholders within HEI to better understand the challenges and barriers within institutions in their implementation and adoption of LA so they can design strategies, prioritise critical factors so resources can be allocated accordingly.

The findings can help non-academic staff to better understand the challenges faced by academics in their use of existing LA tools. This can encourage them to collaborate with academic staff when designing and implementing LA tools to ensure that they meet the needs of the users and that they are effective which will likely encourage adoption.

The findings have highlighted the need for LA policies within institutions and the importance of ensuring that the policy adheres to and supports regulations such as the Protection of Personal Information Act (POPIA), the proposed SA National Data and Cloud Policy (Published on 1 April 2021 by the Dept of Communications and Digital Technologies) as well as The General Data Protection Regulation (GDPR). These policies have a big focus on privacy an area that has continuously been highlighted in the findings as important. The policies should also include ethical issues that are specific in the handling of student data as well as responsibilities in the use of the data.

6.4 Limitations of the Study

At the start of this research, SA went into a lockdown due to the COVID-19 pandemic. This resulted many changes happening across institutions in the country with some institutions putting a hold on researchers requesting participation from their members of staff thus not granting permission for outside research. This study was impacted by the pandemic in that some institutions that were approached to participate declined due to the heavy workload on their staff members while others did not provide any feedback. This resulted in having only four institutions participating in this study. Though attempts were made to cover a broader number of institutions with different profiles, the challenges introduced by the pandemic unfortunately limited the number of participating institutions.

A second limitation, related to the pandemic, was that with the institutions that did provide approval for the study to be conducted, most of their staff members did not have capacity to participate due to the addition to their workload as most institutions had to go fully online thus changing their mode of teaching and support needed by students, while also managing other responsibilities. This also limited the types of roles that participated in the study due to these restrictions. One such role where limited involvement was achieved is from the executive level stakeholders. Access to executive level stakeholders was a major challenge and it is recommended that future research should seek to get participation from more members of the executive to get a balance of the various points of views.

Thirdly, of the institutions that did participate in the study, none had LA policies which would have been evaluated as part of this study to give a documented view of the guidelines in place at the institutions. The field of LA is still in its infancy and many institutions are refining their online learning, data and privacy policies and there is a need to also introduce new policies in line with their use of technologies and tools for LA. This limitation is an opportunity for institutions as they invest more in their online and data capabilities and can incorporate lessons from institutions that have implemented LA policies and where needed customise guidelines to meet their needs.

Lastly, while the study sought to understand use of LA in online learning, only one institution in the study was fully online while the rest had courses that were blended and still others offered Massive Open Online Courses (MOOCs). The COVID-19 pandemic did force all institutions to move to a fully online mode in a very short space of time and it is unclear the extent to which courses were purely in line with online learning guidelines. The study therefore has participants who teach purely online courses, blended and hybrid courses as well as MOOCs. A number of the institutions in SA offer a blended model of teaching and the cases in this study were no exception (Mpungose, 2020).

6.5 Suggestions for Future Research

Future research should continue to understand the state of LA at different HEIs in South Africa to further build on empirical evidence regarding the phenomenon. The study focused primarily on universities and due to limitations in gaining access to a bigger sample of institutions. The selected institutions are leading HEI with a strong LMS, Information Technology (IT) network and data infrastructure and good financial and human resources A wider study that investigates use of LA and maturity levels across different types of HEIs such as technical and vocational colleges, will provide different perspectives and insights through engaging with a more diverse sample of HEIs and will be more representative of HEI in the country. Other institutions may be lower down the Information and Communications Technology (ICT) maturity scale and not (be able to) use LA as well as the ones discussed in the study while others might lack the financial and human resources needed. Furthermore, research in LA in SA is still in its infancy with a large number of studies on the subject being in the global North and institutions being at varying maturity levels in their use of LA (Tsai et al., 2020; West et al., 2016), further studies within a South African context will add value to current body of knowledge.

This study aimed to establish whether LA is being used in SA and to what extent. Future research could explore in detail the effectiveness of LA use, the extent to which institutions use the insights and whether or not recommended interventions are implemented. In addition, future studies should investigate whether the interventions that are implemented lead to improved learner outcomes. The study did not focus on whether the actions that are taken are informed purely by the insights observed or if other factors are taken into consideration. These are areas that are unclear and require further research

A detailed analysis of the tools used at institutions requires focus. Research in this area could evaluate how decisions on which LA tools to use are made, the approach taken when deciding on how said tool will be procured, whether built by the institution or bought off the shelf, customisation of LA tools and the level of consultation if any in the choice and customisation of tools. It will be important to establish the extent to which pedagogy drives the choice of LA tools an institution invests in and the role learning science plays in their design and implementation. This is an area that requires focus as the LA tools play an important role in influencing adoption. Furthermore, evaluating LA implementations at the different types of HEIs, determining whether there is alignment in the goals of LA amongst different stakeholders and alignment on the measures of success with LA requires more study.

LA methods are used to predict student performance, identify students at risk of failing or dropping out based on student interactions with online resources, access to the institution and student performance and engagement. An opportunity for further research lies in understanding the effectiveness of these methods in promoting student retention, helping students improve their performance and engagement as well as supporting students in good academic standing who are at risk of dropping out.

Understanding whether students want to be classified as being at risk and whether there is a risk of algorithmic bias by looking at the variables used in defining at risk students.

Future research could also incorporate the student's perspective to understand the benefits students derive from LA, understand their view on consent, ethical and privacy concerns. Ethical and privacy issues present opportunities for future research with a focus on establishing an institution's attitudes on ethical and privacy issues and how these impact stakeholders in various roles within the institution and students. Principles for ethics and privacy in LA could be an area of further research from the context of SA. This is an area that is being explored in research in Europe, USA and Australia.

6.6 Research Conclusion

The objective of this study was to investigate the current status of LA in online learning at HEIs in SA. The following research questions were posed in meeting the research objective:

RQ1: What type of student data can be and is currently being collected?

RQ2: Is the data used for LA and if so, to what extent?

RQ3: Is LA used by administrators, educators and lecturers to inform decision making?

RQ4: What opportunities and barriers exist for more advanced LA?

This research recommended an integrated model to explain the current state of LA at HEIs in SA. The model provided environmental, technological and organisation factors that impact adoption of LA and furthermore offered reasons regarding the need for using analytics, the challenges that exist in its use along with the opportunities. Incorporating all these factors gave a view of the net benefits that are derived, these can be either positive or negative benefits.

