Patrick Ndayizigamiye ·
Hossana Twinomurinzi ·
Billy Kalema · Kelvin Bwalya ·
Mncedisi Bembe (Eds.)

Communications in Computer and Information Science 1774

Digital-for-Development: Enabling Transformation, Inclusion and Sustainability Through ICTs

12th International Development
Informatics Association Conference, IDIA 2022
Mbombela, South Africa, November 22–25, 2022
Revised Selected Papers

Springer

Communications in Computer and Information Science 1774

Rationale

The CCIS series is devoted to the publication of proceedings of computer science conferences. Its aim is to efficiently disseminate original research results in informatics in printed and electronic form. While the focus is on publication of peer-reviewed full papers presenting mature work, inclusion of reviewed short papers reporting on work in progress is welcome, too. Besides globally relevant meetings with internationally representative program committees guaranteeing a strict peer-reviewing and paper selection process, conferences run by societies or of high regional or national relevance are also considered for publication.

Topics

The topical scope of CCIS spans the entire spectrum of informatics ranging from foundational topics in the theory of computing to information and communications science and technology and a broad variety of interdisciplinary application fields.

Information for Volume Editors and Authors

Publication in CCIS is free of charge. No royalties are paid, however, we offer registered conference participants temporary free access to the online version of the conference proceedings on SpringerLink (http://link.springer.com) by means of an http referrer from the conference website and/or a number of complimentary printed copies, as specified in the official acceptance email of the event.

CCIS proceedings can be published in time for distribution at conferences or as post-proceedings, and delivered in the form of printed books and/or electronically as USBs and/or e-content licenses for accessing proceedings at SpringerLink. Furthermore, CCIS proceedings are included in the CCIS electronic book series hosted in the SpringerLink digital library at http://link.springer.com/bookseries/7899. Conferences publishing in CCIS are allowed to use Online Conference Service (OCS) for managing the whole proceedings lifecycle (from submission and reviewing to preparing for publication) free of charge.

Publication process

The language of publication is exclusively English. Authors publishing in CCIS have to sign the Springer CCIS copyright transfer form, however, they are free to use their material published in CCIS for substantially changed, more elaborate subsequent publications elsewhere. For the preparation of the camera-ready papers/files, authors have to strictly adhere to the Springer CCIS Authors' Instructions and are strongly encouraged to use the CCIS LaTeX style files or templates.

Abstracting/Indexing

CCIS is abstracted/indexed in DBLP, Google Scholar, EI-Compendex, Mathematical Reviews, SCImago, Scopus. CCIS volumes are also submitted for the inclusion in ISI Proceedings.

How to start

To start the evaluation of your proposal for inclusion in the CCIS series, please send an e-mail to ccis@springer.com.

Patrick Ndayizigamiye · Hossana Twinomurinzi ·
Billy Kalema · Kelvin Bwalya · Mncedisi Bembe
Editors

Digital-for-Development: Enabling Transformation, Inclusion and Sustainability Through ICTs

12th International Development Informatics Association Conference, IDIA 2022
Mbombela, South Africa, November 22–25, 2022
Revised Selected Papers

Springer

Editors
Patrick Ndayizigamiye
University of Johannesburg
Johannesburg, South Africa

Billy Kalema
University of Mpumalanga
Mbombela, South Africa

Mncedisi Bembe
University of Mpumalanga
Mbombela, South Africa

Hossana Twinomurinzi
University of Johannesburg
Johannesburg, South Africa

Kelvin Bwalya
University of Johannesburg
Johannesburg, South Africa

ISSN 1865-0929 ISSN 1865-0937 (electronic)
Communications in Computer and Information Science
ISBN 978-3-031-28471-7 ISBN 978-3-031-28472-4 (eBook)
https://doi.org/10.1007/978-3-031-28472-4

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

The 12th edition of the International Development Informatics Association Conference (IDIA 2022) was co-hosted by the University of Johannesburg and the University of Mpumalanga in South Africa under the theme 'Digital-for-development': Enabling transformation, inclusion and sustainability through ICTs. The event was held at the Tsogo Southern Sun Hotel in the city of Mbombela in the Mpumalanga Province of South Africa from 22nd to 25th of November 2022. The conference provided a platform for scholars, academics and IT practitioners to discuss new developments in the context of the paradigm shift from ICT4D to 'digital-for-development'. The conference received 61 full paper submissions, of which 40 full papers were selected for presentation at the conference following a double-blind peer-review process. Each paper submitted to the conference was reviewed at least three times by a panel of national and international reviewers before a final decision was made to publish or reject it. The selected papers were presented at the conference under the following thematic areas: Theories and practices in digital-for-development ecosystems: ICTs and education; Privacy and Security in digital-for-development ecosystems; Emerging technologies for transformation, inclusion and sustainable development: Business applications; Theories and practices in digital-for-development ecosystems: Business applications; Data-driven sustainable development; Emerging technologies for transformation, inclusion and sustainable development: Business applications; Artificial Intelligence (AI) for good; Electronic health for transformation, inclusion and sustainability; Theories and practices in digital-for-development ecosystems; Human-Computer Interaction (HCI) for digital inclusion; Application and architectures of Internet of Things. In the end, only 20 papers (33% of the papers submitted to the conference) were selected to be published in the Springer CCIS volume, based on the reviewers' scores. The remainder of the papers presented at the conference will be published in the normal non-Springer Conference Proceedings.

February 2023 Patrick Ndayizigamiye

Organization

Conference Chairs

Patrick Ndayizigamiye	University of Johannesburg, South Africa
Ntombovuyo Wayi-Mgwebi	University of Mpumalanga, South Africa

Conference Proceedings Chairs

Billy Kalema	University of Mpumalanga, South Africa
Hossana Twinomurinzi	University of Johannesburg, South Africa
Kelvin Bwalya	University of Johannesburg, South Africa
Mncedisi Bembe	University of Mpumalanga, South Africa
Patrick Ndayizigamiye	University of Johannesburg, South Africa

Conference Organizing Committee

Charlie Ndlovu	University of Mpumalanga, South Africa
Duduzile Masuku	University of Mpumalanga, South Africa
Hope Mauwa	University of Mpumalanga, South Africa
Norwell Zhakata	University of Mpumalanga, South Africa
Pardon Blessings Maoneke	University of Mpumalanga, South Africa
Precious Leutle	University of Mpumalanga, South Africa
Rehana Minty	University of Johannesburg, South Africa
Roelien Brink	University of Johannesburg, South Africa
Sinazo Matyila	University of Mpumalanga, South Africa
Siyabonga Mhlongo	University of Johannesburg, South Africa
Stella Bvuma	University of Johannesburg, South Africa

Program Committee

Millicent Agangiba	The University of Mines & Technology, Ghana
Anastasia Griva	University of Galway, Ireland
Farshad Badie	Aalborg University, Denmark
Samantha Buxton	Swansea University, UK
Wallace Chigona	University of Cape Town, South Africa

Christopher Foster	The University of Manchester, UK
Conn Smyth	University of Galway, Ireland
Robert M. Davison	City University of Hong Kong, China
Desireé Cranfield	Swansea University, UK
Devinder Thapa	University of Agder, Norway
Antonio Diaz Andrade	Auckland University of Technology, New Zealand
Efosa C. Idemudia	Arkansas Tech University, USA
Elizabeth Mkoba	Nelson Mandela African Institution of Science and Technology, Tanzania
G. 'Hari' Harindranath	Royal Holloway, University of London, UK
Mamello Thinyane	United Nations University, China
Isabella Rega	Bournemouth University, UK
John Mulyata	Swansea University, UK
John Otieno Oredo	University of Nairobi, Kenya
Joseph Ntabeni Jere	University of KwaZulu-Natal, South Africa
Judy Van Biljon	University of South Africa, South Africa
Caroline Khene	De Montfort University, UK
Macire Kante	Centre National de la Recherche Scientifique et Technologique, Mali
Maria Garrido	University of Washington, USA
Ravishankar Mayasandra-Nagaraja	Loughborough University, UK
Michael Best	Georgia Institute of Technology, USA
Joshua Rumo A. Ndiege	United States International University Africa, Kenya
Petter Nielsen	University of Oslo, Norway
Odoyo Collins Otieno	Masinde Muliro University of Science and Technology, Kenya
P. Vigneswara Ilavarasan	Indian Institute of Technology Delhi, India
Pitso Tsibolane	University of Cape Town, South Africa
Melanie Platz	University of Koblenz-Landau, Germany
Jaco Renken	University of Manchester, UK
Salah Kabanda	University of Cape Town, South Africa
Steven Sam	Brunel University, UK
Tendani Mawela	University of Pretoria, South Africa
Tony Roberts	Institute of Development Studies, UK
Jean-Paul Van Belle	University of Cape Town, South Africa
Gamel Wiredu	Ghana Institute of Management and Public Administration, Ghana

Additional Reviewers

Abiodun Alao
Adedamola Tolani
Appolonia Ilorah
Debas Maren
Elizabeth Mkoba
Emmanuel Howe
Fredrick Kanobe
Gcina Dludla
Kennedy Njenga
Khuliso Sigama
Lean Makhubele
Lucas Gumbi
Michael Moeti
Mmatshuene Anna Segooa
Norwell Zhakata
Ntombiyokusa Nyoni
Rehema Baguma
Sam Ssemugabi
Samwel Mwapwele
Seffi Gebeyehu Mengistu
Sibukele Gumbo
Sinfree Gono
Siyabonga Mhlongo
Tebogo Bokaba
Thulani Shongwe
Tiko Iyamu
Tinofirei Museba
Tobi Fadiji
Tope Adeyelure
Vigneswara Ilavarasan
William Tichaona Vambe

Contents

Theories and Practices in Digital-for-Development Ecosystems

The Value of Theory in Teaching ICT4D at a Graduate Level

Marita Turpin(✉) and Tania Prinsloo

Department of Informatics, University of Pretoria, Pretoria 0002, South Africa
{marita.turpin,Tania.prinsloo}@up.ac.za

Abstract. Theory forms an integral part of teaching at a graduate level. The incorporation of appropriate theories holds the potential to significantly enhance students' understanding and insight into an academic discipline. On the other hand, theories not used in an appropriate manner can act as a barrier to understanding and learning. This study concerns the use of theories as part of teaching a graduate Information and Communication Technology for Development (ICT4D) course. The course made use of a textbook where a wide array of theoretical frameworks was incorporated into the study material. During the course, students had to apply theoretical frameworks during each of their essay-type homework assignments, as well as in their take-home exam assignment. As part of the summative assessment, an exam question was given where students had to reflect on the theories they learnt. This question was added to gauge which theories students found helpful in enabling their understanding of ICT4D issues. The responses of the students constituted the empirical data that was qualitatively analyzed. The Sustainable Livelihoods Framework (SLF) emerged as the framework that most students found helpful. This finding was surprising since the SLF is not a simple framework to interpret and apply. The results indicate that students appreciate the value of looking holistically at an ICT4D problem context. This study recommends that effective use of theory as part of teaching ICT4D needs to include an accessible introduction, examples of application, students' own application and ideally also an opportunity for reflection.

Keywords: ICT4D · Sustainable livelihoods framework · Design-reality gap · DIRT model · Theoretical framework

1 Introduction

Theories in graduate teaching are central to students' understanding, and key to their personal and professional development [1]. However, during a recent search performed by the authors for guidelines or best practices for incorporating theory in graduate teaching, to inform the use of theory in a graduate Information and Communication Technology for Development (ICT4D) course, very few publications were found. Since the authors were already engaged in teaching a graduate ICT4D course that was theory-rich, this course was used as an opportunity to investigate good practices for incorporating theory as part of teaching. Therefore, a research project was initiated based on the following research question:

P. Ndayizigamiye et al. (Eds.): IDIA 2022, CCIS 1774, pp. 3–16, 2023.
https://doi.org/10.1007/978-3-031-28472-4_1

How can theories effectively be used to make it valuable as part of teaching ICT4D at a graduate level?

While the focus of the study was an ICT4D course, we believed that such a study should be informed by the use of theory as part of graduate teaching in general [2], and its findings should be more generally applicable to graduate level teaching. It was therefore anticipated that the study would be able to contribute to the body of knowledge informing the use of theory as part of graduate teaching.

This paper is organized as follows: The literature review first considered the use of theories in graduate teaching, before moving to the use of theories in ICT4D. Following this, an overview is provided of the theories that were covered in the graduate ICT4D course that formed the focus of the study. The literature review is followed by a method section that describes how the study was planned and executed. After this, the findings are presented and discussed. The paper concludes with a summary of the study and a reflection on its findings.

2 Literature Review

In this background section we consider the reason for including theory in graduate teaching. We then consider the context of how theory is generally used in ICT4D research. Following this, we shortly introduce the range of theories that were covered in the ICT4D textbook which was the prescribed text for the ICT4D course that formed the focus of the case study.

2.1 The Use of Theories in Graduate Teaching

Theories in graduate teaching are central to students' understanding, vastly improving their personal and professional development [1]. The value of using theory in graduate teaching lies in the sound underpinning of the current and relevant literature in the specific field or context [3]. Therefore, one of the critical objectives of measuring how students use their theories is to determine if they are well understood and used appropriately [4]. To measure how well theories are understood and used, students need to apply the appropriate theory to a given case study or situation. The interpretation of the theory then needs to be evaluated based on predetermined criteria [5]. The ability of students to apply the theory to the case study and articulate their findings provides a clear picture of how well the subject matter is understood. This enables the lecturer to better guide the students who misinterpreted their case studies and allows for rich discussions in class on the relevant theories as well as their vastly different applications.

2.2 The Use of Theories in ICT4D

The main goal of ICT4D theories is to enable researchers to understand this phenomenon better [6]. Reasons why this phenomenon occurred, how it can be better explained, and why we need to take note of such a theory are all reasons why theories become adopted and used. Further to that, theory can be used to inform action. Sein et al. [7] proposed the following five research agendas that we need to implement in line with the Sustainable Development Goals (SDGs) in an ICT4D context. They are:

1. More focus should be placed on theorizing in ICT4D.
2. We need to look at multiple levels of analysis, not only the group but also the individual level, as well as other factors that apply.
3. It is worthy to interpret and understand a situation, but we need to do more and publish more on intervention studies.
4. It remains crucial to ensure that ICT4D has a philosophical basis.
5. ICT4D is not merely a developing world phenomenon but also needs to be applied in a developed world context.

Some of the more widely used theories in ICT4D are the Technology Acceptance Model, Diffusion of Technology, Actor-Network Theory, the Capabilities Approach, and the Sustainable Livelihood Framework [5]. Still, a lot has changed as the ICT4D field matured in the last 30 years. Walsham [7] describes the periods in decades, from the mid-1980s to the 1990s, where the main focus was on in the IS field, then the mid-1990s to mid-2000s, with inter-disciplinary research becoming more widely published, until the mid-2000s to present, with the absolute "explosion of technology" in the developing world context. He continues to discuss the importance of theory, highlighting the Actor-Network Theory, the Institutional Theory, and the Sustainable Livelihoods Theory [7]. Finally, Maseiro [8] discusses the theoretical crisis that the field of ICT4D is facing and suggests that multiple theories are needed to make sense of real-world issues, including theories that cater more to indigenous case studies.

2.3 Theories Covered in the Prescribed Textbook

The prescribed textbook in the ICT4D graduate course was developed by Richard Heeks [9]. There are nine chapters in the textbook, and there are a variety of theories covered in every chapter. The main theories covered overall in the textbook will be briefly discussed below:

Design-Reality Gap for Analyzing Change. This theory can be used in various ways, one of which is to identify the likelihood of success or failure of projects. The bigger the gaps between the criteria, the larger the chances of project failure [10]. The theory is depicted in Fig. 1 below:

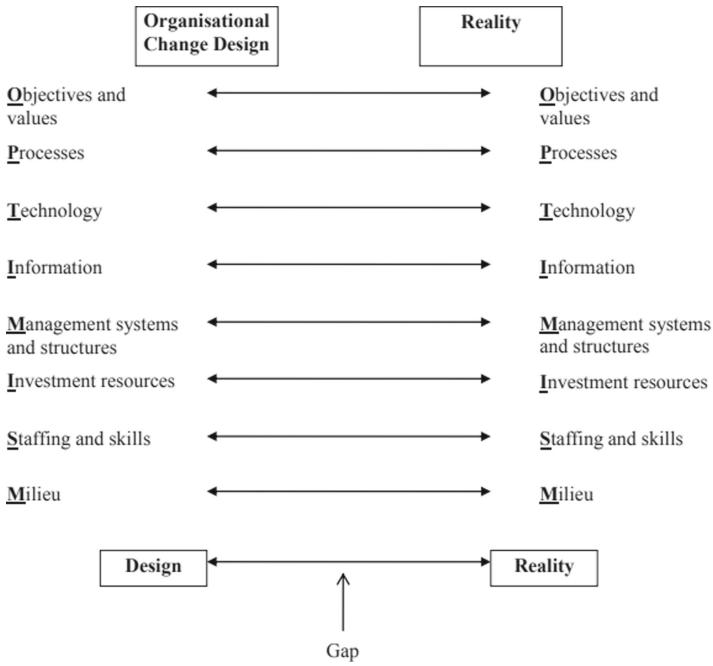


Fig. 1. The design reality gap theory for analyzing change. . Adopted from Bass and Heeks [11]

DIRT Model. The DIRT model is an example of a ladder model, with the stages following on one another. It has four levels of change, and they are:

1. To digitize. This process looks at how the human component can be digitized, without anything else changing.
2. To improve. Here incremental changes are made to either the process itself or different inputs.
3. To reorganize. Here overall changes are made to the process, for example adopting a new technology.
4. To transform. The focus is now on making changes to the overall and broader social structures [10].

Sustainable Livelihoods Framework (SLF). This framework was introduced in 1998 by Scoones [11] and looks at important elements such as the context, the livelihood resources, the livelihood strategies, the outcomes, and institutional processes. It is often criticized as being too complex, but if broken down, it is a beneficial framework to analyze the interplay between various factors in a social system [12]. The framework explores the different forms of capital, namely human, social, physical, financial, and natural capital. It then discusses the vulnerability context, policies and institutions, livelihood outcomes and livelihood strategies [13], shown in Fig. 2.

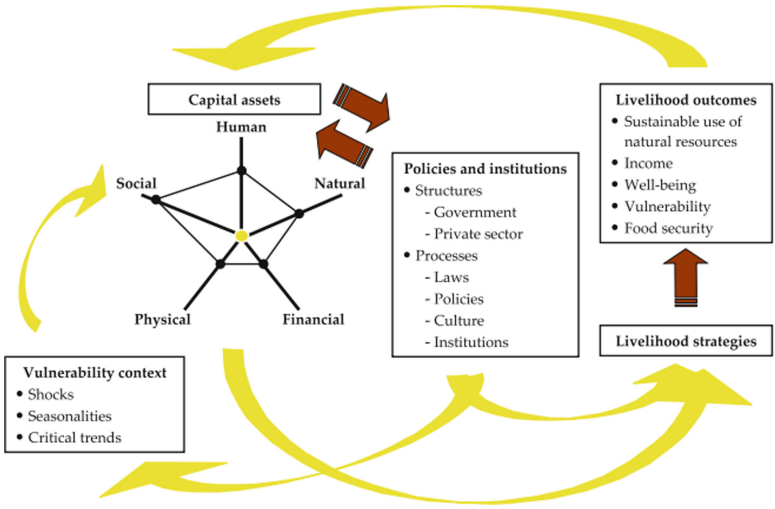


Fig. 2. The sustainable livelihoods framework. . Adopted from Serrat [13]

ICT4D Value Chain. This model focused on input, process, and output [14] in the lifecycle of an ICT4D intervention. It is divided into four different domains, as discussed by Heeks [14], namely readiness, availability, uptake, and impact, with impact looking at the different outputs, outcomes, and developmental impacts, as shown in Fig. 3.

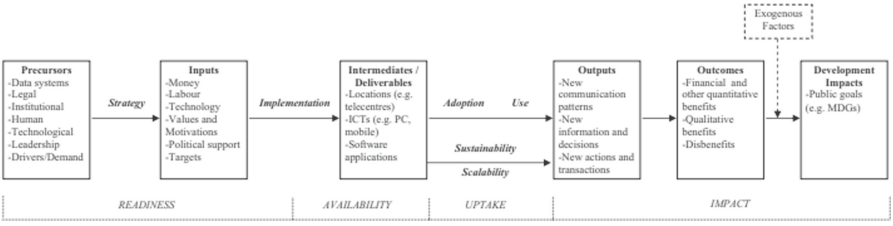


Fig. 3. The ICT4D value chain. . Adopted from Heeks and Molla [15]

CIPSODAR Steps of the Information Value Chain. The CIPSODAR Model shows its value in the fact that it presents itself as a value chain [10], with data becoming information before becoming useful, and can be divided into the following two parts:

1. CIPSO. The CIPSO stands for Capture, Input, Process, Store, and Output.
2. DAR. It stands for Decision, Action, and Result.

The model is depicted in Fig. 4. below:

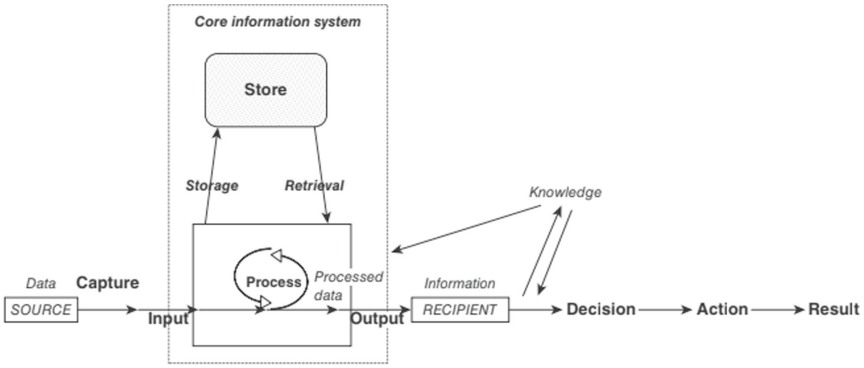


Fig. 4. The CIPSODAR steps of the information value chain [10].