The research found that many different types of student data are currently collected at HEIs in SA. Demographic data as well as a student's previous education data were collected and used for enrolment purposes. Student's financial and residential data was collected and used to support students and provide guidance where assistance was needed. Student's mental health data was collected but only accessible to counsellors and where students provide consent, this data was used by student advisors.

The LMS data was found to be critical in better understanding student engagement, performance, behaviours and in predicting learner outcomes. It also provided educators with key information to identify students at risk early, make decisions on various aspects including course design and the types of initiatives students needed in order to support students in meeting their learning outcomes. The various types of data collected empowered educators in decision-making and presented both opportunities and challenges. The tools used at some of the institutions were found to be a factor in the level of adoption.

The focus of this study was on HEIs in SA and found that some of the challenges and barriers being experienced were similar to those experienced in other countries where adoption is at a more mature level. Challenges around data quality, institutional culture and technological systems in place were similar across institutions. The challenges also presented opportunities with the executive level management seeking to increase the level at which data is used in their institutions as they started seeing the benefits derived from the insights that the data presented.

This study presented the current state of LA at 4 leading HEIs in SA. The findings were presented in line with the research question to establish credibility of the research. The implications of the study along with limitations and recommendations for future research were presented.

References

Acito, F., & Khatri, V. (2014). Business analytics: Why now and what next? *Business Horizons, 57*(5), 565-570.

Agarwal, R., & Dhar, V. (2014). Big data, data science, and analytics: The opportunity and challenge for IS research. *Information Systems Research, 25*(3), 443-448.

Aliyu, A. A., Bello, M. U., Kasim, R., & Martin, D. (2014). Positivist and non-positivist paradigm in social science research: Conflicting paradigms or perfect partners. *Journal of Management & Sustainability, 4*(3), 79-95.

Alom, B. M., & Courtney, M. (2018). Educational data mining: a case study perspectives from primary to university education in australia. *International Journal of Information Technology and Computer Science, 10*(2), 1-9.

Amaratunga, D., Baldry, D., Sarshar, M., & Newton, R. (2002). Quantitative and qualitative research in the built environment: application of "mixed" research approach. *Work study, 51*(1), 17-31.

Anderson, T. (2016). *Theories for learning with emerging technologies*. Edmonton, Canada: AU Press, Athabasca University.

Arnold, K. E., & Sclater, N. (2017). *Student perceptions of their privacy in learning analytics applications.* Paper presented at the Proceedings of the Seventh International Learning Analytics & Knowledge Conference, Vancouver, BC, Canada,.

Avella, J. T., Kebritchi, M., Nunn, S. G., & Kanai, T. (2016). Learning analytics methods, benefits, and challenges in higher education: A systematic literature review. *Online Learning, 20*(2), 13-29.

Bagarukayo, E., & Kalema, B. (2015). Evaluation of elearning usage in South African universities: A critical review. *International Journal of Education and Development Using ICT, 11*(2), 168-183.

Baker, J. (2012). The technology–organization–environment framework. In *Information Systems Theory* (pp. 231-245): Springer.

Baker, R. S., & Inventado, P. S. (2016). Educational Data Mining and Learning Analytics. In G. Veletsianos (Ed.), *Emergence and Innovation in Digital Learning: Foundationas and Applications*. Edmonton, Alberta: AU Press, Athabasca University.

Bakharia, A., Corrin, L., De Barba, P., Kennedy, G., Gašević, D., Mulder, R., Williams, D., Dawson, S., Lockyer, L. (2016, April 2016). *A conceptual framework linking learning design with learning analytics.* Paper presented at the Proceedings of the Sixth International Conference on Learning Analytics & Knowledge, Edinburgh.

Basole, R. C. (2014). Visual Business Ecosystem Intelligence: Lessons from the Field. *IEEE Computer Graphics and Applications, 34*(5), 26-34. doi:10.1109/MCG.2014.104

Bayrak, T. (2015). A Review of Business Analytics: A Business Enabler or Another Passing Fad. *Procedia - Social and Behavioral Sciences, 195*, 230-239. doi:https://doi.org/10.1016/j.sbspro.2015.06.354

Benbasat, I., Goldstein, D. K., & Mead, M. (1987). The case research strategy in studies of information systems. *MIS Quarterly, 11*(3), 369-386.

Bernard, H. R. (2017). *Research methods in anthropology: Qualitative and quantitative approaches*. Oxford, UK: Rowman & Littlefield.

Bhattacherjee, A. (2012). Social science research: Principles, methods, and practices. *Textbooks Collection, 3*.

Boddy, C. R. (2016). Sample size for qualitative research. *Qualitative Market Research: An International Journal, 19*(4), 426-432.

Bodily, R., Graham, C. R., & Bush, M. D. (2017). Online learner engagement: Opportunities and challenges with using data analytics. *Educational Technology, 57*(1), 10-18.

Bornman, E. (2016). Information society and digital divide in South Africa: results of longitudinal surveys. *Information, Communication & Society, 19*(2), 264-278. doi:10.1080/1369118X.2015.1065285

Bose, R., & Luo, X. (2011). Integrative framework for assessing firms' potential to undertake Green IT initiatives via virtualization–A theoretical perspective. *The Journal of Strategic Information Systems, 20*(1), 38-54.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77-101.

Brown, C., Czerniewicz, L., & Noakes, T. (2016). Online content creation: Looking at students' social media practices through a connected learning lens. *Learning, Media and Technology, 41*(1), 140-159.

Brown, M., McCormack, M., Reeves, J., Brooks, C., Grajek, S., Alexander, B., . . . Engelbert, N. (2020). EDUCAUSE Horizon Report. *Teaching and Learning Edition, Louisville, CO: EDUCAUSE*.

Cao, G., Duan, Y., & Li, G. (2015). Linking business analytics to decision making effectiveness: A path model analysis. *IEEE Transactions on Engineering Management, 62*(3), 384-395.

Chen, H.-J. (2010). Linking employees'e-learning system use to their overall job outcomes: An empirical study based on the IS success model. *Computers & Education, 55*(4), 1628-1639.