RABIT Model of Resilience and E-Resilience. RABIT stands for Resilience Assessment Benchmarking and Impact Toolkit [16]. This toolkit focuses strongly on the different aspects of resilience, and fits well in a developmental context, and a diagram depicting the different aspects are shown in Fig. 5 below:

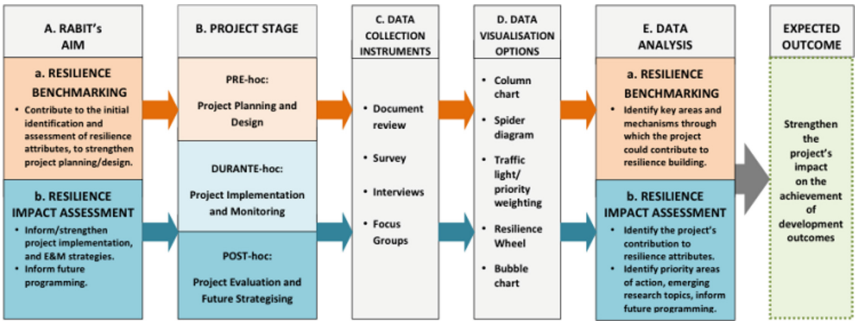


Fig. 5. RABIT: overview of implementation [16].

3 Method

A qualitative case study was conducted from August to December 2021, with a recently introduced ICT4D Honours 4th academic year course at a South African university constituting the case. The university, situated in the Gauteng province of South Africa, occupies a space in the top 1,9% of universities worldwide, ranking highly for Africa across several measurement systems [17]. The university has a multi-faceted way of

approaching teaching and learning, innovation, and achieving community impact. Its Information Systems (IS) department is a significant role player in ICT4D research nationally and internationally. The three-year undergraduate IS degree programme does not have a dedicated subject specifically aimed at ICT4D; however, students need to complete a compulsory one-semester community project. ICT4D is given prominence at the Honours level, with a dedicated module. The semester course is an elective that is part of a coursework Honours degree in Information Systems (IS). The course is presented part-time. The majority of the students completed their Bachelor's degree the previous year and recently started working. The students are a diverse group, who completed their previous degree at one of various Higher Education Institutions (HEI) in South Africa. The number of enrolled students in 2021 was 57. The course consisted of eight contact sessions, two take-home essay-type assignments and an exam assignment. The nine chapters of Heeks' textbook were closely followed during the eight contact sessions, covering a chapter per session, except for session 1 that covered both chapters 1 and 2. Theory only formed part of what was covered; however, it was presented as an important part of the course. During the course, students were exposed to theory as indicated below.:

- Contact sessions: All the theories discussed in Sect. 2.3 were introduced to the students over the course of the eight contact sessions, namely: The Design-Reality Gap for Analyzing Change, the DIRT Model, the SLF, the ICT4D Value Chain, the CIP-SODAR Steps of the Information Value Chain and the RABIT Model of Resilience. As per the textbook, the theories were introduced along with ICT4D topics where the theory could contribute to a better insight of the topic. For example, the SLF was introduced and applied in the chapter on ICT's role in poverty alleviation and livelihoods.
- Take-home assignments: In the two formative take-home assignments, students were presented with problem scenarios where they had to apply theory in alignment with the thinking on the use of theory as part of graduate teaching that was discussed in Sect. 2.1. Assignment 1 covered the theme of ICT4D Implementation. Here, students had to identify an ICT4D project that reached implementation stage and about which sufficient literature existed in order for them to apply the Design-Reality Gap concepts to the project. The topic of Assignment 2 was ICTs and Economic Growth, Poverty and Livelihoods. In this assignment the DIRT model and the SLF were covered. On the DIRT model, students had to answer the following question: *"Compare the DIRT model to other stage/ladder models of ICT application"*. They further had to apply the SLF to an ICT4D case study of their choice, to which the SLF was not previously applied.

The exam assignment was of a similar nature as the take-home assignments but more extensive in scope. The exam assignment had an environmental focus. Along with other questions related to ICT and environmental sustainability, a question was given where students had to apply the RABIT model: *"Use the RABIT model to argue how ICT can help communities to 'go green' sustainably"*. The exam assignment also included a question where students had to reflect on the theories they were taught throughout the semester, namely:

"Which of the frameworks you have learned about in [your ICT4D course], did you find the most helpful in gaining insights into ICT4D issues, and why?"

Students had to respond to the above question in a half to one A4 size page. The responses to the above exam assignment question forms the main source of data for this study. Ethics approval was obtained from the university where the course was presented, to use student assessments as a data source with the students' informed consent. Of the 57 enrolled students, 56 students received exam admission and completed the exam. Of these, 54 consented for their responses to be used for research purposes.

Data analysis was done deductively, specifically looking for which theories students discussed in their responses [18]. The analysis entailed reading through all of the 54 student responses, classifying them according to the theories the students chose to nominate, and doing a thematic analysis of the essay-type data. Common themes were identified among the reasons the students chose a specific theory.

4 Findings

The frameworks or theories that students selected in response to the exam question above, were as follows:

Table 1. Frameworks selected by students in response to exam question

Framework	Number of responses
Sustainable livelihoods framework (SLF)	27 (+4)
Design-reality gap	13 (+2)
DIRT model	7 (+1)
Other: RABIT; CIPSODAR, DOI, ICT4D Value Chain, Capability approach	One each: 3 (+2)

Note that the total in Table 1 does not add up to 54, since 4 students chose to nominate more than one framework. The numbers in brackets refer to cases where that framework was one of two or three that were selected.

As can be seen, the SLF emerged as the framework that was most strongly supported, in terms of what students found most helpful in gaining insights into ICT4D issues. The other frameworks that stood out were the Design-Reality Gap and the DIRT model. Interestingly, these were the three frameworks that students had to apply in their take-home assignments.

Students' reasons for selecting the above frameworks were as follows:

4.1 Sustainable Livelihoods Framework

The most commonly occurring theme mentioned by students was the manner in which the SLF enabled them to understand the ICT4D community context in a holistic and

comprehensive manner, similar to how the community understood themselves: "Lives of the poor must be understood as the poor themselves understand their own lives as a complex of interacting factors"; "…gave me insight and understanding of the livelihoods of the poor"; "…to put yourself in the other persons shoes can be tough and therefore this framework helps you and guides you to do so"; "allows one to view a case study from a holistic point of view"; "broadens my understanding of poor people's livelihoods'; "a comprehensive strategy, which aims to capture and understand the fundamental causes and features of poverty".

A second theme was the benefit of the SLF to assist in the understanding of people's existing capitals base, to build on what they already have and to help with mobilizing existing resources: "…uses what exists in the current setting to help the community even further"; "…building on the assets that the poor has"; "…focus on…the underlying resources (livelihood assets) and capacities needed to break the circle of poverty".

The SLF further helped students recognize that poverty alleviation is not just about money/monetary wealth, and also not just about implementing IT: "not only a matter of throwing money at something and thinking the problems will be solved"; "…although ICT is an essential part of poverty alleviation, it is not a panacea".

Students further stated that the holistic kind of understanding discussed above leads to better insight into the community, and therefore to more informed decision-making around which types of ICT can be best implemented: "…we gain better insight into the current situation of the people within the community… this type of information is vital when determining what types of ICTs can be implemented to improve the livelihoods.."

A few students remarked on the personal value of applying this framework to a case study: "…it shocked me to the bone and gave me new insight…helped me understand why some of these communities in extreme poverty are so helpless and vulnerable"; "Helping the poor community is something which is close to my heart and therefore this model as helped me in every aspect…".

All in all, the students seemed to have obtained a very good succinct grasp of the SLF, perhaps better than several ICT4D scholars. It was encouraging to find that the students appreciated the value of the framework to aid in gaining a holistic understanding of a complex community context, despite the complexity of the framework itself.

4.2 Design-Reality Gap Framework

The main benefit of the Design-Reality Gap Framework that was recognized by the students, was also the aim of the framework, namely to provide a means to assess the gap between an ICT4D implementation plan and the realities of the application context: "This framework actually shows you the gap which might exist between the ICT4D application design and reality"; "The aim of the framework is to analyze the shortfall between design expectations and realities… to evaluate whether it would achieve stakeholders' objectives"; "the… framework can evaluate the risk of projects, reduce the failure rate of projects, and increase the efficiency of resource allocation for projects."

The gap assessment was recognized as a means to guard against failure: "IT projects are notorious for their high failure rate and having a framework to analyze feasibility and risks associated with the project helps to understand in which ways a project can be changed to suit the community"; it helps "…guard against failure".

Students found value in the framework's comprehensiveness in the aspects it considered, its systematic approach as well as the fact that it gave a method in which to quantify the gap: "the framework brings together all required elements that are needed to implement an ICT4D project"; "the… framework is able to measure… and with this model, you can clearly see the gap and the risks that come with it"; "…it takes a systematic approach when looking at ICT4D issues."

Several students commended the framework's simplicity and ease of application: "this model was really straightforward and simple which made it much easier to understand… than… the other frameworks learned in [this course]"; "…easy to use, easy to understand and easy to implement in any ICT4D project. Thus, I am finding it to be the most useful."

In terms of the value they found when applying the framework practically: "When applying the framework to a specific case study, it became evident that the needs of the community were not always met through the implementation of the project." "In Assignment 1… I found gaps that I would not have identified without applying the framework."

One student found that the framework was useful to them beyond the course: "…[this was] the most helpful framework that I took away from this module and I can easily implement this model in my application development career." This student stated that they became "more aware of perception vs reality complexities" in the workplace.

Overall, the students recognized the clear practical value of this framework as a diagnostic tool to assess and assist in responding to the risk of ICT4D project failure, and they appreciated the simplicity and ease of application of the framework. It was encouraging to note that the framework assisted some students to see the application setting in a different way, and that it helped them to discover issues they would not have noticed otherwise.

4.3 DIRT Model of Transformation

The DIRT model provides a ladder-type roadmap to digital transformation. Students found this model helpful "in understanding the extent of change enabled by ICT"; it "explains how ICT is used in the digitizing process in a developing country"; "…it takes you through how ICT improves processes in developing countries to help contribute to making processes efficient and increasing productivity."

Similar to the Design-Reality Gap Framework, students found that "…the DIRT model is easy to remember and simple to apply and understand" while at the same time "it can also give insight in complexities." "…it gives a simple method to help understand the transformation brought by ICTs."

A differentiating factor of this model which the students found beneficial is that it "provides more of a business view on ICT4D", focusing on "making business more productive and make actions more efficient causing economic growth." Further, "this framework focuses on ICTs helping small businesses which is extremely important for developing countries."

When motivating the benefit of the DIRT model, several students referred to Heeks' description in the textbook only, rather than saying how they benefited or gained insight from their own application or even their own reflection on its value. This was perhaps

because students were not required to fully apply the DIRT model at the in-depth level that they were required to apply the SLF and the Design-Reality Gap Framework.

4.4 Other Theoretical Frameworks That Students Found Helpful

One student found the CIPSODAR model helpful because of its "comprehensible and understandable steps for you to explore and analyze ICT4D prevailing issues." Further, "since all firms process data and have decision making procedures it is always helpful to follow some kind of model and the CIPSODAR model has proven to be worthy and valuable."

One student nominated the RABIT framework that deals with resilience. The student found RABIT useful "due to its completeness and inclusiveness that addresses the sustainability of the initiative." It is surprising that only one student chose to mention the RABIT framework, since all students had to apply this framework in the preceding exam question. In the exam, the students performed well overall with the application of the RABIT framework and seemed to have found value in it, as was the case with the SLF and the Design-Reality Gap in earlier assignments.

One student mentioned the Diffusion of Innovation (DOI) Theory [19]. This was an unusual choice since the DOI is not an ICT4D theory/framework. To this student, the DOI theory provided "insights into technology adoption processes".

One student suggested the ICT4D Value Chain Framework and deemed it an important construct that "aids in understanding the distribution of rewards from the chain's many operations."

Lastly, one student suggested the Capability Approach (CA) of Amartya Sen [20], because it was a "holistic lens through which development can be contextualized."

Interestingly, the CA and DOI theory did not receive much attention during the ICT4D lectures, yet one student each still found perceived value in them.

5 Reflection on Findings

While each of the theoretical frameworks discussed in Sect. 2.3 were introduced and discussed during the lectures, the take-home assignments required students to read further on three specific frameworks, namely the Design-Reality Gap Framework, the SLF and the DIRT model. Further, the Design-Reality Gap Framework and SLF had to be applied to ICT4D cases of students' own choice, so students spent relatively more time and attention on these two frameworks. These were also the two frameworks that the students found most helpful. It therefore appears that forcing students to apply a framework to a practical case setting helps them to better appreciate the value of that framework.

It was a surprise to the authors that the SLF was so overwhelmingly popular with the students as a theoretical framework, since it is not simple or easy to understand and apply. However, the students seemed to recognize value in the SLF's comprehensiveness and its holistic way of treating poverty. There is some contradiction in the finding that more than half of the students chose a complex framework as the most helpful, while most of the remaining students chose other frameworks based on the simplicity and ease of application of those frameworks.

Value Gained by Students. Overall, the kinds of value that the students gained from the frameworks included the following:

- It helped them to understand the problem context better;
- It helped them to notice aspects of the application context they would otherwise have missed;
- It improved their understanding and evaluation of the appropriateness of the ICT4D intervention; and
- It gave them a way of thinking that they could apply beyond their study context.

To the lecturers, it was very encouraging to find that the students, to varying degrees, gained the above kinds of value from the theoretical frameworks they were taught and requested to apply.

Learning and Recommendations. From the learning that the lecturers have done through teaching the course and analyzing the data from the student responses, the following learning can be distilled about the teaching of theory during their ICT4D course:

Firstly, it is believed that the textbook provided an excellent first introduction to the theories covered. The discussion of the theories in the textbook were well structured, well informed and comprehensive yet presented in a very accessible manner and with the frequent use of diagrams. Further, effort was taken to have multiple application examples for each theory that was covered. We believe that the quality of the prescribed textbook significantly added to the students' initial understanding of the frameworks.

Second, students were requested to apply selected frameworks during the assignments, to ICT4D cases of their own choice. We believe that giving the students the freedom to choose case settings that were interesting to them, helped to retain the student's attention.

Further, we believe that exposing students to multiple theories helped them to see how each theory or framework aided with insight in a different way, as each framework had its own purpose and angle.

Lastly, we believe that there was value in adding an exam question that required students to reflect on the value of the theoretical frameworks, in order for them to more explicitly recognize their own learning as well as the value that the frameworks added to their learning.

6 Conclusion

This study was guided by the following research question: *How can theories effectively be used to make it valuable as part of teaching ICT4D at a graduate level?*

A qualitative case study was conducted, with an ICT4D Honours course constituting the case. After introducing and exposing students to theoretical frameworks in various manners during the course, of which the most significant was the application of selected frameworks in take-home assignments, students were given the following question as

part of their take-home exam: *Which of the frameworks you have learned about in [your ICT4D course], did you find the most helpful in gaining insights into ICT4D issues, and why?*

Interestingly, the framework that was found most valuable by the students was the SLF, which was also the most complex of the frameworks taught. It was found that students' ability to grasp and apply theory should not be underestimated, provided that they are given sufficient introduction and guidance in its application.

The study ends with recommendations related to the teaching of theory in an ICT4D graduate course. In terms of the contribution of this study, we feel that the learning and insights that were gained, as summarized in Sect. 5, are also applicable to other graduate courses where theory is taught.

A suggestion for future work is to more explicitly plan for achieving the benefits that were emergent in this study. The authors plan to incorporate their own learning to improve their ICT4D course in this manner.

References

1. Hirsh, D.A., Ogur, B., Thibault, G.E., Cox, M.: "Continuity" as an organizing principle for clinical education reform. N. Engl. J. Med. **356**, 858 (2007)
2. Mukhalalati, B.A., Taylor, A.: Adult learning theories in context: a quick guide for healthcare professional educators. J. Med. Educ. Curric. Dev. **6** 2382120519840332 (2019)
3. Kalogiannakis, M., Papadakis, S., Zourmpakis, A.-I.: Gamification in science education. A systematic review of the literature. Educ. Sci. **11**(1), 22 (2021). https://doi.org/10.3390/educsci11010022
4. Soper, D.S., Turel, O.: Identifying theories used in north american is research: a bottom-up, computational approach. In: 2015 48th Hawaii International Conference on System Sciences, pp. 4948–4958. IEEE, (2015)
5. Connolly, P., Keenan, C., Urbanska, K.: The trials of evidence-based practice in education: a systematic review of randomised controlled trials in education research 1980–2016. Educ. Res. **60**, 276–291 (2018)
6. Andersson, A., Hatakka, M.: What are we doing?: theories used in ICT4D research. In: 12th International conference on social implications of computers in developing countries, May 19–22, 2013, Ocho Rios, Jamaica, pp. 282–300. (2013)
7. Walsham, G.: ICT4D research: reflections on history and future agenda. Inf. Technol. Dev. **23**, 18–41 (2017)
8. Masiero, S.: Should we still be doing ICT4D research?. Electron. J. of Inf. Syst. Dev. Ctries. e12215 (2022)
9. Heeks, R.: Information and Communication Technology for Development (ICT4D). Routledge (2017). https://doi.org/10.4324/9781315652603
10. Bass, J.M., Heeks, R.: Changing computing curricula in African universities: evaluating progress and challenges via design-reality gap analysis. Electron. J. Inf. Syst. Dev. Ctries. **48**, 1–39 (2011)
11. Scoones, I.: Sustainable rural livelihoods: a framework for analysis (1998)
12. Levine, S.: How to study livelihoods: bringing a sustainable livelihoods framework to life. Researching livelihoods and Services Affected by Conflict 1–18 (2014)
13. Serrat, O.: The Sustainable Livelihoods Approach. Knowledge Solutions, pp. 21–26. Springer (2017)

14. Heeks, R.: Do information and communication technologies (ICTs) contribute to development? J. Int. Dev. **22**, 625–640 (2010)
15. Heeks, R., Molla, A.: Compendium on impact assessment of ICT-for-development projects (2009)
16. Ospina, A.V., Heeks, R.: Resilience Assessment Benchmarking and Impact Toolkit (RABIT).
17. https://www.up.ac.za/article/2749387/world-rankings-
18. Woiceshyn, J., Daellenbach, U.: Evaluating inductive vs deductive research in management studies: implications for authors, editors, and reviewers. Qual. Re. Organ. Manag.: Int. J. (2018)
19. Dearing, J.W.: Applying diffusion of innovation theory to intervention development. Res. Soc. Work Pract. **19**, 503–518 (2009)
20. Alkire, S.: Why the capability approach? J. Hum. Dev. **6**, 115–135 (2005)

A Systematic Mapping of ICT4D Adoption Research in Developing Countries

Macire Kante and Patrick Ndayizigamiye(✉)

University of Johannesburg, Johannesburg, South Africa
{mkante,ndayizigamiyep}@uj.ac.za

Abstract. While the use of theories and models in Information and Communication Technology for Development (ICT4D) is important and has been studied, the scope of these theories has not been sufficiently researched. With the aid of systematic mapping, this study provides an overview of the research on ICT4D adoption in developing countries to determine the research gaps and trends in the theories used. The study presents research published in ICT4D journals between 2015 and 2019. The results indicate that the understanding of the conceptualisation of the constructs of the theories used in ICT4D research remains sparse. In addition, the statistical techniques (Regression, Partial Least Squares Structural Equation Modeling, amongst others) used to assess the constructs in the selected studies may need more scrutiny. The study recommends that further studies be conducted using other qualitative methods of inquiry to identify any gaps that could have been missed.

Keywords: Theory · Model · Information and communication technology for development · ICT4D · Adoption · Developing countries

1 Introduction

Recently, there has been an increase in the systematic mapping of theories and models used in Information and Communication Technology for Development (ICT4D) research [1–3]. In many instances, ICT4D journals demand that researchers provide a theoretical background for their work. This has compelled ICT4D researchers to embark on research that provides theoretical contributions. However, the literature still points to a lack of in-depth understanding and learning of the change, transformation and impacts in ICT4D research [3] due to, amongst other reasons, the lack of theory building. Theory helps in gaining an in-depth understanding of the issues at hand and provides a solid foundation for further inquiry [4]. It is, therefore, important to investigate the theories and models related to ICT4D studies. While previous studies have attempted to explore ICT4D research using the Activity Theory [3], studies on the systematic mapping [5] of the theoretical foundations of ICT4D research [6] are scant.

This study aimed at presenting a systematic mapping of theories and models used in ICT4D research in developing countries. It was anticipated that the study would provide a better understanding and contextualisation of the theories and models used in ICT4D

P. Ndayizigamiye et al. (Eds.): IDIA 2022, CCIS 1774, pp. 17–34, 2023.
https://doi.org/10.1007/978-3-031-28472-4_2

research in these countries. In doing so it will possibly provide a basis for information and communication technology (ICT) to achieve its development goals as suggested by Sein, Hatakka, and Sæbø [6]. The authors argued that technology needs to be designed to operate in a complex social, political, economic and cultural context. Hence, it is important to understand the theories and models used in designing and implementing such technology.

Mapping studies remain sparse in the ICT4D literature and few studies have been conducted in developing countries. For instance, Zewge and Dittrich [5] conducted a systematic mapping of information technology in developing countries while Kante and Ndayizigamiye [7] investigated the adoption of the Internet of Things in healthcare services. These studies have some gaps which need to be dealt with for a better understanding of ICT4D adoption in developing countries. Firstly, the two studies did not look at the conceptualisation of the variables used to study the adoption. According to Müller-Bloch and Kranz [8], this is a knowledge void which needs to be addressed. Secondly, the mapping studies did not reveal the statistical techniques used in analysing the data and this could be a methodological gap to be dealt with. Finally, the growing literature needs to be further mapped to determine if the recommendations and suggestions of the previous mapping studies have been given due consideration.

A theoretical lens is a basis for "the description, explanation, and prediction of the phenomena it relates to – that is, offer an understanding of the what, how, and why behind the scenes" [9]. The phenomena in this study were technology and development. The understanding, in turn, will help policymakers, researchers, development practitioners and other development stakeholders appreciate the underlying conceptual structures that ultimately enable purposeful and meaningful actions [9].

In this study, developing countries are referred to as low-income, lower-middle-income and upper-middle-income economies as defined by the World Bank [10]. Heeks [11] reported ICT4D as the use of technology "to help deliver on the international development agenda" and that agenda, according to the scholar, comprises the Sustainable Development Goals (SDGs). Thus, in this study, any paper that used ICT linked to any of the SDGs was of interest.

This paper covers research published in ICT4D journals between 2015 and 2019. It is part of an ongoing study and data covering the period of January 2020 to May 2022 are currently being gathered. Furthermore, other criteria such as digital development, among others, are being incorporated to retrieve more studies.

2 Search Procedures and Strategies

Systematic mapping or scoping studies are designed to provide an overview of a research area by classifying and counting the contributions of the categories within each classification. Hence, they can provide an overview of the theories and models used in ICT4D research in developing countries by classifying and counting contributions. Doing so results in structuring the ICT4D research with regard to theories and models (as suggested by Petersen, Vakkalanka and Kuzniarz [12]) thereby providing the basis for ICT to be contextualised and deliver its intended objective of development.

Systematic mapping involves searching the literature to get acquainted with topics that have been covered in a specific research area and to ascertain where the literature has been published [12, 13]. This study followed the guidelines for conducting a systematic mapping as stated by the World Bank [10] and Petersen, Vakkalanka and Kuzniarz [12]. According to these guidelines, systematic mapping is conducted in three phases: a) scoping of the systematic mapping, b) study identification and selection, and c) data extraction and analysis. Each of the phases is described below and the results of the mapping are reported in the "Results and Discussion" section.

2.1 Scoping of the Systematic Mapping

The specific research questions (RQs) of this systematic mapping were:

RQ 1. What are the current models and theories used in ICT4D adoption studies in developing countries?
RQ 2. What are the statistical techniques used to measure technology adoption?

As noted above, this systematic mapping reviewed ICT4D literature published between 2015 and 2019 in a developing country context. The following keywords were generated from the research questions and used as search terms: Theory OR Model OR Adoption OR Use OR Acceptance.

A theory, in general, is a system of constructs (or concepts or variables) and the interrelationships among them that jointly explain the constituent elements of a phenomenon and show how and/or why it occurs [9]. Sometimes a researcher may combine different constructs from various theories and test them in a data-driven study. Such combinations in the field of information systems can be considered models. "Adoption" is the decision of a user (for example, a farmer or a health practitioner) to start using an ICT-based service [14] and, in most cases, studies looking at adoption use adoption and "acceptance" interchangeably to specify that decision. The term "use" of an ICT-based service refers to its continuous use [15], which comes after adoption/acceptance.

2.2 Study Identification and Selection

Three main activities were done during this stage: a) deciding on the search strategies to be followed, b) developing and evaluating the search, and c) applying the inclusion and exclusion criteria.

Search Strategies Followed
A dual search strategy (snowball and a manual search) was developed in relevant ICT4D journals and conferences [5, 12]. Evidence has shown that a manual search can be more effective in identifying relevant studies than an automated one [12]. A snowball search has also been identified as an appropriate search strategy [12] for systematic mapping.

Developing and Evaluating the Search
The manual search to select appropriate journals was based on the ranking of ICT4D journals as defined in the study by Heeks [16] and used in a mapping study by Zewge

and Dittrich [5]. Based on the manual search, the three highest-ranked journals, namely, the Electronic Journal of Information Systems in Developing Countries (EJISDC), the Information Technologies and International Development (ITID) Journal, and the Information Technology for Development (ITD) Journal were targeted. Any paper published in these journals was presumed to be related to ICT4D. In addition, other relevant papers were selected based on the snowball strategy. In snowballing, the reference list and citations of relevant papers are reviewed to identify new papers whereas, in a database search, different databases are searched using predefined search strings to identify new papers [17]. Table 1 displays an overview of the journals and conferences selected in the study.

Table 1. Publications by selected journals/conferences

Acronym	Journal/Conference name	Search strategy	Ranking source	Total number of publications	Total number of selected papers
ITID	*Information Technologies and International Development Journal*	Manual	Richard Heeks (2010)	15	1
EJISDC	*Electronic Journal of Information Systems in Developing Countries*		Richard Heeks (2010)	130	10
ITD	*Information Technology for Development Journal*		Richard Heeks (2010)	20	0
Other	*African Journal of Information Systems (AJIS)*	Snowball		66	6
	Information Processing in Agriculture			1	1
	AIS Transactions on Replication Research			29	1
	Journal of the Midwest Association for Information Systems			22	0
	AMCIS 2016			1	

(*continued*)

Table 1. (*continued*)

Acronym	Journal/Conference name	Search strategy	Ranking source	Total number of publications	Total number of selected papers
	MWAIS 2018			1	
	Education Research International			1	0
	African Journal of Hospitality, Tourism and Leisure			1	0
	International Journal of Business and Management			1	0
	Journal of Management Research			1	0
	European Journals of Business Management			1	0
				290	19

As noted above, the search was performed in each of the journals using the key terms Theory OR Model OR Adoption OR Use OR Acceptance and the date range 2015 - 2019. Papers containing the search terms were downloaded from the three journals and conferences regardless of their relevance to developing countries or ICT4D. It should be noted that only papers available for downloading (open access) were selected for further analysis. Exclusion and inclusion criteria were then applied.

Inclusion and Exclusion of Papers

Any article that did not focus on a developing country was excluded from the review. The distinction between developed and developing countries was drawn from the World Bank [18]. Through a word search, any paper that focused on a developing country was tagged accordingly. As we (the authors) were only interested in papers that collected primary data and analysed data through quantitative methods, a second criterion was applied, that is, any paper with the term "review" in the title was removed from further analysis. Additionally, any paper with an abstract that did not contain the terms "survey" OR "quantitative" was also excluded. The inclusion and exclusion criteria are summarised in Table 2.

Table 2. Inclusion and exclusion criteria

Description	Inclusion	Exclusion
Research paper sources	Ranked [16] OR Identified through the snowball search strategy	Not ranked [16] OR Not identified through the snowball search strategy
Publication year	Published between 2015 and 2019	Published before 2015 and after 2019
Country of focus	Developing countries (lower-income and upper-income) as defined by the World Bank [18]	Developed countries as defined by the World Bank [18]
Words (Strings) in the title of the paper	Does not contain "Review OR Reviews"	Does contain "Review OR Reviews"
Words in the abstract	Quantitative OR Survey	Absence of Quantitative OR Survey

A systematic review of the selected papers was then performed. The Overview Quality Assessment Questionnaire (OQAQ) reported in the study by Lwoga and Sangeda [19] was modified to fit the quality assessment of theory according to Mueller and Urbach [9]. A paper was included for further analysis only if its OQAQ score was six or above (from a range of 0–10). The OQAQ was used to assign points to the paper according to various criteria. The questionnaire consisted of 15 questions divided between 1) definitions, 2) domains, 3) methodology and 4) overview, and the points were distributed amongst these categories. For instance, in the definition subsection of the OQAQ, there were three questions one of which was: "Were many theories stated and used by the study?" If the response was "Yes", the paper received two points and if the response was "No", the paper received zero points. The search strategy and selection process are depicted in Fig. 1.

Data Extraction and Analysis Process

To extract data from the identified primary studies, we developed a template as shown in Table 3. A variable (data extraction), named an "extraction variable, consisted of a data item and a value. The extraction was performed by the first author and reviewed by the second author as suggested by Badampudi, Wohlin and Petersen [17]. The extraction variables were purposefully defined to answer the research questions.

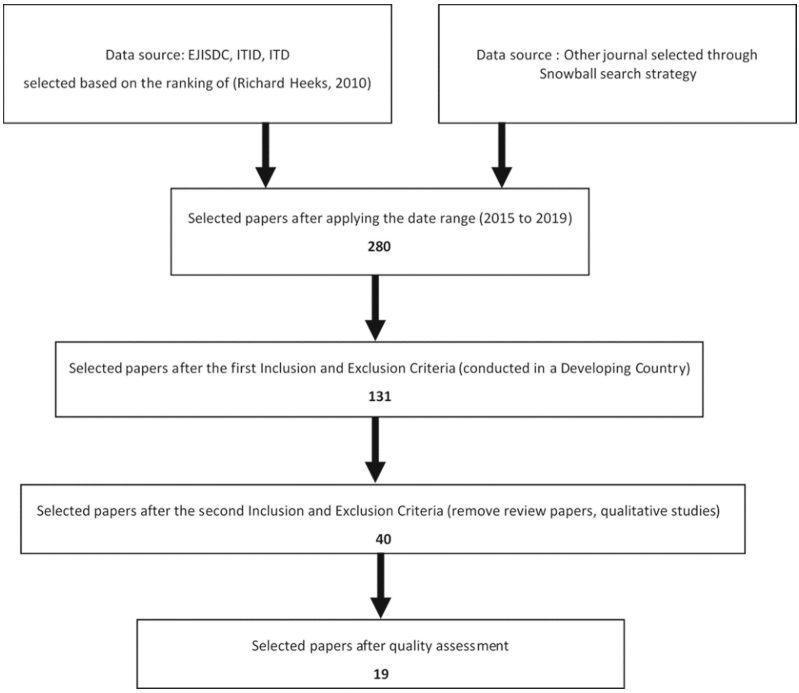


Fig. 1. Search strategy and selection process

Table 3. Data extraction form

Data item	Value	Research question
Article title	Name of the article	
Author name	Set of names of the authors	
Country	In which country has the study been done	
Field of study	What was the field of study	
Population	Whose adoption was studied	
Research questions	What were the research questions	
Main theory	What was the main theory/model of the study	RQ 1
Other theories	What other theories/models have been used	
Statistical technique	What was the statistical technique used	RQ 2

2.3 Tools Used in the Analysis Process

The selection process of the papers was managed using the open-source desktop-based application Mendeley – a database used to classify, tag and reference papers using various attributes. The application can be synchronised with the Cloud Vision API (it works with

a cloud library that is loaded into the add-in function, so there is no need to switch between applications when citing) as used and suggested in the study by Zewge and Dittrich [5]. After the selection and pre-processing phases, articles that were selected for further analysis were exported to a Microsoft Excel 2016 file. The extraction process from Mendeley to Microsoft was done using JabRef, an open-source bibliography reference manager.

3 Results and Discussion

This section reports and discusses the results. It starts by characterising the data in terms of publications per year and per country and the focus areas of the studies. The section then provides an overview of the theoretical lenses (RQ 1) and finally discusses the statistical techniques used to assess the adoption models of the selected studies (RQ 2).

3.1 Publication Per Year and Per Country

Fifty-nine percent of the selected papers were published in the *Electronic Journal of Information Systems in Developing Countries* (*EJISDC*), 5% in the *Information Technologies and International Development Journal* and 36% from other sources (*African Journal of Information Systems*, *Information Processing in Agriculture*, *AIS Transactions on Replication Research*, *Journal of the Midwest Association for Information Systems, AMCIS* 2016, *MWAIS* 2018, *Education Research International*, *African Journal of Hospitality, Tourism and Leisure*, *International Journal of Business and Management*, *Journal of Management Research*, and the *European Journal of Business Management*). Table 4 provides an overview of the publications per year and per country. It was found that there was at least one paper on ICT4D adoption each year from 2015 to 2019 in the *EJISDC*. However, in accordance with the literature [5], there is no consistency of publications related to ICT4D adoption in the selected journals over time.

Table 4. Publications per year

Journal	2015	2016	2017	2018	2019	Total
EJISDC	1	1	2	4	2	10
AJIS		1	2		4	6
ITID		1				1
Snowball				2		2
Total	1	3	4	6	6	19

Concerning the countries where the studies were conducted, the analysis revealed that Tanzania accounted for 18% of the total studies; South Africa and Nigeria accounted for 28% (14% each); while India, Ghana, Fiji, Uganda, Kenya, Burundi, Congo Brazzaville and Vietnam represented approximately 50% of the selected studies. One study could

not be linked to any country as it was an online survey. Over 86% of the studies were conducted in Africa. Thus, in the context of this study, most of the ICT4D adoption research is being conducted in Africa. This can be explained by the fact that Africa's ICT penetration is rising and, consequently, new technologies related to health, agriculture and governance, among others, are being implemented in African countries.

3.2 Focus Areas of the Studies

ICT4D is a multidisciplinary research area [20] that encompasses three study fields: computer science, information systems, and development studies [21]. As this research was concerned with adoption, only studies relevant to the adoption of information systems and, to an extent, development studies were considered. The studies were broken down according to their focus such as e-government, agriculture and others. Table 5 provides a breakdown of the studies according to their focus areas per country and year.

Table 5. Studies' focus area per country and per year

Focus area	2015	2016	2017	2018	2019	Total
Agriculture	No data	No data	1: [22] in Mali	1: [14] in Mali	2: [23] in Nigeria and [24] in Tanzania	4
E-government and e-governance	No data	1: [25] in India	No data	2: [26] in Nigeria and [27] in Ghana	1: [28] in Nigeria	4
Internet services	1: [29] in South Africa	No data	1: [30] online survey	1: [31] in Vietnam	1: [32] in Kenya	4
Education	No data	1: [33] in South Africa	1: [34] in Nigeria	No data	No data	2
Mobile money	No data	No data	1: [35] in Tanzania	1: [36] in Fiji	1: [37] in Uganda	3
Healthcare	No data	No data	1:[38] in Burundi	No data	No data	1
Tourism	No data	No data	No data	No data	1:[40]in Tanzania	1
Total	1	2	5	5	6	19

As shown in Table 5, the subfields of agriculture (including related studies), e-government (including related studies) and internet services (including related studies) were the main focus areas of ICT4D adoption research studies. The fact that many of the papers focused on agriculture is not surprising as agriculture represents the backbone

of the economies of developing countries [40, 41]. Similarly, studies investigating the adoption of e-government and e-governance by rural dwellers are not surprising as many people in developing countries (especially in Africa) live in rural areas and governments need to provide services to them [42]. On the other hand, studies related to internet services tended to look at internet coverage, internet skills and 3G adoption amongst others. These subfields have received much attention in the literature when compared to healthcare, tourism, social media and education. More studies are thus needed in developing countries on the adoption of e-health technologies and the adoption of technology in the tourism industry and education.

3.3 Theories Used and Theoretical Gaps

As pointed out earlier, a (scientific) theory can be defined as a set of variables (constructs) and relationships between these variables [9]. The main theories revealed by the dataset were:

1. The Technology Acceptance Model (TAM). Adapted from the Theory of Reasoned Action, the TAM is an information system theory that models how users come to accept technology and how they use that technology [43].
2. The Unified Theory of Acceptance and Use of Technology (UTAUT). Venkatesh, Thong and Xu [44] formulated a model that integrates and unifies the characteristics and elements of eight models and labelled it as UTAUT.
3. The Diffusion of Innovation Theory (DOI). Also referred to as Innovation Diffusion Theory (IDT) [45], it is one of the theories used in information systems to study the adoption or use of ICT services by users.
4. Technology-Organisation-Environment (TOE) Theory. This theory provides three key factors, namely, technological, organisational, and environmental [25] all of which are vital for implementing technological innovations.
5. The Theory of Planned Behavior (TPB). The TPB implies that individuals make decisions on the basis that they carefully consider all the information available [46].
6. The DIKDAR Model. This model proposes an information needs/mapping approach in which the lack of access to information exposes individuals and communities to vulnerabilities [47].

The following section discusses the theories used, how the constructs were formed and what the theoretical gaps are.

TAM

The results of the study revealed that the TAM was the most used theory. It was used in five studies (27%) as the main theory [23, 30, 31, 35, 36]. The approach of the studies using the TAM was inductive in that they were data-driven. In other words, they were, in the main, testing the theory and some had included constructs from other theories to form a conceptual framework and thus a model.

The TAM was used in the Tanzanian study by Lwoga and Lwoga [35] to investigate user acceptance of mobile payments. In addition to the TAM, the study borrowed constructs from the UTAUT and the DOI Theory/IDT. As in many studies that use the

TAM, the DOI Theory and the UTAUT, some gaps in the conceptualisation of certain constructs in the study were found. One such construct is "compatibility". As suggested by Heeks [11], compatibility should include an item on cost. Therefore, in the case of the study by Lwoga and Lwoga [35], the cost of withdrawing or depositing money, for example, should have been included as a compatibility variable. The authors of the study acknowledged that perceived cost should be included in future studies. Thus, the conceptualisation of the compatibility construct was problematic in the study.

The TAM was also used in the study by Silic, Barlow and Black [30] to evaluate the role of trust in the adoption of an open-source system. The authors added some empirical constructs to the model. These empirical constructs were "perceived reputation", "perceived opportunism", "perceived environmental risk", "perceived structural assurance", and "consumer trust" in mobile payment (M-payment) systems and internet banking. According to Müller-Bloch and Kranz [8], any empirical construct should be supported by a theoretical one which was not the case with the study by Silic, Barlow and Black [30]. This points to a methodological void as the above-mentioned constructs were not grounded in theory. While these "new" constructs did not have any reported theoretical basis, they were supported by the literature. The findings described here shed new light on the use of the TAM in adoption studies and would be relevant for future studies in addressing these highlighted gaps. One such gap could be the conceptualisation of constructs (both theoretical and empirical).

The third study using the TAM was that of Uduji, Okolo-Obasi and Asongu [23] who used the model to investigate the adoption of e-wallet amongst informal farm entrepreneurs in Nigeria. The conceptualisation of the constructs in the study was unclear as none of the constructs used could be linked to the TAM. This suggests that the validity of the study can be queried. Questions such as "Where did the study get the constructs?", "How were they formulated?" and "Were they tested before being used in the study?" need to have been asked and answered.

The TAM was employed in the study by Tapanainen, Dao and Nguyen [31] which investigated the adoption of 3G services in Vietnam. The authors extracted the constructs "intention to adopt" and "social influence" from the UTAUT and the constructs "IT self-efficacy" and "service quality" from the TPB. However, while some empirical factors were added to the study the theoretical evidence of these was not reported. This again poses the question of how the constructs were formed and thus the validity of the constructs.

In Fiji, the TAM was used to study the factors influencing the intention to use mobile value-added services by women-owned microenterprises [36]. Some constructs were borrowed from the UTAUT and TAM 2 (TAM 2 is an extension of TAM). These constructs were "social influence" and "intention to use". We were not able to find a gap in the way the constructs of this study were formed.

In summary, most of the studies using the TAM have gaps in the conceptualisation of the constructs of the theory. This may be attributed to these studies not conducting a pilot or pre-study to operationalise and better capture the intended question. Furthermore, one may link these gaps in the use of the TAM to the philosophical approach that these studies employed, namely, positivism. If they were to have used a pragmatist approach, they may have been forced to conduct a qualitative and quantitative study sequentially.

Doing so may have filled the gap in the conceptualisation of the construct. Furthermore, the TAM itself might be the source of the problem. Perhaps, that is why TAM has been extended to TAM 2 and TAM 3 and later incorporated into the UTAUT.

UTAUT and DOI

The DOI Theory was used in 16% of the selected papers as the main theory. The same applies to the UTAUT. The DOI Theory was used in the context of Mali [14] and Burundi [38]. The former entailed the adoption of ICT-based agricultural (input) information by farmers while the latter pertained to the adoption of mHealth in the country. Again, in the conceptualisation of the "compatibility" construct, the cost factor was not considered. The Mali studies included some constructs from the DIKDAR Model [47] and the Theory of Knowledge [48]. The lack of inclusion of the cost factor in the conceptualisation of the compatibility construct was also observed in the study of Ndayizigamiye and Maharaj [38] in Burundi.

Studies that used the UTAUT as the main theory were those of Chopra and Rajan [25], Malinga and Maiga [37] and Olaleye [28]. Chopra and Rajan [25] investigated the adoption of e-government technologies for food distribution in India. Although no theoretical gaps could be identified, the study did have some methodological gaps. For instance, the authors used the Average Variance Extracted (AVE) of PLS-SEM to measure discriminant validity instead of construct validity.

Malinga and Maiga [37] used the UTAUT to determine the factors that influence the adoption of mobile money by traders in Uganda. The study extended the UTAUT by adding two new factors brought from empirical studies that were not supported by any theory. As suggested by Mueller and Urbach [9], empirical constructs should be supported by theoretical ones.

The third study using the UTAUT was that of Olaleye [28] which investigated the determinants of electronic invoicing in Nigeria. It used "social influence" as well as the "image" as two distinct constructs. However, literature [15, 31, 49] suggests that the image construct from the DOI 2 Theory (an extension of the DOI Theory) is the same as the "social influence" construct of the UTAUT. Moreover, the author did not provide the items used to measure the image construct. Thus, one could not ascertain the face validity of the construct. The findings reported here shed new light on the use of the UTAUT and the DOI Theory in adoption studies and would be relevant for future studies in addressing these gaps.

Other Theories

Other theories that were used included the Institutional Theory in Nigeria [34] and Kenya [32]; the Theory of Reasoned Action in South Africa[33]; the TAM 2 in Tanzania [24]; the Green IT Adoption Model in Tanzania[39]; the Technology-Organisation-Environment (TOE) Framework in Nigeria [26] and South Africa [29]; the Process Virtual Theory [50] in Ghana; the Theory of Information Systems Improvisation [51] in Congo Brazzaville; and the United Nations Development Programme (UNDP) Framework in Tanzania[52]. The conceptualisation of the constructs is the most pertinent gap in the majority of these studies. For example, Ofoeda, Boateng, and Asmah [50] used three variables, that is, "ICT infrastructure", "internet access" and "communication" to measure the technological factors construct. However, ICT infrastructure could have been measured as a latent

variable instead of an observed variable. Another example of a gap comes from the study of Omotayo [34] in which it was hypothesised (page 12) that "there is a significant relationship between Normative Pressures [NPs] and attitude of teachers towards the use of e-learning." However, the author did not specify whether the relationship was positive or negative. Similarly, the same gap was observed in the study conducted by Oredo, Njihia and Iraki [32] in Kenya. Using the Institutional Theory, the authors hypothesised (page 138) that "there is a relationship between coercive pressures and cloud computing adoption by Financial, ICT and Manufacturing firms in Kenya." However, the study did not mention whether the relationship was significant, positive or negative.