Clemmensen, T., Kaptelinin, V., & Nardi, B. (2016). Making HCI theory work: an analysis of the use of activity theory in HCI research. *Behaviour & Information Technology, 35*(8), 608-627.

Creswell, J. W. (2003). *Research design: Qualitative, quantitative, and mixed methods approaches* (2 ed.). Thousand Oaks, California: Sage publications.

Creswell, J. W. (2007). *Qualitative inquiry and research design: Choosing among five approaches* (2 ed.). Thousand Oaks, California: Sage publications.

Creswell, J. W., Hanson, W. E., Clark Plano, V. L., & Morales, A. (2007). Qualitative research designs: Selection and implementation. *The Counseling Psychologist, 35*(2), 236-264.

Cronin, P., Ryan, F., & Coughlan, M. (2008). Undertaking a literature review: a step-by-step approach. *British Journal of Nursing, 17*(1), 38-43.

Darke, P., Shanks, G., & Broadbent, M. (1998). Successfully completing case study research: combining rigour, relevance and pragmatism. *Information Systems Journal, 8*(4), 273-289.

Datnow, A., & Hubbard, L. (2016). Teacher capacity for and beliefs about data-driven decision making: A literature review of international research. *Journal of Educational Change, 17*(1), 7-28.

Dawson, S., Gašević, D., Siemens, G., & Joksimovic, S. (2014). *Current state and future trends: A citation network analysis of the learning analytics field.* Paper presented at the Proceedings of the fourth international conference on learning analytics and knowledge, Indianapolis Indiana USA.

Dawson, S., Joksimovic, S., Poquet, O., & Siemens, G. (2019). *Increasing the Impact of Learning Analytics*. Paper presented at the Proceedings of the 9th International Conference on Learning Analytics & Knowledge, Tempe, AZ, USA.

Dawson, S., & Siemens, G. (2014). Analytics to literacies: The development of a learning analytics framework for multiliteracies assessment. *The International Review of Research in Open and Distributed Learning, 15*(4).

Dede, C. J. (2016). Next steps for "Big Data" in education: Utilizing data-intensive research. *Educational Technology, 56*(2), 37-42.

Delen, D., & Ram, S. (2018). Research challenges and opportunities in business analytics. *Journal of Business Analytics, 1*(1), 2-12. doi:10.1080/2573234X.2018.1507324

DeLone, W. H., & McLean, E. R. (1992). Information systems success: The quest for the dependent variable. *Information Systems Research, 3*(1), 60-95.

DeLone, W. H., & McLean, E. R. (2003). The DeLone and McLean model of information systems success: a ten-year update. *Journal of Management Information Systems, 19*(4), 9-30.

DeLone, W. H., & Mclean, E. R. (2004). Measuring e-commerce success: Applying the DeLone & McLean information systems success model. *International Journal of Electronic Commerce, 9*(1), 31-47.

Deming, D. J., Goldin, C., Katz, L. F., & Yuchtman, N. (2015). Can online learning bend the higher education cost curve? *American Economic Review, 105*(5), 496-501.

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

Dewan, S., Michael, S. C., & Min, C.-K. (1998). Firm characteristics and investments in information technology: Scale and scope effects. *Information Systems Research, 9*(3), 219-232.

Duan, Y., Cao, G., & Edwards, J. S. (2018). Understanding the impact of business analytics on innovation. *European Journal of Operational Research, 281*(3), 673-686. doi:https://doi.org/10.1016/j.ejor.2018.06.021

Duin, A. H., & Tham, J. (2020). The current state of analytics: Implications for learning management system (LMS) use in writing pedagogy. *Computers and Composition, 55*, 1-23.

Ferguson, R., Brasher, A., Clow, D., Cooper, A., Hillaire, G., Mittelmeier, J., . . . Vuorikari, R. (2016). Research evidence on the use of learning analytics: Implications for education policy. *Joint Research Centre*.

Gangwar, H., Date, H., & Ramaswamy, R. (2015). Understanding determinants of cloud computing adoption using an integrated TAM-TOE model. *Journal of Enterprise Information Management, 28*(1), 107-130.

Gasevic, D., Dawson, S., Mirriahi, N., & Long, P. D. (2015). Learning analytics - A growing field and community engagement. *Journal of Learning Analytics, 2*(1), 1-6.

Gašević, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. *TechTrends, 59*(1), 64-71. doi:10.1007/s11528-014-0822-x

Gasevic, D., Tsai, Y.-S., Dawson, S., & Pardo, A. (2019). How do we start? An approach to learning analytics adoption in higher education. *The International Journal of Information and Learning Technology, 36*(4), 342-353. doi:http://dx.doi.org.ezproxy.uct.ac.za/10.1108/IJILT-02-2019-0024

Gillham, B. (2000). *Case study research methods*. London: Bloomsbury Publishing.

Goldie, J. G. S. (2016). Connectivism: A knowledge learning theory for the digital age? *Medical Teacher, 38*(10), 1064-1069.

Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. *Handbook of Qualitative Research, 2*(163-194), 105.

Guillemin, M., & Gillam, L. (2004). Ethics, reflexivity, and "ethically important moments" in research. *Qualitative Inquiry, 10*(2), 261-280.

Halawi, L., & McCarthy, R. (2006). Which theory applies: An analysis of information systems research. *Issues in information systems, 7*(2), 252.

Hennink, M. M., Kaiser, B. N., & Marconi, V. C. (2016). Code Saturation Versus Meaning Saturation: How Many Interviews Are Enough? *Qualitative Health Research, 27*(4), 591-608. doi:10.1177/1049732316665344

Holsapple, C., Lee-Post, A., & Pakath, R. (2014). A unified foundation for business analytics. *Decision Support Systems, 64*, 130-141.

Hwang, G.-J., Spikol, D., & Li, K.-C. (2018). Guest Editorial: Trends and Research Issues of Learning Analytics and Educational Big Data. *Journal of Educational Technology & Society, 21*(2), 134-136.

Ifenthaler, D. (2017). Are higher education institutions prepared for learning analytics? *TechTrends, 61*(4), 366-371.

Ifenthaler, D., Mah, D.-K., & Yau, J. Y.-K. (2019). Utilising learning analytics for study success: Reflections on current empirical findings. In *Utilizing Learning Analytics to Support Study Success* (pp. 27-36): Springer.