A further theoretical gap comes from the study of Gareeb and Naicker [29]. This study used the TOE Framework as the main theory to determine the factors that influence the adoption of broadband internet technologies in South Africa. The study grouped distinct constructs, that is, "relative advantage", "perceived ease of use", "service quality", "availability", "compatibility" and "observability" under a single construct called "technology context". However, relative advantage, for example, is a construct on its own and is measured using many items (observable variables). Hence, as the study did not provide the items used to measure relative advantage, the face validity could not be assessed.

Regarding the qualitative/mixed studies [51, 53], no theories were identified as having been used. Rather, the studies gathered empirical factors from the literature and conducted interviews around these factors. The study of Momo and Twum-Darko [53] for instance, examined the "acceptance factors" of health information systems in the Congo-Brazzaville with 18 key informants. The authors could have matched their empirical factors with theoretical ones from the TAM or the UTAUT. This may be a gap as they could have grounded these factors in theory as recommended in the field of information systems research [6, 9].

3.4 Methodological Gaps

A methodological gap appears when new research methods are necessary to provide new insights or to avoid distorted findings [8]. In this study and following the framework suggested by Müller-Bloch and Kranz [8] on how to identify methodological gaps, we only assessed the statistical techniques that the selected studies used to generate their findings.

Approximately 45% of the selected studies used Structural Equation Modeling (SEM) (Partial Least Square (PLS) component) to assess relationships between variables. This is not surprising as the literature [54–56] reports that SEM (and mostly its PLS component) is the statistical technique most used in information systems research. SEM allows researchers to answer a set of interrelated research questions by modeling the relationships among many variables in a single, systematic, and comprehensive analysis [57]. The capability of PLS-SEM to simultaneously analyse a set of variables differs greatly from most first-generation regression models such as linear regression, LOGIT, ANOVA and MANOVA which can only analyse one layer of linkages between independent and dependent variables at a time [57, 58]. The use (and overuse) of PLS as a method to test one's hypotheses may have resulted in the publication of a large

number of studies the findings of which are invalid [59]. However, this contention has been countered in the study by Henseler et al. [60].

Another methodological gap emanated from the fact that most of the selected studies used either the Fornell-Larcker or the cross-loadings criteria to establish the discriminant validity of their constructs. While these criteria can be used, the literature [61, 62] suggests that the heterotrait-monotrait ratio (HTMT) better detects the existence or lack of discriminant validity than the other two criteria (Fornell-Larcker or the cross-loadings). Hence, this is a methodological gap as a second study with the same data may indeed find that the constructs' discriminant validity may not be established using the HTMT. Furthermore, some gaps in a composite variable were measured in a few studies. For example, the variable "green ebusiness adoption and use" was measured as a single latent variable in the study by Masele [39]. However, adoption and use are two variables that should not be measured as a single variable. Consequently, the R^2 (the measure of the extent of the variance of the dependent variable explained by the independent variables) generated from Masele's study may not reflect adoption and use at the same time but rather adoption or use. Hence, further analysis of the data generated through this study may lead to new insights.

The remaining studies used regression analysis [26, 33, 37]; principal component analysis[38]; multiple regression [27, 34]; the Bivariate Probit Model [23]; content analysis [51]; descriptive statistics [51, 52]; and ANOVA statistical tests [24, 53].

A methodological gap in the study by Omotayo [34] was identified. While the study presented some controlled variables (age, gender and income), a closer look at the model generated by the study revealed that these variables were independent ones having a direct effect on the dependent variable and thus were not controlling variables. A controlled (monitored or constant) variable is maintained in the model to observe if it has a minimal impact on the relationship between the independent variable and the dependent variable [63]. The controlled variable is not usually hypothesised, which was not the case in Omotayo's [34] study.

The findings reported here shed new light on the use of PLS-SEM and other statistical techniques in adoption studies and would be relevant for future studies addressing these gaps.

4 Conclusion

The purpose of this study was to identify the gaps in the theories and methodologies used in ICT4D adoption research in the context of developing countries. The study found that TAM was the most used theory (27%) in ICT4D adoption research, followed by the DOI Theory (16%), the UTAUT (16%) and lastly the Institutional Theory (11%). The remaining theories used were the TAM 2 (6%), the Green IT Adoption Model (6%), the TOE Framework (6%), the Process Virtual Theory (6%) [50], and the United Nations Development Programme (NDP) Framework (6%) [52]. The major theoretical gaps identified in this systematic mapping were related to the conceptualisation of the theories' constructs. The identified methodological gaps were related to the way variables were grouped and the type of statistical tests used to establish the discriminant validity. Overall, the results from this mapping study suggest that there is a need to undertake

more studies on the contextualisation and conceptualisation of the established theories such as the TAM, the DOI Theory and the UTAUT so that ICT4D adoption research can deliver its intended goal in the context of developing countries.

The authors acknowledge that the selection process may have left out papers that qualified to be analysed. In addition, some gaps could have been missed during the analysis process. It is, therefore, recommended that further studies using other qualitative or quantitative methods of inquiry be conducted to identify any gaps that could have been missed. Finally, as noted, this paper is part of an ongoing study and data from studies published from January 2020 to May 2022 are currently being analysed.

References

1. Heeks, R.: Theorizing ICT4D research. Inf. Technol. Int. Dev. **3**(3), 1–4 (2007). https://doi.org/10.1162/itid.2007.3.3.1
2. S. Lim, T. Saldanha, S. Malladi, Melville, N.P.: Theories used in information systems research: identifying theory networks in leading IS journals. In: International Conference on Information Systems, pp. 10. Available: http://aisel.aisnet.org/icis2009 (2009)
3. Karanasios, S.: Framing ICT4D research using activity theory : a match between the ICT4D field and theory? Inf. Technol. Int. Dev. **10**(2), 1–17 (2014)
4. Heeks, R.: Using competitive advantage theory to analyze IT sectors in developing countries: a software industry case analysis. Inf. Technol. Int. Dev. **3**(3), 5–34 (2006). https://doi.org/10.1162/itid.2007.3.3.5
5. Zewge, A., Dittrich, Y.: Systematic mapping study of information technology for development in agriculture (the case of developing countries). Electron. J. Inf. Syst. Dev. Ctries. **82**(2), 1–25 Available: http://www.is.cityu.edu.hk/staff/isrobert/ejisdc/82-2.pdf (2017)
6. Sein, M.K., Thapa, D., Hatakka, M., Sæbø, Ø.: A holistic perspective on the theoretical foundations for ICT4D research. Inf. Technol. Dev. **25**(1), 7–25 (2019). https://doi.org/10.1080/02681102.2018.1503589
7. Kante, M., Ndayizigamiye, P.: A systematic mapping of the adoption of Internet of Things to provide healthcare services in developing countries. In: Ndayizigamiye, P., Barlow-Jones, G., Brink, R., Bvuma, S., Minty, R., Mhlongo, S. (eds.) Perspectives on ICT4D and Socio-Economic Growth Opportunities in Developing Countries, pp. 99–126. IGI Global, Pennsylvania (2020)
8. Müller-Bloch, C., Kranz, J.: A Framework for rigorously identifying research gaps in qualitative literature reviews. In: Thirty Sixth International Conference on Information Systems, pp. 1–19 (2015)
9. Mueller, B., Urbach, N.: The why, what, and how of theories in IS research. In: Thirty Fourth International Conference on Information Systems, vol. 192, no. 111, pp. 181–192 (2013)
10. (World Bank): WDI - classifying countries by income. World Development Indicator (WDI). https://datatopics.worldbank.org/world-development-indicators/stories/the-classification-of-countries-by-income.html (2019). Accessed 25 Feb 2020
11. Heeks, R.: Information and Communication Technology for Development (ICT4D). Routledge (2017). https://doi.org/10.4324/9781315652603
12. Petersen, K., Vakkalanka, S., Kuzniarz, L.: Guidelines for conducting systematic mapping studies in software engineering: An update. Inf. Softw. Technol. **64**, 1–18 (2015). https://doi.org/10.1016/j.infsof.2015.03.007
13. Thiébaut, J.: Deuxième note sur quelques plantes de la flore libano-syrienne. Bulletin de la Societe Botanique de France **82**(2), 189–196 (1935). https://doi.org/10.1080/00378941.1935.10832973

14. Kante, M., Oboko, R., Chepken, C.: An ICT model for increased adoption of farm input information in developing countries: a case in Sikasso, Mali. Inf. Process. Agric. **6**(1), 26–46 (2019). https://doi.org/10.1016/j.inpa.2018.09.002
15. Venkatesh, V., Morris, M.G., Davis, G.B., Davis, F.D.: User acceptance of information technology: Toward a unified view. MIS Q. **27**(3), 425–478 (2003)
16. Richard Heeks: ICT4D Journal Ranking Table. ICTs for Development. https://ict4dblog.wordpress.com/2010/04/14/ict4d-journal-ranking-table/ (2010). Accessed 30 Jul 2019
17. Badampudi, D., Wohlin, C., Petersen, K.: Experiences from using snowballing and database searches in systematic literature studies. In: ACM International Conference Proceeding Series, vol. 27–29 (Apr 2015). https://doi.org/10.1145/2745802.2745818
18. ISGE: List of Developing Countries. ISGE 2018. https://isge2018.isgesociety.com/registration/list-of-developing-countries/ (2018). Accessed 17 Feb 2020
19. Lwoga, E.T., Sangeda, R.Z.: ICTs and development in developing countries: a systematic review of reviews. Electron. J. Inf. Syst. Dev. Ctries. **85**(1), e12060 (2019). https://doi.org/10.1002/isd2.12060
20. Unwin, T.: s (2007) edited volume on", pp. 1–16 (2010). https://doi.org/10.1002/j.1681-4835.2010.tb00313.x
21. Thapa, D., Sæbø, Ø.: Exploring the link between ICT and development in the context of developing countries: a literature review. Electron. J. Inf. Syst. Dev. Ctries. **64**(1), 1–15 (2014). https://doi.org/10.1002/j.1681-4835.2014.tb00454.x
22. Kante, M., Oboko, R., Chepken, C.: Influence of perception and quality of ICT-based agricultural input information on use of ICTs by farmers in developing countries: case of Sikasso in Mali. Electron. J. Inf. Syst. Dev. Ctries. **83**(1), 1–21 (2017). https://doi.org/10.1002/j.1681-4835.2017.tb00617.x
23. Uduji, J.I., Okolo-Obasi, E.N., Asongu, S.A.: The impact of e-wallet on informal farm entrepreneurship development in rural Nigeria. Electron. J. Inf. Syst. Dev. Ctries. **85**(3), e12066 (2019). https://doi.org/10.1002/isd2.12066
24. Lubua, E.W., Prof, M.E.K. : The influence of socioeconomic factors to the use of mobile phones in the agricultural sector of Tanzania. African J. Inf. Syst. **11**(4) (2019)
25. Chopra, S., Rajan, P.: Modeling intermediary satisfaction with mandatory adoption of e-government technologies for food distribution. Inf. Technol. Int. Dev. **12**(1), 15–34. Available: https://lib.dr.iastate.edu/cgi/viewcontent.cgi?article=2010&context=abe_eng_pubs (2016)
26. Olumoye, M.Y., Govender, I.: An empirical investigation of factors influencing integrated e-Government implementation in Nigeria: a case of housing and urban development agency. Electron. J. Inf. Syst. Dev. Ctries. **84**(1), e12012 (2018). https://doi.org/10.1002/isd2.12012
27. Ofoeda, J., Boateng, R., Asmah, A.: Virtualization of government-to-citizen engagement process: enablers and constraints. Electron. J. Inf. Syst. Dev. Ctries. **84**(5), e12037 (2018). https://doi.org/10.1002/isd2.12037
28. Olaleye, S.A.: The need for green companies in Nigeria: a study of electronic invoicing. African J. Inf. Syst. **11**(1), 26 (2019)
29. Gareeb, P.P., Naicker, V.: Determinants for South African SMEs to adopt broadband internet technologies. Electron. J. Inf. Syst. Dev. Ctries. **68**(1), 1–24 (2015). https://doi.org/10.1002/j.1681-4835.2015.tb00491.x
30. Silic, M., Barlow, J., Back, A.: Evaluating the role of trust in adoption: a conceptual replication in the context of open source systems. AIS Trans. Replication Res. **4**(March), 1–17 (2018). https://doi.org/10.17705/1atrr.00021
31. Tapanainen, T.J., Dao, K.T., Nguyen, H.T.T.: Investigating adoption factors of 3G services: an empirical study in Vietnam. Electron. J. Inf. Syst. Dev. Ctries. **84**(2), e12022 (2018). https://doi.org/10.1002/isd2.12022
32. Oredo, J.O., Njihia, J.: Adoption of cloud computing by firms in Kenya : the role of institutional pressures. African J. Inf. Syst. **11**(3), 25 (2019)

33. Kalema, B.M.M., Motsi, L., Motjolopane, I.M.: Utilizing IT to enhance knowledge sharing for school educators in developing countries. Electron. J. Inf. Syst. Dev. Ctries. **73**(1), 1–22 (2016). https://doi.org/10.1002/j.1681-4835.2016.tb00533.x
34. Omotayo, F.O.: Influence of socio-environmental forces on use of e-learning by teachers in selected tertiary institutions in Oyo state, Nigeria. African J. Inf. Syst. **9**(1), 36 (2017)
35. Lwoga, E.T., Lwoga, N.B.: User acceptance of mobile payment: the effects of user-centric security, system characteristics and gender. Electron. J. Inf. Syst. Dev. Ctries. **81**(1), 1–24 (2017). https://doi.org/10.1002/j.1681-4835.2017.tb00595.x
36. Sathye, S., Prasad, B., Sharma, D., Sharma, P., Sathye, M.: Factors influencing the intention to use of mobile value-added services by women-owned microenterprises in Fiji. Electron. J. Inf. Syst. Dev. Ctries. **84**(2), e12016 (2018). https://doi.org/10.1002/isd2.12016
37. Malinga, R.B., Maiga, G.: A model for mobile money services adoption by traders in Uganda. Electron. J. Inf. Syst. Dev. Ctries. **86**(2), e12117 (2019). https://doi.org/10.1002/isd2.12117
38. Ndayizigamiye, P., Maharaj, N.: Determinants of mobile health adoption in Burundi. African J. Inf. Syst. **9**(3), 171–191 (2017)
39. Masele, J.J.: Modeling green ebusiness adoption among small and medium tourism enterprises in Tanzania. African J. Inf. Syst. **11**(3), 24 (2019)
40. Aker, J.C., Ghosh, I., Burrell, J.: The promise (and pitfalls) of ICT for agriculture initiatives. Agric. Econ. (United Kingdom) **47**, 35–48 (2016). https://doi.org/10.1111/agec.12301
41. Kante, M., Oboko, R., Chepken, C., Hamunyela, S.: Farmers' perceptions of ICTs and its effects on access and use of agricultural input information in developing countries: Case of Sikasso, Mali. In: 2017 IST-Africa Week Conf., pp. 1–8 (May 2017). https://doi.org/10.23919/ISTAFRICA.2017.8101973
42. Furuholt, B., Sæbø, Ø.: The role telecentres play in providing e-government services in rural areas. Electron. J. Inf. Syst. Dev. Ctries. **84**(1), e12006 (2018). https://doi.org/10.1002/isd2.12006
43. Kante, M., Oboko, R., Chepken, C., Hamunyela, S.: Farmers' perceptions of ICTs and its effects on access and use of agricultural input information in developing countries: Case of Sikasso, Mali. In: IST-Africa Week 2017, pp. 1–8 (2017). [Online]. Available: www.IST-Africa.org/Conference2017
44. Venkatesh, V., Thong, J.Y.L., Xu, X.: Unified theory of acceptance and use of technology: a synthesis and the road ahead. J. Assoc. Inf. Syst. **17**(5), 328–376 (2016). https://doi.org/10.17705/1jais.00428
45. Rogers, E.M., Singhal, A., Quinlan, M.M.: Diffusion of innovations 1. In: Stacks, D.W., Salwen, M.B., Eichhorn, K.C., Stacks, D.W., Salwen, M.B., Eichhorn, K.C. (eds.) An Integrated Approach to Communication Theory and Research, pp. 415–434. Routledge (2019). https://doi.org/10.4324/9780203710753-35
46. Beck, L., Ajzen, I.: Predicting dishonest actions using the theory of planned behavior. J. Res. Pers. **25**(3), 285–301 (1991). https://doi.org/10.1016/0092-6566(91)90021-H
47. Heeks, R., Molla, A.: Development informatics compendium of approaches. Manchester. Available: http://www.sed.manchester.ac.uk/idpm/research/publications/wp/di/index.htm (2009)
48. Foucault, M.: Power/Knowledge. Pantheon Books, New York (1977)
49. March, S., et al.: Attitudes toward e-mental health services in a community sample of adults: online survey. J. Med. Internet Res. **20**(2), e59 (2018). https://doi.org/10.2196/jmir.9109
50. Ofoeda, J., Boateng, R., Asmah, A.: Virtualization of government-to-citizen engagement process: enablers and constraints. Electron. J. Inf. Syst. Dev. Ctries. **84**(5), 1–16 (2018). https://doi.org/10.1002/isd2.12037
51. Momo, A.M., Twum-Darko, M.: Usability of online business registration improvisation as Congo-Brazzaville re-branding tool. Electron. J. Inf. Syst. Dev. Ctries. **83**(1), 1–20 (2017). https://doi.org/10.1002/j.1681-4835.2017.tb00611.x

52. Barakabitze, A.A., Fue, K.G., Sanga, C.A.: The use of participatory approaches in developing ICT-based systems for disseminating agricultural knowledge and information for farmers in developing countries: The case of Tanzania. Electron. J. Inf. Syst. Dev. Ctries. **78**(1), 1–23 (2017). https://doi.org/10.1002/j.1681-4835.2017.tb00576.x
53. Duffett, R.G.: Influence of Facebook commercial communications on Generation Z's attitudes in South Africa. Electron. J. Inf. Syst. Dev. Ctries. **81**(1), 1–22 (2017). https://doi.org/10.1002/j.1681-4835.2017.tb00600.x
54. Evermann, J., Tate, M.: Comparing the predictive ability of pls and covariance models. In: Thirty-fifth International Conference on Information Systems (ICIS), pp. 1–18 (2014)
55. Henseler, J., Hubona, G., Ray, P.A.: Using PLS path modeling in new technology research : updated guidelines. **116**(1), 2–20 (2016), https://doi.org/10.1108/IMDS-09-2015-0382
56. Kante, M., Chepken, C., Oboko, R.: Partial Least Square Structural Equation Modelling' use in information systems: An updated guideline of practices in exploratory settings. Kabarak J. Res. Innov. **6**(1), 49–67 (2018). Available: http://eserver.kabarak.ac.ke/ojs/
57. Gefen, D., Straub, D.W., Boudreau, M.-C.: Structural equation modeling and regression : Guidelines for research practice. Commun. Assoc. Inf. Syst. **4**(October), 7 (2000). https://doi.org/10.1.1.25.781
58. Urbach, N., Ahlemann, F.: Structural equation modeling in information systems research using partial least squares. J. Inf. Technol. Theory Appl. **11**(2), 5–40 (2010). https://doi.org/10.1037/0021-9010.90.4.710
59. Rönkkö, M., Parkkila, K., Ylitalo, J.: Use of partial least squares as a theory testing tool - an analysis of information systems papers. In: European Conference on Information Systems (ECIS), no. 2012, pp. 145 (2012)
60. Henseler, J., et al.: Common beliefs and reality about PLS: comments on Ronkko and Evermann (2013). Organ. Res. Methods **17**(2), 182–209 (2014). https://doi.org/10.1177/1094428114526928
61. Garson, G.D.: Partial Least Squares: Regression and Structural Equation Models, 2016th edn. Statistical Associates Publishing, Asheboro (2016)
62. Hair, J.F., Risher, J.J., Sarstedt, M., Ringle, C.M.: When to use and how to report the results of PLS-SEM. Eur. Bus. Rev. **31**(1), 2–24 (2019). https://doi.org/10.1108/EBR-11-2018-0203
63. Fung, H.P.: Is there any difference between 'control variable' and 'moderating variable?. www.researchgate.net/post/Is_there_any_different_between_control_variable_and_moderating_variable2 (2015). Accessed 25 Jul 2017

Determinants of Attitudes and Intentions to Use a Digital Library System: The Role of Meta-cognitive Strategies Amongst End-Users at a Historically Disadvantaged University in South Africa

Niki Kunjuzwa, Willie Chinyamurindi(✉), and Liezel Cilliers

University of Fort Hare, Alice Campus, Alice, Eastern Cape Province, South Africa
wchinyamurindi@ufh.ac.za

Abstract. Universities globally are adopting Information Communication Technology to aid education delivery. Digital libraries emerge as a popular tool in achieving this. A challenge often experienced with using digital library systems consist of end-user technology adoption and usage issues. The study was aimed at exploring the role of meta-cognitive strategies and user attitudes on the intention to use a digital library system. A quantitative research approach using a survey method was used amongst 316 students as end-users at a rural historically disadvantaged university campus in South Africa. The findings show that meta-cognitive strategies: a) awareness, b) self-checking and debugging, c) planning, and d) cognitive strategy influence end-user attitudes. Further, user attitudes were found to inform intention to use a digital library system. Based on the findings, we proffer implications that affect end-users and stakeholders within the university system on how to improve aspects related to digital library system usage.

Keywords: Meta-cognitive strategies · Historically disadvantaged institution · End-user · Digital library systems

1 Introduction

1.1 The Higher Education Context

The higher education globally is experiencing change [1]. One notable driver of such a change is the growth in technology. This trend in education has resulted in new enabling technologies that have facilitated the virtual delivery of academic programs in universities [2]. The impact of technology in higher education is noted in teaching and learning as well as research [3]. It is argued that institutions must develop policies that support how they want to educate young technology users [4]. Research also shows that there are significant barriers to technology use in higher education institutions. Calls have been made for research that addresses end-user barriers to technology usage and adoption [5].