Kaliisa, R., & Picard, M. (2017). A Systematic Review on Mobile Learning in Higher Education: The African Perspective. *Turkish Online Journal of Educational Technology-TOJET, 16*(1), 1-18.

Klein, C., Lester, J., Rangwala, H., & Johri, A. (2019). Learning analytics tools in higher education: Adoption at the intersection of institutional commitment and individual action. *The Review of Higher Education, 42*(2), 565-593.

Knight, S., & Shum, S. B. (2017). Theory and learning analytics. *The Handbook of Learning Analytics*, 17-22.

Knight, S., Shum, S. B., & Littleton, K. (2014). Epistemology, assessment, pedagogy: where learning meets analytics in the middle space. *Journal of Learning Analytics, 1*(2), 23-47-23-47.

Kotoua, S., Ilkan, M., & Kilic, H. (2015). *The Growing of Online Education in Sub Saharan Africa: Case Study Ghana* (Vol. 191).

Kozulin, A., Ageyev, V. S., Gindis, B., & Miller, S. M. (2003). *Vygotsky's Educational theory in cultural context*. New York: Cambridge University Press.

Krishnamoorthi, S., & Mathew, S. K. (2018). Business analytics and business value: A comparative case study. *Information & Management, 55*(5), 643-666.

Kruss, G., McGrath, S., Petersen, I.-h., & Gastrow, M. (2015). Higher education and economic development: The importance of building technological capabilities. *International Journal of Educational Development, 43*, 22-31.

Lester, J., Klein, C., Rangwala, H., & Johri, A. (2017). Learning Analytics in Higher Education. *ASHE Higher Education Report, 43*(5), 1-145. doi:10.1002/aehe.20121

Lin, A. C. (1998). Bridging positivist and interpretivist approaches to qualitative methods. *Policy Studies Journal, 26*(1), 162-180.

Macfadyen, L. P., Dawson, S., Pardo, A., & Gaševic, D. (2014). Embracing big data in complex educational systems: The learning analytics imperative and the policy challenge. *Research & Practice in Assessment, 9*, 17-28.

Marsh, J. A., & Farrell, C. C. (2015). How leaders can support teachers with data-driven decision making: A framework for understanding capacity building. *Educational Management Administration & Leadership, 43*(2), 269-289.

Marshall, M. N. (1996). Sampling for qualitative research. *Family Practice, 13*(6), 522-526.

Mason, R. O. (1978). Measuring information output: A communication systems approach. *Information & Management, 1*(4), 219-234.

Melville, N., Kraemer, K., & Gurbaxani, V. (2004). Information technology and organizational performance: An integrative model of IT business value. *MIS Quarterly, 28*(2), 283-322.

Mohammadi, H. (2015). Investigating users' perspectives on e-learning: An integration of TAM and IS success model. *Computers in Human Behavior, 45*, 359-374.

Morgan, D. L. (2013). *Integrating qualitative and quantitative methods: A pragmatic approach*. Thousand Oaks, California: Sage publications.

Mpungose, C. B. (2020). Emergent transition from face-to-face to online learning in a South African University in the context of the Coronavirus pandemic. *Humanities and Social Sciences Communications, 7*(1), 1-9.

Mtebe, J. (2015). Learning Management System success: Increasing Learning Management System usage in higher education in sub-Saharan Africa. *International Journal of Education and Development using ICT, 11*(2).

Ng'ambi, D., Brown, C., Bozalek, V., Gachago, D., & Wood, D. (2016). Technology enhanced teaching and learning in South African higher education–A rearview of a 20 year journey. *British Journal of Educational Technology, 47*(5), 843-858.

Oliveira, T., & Martins, M. F. (2011). Literature review of information technology adoption models at firm level. *Electronic Journal of Information Systems Evaluation, 14*(1), 110.

Orlikowski, W. J., & Baroudi, J. J. (1991). Studying information technology in organizations: Research approaches and assumptions. *Information Systems Research, 2*(1), 1-28.

Ormston, R., Spencer, L., Barnard, M., & Snape, D. (2014). The foundations of qualitative research. *Qualitative Research Practice: A Guide for Social Science Students and Researchers, 2*, 52-55.

Oyo, B., & Kalema, B. M. (2014). Massive open online courses for Africa by Africa. *International Review of Research in Open and Distance Learning, 15*(6), 1-13.

Padilla-Díaz, M. (2015). Phenomenology in educational qualitative research: Philosophy as science or philosophical science. *International Journal of Educational Excellence, 1*(2), 101-110.

Palvia, S., Aeron, P., Gupta, P., Mahapatra, D., Parida, R., Rosner, R., & Sindhi, S. (2018). Online Education: Worldwide Status, Challenges, Trends, and Implications. *Journal of Global Information Technology Management, 21*(4), 233-241. doi:10.1080/1097198X.2018.1542262

Papamitsiou, Z., & Economides, A. A. (2014). Learning analytics and educational data mining in practice: A systematic literature review of empirical evidence. *Journal of Educational Technology & Society, 17*(4), 49-64.

Pardo, A., Jovanovic, J., Dawson, S., Gašević, D., & Mirriahi, N. (2019). Using learning analytics to scale the provision of personalised feedback. *British Journal of Educational Technology, 50*(1), 128-138.

Paré, G. (2004). Investigating information systems with positivist case research. *Communications of the association for information systems, 13*(1), 18.

Patton, M. Q. (2005). Qualitative research. *Encyclopedia of Statistics in Behavioral Science, 3*, 1633–1636.

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of Management Information Systems, 24*(3), 45-77.

Petter, S., DeLone, W., & McLean, E. R. (2012). The past, present, and future of" IS Success". *Journal of the Association for Information Systems, 13*(5), 341.

Picciano, A. G. (2012). The evolution of big data and learning analytics in American higher education. *Journal of Asynchronous Learning Networks, 16*(3), 9-20.

Poitras, E. G., Lajoie, S. P., Doleck, T., & Jarrell, A. (2016). Subgroup Discovery with User Interaction Data: An Empirically Guided Approach to Improving Intelligent Tutoring Systems. *Journal of Educational Technology & Society, 19*(2).

Power, D., Heavin, C., McDermott, J., & Daly, M. (2018). Defining business analytics: an empirical approach. *Journal of Business Analytics, 1*(1), 40-53.

Prinsloo, P., & Slade, S. (2017). Ethics and Learning Analytics: Charting the (Un)Charted. *In: Lang, C.; Siemens, G.; Wise, A. & Gašević, D. eds. Handbook of Learning Analytics. SOLAR,* 49-57. doi:http://dx.doi.org/doi:10.18608/hla17

Prinsloo, P., Slade, S., & Khalil, M. (2018, June). *Stuck in the middle? Making sense of the impact of micro, meso and macro institutional, structural and organisational factors on implementing learning analytics.* Paper presented at the European Distance and E-Learning Network Annual Conference, Genova, Italy.

Reyes, J. A. (2015). The skinny on big data in education: Learning analytics simplified. *TechTrends, 59*(2), 75-80.

Riggins, F. J., & Wamba, S. F. (2015, 5-8 Jan. 2015). *Research Directions on the Adoption, Usage, and Impact of the Internet of Things through the Use of Big Data Analytics.* Paper presented at the 2015 48th Hawaii International Conference on System Sciences.

Rogers, T., Gašević, D., & Dawson, S. (2016). Learning analytics and the imperative for theory driven research. *The SAGE Handbook of E-Learning Research*, 232-250.

Rubel, A., & Jones, K. M. L. (2016). Student privacy in learning analytics: An information ethics perspective. *The Information Society, 32*(2), 143-159. doi:10.1080/01972243.2016.1130502

Ruipérez-Valiente, J. A., Muñoz-Merino, P. J., Leony, D., & Kloos, C. D. (2015). ALAS-KA: A learning analytics extension for better understanding the learning process in the Khan Academy platform. *Computers in Human Behavior, 47*, 139-148.

Saggi, M. K., & Jain, S. (2018). A survey towards an integration of big data analytics to big insights for value-creation. *Information Processing & Management, 54*(5), 758-790.

Saunders, M., Lewis, P., & Thornhill, A. (2016). *Research Methods for Business students.* (7 ed.). Harlow, Essex: Pearson Education Limited.

Scheffel, M., Drachsler, H., Stoyanov, S., & Specht, M. (2014). Quality indicators for learning analytics. *Journal of Educational Technology & Society, 17*(4), 117-132.

Schoenherr, T., & Speier-Pero, C. (2015). Data science, predictive analytics, and big data in supply chain management: Current state and future potential. *Journal of Business Logistics, 36*(1), 120-132.

Sclater, N., Peasgood, A., & Mullan, J. (2016). *Learning analytics in higher education: A review of UK and international practice*. Retrieved from https://www.jisc.ac.uk/sites/default/files/learning-analytics-in-he-v2_0.pdf

Scott, I. (2018). Designing the South African higher education system for student success. *Journal of Student Affairs in Africa, 6*(1), 1-17.

Seddon, P. B. (1997). A respecification and extension of the DeLone and McLean model of IS success. *Information Systems Research, 8*(3), 240-253.

Seddon, P. B., Constantinidis, D., Tamm, T., & Dod, H. (2017). How does business analytics contribute to business value? *Information Systems Journal, 27*(3), 237-269.

Shanks, G. (2002). Guidelines for conducting positivist case study research in information systems. *Australasian Journal of Information Systems, 10*(1), 76-85.

Shannon, C. E., & Weaver, W. (1998). *The mathematical theory of communication*. Illinois: University of Illinois press.

Sharma, R., Mithas, S., & Kankanhalli, A. (2014). Transforming decision-making processes: a research agenda for understanding the impact of business analytics on organisations. In: Taylor & Francis.

Siemens, G., & Long, P. (2011). Penetrating the fog: Analytics in learning and education. *EDUCAUSE Review, 46*(5), 30.

Sin, K., & Muthu, L. (2015). Application of big data in education data mining and learning analytics: A literature review. *ICTACT Journal on Soft Computing, 5*(4), 1035-1049.

Stefanovic, D., Marjanovic, U., Delić, M., Culibrk, D., & Lalic, B. (2016). Assessing the success of e-government systems: An employee perspective. *Information & Management, 53*(6), 717-726.

Straub, D., Boudreau, M.-C., & Gefen, D. (2004). Validation guidelines for IS positivist research. *Communications of the Association for Information Systems, 13*(1), 24.

Swartz, R., Ivancheva, M., Czerniewicz, L., & Morris, N. P. (2019). Between a rock and a hard place: dilemmas regarding the purpose of public universities in South Africa. *Higher Education, 77*(4), 567-583.

Thanh, N. C., & Thanh, T. (2015). The interconnection between interpretivist paradigm and qualitative methods in education. *American Journal of Educational Science, 1*(2), 24-27.

Torraco, R. J. (2016). Writing integrative literature reviews: Using the past and present to explore the future. *Human Resource Development Review, 15*(4), 404-428.

Tsai, Y.-S., & Gasevic, D. (2017, Mar 13-17 2017). *Learning analytics in higher education --- challenges and policies: a review of eight learning analytics policies.* Paper presented at the Learning Analytics & Knowledge conference, Simon Fraser University, Vancouver, Canada.

Tsai, Y.-S., Rates, D., Moreno-Marcos, P. M., Muñoz-Merino, P. J., Jivet, I., Scheffel, M., . . . Gašević, D. (2020). Learning analytics in European higher education–trends and barriers. *Computers & Education*, 103933.

Tshabalala, M., Ndeya-Ndereya, C., & van der Merwe, T. (2014). Implementing Blended Learning at a Developing University: Obstacles in the Way. *The Electronic Journal of e-Learning, 12*(1), 101-110.

Urbach, N., & Müller, B. (2012). The updated DeLone and McLean model of information systems success. In *Information Systems Theory* (pp. 1-18): Springer.

Veletsianos, G. (2016). *Emergence and innovation in digital learning: Foundations and applications*. Edmonton: Athabasca University Press.

Viberg, O., Hatakka, M., Bälter, O., & Mavroudi, A. (2018). The current landscape of learning analytics in higher education. *Computers in Human Behavior, 89*, 98-110. doi:10.1016/j.chb.2018.07.027

Vidgen, R., Shaw, S., & Grant, D. B. (2017). Management challenges in creating value from business analytics. *European Journal of Operational Research, 261*(2), 626-639. doi:https://doi.org/10.1016/j.ejor.2017.02.023

Walliman, N. (2011). *Research methods: The basics*. Milton Park, Abingdon, Oxon: Routledge.