P. Ndayizigamiye et al. (Eds.): IDIA 2022, CCIS 1774, pp. 35–45, 2023.
https://doi.org/10.1007/978-3-031-28472-4_3

Technological advancements have also resulted in the modernization of library services [6]. Saliently this has led to the growth in popularity of digital library systems. A digital library system consists of a software system that is based on a defined architecture and provides functionality to access information [7]. A digital library system enables the users to interact effectively with the information distributed across the network and thus provides universal access to digitized information. The benefit here is for a system that can host vast online databases with a wealth of information to guide learning and research [8].

In 2013, the South African government, through the Department of Higher Education and Training (DHET) and the Department of Science and Innovation (DSI), embarked on a strategy to assist the historically disadvantaged institutions (HDIs) to increasing their research capacity through bilateral engagements [9]. HDIs are a common term in South Africa referring to a cluster of universities created under apartheid to cater to Africans and other non-white populations. These are universities commonly found in the former homeland, most located in rural settings [10]. Calls exist for more research within the context of HDIs especially around technology usage as this area has received scant empirical focus [11].

The use of Information Communication Technology (ICT) is thriving in various fields. The South African government has provided opportunities for the use of ICT, especially for education. This has potentially assisted students especially from disadvantaged backgrounds in accessing technology for their career advancement [12]. This has also led to calls to improve the institutional factors that can potentially affect the uptake of technology including digital library systems [13]. Potentially this can help increase the adoption of technology and minimize resistance [14].

Academics are encouraged to transform teaching practices by integrating technology to align with the 21st century needs [15, 16]. The process of embracing this new shift has posed challenges, one of the possible challenges is user resistance. Researchers identify user resistance to be one of the causes of the failure of technology penetration. Resistance means the opposite of acceptance [17]. Focus has been on the adoption of technology, yet an often-neglected angle is understanding the role of social and psychological factors amongst end-users. This study magnifies the role of meta-cognitive strategies as important social and psychological issues that can affect end-user usage.

Flavell introduced the concept of meta-cognition as a concept that concerns aspects of intelligence with emphasis on how information is monitored, stored and retrieved in making a decision [18]. In essence, addressing meta-cognitive issues concerns addressing: a) an awareness of what one knows - "metacognitive knowledge," b) an awareness of what one can do - "metacognitive skills," and c) an awareness of what one knows about their cognitive abilities - "metacognitive experience [19]. Given the presented study context, the following section explores the problem statement.

2 Problem Statement

Calls exist in the literature to understand issues on end-user adoption of technology [20]. This potentially includes the role that meta-cogntive strategies can have on technology adoption. Further, the role of social context to all this becomes important. This research

narrows focus to those end-users classified as historically disadvantaged. In South Africa, such a context becomes important especially given changes happening within the higher education landscape [21, 22]. Also emerging from the COVID-19 pandemic higher education systems have adopted the online mode of learning as response to the challenges posed by the pandemic [23]. Given this a need exists to understand how barriers that impact technology adoption can be addressed, especially from an end-user perspective [24]. One such issue requiring attention is the role of meta-cognitive strategies.

3 Aim of the Study

The study investigates determinants of end-user attitudes and intentions to use a digital library system considering the role of meta-cognitive strategies at a historically disadvantaged institution in South Africa.

4 Objectives of the Study

This study aims to examine the influence of meta-cognitive strategies (awareness, self-checking, debugging, planning, and cognitive) on users' attitudes toward the intentions to use digital library systems.

5 Literature Review

In order to gain an overall understanding of the digitalization of library systems, this paper will review the existing literature considering both the theoretical and empirical literature.

5.1 Underpinning Theories

The benefits of any technology can only be realized when it is genuinely accepted and utilized. Therefore, research has focused on the user acceptance of technologies. Davis introduced the Technology Acceptance Model (TAM), which focuses on perceived usefulness and perceived ease-of-use and their influences on user attitudes [25]. The TAM has been pioneering in assisting our understanding of technology acceptance [26]. The priority here is on the importance of user attitudes also as precursors to technology adoption. A plethora of work lauds the role of the TAM [27–29]. Potentially this study extends understanding of the TAM within a South African digital library system.

Linked to the TAM is also the Theory of Planned Behaviour (TPB). The TPB shows the various factors influencing behavior including intentions [30]. Intentions are assumed to capture the motivational factors that influence behaviour and indications of the extent individuals are willing to go in improving their performative behaviour. As a general rule, the stronger the intention to engage in a behavior, the more likely its performance should be [31]. The theory states that your behavior is a function of attitudes, subjective norms, and perceived behavioural control.

5.2 Metacognitive Strategies and User Attitudes

People have a fundamental need to understand how things work inside their minds. Potentially this can influence the behaviours ensued [32]. A sense of awareness helps individuals to understand their thought processes allowing them a better understanding of the self. This places importance on the development of meta-cognitive strategies [33]. Such a focus on meta-cognitive strategies potentially can also lead to positive end-user attitudes including individual action to the task being performed [34].

The literature attributes some meta-cognitive strategies to be in existence. These include self-checking, debugging, and meaning taking [35]. All these allow the individual to be able to reflect on their intended action, the situation, and the expected outcome [35]. The user thinks about whether the proposed action is the right action for the situation, thus potentially forming a positive attitude [33]. Planning is another crucial meta-cognitive strategy, whereby the more you can engage on a task using the meta-cognitive strategy, the more relaxed an individual can be in performing the task [35]. In essence, planning as a meta-cognitive strategy enables an individual to be innovative also to the task at hand [36]. The meta-cognitive component is strongly believed to influence beliefs, knowledge, and thoughts that influence attitudes [34]. Subsequently, allowing for individuals to be able to self-correct and learn from mistakes [36]. Conversely, not fully utilizing meta-cognitive strategies can be attributed to impulsive behaviours [37].

5.3 Meta-cognitive Strategies and User Intentions

Considerable effort has been invested in innovative practices regarding digital library systems. Although the use of meta-cognitive strategies in the learning process is acknowledged [35], calls exist for more research related to this [37]. Given the importance of meta-cognitive strategies, this study seeks to examine the effect of such strategies on technology adoption. Emphasis on the technology is given to digital library systems.

5.4 User Attitude and Intention to Adopt Digital Library Systems

Positive attitudes are needed in an individual for them to be motivated and engaged in a task. Previous research appears to be supporting this [38, 39]. Attitudes are crucial in informing usage behaviour [40]. For digital library systems, the thinking here is that paying attention to meta-cognitive strategies is crucial in informing attitudes and subsequently stimulating the user's intention to use the digital library system. Based on the presented literature, it can be expected that:

H1: Meta-cognitive strategy have a positive effect on user's attitude toward digital library system usage.
H2: User attitude towards digital library systems have a positive effect on the intention to use digital library systems.

Figure 1 presents the research model for the study:

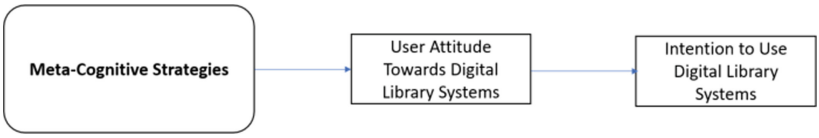


Fig. 1. Research model for the study.

6 Research Approach

The research adopted a positivist research philosophy utilizing a quantitative survey design technique. The aim was to test for association between the variables in the study [41]. A survey was deemed the best technique for sampling a large respondent pool [42]. The research adopted a convenience sampling approach using a sample of students as potential end-users to a digital library system [43].

6.1 Research Instrument

The instrument used a five-point Likert scale (ranging from strongly disagree to strongly agree) with 22 items. The study's independent variable was the Meta-cognitive Assessment Inventory [44]. The 27-item scale reflects components of meta-cognitive strategies to include: a) *knowledge of cognition* and b) *regulation of cognition.* The scale was noted to have good ratings of reliability and validity (Cronbach Alpha r = 0.96). An example item: "*I always know how to organize time for learning*". User attitudes towards using digital library systems were measured using a 3-item scale measuring attitudes towards using digital library systems by end users [45]. An example item for the user attitudes towards digital library systems is: "*I believe it is beneficial for me to use a digital library system.*" The scale was also noted to have good ratings of reliability and validity (Cronbach Alpha r = 0.81). The final scale measured the intention to use digital library systems [46]. An example from the three-item scale was "*I am willing to adopt digital library systems.*" This scale also had good ratings of reliability and validity (Cronbach Alpha r = 0.85).

The study was conducted at a historically disadvantaged institution in the Eastern Cape Province of South Africa. Ethical clearance was applied for and granted through the University Research Ethics Committee. As mentioned earlier, a convenience sampling technique was used. Five hundred questionnaires were distributed to undergraduate and postgraduate students at the participating institution. From this, 316 students filled out the survey during the eight months data collection period between November 2021 and 2022 July. The response rate to the survey as at 63.2%.

7 Results

7.1 Biographical Results

Table 1 presents the biographical results of the participants in the study.

Table 1 shows most respondents to be female 64% and mostly Black Africans (92%). In terms of study levels, undergraduate students comprised 69% of the sample, and

Table 1. Demographic profile of respondents

Demographic profile	Frequency	%
Gender		
Female	201	64
Male	115	36
Age		
20–25	228	72
Above 25	88	28
Ethnicity		
Black (African)	290	92
Coloured	26	8
Level of Study		
Undergraduate	217	69
Postgraduate	99	31
Where do you access the internet the most?		
Mobile Phone	201	67
Computer Lab	115	33

n = 316
Source: Authors' own work.

postgraduate studies at 31%. In terms of access to the internet, most of the participants cited accessing the internet from a mobile phone at 71%.

7.2 Reliability and Validity

Tests for reliability and validity were conducted. Table 2 shows the results of the Cronbach Alpha coefficients tests conducted.

Table 2. Cronbach's alpha coefficients of instrument used for this research.

Scale	Cronbach's alpha coefficient
Meta-cognitive strategy	0.727
User attitude toward digital library system	0.813
Intention to use digital library system	0.789

n = 316

Based on the results in Table 2, there is sufficient evidence to show that the study scales have good internal consistency ratings. All Cronbach Alpha coefficients were above the required threshold of 0.70 (Nunnally, 1978).

7.3 Regression Analysis

Table 3 reveals the results of the regression analysis. Concerning hypothesis 1, the resulting F-value (F = 21.91; Pr > F  0.000) showed that the model was highly significant. However, only 23.9% (R-square = 0.239) of the variation in the role of metacognitive strategy on user-attitude toward digital library use is being explained by the explanatory variable. The parameter estimates output showed that meta-cognitive strategy (estimate = 0.51; t value = 5.91; Pr > |t|  0.0001) has a has a positive effect on user's attitude toward digital library system usage. Hypothesis 1 was therefore supported.

Table 3. Regression analysis results.

Hypotheses	Relationship	Estimate	*F* value	Pr > *F*	*t* value	Pr > |t|	*R* square
1	M-CS -> UATDLS	0.51	21.91	<0.0001	5.91	<0.0001	0.239
2	UATDLS -> IUDLS	0.50	20.12	<0.0001	6.12	<0.0001	0.291

M−CS − Meta−Cognitive Strategy; UATDLS − User Attitude Toward Digital Library System; IUDLS − Intention to Use Digital Library System
Source: Authors' own work.

Concerning hypothesis 2 (see Table 3), the resulting F-value (F = 20.12; Pr > F  0.000) showed that the model was highly significant. In this, the explanatory variable explains 29.1% (R-square = 0.291) of the variation in the role of user's attitude toward digital library system usage on intention to use digital library systems. The parameter estimates output showed that user's attitude toward digital library system usage (estimate = 0.50; t value = 6.12; Pr > |t|  0.0001) has a has a positive effect on intention to use digital library systems. Hypothesis 2 was therefore supported.

Concerning correlation analysis, Table 4 presents the results of the analysis conducted.

Table 4. Correlation analysis results.

Hypothesis	Spearman's correlation	p	Supported
1	0.43678	0.0001*	Yes
2	0.33849	0.0001*	Yes

Items marked with (*) are significantly correlated at 5% sig level.

In summary based on Table 4, meta-cognitive strategy has a positive effect on users' attitudes toward digital library system usage. Furthermore, user attitudes towards digital library systems positively affect the intention to use digital library systems.

8 Discussion

The study was aimed at exploring the role of meta-cognitive strategies and user attitudes on the intention to use a digital library system. The findings of the study show that meta-cognitive strategy to have a positive effect on users' attitudes toward digital library system usage. Furthermore, user attitudes towards digital library systems positively affects the intention to use digital library systems. The study places the importance of technology as an important driver in human capital development [1]. Digital library systems appear a useful tool allowing for such human capital development. This becomes important especially within the higher education context [3]. The findings of the study potentially can assist academics in developing meta-cognitive strategies that can be used in the utilization of digital library systems. In essence, the findings of the study potentially answer the call for academics to be responsive using technology in addressing challenges faced by students [5]. The focus of this research contributes especially given the popularity of digital library systems [6]. Given the use of digital library systems, a need is noted around improving the functionality of digital library systems [7]. Within the context of HDIs given the noted challenges [14], paying attention of the development of metacognitive strategies can assist in improving digital library systems usage attitudes. The development of meta-cognitive strategies can potentially assist to address the challenges such as resistance to system adoption [17]. Importantly, the starting point for the student is an awareness of one-self and knowledge processes around what is being learnt [19]. Potentially addressing digital library user attitudes has been found to be an important precursor towards systems adoption [27–29]. As found in this study, meta-cognitive strategies such as assisting students to have an awareness of their thought processes [33] and the taking of coherent individual actions to learning [34] potentially can assist in the formation of user attitudes in the adoption of technology.

8.1 Practical Implications

A unified stakeholder engagement approach can assist in the development of student meta-cognitive strategies. Emphasis can be placed on aspects such as goal setting, planning, self-monitoring and knowledge acquisition. These have been identified as core to meta-cognitive capabilities needed by students. Interventions here could include supporting services offered to students through training by lecturers and university stakeholders. An added implication could be to incentivize students to take courses in which they can be rewarded for taking online courses that aid the development metacognitive tasks. Given the focus of the study was on digital library systems, the university library stakeholders can also make linkages between their systems to be in line with student learning outcomes. The appetite to use digital library systems can be enhanced through such a link between what the student is learning, and tool being used to enhance that learning experience.

8.2 Limitations and Future Research

Some limitations can be drawn from the study. First, the study was only conducted at one historically disadvantaged institution in South Africa. The findings are not generalizable

to the entire list of historically disadvantaged institutions in South Africa. Second, the sample was mostly skewed towards females and the Black African ethnic group. Future research could strive for a more balanced sample in terms of the demographic group to the study. We therefore plead for caution when interpreting these results given these sampling challenges.

Future research could prioritize testing the model against other student related factors and ascertain how they influence user attitudes towards digital library use including their adoption. The study magnified the role of metacognitive strategies as a starting point. Further, a qualitative inquiry can be used. This could entail interviews with students who use digital library systems to subjectively understand their experiences of interacting with such systems. From such a qualitative inquiry, interventions can be proposed at two levels. First, the need to enhance the human-computer interaction especially concerning digital library use. Second, to proffer solutions that enhance the student learning experience while using digital library systems.

9 Conclusion

The study places importance on the role of metacognitive strategies in affecting both user attitudes and intention to adopt digital library systems. Specially, the argument is for the development of an environment that encourages students to develop such strategies as this has the propensity to influence aspects related to system use and adoption. Students in institutions of higher learning may need support mechanisms to be put in place to support the development of metacognitive strategies. The focal point is not just empowering institutions with technologies such as digital library systems, focus should also be assisting the end-users of such technology. This study gave focus to the important role of metacognitive strategies as an end-user factor to digital library system attitudes and usage.

Acknowledgements. The researchers acknowledge support of The Govan Mbeki Research and Development Centre (GMRDC) at the University of Fort Hare under the Postdoctoral Support Service. The researchers also acknowledge The Council for Scientific and Industrial Research (CSIR) for the funding on e-Research and cyberinfrastructure in Historically Disadvantaged Institutions (HDIs) under the National Integrated Cyberinfrastructure Initiative (NICIS).

References

1. Mense, E., Lemoine, P., Garretson, C., Richardson, M.: The Development of Global Higher Education in a World of Transformation. J. Educ. Dev. **2**, 47–60 (2018)
2. Cilliers, L., Chinyamurindi, W.T., Van Rhyne, Z.: Social network addiction and advertising on social networks: a case study of rural students in South Africa. S. Afr. J. Inf. Manag. **21**(1), 1–7 (2019)
3. Sai, N.R., Kumar, G.S.C., Prasad, K.D.: Impact of technology on higher education in India. J. Softw. Proj. Manag. Qual. Assur. **1**(1), 5–9 (2011)
4. Tikly, L.: Education for sustainable development in Africa: a critique of regional agendas. Asia Pac. Educ. Rev. **20**(2), 223–237 (2019). https://doi.org/10.1007/s12564-019-09600-5

5. Lillejord, S., Børte, K., Nesje, K., Ruud, E.: Learning and Teaching with Technology in Higher Education – A Systematic Review. Knowledge Centre for Education, Oslo (2018)
6. Jharotia, A.: Innovative technologies and services in libraries: an overview. Innovative Research in Library and Information Science 45–63 (2018)
7. Hartog, P.: Xie and Matusiaks' "Discover digital libraries: theory and practice" (book review). The Christian Librarian: **61**(2), 20 (2018)
8. Delaney, G., Bates, J.: Envisioning the academic library: a reflection on roles, relevancy and relationships. New Rev. Acad. Librariansh. **21**, 30–51 (2015)
9. Mohuba, D.K., Govender, K.: Cogent business & management. The merger of historically disadvantaged tertiary institutions in South Africa: a case study of the University of Limpopo. **3**(1), 1–12 (2016)
10. Nyoni, P.N.: Pedagogy and Agency in Postgraduate Student Supervision in a Rural South African University. Postgraduate Research Engagement in Low Resource Settings, IGI Global (2020)
11. Harry, T., Chinyamurindi, W.T.: Illustrating school-to-work transition through drawings: an exploratory study with a sample of South African students. SA J. Hum. Resour. Manag. **18**, 1–10 (2020)
12. van der Westhuizen, D., Barlow-Jones, G.: The role of digital literacy in the academic performance of first-year students in an information technology course. In: Bastiaens, T., Ebner, M. (eds.) Proceedings of ED-MEDIA 2011-World Conference on Educational Multimedia, Hypermedia & Telecommunications, pp. 873–878. Association for the Advancement of Computing in Education (AACE), Lisbon, Portugal (2011)
13. Basak, S.K., Wotto, M., Bélanger, P.: Factors affecting to E-learning in continuing education in Africa: a review of literature. Int. J. Eng. Sci. Manag. Res. **4**(1), 86–97 (2017)
14. Chaminade, C., Padilla-pérez, R: The challenge of alignment and barriers for the design and implementation of science, technology and innovation policies for innovation systems in developing countries. In: Kuhlmann, S., Ordóñez-Matamoros, G. (eds.) Research Handbook on Innovation Governance for Emerging Economies: Towards Better Models, pp.181–204. Social and Political Science (2017)
15. Stosic, L.: The importance of educational technology in teaching. International Journal of Cognitive Research in Science Engineering and Education (IJCRSEE) **3**(1), 111–114 (2015)
16. Kariippanon, K., Cliff, D.P., Lancaster, S., Parrish, A.: Perceived interplay between flexible learning spaces and teaching, learning and student wellbeing. Learning Environ. Res. **21**(3), 301–320 (2018)
17. Mahmud, I., Ramayah, T., Kurnia, S.: To use or not to use: modelling end user grumbling as user resistance in pre-implementation stage of enterprise resource planning system. Inf. Syst. **69**, 164–179 (2017)
18. Flavell, J.H.: First discussant's comments: what is memory development the development of? Hum. Dev. **14**(4), 272–278 (1971)
19. Flavell, J.H.: Metacognition and cognitive monitoring: a new area of cognitive developmental inquiry. Am. Psychol. **34**, 906–911 (1979)
20. Ghavifekr, S., Rosdy, W.A.W.: Teaching and learning with technology: effectiveness of ICT integration in schools. Int. J. Res. Educ. Sci. (IJRES) **1**(2), 175–191 (2015)
21. Dutta, S., Smita, M.K.: The impact of COVID-19 pandemic on tertiary education in Bangladesh: students' perspectives. Open J. Soc. Sci. **8**, 53–68 (2020)
22. Harry, T., Chinyamurindi, W.T: Still haven't found what I am looking for. Rural black students' perceived work readiness and assessment of labour market access. Education + Training **64**(2), 276–289 (2021)
23. Al-Tammemi, A.B., Akour, A., Alfalah, L.: The impact of COVID-19 pandemic on tertiary education in bangladesh: students' perspectives. Open J. Soc. Sci. **8**(9) (2020)