Wamba, S. F., Akter, S., Edwards, A., Chopin, G., & Gnanzou, D. (2015). How 'big data'can make big impact: Findings from a systematic review and a longitudinal case study. *International Journal of Production Economics, 165*, 234-246.

Wamba, S. F., Gunasekaran, A., Akter, S., Ren, S. J.-f., Dubey, R., & Childe, S. J. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Business Research, 70*, 356-365. doi:https://doi.org/10.1016/j.jbusres.2016.08.009

Wang, Y., Kung, L., & Byrd, T. A. (2018). Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations. *Technological Forecasting and Social Change, 126*, 3-13.

West, D., Heath, D., & Huijser, H. (2016). Let's talk learning analytics: A framework for implementation in relation to student retention. *Online Learning, 20*(2), 1-21.

Wilder, C. R., & Ozgur, C. O. (2015). Business analytics curriculum for undergraduate majors. *INFORMS Transactions on Education, 15*(2), 180-187.

Wilson, A., Watson, C., Thompson, T. L., Drew, V., & Doyle, S. (2017). Learning analytics: Challenges and limitations. *Teaching in Higher Education, 22*(8), 991-1007.

Wise, A. F., & Shaffer, D. W. (2015). Why theory matters more than ever in the age of big data. *Journal of Learning Analytics, 2*(2), 5-13.

Wong, B. T. M. (2017). Learning analytics in higher education: an analysis of case studies. *Asian Association of Open Universities Journal, 12*(1), 21-40. doi:doi:10.1108/AAOUJ-01-2017-0009

Wright, G., Cillers, L., Van Niekerk, E., Seekoe, E., & International Association for Development of the Information, S. (2017). *The Next Stage of Development of eLearning at UFH in South Africa*. Retrieved from http://ezproxy.uct.ac.za/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=ED579391&site=ehost-live

Wu, J.-H., & Wang, Y.-M. (2006). Measuring KMS success: A respecification of the DeLone and McLean's model. *Information & Management, 43*(6), 728-739.

Xiang, Z., Schwartz, Z., Gerdes Jr, J. H., & Uysal, M. (2015). What can big data and text analytics tell us about hotel guest experience and satisfaction? *International Journal of Hospitality Management, 44*, 120-130.

Zhang, C., & Dhaliwal, J. (2009). An investigation of resource-based and institutional theoretic factors in technology adoption for operations and supply chain management. *International Journal of Production Economics, 120*(1), 252-269.

Zhang, J., Zhang, X., Jiang, S., Ordóñez de Pablos, P., & Sun, Y. (2018). Mapping the study of learning analytics in higher education. *Behaviour & Information Technology, 37*(10-11), 1142-1155.

Zhang, J.-H., Zhang, Y.-X., Zou, Q., & Huang, S. (2018). What learning analytics tells us: Group behavior analysis and individual learning diagnosis based on long-term and large-scale data. *Journal of Educational Technology & Society, 21*(2), 245-258.

Appendices

Appendix A Ethics Committee Approval Letter

Faculty of Commerce
Private Bag X3, Rondebosch, 7701
2.26 Leslie Commerce Building, Upper Campus
Tel: +27 (0) 21 650 4375/ 5748 Fax: +27 (0) 21 650 4369
E-mail: com-faculty@uct.ac.za
Internet: www.uct.ac.za
@Commerce UCT **UCT Commerce Faculty Office**

We are pleased to inform you that your ethics application has been approved. Unless otherwise specified this ethical clearance is valid until

Your clearance may be renewed upon application.

Please be aware that you need to notify the Ethics Committee immediately should any aspect of your study regarding the engagement with participants as approved in this application, change. This may include aspects such as changes to the research design, questionnaires, or choice of participants.

The ongoing ethical conduct throughout the duration of the study remains the responsibility of the principal investigator.

We wish you well for your research.

Commerce Research Ethics Chair
University of Cape Town
Commerce Faculty Office
Room 2.26 | Leslie Commerce Building

Office Telephone: +27 (0)21 650 2695 / 4375
Office Fax: +27 (0)21 650 4369
E-mail: com-faculty@uct.ac.za
Website: https://www.commerce.uct.ac.za/Pages/Ethics-in-Research

"Our Mission is to be an outstanding teaching and research university, educating for life and addressing the challenges facing our society."

Appendix B Interview Protocol

Department of Information Systems
Leslie Commerce Building
Engineering Mall, Upper Campus
OR
Private Bag X3 - Rondebosch - 7701
Tel: +27 (0) 21 650 2261 Fax: +27 (0) 21650 2280
Internet: http://www.commerce.uct.ac.za/informationsystems/

Personal Information	
1.	What is your role and responsibilities in your institution?
Student data collection	
1	What type of student data is collected by the institution/faculty/department?
2	o Are there specific times and events used to collect student & institution level data
3	What is the purpose of collecting student data? How is the data used?
4	Are there any challenges in collecting student data? o Are any of these challenges influenced by how the institution functions/operates?
5	What are the constraints or issues of collecting such data?
6	What benefits does the data provide? If any?
7	Have there been any relevant findings in the last 4/5 years? o What highlights and key observations have stood out?
8	Have the findings resulted in any changes or interventions within the faculty/institution? If so what are some of the changes?
9	Do any barriers or challenges exist in the implementation of the changes?
10	Have the changes or interventions proven to be effective? o How has this been measured?
11	Is the data shared or shareable across the institution?
Technology Use	
1	What are the objectives of analytics at the institution?
2	What analytics systems are used? o What systems are used for collecting, analysing, interpreting and reporting on the data?
3	Are analytics tools and dashboards used at the institution? If so what are they and are they easily accessible and usable?
4	Are the data collection tools or systems easy to use and understand?
5	What are the challenges in using these systems?
6	Is the data sufficient for the intended purpose?
7	Do you trust the integrity of the data?
Organisation/environment	
1	Are there outcomes with regards to incorporating data at the institution?
2	Are there future analytics driven innovations at the institution that aim to use student data?
3	Are these championed by the institution's senior management or run independently at a departmental level?
4	Does the institution's culture support use of analytics in decision making and bringing about change?

"Our Mission is to be an outstanding teaching and research university, educating for life and addressing the challenges facing our society."