24. Wiegert, T.: Barriers to effective technology adoption and integration. All Electronic Theses and Dissertations. 658 (2018)
25. Davis, F.D: A technology acceptance model for empirically testing new end-user information systems: theory and results. Sloan School of Management, Massachusetts Institute of Technology (1986)
26. Venkatesh, V., Davis, F.D.: A theoretical extension of the technology acceptance model: Four longitudinal field studies. Manage. Sci. **46**, 186–204 (2000)
27. Thong, J.Y., Hong, W., Tam, K.Y.: What leads to user acceptance of digital libraries? Commun. ACM **47**, 78–83 (2004)
28. Vaidyanathan, G., Sabbaghi, A., Bargellini, M.: User acceptance of digital library: an empirical exploration of individual and system components. Issues in Information Systems **6**(2), 279–285 (2005)
29. Park, N., Roman, R., Lee, S., Chung, J.E.: User acceptance of a digital library system in developing countries: an application of the technology acceptance model. Int. J. Inf. Manage. **29**(2), 196–209 (2019)
30. Ajzen, I.: From intentions to actions: a theory of planned behavior. In: Kuhl, J., Beckmann, J. (eds.) Action control. SSSSP, pp. 11–39. Springer, Heidelberg (1985). https://doi.org/10.1007/978-3-642-69746-3_2
31. Ajzen, I.: The theory of planned behavior. Organ. Behav. Hum. Decis. Process. **50**(2), 179–211 (2004)
32. Segedy, J.R.: Adoptive scaffolds in open-ended computer-based learning environments. Diss. Abstr. Int. **76**(1), 310 (2014)
33. Zhang, Y., et al.: The evolution of metacognitive strategy use in an open-ended learning environment: do prior domain knowledge and motivation play a role? Contemp. Educ. Psychol. **69**(2), 10206 (2022)
34. Segedy, J.R., Kinnebrew, J.S., Biswas, G.: Using coherence analysis to characterize self-regulated learning behaviours in open-ended learning environments. J. Learn. Anal. **2**(1), 13–48 (2015)
35. Zhang, Y., et al.: The relationship between confusion and meta-cognitive strategies in Betty's brain. In: Rensing, C., Drachsler, H. (eds.) Proceedings of the Tenth International Conference on Learning Analytics and Knowledge, pp. 276–284. Association for Computing Machinery (2020)
36. Kim, D., Lee, D.: Impacts of meta-cognition on innovative behaviors: focus on the mediating effects of entrepreneurship. J. Open Innov. Technol. Mark. Complex. **4**(2), 1–9 (2018)
37. Kurtz, B.E., Borkowski, J.G.: Development of strategic skills in impulsive and reflective children: a longitudinal study of metacognition. J. Exp. Child Psychol. **43**(1), 129–148 (1987)
38. Moorthy, K., et al.: Behavioural intention to adopt digital library by the undergraduates. Int. Inf. Libr. Rev. **51**(2), 128–144 (2019)
39. Shava, H., Chinyamurindi, W.T.: Determinants of social media usage among a sample of rural South African youth. SA J. Infor. Manag. **20**(1), 1–8 (2018)
40. Alokluk, J.: Attitude of students towards the use of library facilities: a case study. Int. J. Humani. Soc. Sci. Educ. (IJHSSE) **7**(1), 24–36 (2020)
41. Creswell, J.W.: Research Design: Qualitative, Quantitative and Mixed Methods Approaches, 4th edn. Sage, Thousand Oaks, CA (2014)
42. Chinyamurindi, W.T., Louw, G.C.J.: Gender differences in technology acceptance in selected South African companies: implications for electronic learning. SA J. Hum. Resour. Manag. **8**(1), 1–7 (2010)
43. Churchill, G.A., Brown, J.T.: Basic Marketing Research. Thompson, South West City, MO (2010)

Mobile Cloud Based Enterprise Resource Planning Systems for Small Medium and Micro Enterprises: A Systematic Literature Review

Nametshego Gumbi, Lucas Gumbi(✉), and Hossana Twinomurinzi

University of Johannesburg, Auckland Park 17, Johannesburg, South Africa
Nametshego.m@gmail.com, hossanat@uj.ac.za

Abstract. Small Medium and Micro Enterprises (SMMEs) are significant contributors to economic growth and employment and therefore playing a vital role in wealth creation and human welfare. Yet despite their critical importance, SMMEs are yet to realize their full potential due to numerous challenges presented by the nature of operations in a global market, including the emergent hyper digital contexts as presented by COVID-19. Mobile Cloud Based Enterprise Resource Planning (MCB-ERP) systems present opportunities such as on-demand use, low upfront capital expenditure, security and stability, easy upgrade, and low operational and maintenance costs for SMMEs operating in these global markets. This paper reports on a systematic review of 35 peer reviewed journals and conference papers to identify MCB-ERP research that has been undertaken and to identify opportunities, gaps, and challenges in relation to SMMEs. The key findings suggest that researchers have mainly focused on Cloud-Based ERP adoption and implementation factors; the use, benefits, and impact for SMMEs received limited attention; and that SMME readiness for adoption is impacted by a lack of MCB-ERP vendor presence, awareness, and experience. Moreover, most of the research is not contextualized and treats SMMEs as homogenous entities operating in sectors with similar environments and preconditions. Future context specific research investigating MCB-ERPs in specific contexts, and non-ICT based sectors such as agriculture for SMMEs is recommended.

Keywords: Enterprise resource planning system · Cloud based enterprise resource planning system · Mobile cloud based enterprise resource planning system · Small medium and micro enterprises · Benefits · Use · Impact

1 Introduction

The importance of Small Micro and Medium Enterprises (SMMEs) to economies of many countries has been well documented in the literature, government policies, guidelines, and reports. SMMEs are an important means to improving the socio-economic conditions of any country [1–5] and play an important role as a source of employment, creation of new markets and significant contribution to Gross Domestic Product (GDP). Despite their (SMMEs) importance, their high failure rate [4–7] means that their full

P. Ndayizigamiye et al. (Eds.): IDIA 2022, CCIS 1774, pp. 46–69, 2023.
https://doi.org/10.1007/978-3-031-28472-4_4

potential to economic contribution remains untapped. The advent of the 4th Industrial Revolution (4IR) [8] and the COVID-19 pandemic has changed the global economy landscape giving more prominence to digital technologies and digital innovation [9–11]. This means that SMMEs are compelled to innovate more with digital technologies to become more efficient and competitive in the hyper digital contexts [1, 12–14]. For example, the operational costs of SMMEs are often too high due to inefficiencies in managing their business processes [15] yet digital technologies offer an efficient means to automate some of these processes. Enterprise Resource Planning (ERP) systems are a viable solution for such SMME problems as they enable the simplification of internal processes, seamless information retrieval, efficient performance management, production efficiency, business intelligence and knowledge management [16–18]. However, SMMEs experience implementation complexities and a lack of both financial and human resources to adopt traditional ERP systems [18, 19]. Mobile Cloud Based ERPs (MCB-ERPs) have been identified as alternatives for SMMEs as a result of the proliferation of mobile devices [20–24]. MCB-ERPs enable both the enterprise software application and the database system to be hosted on a service provider's infrastructure, and to be rented by SMMEs at a fraction of the cost, and have less deployment complexities as compared to a traditional on-premise solutions [12, 23, 25]. SMMEs then access the cloud server over the Internet using smart phones, notebook computers, tablet computers, and desktop computers to obtain the ERP application. This type of business model provides advantages such as on-demand use, low cost access, security and stability, easy upgrade, and maintenance cost savings for SMMEs [24]. For example, manufacturing SMMEs that had adopted MCB-ERP systems by the time COVID-19 hit were able to work from home by minimizing the number of personnel coming into plants and offices, and even predicting demand for an effective response to unforeseen changes and risk to supply chains [26–29].

The importance of cloud-based solutions for manufacturing environments has also been emphasized [30]. MCB-ERPs are also now being used during COVID-19 to remotely track job quotes, schedules, orders, automate reports, exchange data, and monitor production to adjust it when the need for meeting the rising demand arises. Yet, adoption of MCB-ERPs by SMMEs has remained low [21, 24]. This paper therefore sought to investigate how MCB-ERPs have been researched regarding SMMEs. The investigation is carried out through a Systematic Literature Review (SLR) and has covered the research done from 2015 to 2020. The SLR sought to find answers for the following Research Questions (RQ);

RQ1: What research has been undertaken in relation to Mobile Cloud Based Enterprise Resource Planning systems for SMMEs?
RQ2: How is Mobile Cloud Based Enterprise Resource Planning system being used by SMMEs?
RQ3: What are the benefits of a Mobile Cloud Based Enterprise Resource Planning system?
RQ4: What is the possible impact of Mobile Cloud Based Enterprise Resource Planning system on SMMEs?

The rest of the paper is organized as follows; Sect. 2 presents the methodology followed to carry out the SLR. Section 3 presents the results of the study. Lastly, Sect. 4 concludes the study and provides proposed future research work.

2 Methodology

To establish a structured approach and a guiding framework for the analysis of the research, this study followed [31] systematic review protocol shown in Fig. 1. The protocol is similar to that proposed by Lage Junior and Godinho Filho [32].

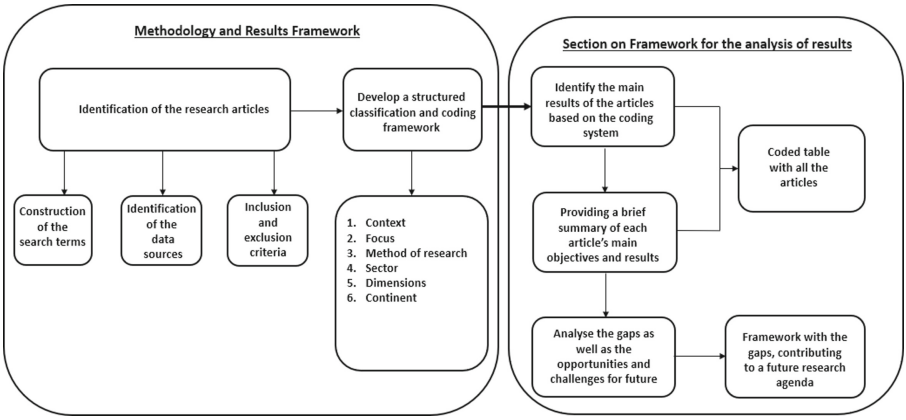


Fig. 1. A systematic review protocol for this study

2.1 Identification of the Research Articles

The first steps are (1) the construction of the search terms, (2) identification of the relevant data sources and (3) inclusion and exclusion criteria as indicated in Fig. 1.

2.1.1 Construction of the Search Terms

The search terms were derived from RQ1, RQ2, RQ3 and RQ4. The search strings were constructed using the unit of analysis (SMMEs); technology artefact (Mobile Cloud Based Enterprises Resource Planning System); and phenomenon of interest (Use, benefits, and impact). [Unit of Analysis] AND [Technology Artefact] AND [Phenomenon of Interest]. The full details on specific terms used for the unit of analysis, technology artefact, phenomenon of interest, and of the 25 constructed search strings that were used to identify suitable research articles are presented in Appendices A.1.

2.1.2 Identification of the Data Source

The study falls under the interdisciplinary field of Information Systems (IS) and therefore used the following multidisciplinary databases as the data source, Web of Science, Academic Search Premier, Science Direct and Google Scholar. These databases are consistently utilized by many peer reviewed systematic literature review studies [30, 33–40].

2.1.3 Inclusion and Exclusion Criteria

After using the same constructed search terms through all the four databases a total of 1365 articles were found from these databases (Fig. 2). The first stage of selection involved a selection of papers based on the relevancy of title, abstract and keywords. This resulted in a total of 470 articles. The second selection was based on the following inclusion criteria to include journals or conference articles that; mentioned the specific terms for the unit of analysis, technology artefact and phenomenon of Interest (Appendices A.1) within their content or title; are written in English language; were academically peer reviewed; were less than six years (2015 to 2020); and focused at a firm level and not at an individual level, focusing on all sectors or industries. In the second sorting, after removing duplicates in each database separately using title and abstract, a total of 895 articles were excluded from the found articles.

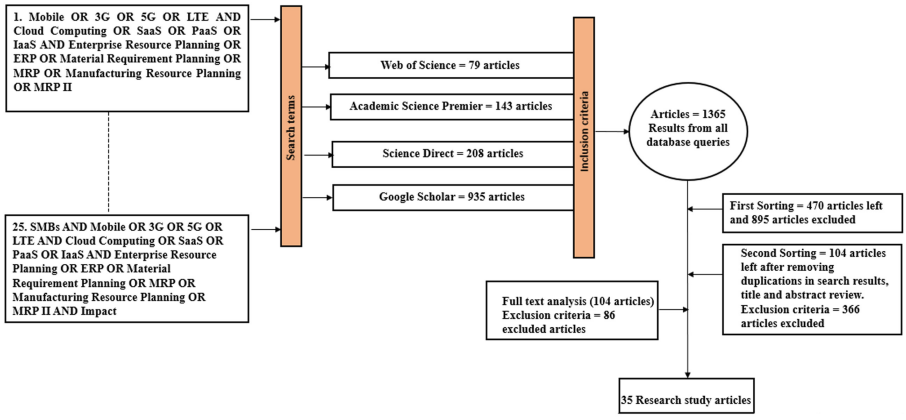


Fig. 2. A schematic representation of the inclusion and exclusion process

In the third iteration exclusion criteria was used to exclude articles that; merely mentioned some of the specific search terms but do not solely focus on the research topic Cloud-based ERP neither do they focus on MCB-ERP; were practitioner based and not based on any scientific research method; do not focus on SMMEs, and articles that merely mention SMMEs. After combining all articles from the four database sources, a further sorting was made to remove any duplicates; this resulted in a further 366 articles being excluded. In the final sorting, 86 articles were excluded based on the analysis of the full text of the articles and only 35 peer reviewed journals and conference papers

remained as shown in Fig. 2. The primary exclusion criteria were articles that do not discuss any of the three study dimensions (use, benefits, and impact) and the articles that did not mention Cloud-based ERP or Mobile Cloud Based ERP.

2.2 Coding and Classification Framework

The systematic review is analyzed using the coding and classification framework developed by Amui et al. [31] based on other works [32, 41–44]. The classification was done using the following categories: context, focus, method, sector, dimensions or practices, and the global locations in the form of origin or continents as shown in Table 1.

Table 1. Classification and coding framework

Classification	Description	Codes
Context	Low-income countries	1A
	Lower-middle income countries	1B
	Upper-middle income countries	1C
	High income countries	1D
Focus	Mobile cloud based enterprise resource planning as the main theme	2A
	Cloud based enterprise resource planning as the main theme	2B
	SaaS enterprise resource planning as the main theme	2C
Method	Qualitative	3A
	Quantitative	3B
	Theoretical	3C
	Empirical	3D
	Case studies/interviews	3E
	Survey	3F
	Design	3G
Sector	Agriculture	4A
	Manufacturing	4B
	Information and communication technology	4C
	Retail	4D
	Finance	4E
	Health	4F
	General	4G

(*continued*)

Table 1. (*continued*)

Classification	Description	Codes
Dimensions	Use	5A
	Benefits	5B
	Impact	5C
Origin (Continents)	America	6A
	Europe	6B
	Asia	6C
	Africa	6D
	Oceania	6E

3 Analysis and Results of the Literature Review

The 35 selected research articles were coded according to the coding and classification framework in Table 1 with the resultant coding results shown in Appendices A.2. A summary of the findings for the systematic review is shown in Fig. 3.

The majority of the articles (54%) in this systematic review researched the *Benefits of Cloud Based ERP* (2B + 5B) as shown in Fig. 3. There were no studies that researched the *Benefits of MCB-ERP* (2A + 5B) answering RQ3. There were also no studies that researched either the *Use of MCB-ERP* (2A + 5A) or *the Impact of MCB-ERP* (2A + 5C) answering RQ2 and RQ4, respectively. Furthermore, the current research on *MCB-ERP* (2A) is very scant (Fig. 3) and thus limiting the response to RQ1. A detailed analysis of the findings in accordance with coding classification (Context, Focus, Method, Sector, Dimensions and Origins) is provided in the next sub-sections.

3.1 Context

The context analysis was based on the country classifications used by World Bank [45]. There were (14) articles from upper-middle income countries (1C), (10) articles from high income countries (1D), (11) articles from lower-middle income (1B) and no articles from low-income countries (1A) (Appendices A.2). These findings suggest that the intensity of research in relation to Cloud-based ERP digital innovation for SMMEs in the context of lower-middle, upper-middle- and high-income countries is very high. This presents a significant opportunity for SMMEs in such contexts to leverage MCB-ERPs for sustainability, competitive advantage, and market growth. Even though there is almost an equal share of studies amongst the lower-middle, upper-middle- and high-income countries, many (17) did not mention explicitly where the studies were conducted. The findings suggest that SMMEs are studied in general and there is no contextualization when compared to localization (institutions or country of origin). Further studies that explicitly considers the context of the research are needed on MCB-ERPs for SMMEs.

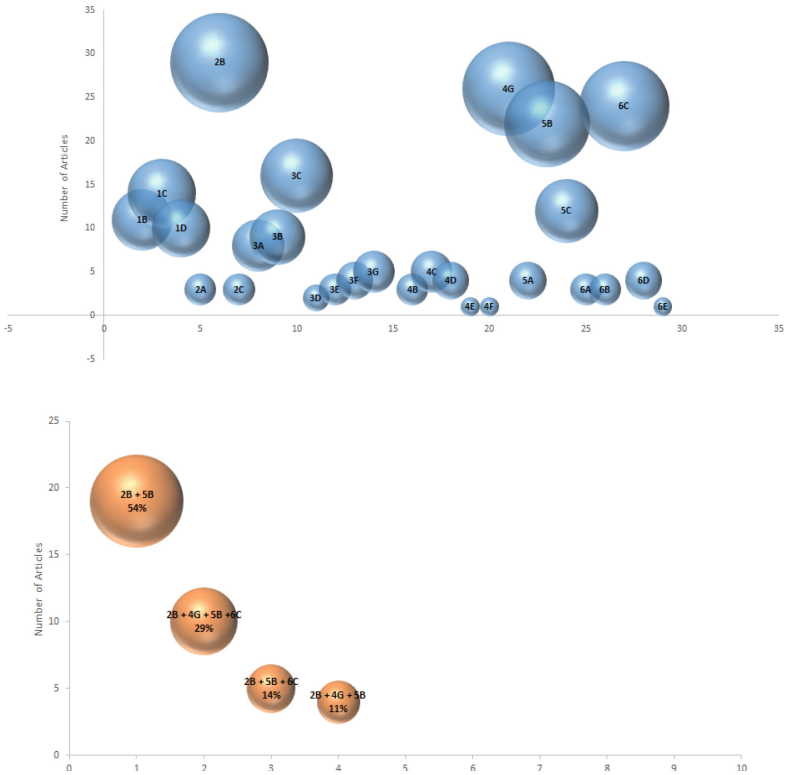


Fig. 3. Summary of the findings for the systematic review

3.2 Origin

Eleven (11) articles were from lower-middle income countries, (14) articles from upper-middle income countries while (11) articles were from high income countries (Appendices A.2). The articles (11) that came from lower-middle income countries originated from South Asia, Middle East and North America, and Sub-Saharan Africa (India, Sri Lanka, Egypt, Pakistan and Nigeria). The studies that came from upper-middle income countries were from East Asia & Pacific (China, Indonesia, Malaysia), Latin America & Carribian (Brazil, Equador), Middle East & North America (Irag, Jordan), and Sub-Saharan Africa (Namibia, South Africa). The articles that came from high income countries originated from East Asia and Pacific (Australia, New Zealand, Taiwan), Europe and Central Asia (Germany, Norway), Middle East and North America (United Arab Emirates, Saudi Arabia), and North America (USA). These findings suggest that even within the context of upper-middle income countries, there is a need for researchers to contextualize these studies in accordance with different country specific SMMEs needs and requirements [46]. There were only 2 studies done in Sub-Saharan Africa (Namibia & South Africa) (Table 2). The specific needs and requirements for SMMEs are different in each social context and are unlikely to sufficiently reflect the social norms. The studies are geographically limited in context and the findings are limited in terms of universal

validity [46, 47]. For example, [48] identified that SMMEs in Sub-Saharan Africa are facing systemic challenges that are exclusive to them and not shared by enterprises in other countries, these include limitations in accessing financial resources, markets and marketing skills, and the business environment. It is therefore important to make theoretical and empirical research which consider local contexts using geographical location, culture, and profiles of enterprises as categorized by [46].

Table 2. Origin of the studies

Study conducted in	No. of articles	Research context code
Brazil	1	1C
China	1	1C
India	2	1B
Iraq	1	1C
Indonesia	1	1C
Jordan	1	1C
Malaysia	1	1C
Norway	1	1D
Namibia	1	1C
Pakistan	1	1B
South Africa	1	1C
Saudi Arabia	3	1D
Sri Lanka	1	1B
Taiwan	2	1D
Not discussed/mentioned	17	

3.3 Focus on Mobile Cloud Based Enterprise Resource Planning System

In this study, only 3 articles used MCB-ERPs (2A) as the main theme (Appendices A.2). This may be that researchers considered mobility to be inherent in Cloud-based ERP systems or cloud computing application to be inherent in MCB-ERP systems. However, this was neither implicitly nor explicitly indicated. The authors only provided a definition to Cloud-based ERP as an approach to ERP which relies on the platforms and services of cloud computing to modify business processes, and it is accessed over the Internet without any installation or configuration at the user side [16, 49–51]. Mehta et al. [17] and Ziani and Alfaadhel [52] defined Cloud-based ERP as a combination of ERP and cloud computing, provided as SaaS. The 3 articles focused on the design of the application for MCB-ERPs system on SaaS cloud and on criteria that researchers in this area should use when selecting the ERP system for research purposes. The findings indicate that MCB-ERPs are still a nascent and emerging research area with very limited

specific and focused research. Similar to the findings of these systematic review, other authors found that there was limited research on Cloud-based ERPs and recommended further studies which focus on SMMEs from different contexts, industries and/or sectors [12, 23, 53–57]. Two studies [58, 59] show that the use of ERP with mobile or wireless devices is a limited research area and in its infancy stage, and make recommendations for future research to include enterprises from different sectors. This is a research gap that presents significant opportunities that needs to be further explored given MCB-ERPs' potential benefits to SMMEs. Further, the findings reveal a confusion between Mobile Cloud and Cloud-based ERP definitions or descriptions, with the mobile element taken as a given in relation to Cloud-based ERP. Further research that explicitly defines or describes mobile when researching Cloud-based ERP is recommended especially because of the proliferation of mobile devices such as cell phones, smartphones, laptop, Personal Digital Assistants and tablet, that have made mobile applications a part of daily life [58, 59]. The physical and functional features of mobile devices allow for interaction, information processing and collaboration while reducing some costs for the enterprise. On the other hand, Mobile ERPs have a challenge of the user interface which hampers the working model.