Appendix C Consent Letter

Department of Information Systems
Leslie Commerce Building
Engineering Mall, Upper Campus
OR
Private Bag X3 - Rondebosch - 7701
Tel: +27 (0) 21 650 2261 Fax: +27 (0) 21650 2280
Internet: http://www.commerce.uct.ac.za/informationsystems/

dd mmm yyyy

<u>Request to conduct research and interview participation consent form</u>

Dear Sir/Madam,

In terms of the requirements for completing a Masters Degree in Information Systems at the University of Cape Town a research study is required.

The researcher, in this case Palesa Maralitle Molokeng, has chosen to conduct a case study entitled "Investigate the use of learning analytics in online learning at South Africa's higher education institutions (HEI)". The objective of the research is to investigate the current status of LA in online learning at HEI in South Africa, with the aim of identifying future potential of the use of LA in informing decision making, predicting learner outcomes and improving the overall learning experience for students.

Participation in this study will contribute towards gaining insight in the level of adoption of LA, an understanding of the drivers for use of LA, perceived benefits of use and extent to which LA are used to improve learning in South Africa. This study will provide you, as a participant an awareness of the current LA adoption landscape in South Africa. Understanding your institution as well as other institution's level of adoption will create opportunities of collaboration and learnings across the industry.

Your participation in this research is voluntary. All information will be treated in a confidential manner and used exclusively for the purpose of this study. No individual names will be recorded or published. Identifiable information will not be required for this study and all responses will be confidential and used for the purposes of this research only. Collected data will be securely stored for a minimum of five years and access to the data will be limited to the research team only. You can choose to withdraw from the research at any time for whatever reason, in accordance with ethical research requirements. Findings of the study will be made available on request once the study has been completed.

The data collection method will be document analysis as well as one-on-one interviews with a small group of the staff responsible for monitoring student and course performance, decision making on the type of data collected and tools used within the institution. The document analysis and interviews will be conducted at the institution's premises and will last 60 minutes. If you are willing to participate in this study, kindly sign the attached form and return to me at your earliest convenience.

Should you have any questions regarding this research, please feel free to contact me on 0835893397 or email: mlkpal002@myuct.ac.za.

Your participation in this study would be greatly appreciated but is entirely voluntary.

University of Cape Town • iYunivesithi yaseKapa • Universiteit van Kaapstad

SPES BONA

Department of Information Systems

Leslie Commerce Building
Engineering Mall, Upper Campus
OR
Private Bag X3 - Rondebosch - 7701
Tel: +27 (0) 21 650 2261 Fax: +27 (0) 21650 2280
Internet: http://www.commerce.uct.ac.za/informationsystems/

Sincerely,

Palesa Maralitle Molokeng

Maralitle

M.Com Student, (UCT)
Department of Information Systems
University of Cape Town
Email: mlkpal002@myuct.ac.za

Professor Jean-Paul Van Belle

Research Supervisor
Department of Information Systems
University of Cape Town
Email: jean-paul.vanbelle@uct.ac.za

Department of Information Systems
Leslie Commerce Building
Engineering Mall, Upper Campus
OR
Private Bag X3 - Rondebosch - 7701
Tel: +27 (0) 21 650 2261 Fax: +27 (0) 21650 2280
Internet: http://www.commerce.uct.ac.za/informationsystems/

Research Participant Consent Form

I, ____________________________________, consent to participate in the research on Investigate the use of learning analytics in online learning at South Africa's higher education institutions

I am aware that participation is voluntary and that I may choose to withdraw from this study at any time, should I choose to do so.

__________________________ __________________________

Signature Date

Appendix D Excerpt of a transcribed interview

Duration	70min	Participant Title	Deputy Director E-Learning
Case	InstD	Date	31/08/2020

SPEAKERS

Interviewer – PM

Participant – JC

PM

What type of student data is collected?

JC

So easier to show to show and then talk through it. Okay so this is this is an approach where, it will answer your question. So you can see predominantly, we've got two data sources, we've got the LMS data which be use in [LMS system] then we've got data from [SIS system]. From the LMS we derived other data like grade marks and so forth then we've got also data in our student readiness survey in our Learner Case Management System, in our lecture evaluations system so the data collected therefore, is specifically related to the. Then there's data specifically related to biographical data, course demographics, course trends, course activity data and so it's real time and cohort analysis. So we also integrate our HEDA system, you also got the HEDA system, that is the institutional data analysis system so that is the the overview of data. So in layman's terms it is your student information system, it is your LMS system data and therefore it's derived data from the two sources that's including predictive data and descriptive data.

PM

And when you say predictive and descriptive data, may you please elaborate on that?

JC

Yes Okay. So we using a predictive model, including again resources, data sources from student historical data and student current activity and Through random forest machine learning algorithm, we derived on a daily basis, the predictive possibility, which will indicate what is the probability of the student being able to pass that specific course with a percentage of 51% and higher. The descriptive data includes thefore the data derived from the LMS activity, and from the HEDA cohort analysis data. So, there's learning analytics and academic analytics involved here.

PM

So, who uses that

JC

okay who use the data. Every every lecturer has got access to a wealth of resources in his or her LMS module, including detailed descriptive data, comparing student performance against peers in the module and the predictive data analysing in the module that student probability of passing, which indicate the trend as the week's progress. So if it's in the third week, like last night was this week, week 35, was ETL went through the ETL process, and this morning everybody has got therefore kind of a 24/12 hour delayed prediction model of their students probability of passing. Now that you need to see in the perspective of the focus of the data, it does not include emotional or other sources such as library access or campus access. It's purely focusing on a biographical data and the current activity data in the LMS. There are elements of socio economic situation elements included in that data but the array of sources we use to pull into that prediction model is quite scary or fields we use to pull into that prediction model.

PM

Okay, I'm just, I just need to clarify two things. So, so you said the focus of the data does not include social aspects?

JC

Well, it includes but based on the the sources we have available, for example, the NSFAS category is delayed. I mean, in the beginning of the year, we only get the NSFAS data in March, I think, if I'm not mistaken, so it's delayed. So it depends where we are in the year what type of information we have. I mean, yeah, so that's, and you asked me actually who's using it, so it's the lecturers, it's a student, faculty administrators, it is deputy Dean's, it is HODs and it's the executive. So but the executive purely for informed decision-making purposes and on that aspect I'm more referring to teaching analytics and learning analytics.