3.4 Research Methods

The majority of the articles (16) utilized theoretical (3C) research methods (Table 4). Very few articles (2) considered empirical (3D) and or case studies (3 articles) (3E) as research methods (Table 3). The design research approach had 5 articles and survey, 4. Qualitative research had 8 articles and quantitative 9. Thirteen of the articles (13) using theoretical research methods to understand the developing behavior of Cloud-based ERP did not mention the context of the studies. The studies have contextual inadequacies and therefore pose a limitation on the scope of validity of their research findings. The 11 articles utilizing theoretical research focused on factors of adoption and implementation, 1 article on the impact of adoption, 2 on integration of on-premise ERP with cloud, 1 article on design, and 1 article on challenges with the implementation of Cloud-based ERP. Fifteen articles (15) employed the theoretical research method in a generalized manner, without any specific focus on any economic sector. Only 1 article focused on retail. The findings also reveal that few articles (3) used case studies (2 retail and 1 ICT) which limits the generalizability of this cases' findings to other industries, sectors, and contexts. Case studies are suitable for learning more about a little known or poorly understood situation and are likely to provide valuable knowledge, using few cases poses limitations on generalizability of findings [25].

Further study is needed across different industries and sectors to gain more knowledge on MCB-ERPs as it applies to SMMEs to generalize the findings. The 5 articles that utilized design research to develop new artefacts about the MCB-ERPs, 3 originated from Asia (China, Taiwan & Indonesia), one from Saudi Arabia and 1 from Europe (Germany). Other 2 articles were not explicit on the context. The 4 articles that used surveys, focused on Cloud-based ERP adoption and implementation factors. It is observed that even though surveys might indicate maturity in a field, the focus was on the ICT sector or SMMEs that already adopted these digital technologies. The SMMEs that do not have the know-how are left out and contextualization is lacking, and it limits universal

validity of the results to such SMMEs. The results indicate that maturity is yet to be reached and that research in this area is still in its infancy, and needs to be furthered by covering different sectors, contexts, and mobile element. In the survey researches, 3 articles came from Asia (India, Taiwan & Pakistan) and 1 from Europe (Norway). In African countries there was no survey research that had been done on MCB-ERPs. Out of 8 articles that utilized qualitative research to develop a deeper understanding of Cloud-based ERP for SMMEs, 6 researched factors of adoption and implementation, 1 studied impact of these artefact, and 1 studied the impact of mobile ERP. The qualitative research articles (2) focusing on the ICT, 2 on retail, and 1 article on manufacturing, finance, and health sectors. In terms of context, 3 articles did not mention were studies were conducted. Only few studies (2) were conducted on SMMEs in Africa (Namibia & South Africa). America had 2 articles with one study conducted on SMMEs in Brazil, and the other not mentioning were studies were conducted. The 3 articles from Asia with studies conducted in Sri Lanka, Malaysia, and Saudi. The qualitative research done is indicating that the research and theories on Cloud-based ERP are at their early, formative stages.

The articles (9) using quantitative research, 1 studied factors of Cloud-based ERP adoption, 4 studied integration of ERP with cloud environment, and 3 articles on implementation, use and challenges related to Cloud-based ERP. Most of the articles (7) originated from Asia with 6 studies conducted on SMMEs from India, Indonesia, Jordan, Saudi Arabia, Sri Lanka and Taiwan, and 1 study did not mention the context. One article originated from Europe with the study conducted on SMMEs from Norway. The last article (1) originated from Oceania without mentioning the context of the study. These resulted in 2 articles out of 9 not mentioning the context of the study. The findings indicate the need to diversify and use more research methods that can contextualize MCB-ERPs research to SMMEs specific context. Specifically, qualitative research methods as they bring out context specific challenges, constraints, issues, and pre-conditions.

Table 3. Research methods analysis

Research methods	Qualitative 3A	Quantitative 3B	Theoretical 3C	Empirical 3D	Case studies 3E	Survey 3F	Design 3G	No. of articles
Single research methods	x	x	16	x	x	x	x	16
Mixed research methods	x	x	x	x	x	x	5	5
Qualitative	8	x	x	x	3	x	x	11
Quantitative	x	9	x	2	x	4	x	15

3.5 Economic Sector

Majority of the articles (26) studied Cloud-based ERP in relation to SMMEs in a generalized (4G) manner (Table 4). Few articles focused on manufacturing (3), Information and Communication Technology (ICT) (5), retail (4), finance, and health (1 each) sectors.

Table 4. Economic sector analysis

Economic sector	Code	No. of articles
Manufacturing	4B	3
ICT	4C	5
Retail	4D	4
Finance	4E	1
Health	4F	1
General	4G	26

There were no studies on Cloud Based ERPs or MCB-ERPs focusing on the agriculture sector. The benefits of MCB-ERPs system for each SMMEs sector are different, and as such the needs and requirements for each sector in relation to this technology artefact will be different and sector specific. One such sector which was not researched by any of the reviewed studies, is agriculture. The agriculture sector is undergoing a transformation towards precision agriculture to automate the sector. Pandemics such as COVID-19, 4IR and the future deployment of the Fifth Generation (5G) wireless technology network requires the agriculture sector to use digital technologies such as MCB-ERPs to remain in operation. [51] recommended that among other characteristics of SMMEs, future studies should include industry and/or sectors in relation to Cloud-based ERP. [60] and [25] also recommended studies across different sectors and/or industries. Although SMMEs share a number of similarities, they are heterogeneous in nature.

3.6 Current Research Done on MCB-ERPs

The articles were analyzed to determine the current research status on MCB-ERPs digital innovation (Appendices A.2) to further answer RQ1. The majority of the articles (32) (2B + 2C) focused on Cloud-based ERP or SaaS ERP and not MCB-ERPs. Only 3 articles (2A) focused explicitly on Mobile ERP and explicitly on MCB-ERP system. The 32 articles on Cloud-based ERP focused mainly on factors of adoption (17) such as perceived cost reduction, data security risks, privacy, relative advantage, system performance and availability, top management support, compatibility, scalability, subsidy, end-users and/or suppliers, sunk cost and satisfaction with existing system, improved IT reliability, loss of control and vendor trust, functionality fit, limitations on customizations, attitude towards change, internet connection, and lack of knowledge. Integration of on-premise ERP with cloud (6); factors of implementation (4); design (3); impact of adoption and implementation (2); challenges of implementation (2); and factors of use

(1) such as compatibility, output quality, perceived usefulness, perceived ease of use, expectation confirmation, and satisfaction. This indicates that although current research on MCB-ERPs is limited, it is evident that most research is focused on the architectural design and methodologies to evaluate factors of adoption and implementation.

3.7 Dimensions

To further answer the research questions RQ2, RQ3 and RQ4 posed in the Introduction, the three dimensions (1) use (utilization), (2) benefits (relative advantage), (3) impact (positive consequence) shown in Table 1, were analyzed in accordance with the codification results in Appendices A.2.

3.7.1 Mobile Cloud Based Enterprise Resource Planning System: Use

The "Use" dimension sought to establish the use of MCB-ERPs system by SMMEs. While the majority of the studies (34) did not mention or discuss the mobility element in Cloud-Based ERP, they did not look explicitly into the use of Cloud-based ERP (Table 5). One (1) "use" article focused on sales force automation [58]. The findings indicate that there is little research investigating the use of MCB-ERPs systems by SMMEs. Research focusing on the "use" of MCB-ERPs would provide opportunities to assess the applicability and relevance of digital innovation (MCB-ERPs) for SMMEs.

Table 5. Use

Dimension: use	Code	No. of articles
Use of MCB-ERPs by SMMEs explicitly identified	5A	1
Not discussed/mentioned	5A	34

3.7.2 Mobile Cloud Based Enterprise Resource Planning System: Benefits

The "Benefit" dimension seeks to identify the benefits of MCB-ERPs system. Majority of the research studies (32) focused on Cloud-based ERP and not MCB-ERPs (Table 6). Only one of the 3 articles identified benefits of MCB-ERPs. The identified benefit of MCB-ERPs was on-demand use, low cost, security and stability, easy upgrade, and maintenance, and helps SMMEs achieve digital management [61]. The 16 articles that focused on Cloud-based ERP revealed lower costs, less human resource, scalability, flexibility, accessibility, easy to use, upgradable, streamlined and accelerated process, core business focus, mobility, easier integration, fast implementation, reliability, pay per use, fast-track time to market, organize and maintain data remotely, and elasticity as potential benefits.

Table 6. Benefits

Dimension: benefits	Code	No. of articles
Potential benefits of MCB-ERPs explicitly identified	5B	1
Potential benefit of Cloud-based ERP explicitly identified	5B	16
Not discussed/mentioned	5B	18

The other 18 articles did not investigate the benefits of Cloud-based ERP or MCB-ERPs. The findings point out to the need for research studies to focus on explicitly identifying benefits of MCB-ERPs system. Future research should investigate whether combining the identified benefits of Cloud-based ERP system with mobile benefits could translate into potential relative advantages to SMMEs.

3.7.3 Mobile Cloud Based Enterprise Resource Planning System: Impact

The "Impact" dimension sought to identify the impact (positive consequence) of the use of MCB-ERPs system on SMMEs (Table 7). Only 2 articles identified the positive consequence of using MCB-ERPs system as increased productivity, efficiency and effectiveness, quality of user's life improved, reduced workload of business personnel, and improved the efficiency of real time dynamic management of enterprises [58, 61]. Majority of the research studies (24) did not look into the positive consequence to identify explicitly or implicitly any impact. Only 9 articles of the ones that focused on Cloud-based ERP identified explicitly its impact. The findings point to the need for research studies to focus on the issue of understanding both positive and negative consequences of using MCB-ERPs systems by SMMEs. This will assist in identifying threats such as security bridges, data privacy compliance and confidentiality requirements including other related issues that may be a threat to SMME businesses. Consequently, assist SMMEs when making decisions on adopting, implementing, and using such systems.

Table 7. Impact (positive consequence)

Dimension: impact (positive consequence)	Code	No. of articles
Impact of MCB-ERPs on SMMEs explicitly identified	5C	2
Impact of cloud-based ERP on SMMEs explicitly identified	5C	9
Not discussed/mentioned	5C	24

4 Conclusion, Limitations and Future Work

SMMEs are important to national social and economic development. The current global markets and the digital context (4IR) requires SMMEs to increasingly adopt digital

innovations. MCB-ERPs are recommended as a solution for SMMEs in order to remain relevant and to succeed. MCB-ERP offer a solution that eliminates the challenges of traditional ERPs for SMMEs. However, the findings suggest that MCB-ERPs research on SMMEs is limited and is not contextualized to sector or industry specific challenges and preconditions. Accordingly, future context specific research investigating MCB-ERPs in non-ICT or high tech SMME sectors such as Agriculture is recommended. Most of the research articles focused on Cloud-based ERP adoption and implementation.

The limited studies did not demonstrate the evaluation and adoption of Cloud-Based ERP when combined with mobile devices such as smart phone, giving the impression that mobility is considered inherent to Cloud-based ERPs. This creates difficulties in determining whether the artefact under study is Cloud-Based ERP or MCB-ERP. Future research that explicitly defines or describes the mobile aspect when researching Cloud-based ERP is recommended as this has significant implications on the findings of the studied artefact (MCB-ERP or Cloud-Based ERP). Even though there is a relationship between awareness of MCB-ERPs digital innovations, experience in using such technologies, and vendor availability to the readiness of SMMEs' adoption and use of MCB-ERPs [12, 62], the issues of MCB-ERPs utilization, relative advantage, positive impact and how they relate to SMMEs received very limited research attention.

More research to evaluate the use, benefits, and impact of MCB-ERPs for SMMEs should be considered for future studies. The use of MCB-ERPs had negative consequences such as increased surveillance and control over employees. This indicates a need for more awareness or education on MCB-ERPs to eliminate the fear of employee surveillance and control by the SMMEs owner or employer. The paper only utilized four multidisciplinary databases as the data sources which might have affected the number of papers found that were relevant for the systematic review study. More multidisciplinary databases should be considered for future systematic review studies on MCB-ERPs for SMMEs.

A. Appendices

A.1. Specific Terms and Search Strings

The specific terms for the [**Unit of analysis**] were:

- Small Micro and Medium Enterprises OR
- SMMEs OR
- Small Business OR
- Small and Medium sized Enterprises OR
- SMEs OR
- Small and Medium Business OR
- SMBs

The specific terms for the [**Technology Artefact**] are:

- Mobile Cloud Based Enterprise Resource Planning System OR
- Mobile OR 3G OR 5G OR Long-Term Evolution (LTE)

- Cloud OR SaaS OR PaaS OR IaaS
- Enterprise Resource Planning OR ERP OR Material Requirement Planning OR MRP OR Manufacturing Resource Planning OR MRP II

The specific terms for the [**Phenomenon of Interest**] are:

- Use OR
- Benefits OR
- Impact

The [Technology Artefact]

1. Mobile OR 3G OR 5G OR LTE AND Cloud Computing OR SaaS OR PaaS OR IaaS AND Enterprise Resource Planning OR ERP OR Material Requirement Planning OR MRP OR Manufacturing Resource Planning OR MRP II

The [Technology Artefact] AND [**Phenomenon of Interest**]

2. Mobile OR 3G OR 5G OR LTE AND Cloud Computing OR SaaS OR PaaS OR IaaS AND Enterprise Resource Planning OR ERP OR Material Requirement Planning OR MRP OR Manufacturing Resource Planning OR MRP II AND Use
3. Mobile OR 3G OR 5G OR LTE AND Cloud Computing OR SaaS OR PaaS OR IaaS AND Enterprise Resource Planning OR ERP OR Material Requirement Planning OR MRP OR Manufacturing Resource Planning OR MRP II AND Benefits
4. Mobile OR 3G OR 5G OR LTE AND Cloud Computing OR SaaS OR PaaS OR IaaS AND Enterprise Resource Planning OR ERP OR Material Requirement Planning OR MRP OR Manufacturing Resource Planning OR MRP II AND Impact

The [**Unit of Analysis**] AND [**Technology Artefact**] AND [**Phenomenon of Interest**]

5. Small Micro and Medium Enterprises AND Mobile OR 3G OR 5G OR LTE AND Cloud Computing OR SaaS OR PaaS OR IaaS AND Enterprise Resource Planning OR ERP OR Material Requirement Planning OR MRP OR Manufacturing Resource Planning OR MRP II AND Use
6. Small Micro and Medium Enterprises AND Mobile OR 3G OR 5G OR LTE AND Cloud Computing OR SaaS OR PaaS OR IaaS AND Enterprise Resource Planning OR ERP OR Material Requirement Planning OR MRP OR Manufacturing Resource Planning OR MRP II AND Benefits
7. Small Micro and Medium Enterprises AND Mobile OR 3G OR 5G OR LTE AND Cloud Computing OR SaaS OR PaaS OR IaaS AND Enterprise Resource Planning OR ERP OR Material Requirement Planning OR MRP OR Manufacturing Resource Planning OR MRP II AND Impact

8. SMMEs AND Mobile OR 3G OR 5G OR LTE AND Cloud Computing OR SaaS OR PaaS OR IaaS AND Enterprise Resource Planning OR ERP OR Material Requirement Planning OR MRP OR Manufacturing Resource Planning OR MRP II AND Use
9. SMMEs AND Mobile OR 3G OR 5G OR LTE AND Cloud Computing OR SaaS OR PaaS OR IaaS AND Enterprise Resource Planning OR ERP OR Material Requirement Planning OR MRP OR Manufacturing Resource Planning OR MRP II AND Benefits
10. SMMEs AND Mobile OR 3G OR 5G OR LTE AND Cloud Computing OR SaaS OR PaaS OR IaaS AND Enterprise Resource Planning OR ERP OR Material Requirement Planning OR MRP OR Manufacturing Resource Planning OR MRP II AND Impact
11. Small Business AND Mobile OR 3G OR 5G OR LTE AND Cloud Computing OR SaaS OR PaaS OR IaaS AND Enterprise Resource Planning OR ERP OR Material Requirement Planning OR MRP OR Manufacturing Resource Planning OR MRP II AND Use
12. Small Business AND Mobile OR 3G OR 5G OR LTE AND Cloud Computing OR SaaS OR PaaS OR IaaS AND Enterprise Resource Planning OR ERP OR Material Requirement Planning OR MRP OR Manufacturing Resource Planning OR MRP II AND Benefits
13. Small Business AND Mobile OR 3G OR 5G OR LTE AND Cloud Computing OR SaaS OR PaaS OR IaaS AND Enterprise Resource Planning OR ERP OR Material Requirement Planning OR MRP OR Manufacturing Resource Planning OR MRP II AND Impact
14. Small and Medium Enterprises AND Mobile OR 3G OR 5G OR LTE AND Cloud Computing OR SaaS OR PaaS OR IaaS AND Enterprise Resource Planning OR ERP OR Material Requirement Planning OR MRP OR Manufacturing Resource Planning OR MRP II AND Use
15. Small and Medium Enterprises AND Mobile OR 3G OR 5G OR LTE AND Cloud Computing OR SaaS OR PaaS OR IaaS AND Enterprise Resource Planning OR ERP OR Material Requirement Planning OR MRP OR Manufacturing Resource Planning OR MRP II AND Benefits
16. Small and Medium Enterprises AND Mobile OR 3G OR 5G OR LTE AND Cloud Computing OR SaaS OR PaaS OR IaaS AND Enterprise Resource Planning OR ERP OR Material Requirement Planning OR MRP OR Manufacturing Resource Planning OR MRP II AND Impact
17. SMEs AND Mobile OR 3G OR 5G OR LTE AND Cloud Computing OR SaaS OR PaaS OR IaaS AND Enterprise Resource Planning OR ERP OR Material Requirement Planning OR MRP OR Manufacturing Resource Planning OR MRP II AND Use
18. SMEs AND Mobile OR 3G OR 5G OR LTE AND Cloud Computing OR SaaS OR PaaS OR IaaS AND Enterprise Resource Planning OR ERP OR Material Requirement Planning OR MRP OR Manufacturing Resource Planning OR MRP II AND Benefits

19. SMEs AND Mobile OR 3G OR 5G OR LTE AND Cloud Computing OR SaaS OR PaaS OR IaaS AND Enterprise Resource Planning OR ERP OR Material Requirement Planning OR MRP OR Manufacturing Resource Planning OR MRP II AND Impact
20. Small and Medium Businesses AND Mobile OR 3G OR 5G OR LTE AND Cloud Computing OR SaaS OR PaaS OR IaaS AND Enterprise Resource Planning OR ERP OR Material Requirement Planning OR MRP OR Manufacturing Resource Planning OR MRP II AND Use
21. Small and Medium Businesses AND Mobile OR 3G OR 5G OR LTE AND Cloud Computing OR SaaS OR PaaS OR IaaS AND Enterprise Resource Planning OR ERP OR Material Requirement Planning OR MRP OR Manufacturing Resource Planning OR MRP II AND Benefits
22. Small and Medium Businesses AND Mobile OR 3G OR 5G OR LTE AND Cloud Computing OR SaaS OR PaaS OR IaaS AND Enterprise Resource Planning OR ERP OR Material Requirement Planning OR MRP OR Manufacturing Resource Planning OR MRP II AND Impact
23. SMBs AND Mobile OR 3G OR 5G OR LTE AND Cloud Computing OR SaaS OR PaaS OR IaaS AND Enterprise Resource Planning OR ERP OR Material Requirement Planning OR MRP OR Manufacturing Resource Planning OR MRP II AND Use
24. SMBs AND Mobile OR 3G OR 5G OR LTE AND Cloud Computing OR SaaS OR PaaS OR IaaS AND Enterprise Resource Planning OR ERP OR Material Requirement Planning OR MRP OR Manufacturing Resource Planning OR MRP II AND Benefits
25. SMBs AND Mobile OR 3G OR 5G OR LTE AND Cloud Computing OR SaaS OR PaaS OR IaaS AND Enterprise Resource Planning OR ERP OR Material Requirement Planning OR MRP OR Manufacturing Resource Planning OR MRP II AND Impact

Authors		Context			Focus			Method							Sector						Dimensions			Origin/continents				
1	[19]	1B				2B			3B		3D		3F							4G						6C		
2	[63]		1C			2B				3C										4G						6C		
3	[16]	1B				2B				3C										4G		5B				6C		
4	[64]	1B				2B				3C										4G		5B					6D	
5	[65]			1D		2B			3B				3F							4G	5A					6C		
6	[21]		1C				2C	3A							4B			4E	4F								6D	
7	[57]		1C				2C	3A												4G							6D	
8	[54]			1D		2B			3B				3F							4G		5B			6B			
9	[25]			1D		2B		3A				3E					4D							6A				
10	[51]	1B				2B				3C										4G	5A	5B	5C			6C		
11	[59]			1D	2A					3C				3G						4G	5A	5B			6B			
12	[23]			1D		2B			3B						4B											6C		
13	[49]	1B				2B				3C										4G		5B					6D	
14	[58]		1C		2A			3A				3E								4G	5A		5C	6A				
15	[50]	1B				2B		3A	3B	3C										4G			5C			6C		
16	[52]			1D		2B								3G						4G		5B				6C		
17	[66]			1D		2B		3A				3E				4C	4D					5B	5C			6C		
18	[17]	1B				2B				3C										4G		5B				6C		
19	[67]			1D		2B				3C										4G					6B			
20	[68]		1C			2B		3A								4C				4G						6C		
21	[63]		1C			2B				3C										4G		5B				6C		
22	[69]			1D		2B								3G	4B											6C		
23	[70]	1B				2B							3F			4C						5B				6C		
24	[56]			1D		2B		3A	3B		3D					4C				4G								6E
25	[61]		1C		2A									3G						4G		5B	5C			6C		
26	[55]		1C			2B				3C										4G		5B	5C			6C		

(*continued*)

(*continued*)

Authors		Context			Focus			Method							Sector						Dimensions			Origin/continents				
27	[71]		1C			2B				3C							4D					5B				6C		
28	[15]		1C			2B								3G			4D					5B	5C			6C		
29	[72]	1B					2C		3B											4G		5B	5C			6C		
30	[20]		1C			2B				3C										4G		5B				6C		
31	[73]		1C			2B				3C										4G		5B	5C			6C		
32	[22]	1B				2B				3C										4G		5B				6C		
33	[60]		1C			2B			3B											4G		5B	5C			6C		
34	[74]	1B				2B			3B							4C						5B	5C			6C		
35	[75]		1C			2B				3C										4G		5B	5C	6A				