PM

You mentioned that the purpose of collecting the data is for predict for predictive reasons and descriptive reasons. Am I correct in my understanding?

JC

So it's actually linked to our Student Success Initiative. So we want to empower every role player in a teaching and learning environment with additional data sources to make informed decisions. In other words, let me use the last few months as an example. One of the key variables which made a significant impact on our student support initiatives was for example, and it boils down to descriptive data and nothing with predictive data, boils down to student access to technology. So in this case, we use the descriptive data to analyse what devices the students are using to access the LMS. From that we derived the lack of access to devices and that therefore impacting our support. Taking it more closely to teaching and learning it allows the lecturer in the beginning of the course and as the course progress, to monitor student activity and having the additional benefit of having this significant predictive model at the back end providing additional information. So while the student may be performing well, and of course in predictive analytics you always got your false positive and false negatives but that doesn't matter, while this the course progress, the lecturer therefore can track then the student's are risk indicator. At the institutional level, the faculty student advisors have access to course prediction and cumulative prediction, in other words, combined prediction. So there can be two predictive indicators for the student to pass the course and for the student aggregated prediction in all his modules. He may have a medium risk in a course but he might have a high risk across all these other modules. That's that's all it takes.

PM

And are there any challenges in terms of collecting the data or using?

JC

Yeah, I think, the as you know it relies on the whole thing of garbage in garbage out, so we must make sure that when the data is collected. Let's say I'm looking at an email here where I'm just requiring an additional source of data to be added to our data source but the bottom line is if we lack information, and I'm going to make a very simple explanation again, I mean, the mere fact that we lacking in the beginning of the term the financial status data of a student, we may have indication on our student academic readiness survey of perceived social economic status, but it may not rely on truth and therefore if we have bursary data coming in late we have in as far as data coming in late that impacting us. It also imply that our our [SIS] system which we use for student information system, must have accurate data. So those people who capture and luckily for us, we've got online enrolment like the majority of universities, so therefore, the student capturing the data or withholding any data may have an impact. But as you listen to the sources of data, and majority of sources have got a high percent of accuracy because it is compulsory for these sources to be captured.

PM

And are there any challenges experienced by the various users of the data in terms of them understanding it or using it?

JC

Yes. Well, I think that the first one is applicable. So it's contextual understanding of what does this mean and look, the way the data is derived and the data is described have got certain meanings. If I'm looking at a specific report and there's a specific measurement which means something so for example, just a simple measurement of course activity or course interactions, or something like that for your, for your normal teaching and learning professional support person, they will immediately understand but your academic may want to know for example, if you have a measurement such as course accesses, you know, then they want to understand the normal question we get is but does this include time away from the computer? You know, when it's course minutes for example, you know, what if I switch on my computer and I go for a walk or something like that, you know what, what is included and what's excluded? Then you have to have a data dictionary to provide that kind of explanation but even data dictionaries have the tendency to be very complex in their description of what is described in the data dictionary and therefore you need to have interventions to explain what they are looking at. Unfortunately, that is a challenge. So that is a challenge and it will be a challenge across the world and across University. We have a staff development course, which we specifically focusing on all the data sources and reports available within our LMS that allow a lecturer to understand what each of those reports means and which report they can use for which purpose. So that's that's how we try to address that.

Appendix E Challenges in the use of LA

Challenges	Description	Institution A	Institution B	Institution C	Institution D
Correct interpretation of data	Understanding what the data means and knowing the type of data that is available and being able to translate it into something meaningful			*"any interaction that the student has with the university...is stored somewhere but the challenge is seeing it all in one place and understanding what it all means." CI*	*"So it's contextual understanding of what does this mean, the way the data is derived and the data is described, have got certain meanings… what is included and what's excluded?" JC*
Common definition of data variables	Understanding of how data is structured to aid in its interpretation			*"The main constraint is some some kind of mediation between the data and the person that needs it…there's a big gap between the data that's collected and what you can actually use at the end." HD*	*"Then you have a data dictionary to provide that kind of explanation but even data dictionaries have the tendency to be very complex in their description of what is described in the data dictionary and therefore you need to have interventions to explain what they are looking at." JC*
Unrecorded student interactions	Accounting for learning that happens outside of the institution's systems	*"with it being an online course, there's nothing stopping someone downloading content and working on it in their own time... So you need to account for things like that." QE*	*"we working with a lot of gaps in the data…we don't know how many hours is the student spending on their studies...and unless they do come into the LMS we sit with fairly big gaps in our system. So the pedagogy really shapes the analytics as well." TR*		

Challenges	Description	Institution A	Institution B	Institution C	Institution D
Consistency in the data	Inconsistencies in the way the data is captured and stored and changes in data structures that do not get applied to existing data		*"When I draw the data, I look at the consistency and integrity of the data…and mostly it is not consistent....you look for some data and only half of it pops up…you see there was a dimensional data change at some point but...they didn't convert the old data to the new format" PJ*		
Lack of access to student data	A lack of or limited access to certain types of student data is a challenge in the use of LA. Access to too much raw student data is also a challenge when users do not have the capacity or skills to analyse and interprete it			*"as a course convenor you might never have seen the demographic breakdown of students in your course, nobody gave you that information and you didn't know how to get it." CI "so we begin each year by finding out what their support needs are…so we are very attuned and aware of support needs. And basically using whatever we have…and just drawing out the analytics…as is needed." HD*	*"It's not open for all to use, I think that's a challenge. It's usually one or two people who are administrators on the tools and if you want something or need something, you have to email them with your request or your question and then they go into that system and they draw the reports for you." KT*

Challenges	Description	Institution A	Institution B	Institution C	Institution D
Incorrectly classifying institutional reporting as LA	Associating general performance of the institution with LA		*"previously when we were looking at learning analytics, it was purely from the education standpoint, whereas now, it can also be from a market research side of it and bring in other elements into that: And then the other reason is, if learning analytics doesn't flow from the pedagogy and the structure of the course there is no reason why I should be interested. If it is not a compulsory ingredient in my pedagogy, then why should I click on that but to look at the dashboard?" EO*		
Implementing changes based on the insights	One of the challenges highlighted is around the ability to act upon the insights that the data is illustrating and make changes based on those.			*"the challenge, the challenge is that sometimes we might not have the mechanisms to be able to fully implement the changes which we want." HI*	