A.2. Results of the Codification of the Articles

References

1. Chimucheka, T., Mandipaka, F.: Challenges Faced By Small, Medium And Micro Enterprises In The Nkonkobe Municipality. Int. Bus. Econ. Res. J. **14**, 309–316 (2015). https://doi.org/10.19030/iber.v14i2.9114
2. Jili, N.N., Masuku, M.M., Selepe, B.M.: SMMEs promoting local economic development (LED) in UMlalazi local municipality, KwaZulu-Natal. Afr. J. Hosp. Tour. Leis. **6**, 1–10 (2017)
3. Lumbwe, A.K., Nwobodo-Anyadiegwu, E.N., Mbohwa, C.: The impact of location decision on small, micro, and medium enterprises' performance in Johannesburg. In: Proceedings of the International Conference on Industrial Engineering and Operations Management, pp. 1205–1216 (2018)
4. Zafar, A., Mustafa, S.: SMEs and its role in economic and socio-economic development of Pakistan. Int. J. Acad. Res. Account. Financ. Manage. Sci. **6**, 1–16 (2017). https://doi.org/10.6007/ijarafms/v7-i4/3484
5. Gumbi, L., Mnkandla, E.: Investigating South African vendors' cloud computing value proposition to small, medium and micro enterprises: a case of the city of Tshwane metropolitan municipality. Afr. J. Inf. Syst. **7**, 1 (2015)
6. Bvuma, S., Marnewick, C.: Sustainable livelihoods of township small, medium and micro enterprises towards growth and development. Sustainability **12**, 1–17 (2020). https://doi.org/10.3390/SU12083149
7. Pienaar, L., Traub, L.N.: Understanding the smallholder farmer in South Africa: towards a sustainable livelihoods classification. In: 29th International Conference of Agricultural Economists, pp. 1–36 (2015)
8. Safavi, N.S., Amini, M., Javadinia, S.A.: The determinant of adoption of enterprise resource planning for small and medium enterprises in Iran. Int. J. Adv. Res. IT Eng. **3**, 1–8 (2014)
9. Bragazzi, N.L.: Digital technologies-enabled smart manufacturing and industry 4.0 in the post-COVID-19 era: lessons learnt from a pandemic. Int. J. Environ. Res. Public Health **17**, 1–2 (2020). https://doi.org/10.3390/ijerph17134785
10. Kumar, S., Raut, R.D., Narwane, V.S., Narkhede, B.E.: Applications of indusrty 4.0 to overcome the COVID-19 operational challenges. Diabetes Metab. Syndr. Clin. Res. Rev. **14**, 1283–1289 (2020). https://doi.org/10.1016/j.dsx.2020.07.010
11. Tairas, D.R.: COVID-19 pandemic and MSMEs: impact and mitigation. J. Ekon. Indones. **9**, 67–80 (2020)
12. Faasen, J., et al.: SaaS ERP adoption intent : explaining the South African SME perspective to cite this version: HAL Id : hal-01484690 SaaS ERP adoption intent: explaining the South African SME perspective (2017)
13. Assante, D., Castro, M., Hamburg, I., Martin, S.: The use of cloud computing in SMEs. In: Procedia Computer Science, pp. 1207–1212. Elsevier Masson SAS (2016)
14. Muriithi, S.M.: African small and medium enterprises (SMES): contributions, challenges and solutions. Eur. J. Res. Reflect. Manage. Sci. **5**, 36–48 (2017)
15. Gunawan, J., Kosala, R.R.: Genie enterprise resource planning for small medium enterprises implementing single page web application. In: IOP Conference Series: Earth and Environmental Science, pp. 1–10. Institute of Physics Publishing (2020)
16. Ramasamy, M., Singh, R.P.: Benefits, challenges and selection of cloud ERP systems used in SMEs. Asian J. Sci. Technol. **8**, 5966 (2017)
17. Mehta, S., Oberoi, A., Singh Bhatia, S.: A cloud based assurance framework for implementation of ERP in SMEs: a literature survey. Int. J. Comupt. Sci. Eng. **6**, 819–829 (2018)

18. Antoniadis, I., Tsiakiris, T., Tsopogloy, S.: Business intelligence during times of crisis: adoption and usage of ERP systems by SMEs. In: Procedia – Social and Behavioral Sciences, pp. 299–307 (2015)
19. Gupta, V., Singh Bhatia, S.: Developing assurance framework of cloud computing in the implementation of ERP: a literature survey. Int. J. Comput. Sci. Inf. Technol. Secur. **6**, 84–92 (2016)
20. Aulia, R., Putri, A.N., Raihan, M.F., Ayub, M., Sulistio, J.: The literature review of cloud-based enterprise resource planning. In: Annual Conference on Industrial and System Engineering, pp. 1–8. Institute of Physics Publishing (2019)
21. Faasen, J., Seymour, L.F., Schuler, J.: SaaS ERP adoption intent: explaining the South African SME perspective. In: Poels, G. (ed.) CONFENIS 2012. LNBIP, vol. 139, pp. 35–47. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-36611-6_3
22. Gupta, V., Bhatia, S.S.: Cloud computing: an operational framework in the implementation of ERP. Int. J. Adv. Res. Comput. Sci. Softw. Eng. **7**, 164–169 (2017). https://doi.org/10.23956/ijarcsse/V7I2/0109
23. Mezghani, K.: Effects of personal innovativeness on is managers' intentions to switch toward cloud ERP in Saudi SMEs. Electron. J. Inf. Syst. Eval. **21**, 46–61 (2018)
24. Wen, Y.: Research and design of ERP system for small and medium-sized enterprises under great intelligence mobile cloud. In: IOP Conference Series: Materials Science and Engineering. International Conference on Artificial Intelligence Applications and Technologies, pp. 1–6. Institute of Physics Publishing (2019)
25. Zadeh, A.H., Akinyemi, B.A., Jeyaraj, A., Zolbanin, H.M.: Cloud ERP systems for small-and-medium enterprises: a case study in the food industry. J. Cases Inf. Technol. **20**, 53–70 (2018). https://doi.org/10.4018/JCIT.2018100104
26. Uphoff, T.: COVID-19-Related Resource Planning Causes ERP Software Surge (2020). https://www.thomasnet.com/insights/covid-19-related-resource-planning-causes-erp-software-surge/
27. Bieszczat, S.: ERP Sustains Manufacturing During COVID (2020). https://erpblog.iqms.com/erp-sustains-manufacturing-during-covid-19/
28. Neil, S.: The Potential of ERP Amid a Pandemic (2020). https://www.automationworld.com/covid-19/article/21124819/the-potential-of-erp-amid-a-pandemic
29. Spanos, J.: 5 Ways COVID-19 is changing the food and beverage industry (2020). https://www.nextecgroup.com/5-ways-covid-19-is-changing-the-food-and-beverage-industry/
30. Gumbi, L., Twinomurinzi, H.: SMME readiness for smart manufacturing (4IR) adoption: a systematic review. In: Hattingh, M., Matthee, M., Smuts, H., Pappas, I., Dwivedi, Y.K., Mäntymäki, M. (eds.) I3E 2020. LNCS, vol. 12066, pp. 41–54. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-44999-5_4
31. Amui, L.B.L., Jabbour, C.J.C., de Sousa Jabbour, A.B.L., Kannan, D.: Sustainability as a dynamic organizational capability: a systematic review and a future agenda toward a sustainable transition. J. Clean. Prod. **142**, 308–322 (2017). https://doi.org/10.1016/j.jclepro.2016.07.103
32. Lage Junior, M., Godinho Filho, M.: Variations of the Kanban system: literature review and classification. Int. J. Prod. Econ. **125**, 13–21 (2010). https://doi.org/10.1016/j.ijpe.2010.01.009
33. Adam, I.O., Musah, A.: Small and medium enterprises (SMEs) in the cloud in developing countries: a synthesis of the literature and future research directions. J. Manag. Sustain. **5**, 115–139 (2015). https://doi.org/10.5539/jms.v5n1p115
34. Asatiani, A.: Why cloud? A review of cloud adoption determinants in organizations. In: Twenty-Third European Conference on Information Systems (ECIS), pp. 1–17 (2015)
35. Asenso-Okyere, K., Mekonnen, D.A.: The Importance of ICTs in the Provision of Information for Improving Agricultural Productivity and Rural Incomes in AfricaIUNDP in Africa (2012)

36. Attard, J., Orlandi, F., Scerri, S., Auer, S.: A systematic review of open government data initiatives. Gov. Inf. Q. **32**, 399–418 (2015). https://doi.org/10.1016/j.giq.2015.07.006
37. El-Gazzar, R.F.: A literature review on cloud computing adoption issues in enterprises. In: Bergvall-Kåreborn, B., Nielsen, P.A. (eds.) TDIT 2014. IAICT, vol. 429, pp. 214–242. Springer, Heidelberg (2014). https://doi.org/10.1007/978-3-662-43459-8_14
38. Jacsó, P.: The plausibility of computing the h-index of scholarly productivity and impact using reference-enhanced databases. Online Inf. Rev. **32**, 266–283 (2008). https://doi.org/10.1108/14684520810879872
39. Shah, S.R.U., Mahmood, K., Hameed, A.: Review of google scholar, web of science, and scopus search results: the case of inclusive education research. Libr. Philos. Pract. (2017)
40. Yang, H., Tate, M.: A descriptive literature review and classification of cloud computing research. Commun. Assoc. Inf. Syst. **31**, 35–60 (2012). https://doi.org/10.17705/1cais.03102
41. Jabbour, C.J.C.: Environmental training in organisations: from a literature review to a framework for future research. Resour. Conserv. Recycl. **74**, 144–155 (2013). https://doi.org/10.1016/j.resconrec.2012.12.017
42. Fahimnia, B., Sarkis, J., Davarzani, H.: Green supply chain management: a review and bibliometric analysis. Int. J. Prod. Econ. **162**, 101–114 (2015). https://doi.org/10.1016/j.ijpe.2015.01.003
43. Jabbour, C.J.C., Jugend, D., De Sousa Jabbour, A.B.L., Gunasekaran, A., Latan, H.: Green product development and performance of Brazilian firms: measuring the role of human and technical aspects. J. Clean. Prod. **87**, 442–451 (2015). https://doi.org/10.1016/j.jclepro.2014.09.036
44. Mariano, E.B., Sobreiro, V.A., do Nascimento Rebelatto, D.A.: Human development and data envelopment analysis: a structured literature review. Omega (United Kingdom). **54**, 33–49 (2015). https://doi.org/10.1016/j.omega.2015.01.002
45. World Bank: World Bank List of Economies (June 2020). https://www.ilae.org/files/dmfile/World-Bank-list-of-economies-2020_09.pdf
46. Davison, R.M., Martinsons, M.G.: Context is king! Considering particularism in research design and reporting. J. Inf. Technol. **31**, 241–249 (2016). https://doi.org/10.1057/jit.2015.19
47. Priandi, M., Fernandez, W., Sandeep, M.S.: The role of context in IS research: a review of E-government research in developing economies. In: Proceedings of the 12th International Conference on Theory and Practice of Electronic Governance, pp. 387–390 (2019)
48. Alemayehu, B.Z., Van Vuuren, J.: Munificence contingent small business growth model (special emphasis to African SMEs' context). J. Small Bus. Entrep. **29**, 251–269 (2017). https://doi.org/10.1080/08276331.2017.1327565
49. Elmonem, M.A.A., Nasr, E.S., Geith, M.H.: Benefits and challenges of cloud ERP systems – a systematic literature review. Futur. Comput. Inform. J. **1**, 1–9 (2016). https://doi.org/10.1016/j.fcij.2017.03.003
50. Nakkeeran, A., Niranga, M., Wickramarachchi, R.: A model for on-premises ERP system and cloud. In: Proceedings of the International Conference on Industrial Engineering and Operations Management, pp. 1381–1392 (2020)
51. Ramasamy, M., Periasamy, J.: Explore the impact of cloud computing on ERP systems used in small and medium enterprises. Int. J. Adv. Res. Comput. Sci. Manag. Stud. **5**, 87–91 (2017)
52. Ziani, D., Alfaadhel, N.: Web services as a solution for cloud enterprise resource planning interoperability. Int. J. Comput. Sci. Inf. Technol. **12**, 25–41 (2020). https://doi.org/10.5121/ijcsit.2020.12102
53. Gupta, S., Misra, S.C.: Compliance, network, security and the people related factors in cloud ERP implementation. Int. J. Commun. Syst. **29**, 1395–1419 (2016). https://doi.org/10.1002/dac.3107

54. Haddara, M., Fagerstrøm, A., Mæland, B.: Cloud ERP systems: anatomy of adoption factors and attitudes. J. Enterp. Resour. Plan. Stud. **2015**, 1–24 (2015). https://doi.org/10.5171/2015.521212
55. Mohammed, G., Burhanuddin, M.A.: The influence of factors to adopt cloud-based ERP for Iraqi SMEs. J. Adv. Res. Dyn. Control Syst. **10**, 2698–2707 (2018)
56. Tongsuksai, S., Sanjay, M., Nazim, T.: A framework to assess the CSFs for CERP adoption in SMEs. In: International Conference on Information Resource Management (CONF-IRM), pp. 1–8 (2019)
57. Hasheela, V., Smolander, K., Mufeti, T.: An investigation of factors leading to the reluctance of SaaS ERP adoption in Namibian SMEs. Afr. J. Inf. Syst. **8**, 1–13 (2016)
58. Pavin, R., Klein, A.: Organizational consequences of the adoption of mobile ERP systems: case studies in Brazil. J. Inf. Syst. Technol. Manage. **12**, 219–232 (2015). https://doi.org/10.4301/s1807-17752015000200002
59. Omar, K., Gomez, J.M.: A selection model of ERP system in mobile ERP design science research: case study: mobile ERP usability. In: Proc. IEEE/ACS Int. Conf. Comput. Syst. Appl., pp. 1–8 (2016). https://doi.org/10.1109/AICCSA.2016.7945791
60. Atwah Al-Ma'aitah, M.: The drivers of ERP cloud computing from an institutional perspective. J. Theor. Appl. Inf. Technol. **95**, 5142–5157 (2017)
61. Wen, Y.: Research and implementation of intelligent ERP platform for SMEs based on cloud computing. In: IOP Conference Series: Materials Science and Engineering. Institute of Physics Publishing (2019)
62. Musa Zakari Usman, U., Nazir Ahmad, M., Hidayati Zakaria, N., Alaa Hani Alkurdi, A.: A review of key factors of cloud enterprise resource planning (ERP) adoption by SMEs. J. Theor. Appl. Inf. Technol. **95**, 3884–3901 (2017)
63. Usman, U.M.Z.U., Ahmad, M.N.M., Zakaria, N.N.H.: Antecedents of cloud ERP adoption in manufacturing industry: Nigerian SMEs context. In: Proceedings of Knowledge Management International Conference (KMICE), pp. 339–344 (2016)
64. Assad, C.A.R.: Building a cloud ERP framework for modern business in developing countries. Int. J. Emerg. Technol. Adv. Eng. **5**, 15–20 (2015)
65. Cheng, Y.M.: What drives cloud ERP continuance? An integrated view. J. Enterp. Inf. Manag. **31**, 724–750 (2018). https://doi.org/10.1108/JEIM-02-2018-0043
66. AlGhazzawi, N.A., Katooa, N.E.: Vision 2030 of cloud enterprise resource planning systems (CERPS) in the Kingdom of Saudi Arabia. Int. J. Manage. Innov. Syst. **4**, 1–19 (2019). https://doi.org/10.5296/ijmis.v4i1.14644
67. Sørheller, V.U., Høvik, E.J., Hustad, E., Vassilakopoulou, P.: Implementing cloud ERP solutions: a review of sociotechnical concerns. Procedia Comput. Sci. **138**, 470–477 (2018). https://doi.org/10.1016/j.procs.2018.10.065
68. Salum, K.H., Rozan, M.Z.A.: Exploring the challenge impacted SMEs to adopt cloud ERP. Ind. J. Sci. Technol. **9**(45), 1–8 (2016). https://doi.org/10.17485/ijst/2016/v9i45/100452
69. Liu, J.-L., Wang, L.-C., Chu, P.-C.: Development of a cloud-based advanced planning and scheduling system for automotive parts manufacturing industry. Procedia Manuf. **38**, 1532–1539 (2019). https://doi.org/10.1016/j.promfg.2020.01.133
70. Ul Hassan, Z., Raza Sattar, A., Nasim, R., Arif, M.: Impact of merging cloud with customary ERP in small and medium scale enterprises in Pakistan. J. Inf. Eng. Appl. **6**, 42–49 (2016)
71. Zhang, L., Fan, H., Xue, C.: The development trend of the new retail format integration application cloud ERP System. Adv. Econ. Bus. Manag. Res. **80**, 55–60 (2019)
72. Sandu, N.: A strategic framework for successful adoption of cloud-based services in Indian SMEs. Int. J. Innov. Manage. Technol. **10**(5), 191–196 (2019). https://doi.org/10.18178/ijimt.2019.10.5.860
73. Haji Salum, K., Zaidi, M., Rozan, A.: Conceptual model for cloud ERP adoption for SMEs. J. Theor. Appl. Inf. Technol. **95**, 743–756 (2017)

74. Jain, D., Sharma, M.Y.: Cloud computing with ERP-A push business towards higher efficiency. Annu. Res. J. Symbiosis Cent. Manag. Stud. **4**, 140–155 (2016). https://doi.org/10.2139/ssrn.2755457
75. Saa, P., Costales, A.C., Moscoso-Zea, O., Lujan-Mora, S.: Moving ERP systems to the cloud – data security issues. J. Inf. Syst. Eng. Manage. **2**, 1–9 (2017). https://doi.org/10.20897/jisem.201721

Towards the Success of DevOps Environments in Software Organizations: A Conceptual Model Approach

Ashley Gwangwadza and Ridewaan Hanslo(✉)

University of Pretoria, Gauteng, South Africa
{ashley.gwangwadza,ridewaan.hanslo}@up.ac.za

Abstract. This research assesses the aspects of software organizations' DevOps environments and identifies the factors contributing to these environments' success. DevOps is a recent concept, and many organizations are moving from old-style software development methods to agile approaches such as DevOps. However, there is no comprehensive information on what factors impact the success of the DevOps environment once organizations adopt it. This research focused on addressing this gap through a systematic literature review. The systematic review consisted of 33 articles from five selected search systems and databases from 2015 to 2021. Based on the included articles, 15 factors were identified and grouped into four categories: Collaborative Culture, Organizational Aspects, Tooling and Technology, and Continuous Practices. In addition, this research proposes a DevOps environment success factors model to potentially contribute to DevOps research and practice. Recommendations are made for additional research on the effectiveness of the proposed model and its success factors.

Keywords: DevOps · Continuous deployment · Continuous integration · Continuous delivery · Success factors · Software organizations · Conceptual model

1 Introduction

Digital transformation presents new challenges for software organizations with the increasing demand for customised IT solutions [23]. Furthermore, to sustain a competitive advantage, software organizations need to deliver valuable and quality IT solutions in short cycles [30]. To address this need, many software organizations implement continuous practices: continuous delivery, continuous deployment and continuous integration [1]. In addition to continuous practices, recent studies show that DevOps adoption also improves delivery cycle times and quality [30].

The term DevOps comes from the terms of two of the core software development teams: development and operations [6]. There is no clear definition for the DevOps term. Still, the most shared interpretation is DevOps is a set of practices based on a culture that emphasizes integration and collaboration between the operations teams, the developer

P. Ndayizigamiye et al. (Eds.): IDIA 2022, CCIS 1774, pp. 70–88, 2023.
https://doi.org/10.1007/978-3-031-28472-4_5

teams, and all persons who participate in the design, development, and deployment of software [27].

DevOps takes a continuous approach that extends outside the borders of old-style software development and extends to operations [5]. Thus, DevOps implementation has been increasing as more software organizations are opting to move from the old-style software development approaches such as the Waterfall Model and Spiral Model to agile approaches as a way to acquire a competitive advantage [32]. The competitive advantage stems from the benefits of DevOps, which include a shorter time-to-market cycle, improved customer satisfaction, greater developer efficiency and productivity, continuous quick feedback and higher software quality and reliability [5].

However, adopting DevOps remains challenging because although there is an excess of information, DevOps tools, and DevOps practices, there is a lack of strategy to organize the information for correct DevOps adoption [21]. The adoption of DevOps affects different aspects of a software organization, such as organizational structures, company culture, products, processes, and related technologies used in software development [19]. Thus, many factors, such as environmental constraints and the need for clarity on the meaning of the term DevOps, may lead to unsuccessful DevOps environments [22].

Therefore, the research problem is that there is comprehensive information available on tools, practices, benefits, and challenges of DevOps but no comprehensive information for software organizations on the factors that contribute to a successful DevOps environment. Because of this research gap, it is poised that there is a need to research the factors contributing to the success of DevOps environments within a software organization context.

This research aims to identify and describe the factors that contribute to the success of DevOps environments within software organizations. This is done through a systematic literature review.

The paper is structured as follows; Sect. 2 explains the research method. Section 3 details the results of this study and Sect. 4 discusses the findings. Section 5 concludes the research, providing research implications, limitations and recommendations for future research.

2 Research Method

The research is conducted using a systematic literature review (SLR), defined as "*a means of identifying, evaluating and interpreting all available research relevant to a particular research question, or topic area, or phenomenon of interest*" [15]. Simply put, an SLR is a review of primary studies. This study follows the SLR guidelines by [15], which are: identifying resources, study selection, data extraction, data synthesis and writing up the study as a report.

2.1 Research Question

This research aims to answer the following question:

What are the factors that contribute to the success of DevOps environments within software organizations?

2.2 Search Terms

"DevOps" AND ("continuous deployment" OR "continuous development" OR "environments" OR "development" OR "software engineering" OR "methodology" OR "software organization" OR "software" OR "factors" OR "success").

2.3 Selection Criteria

The selection of research material for inclusion in this systematic review is based on the following inclusion and exclusion criteria.

Inclusion Criteria. For a source to be included in the research, it had to meet the following criteria:

- Papers that identify the factors that influence a DevOps Environment
- Papers that discuss DevOps environments within a software organization
- Papers that discussed practices or factors that contribute to the success of DevOps environments
- Journal articles, conference papers, book chapters, dissertations and thesis are considered for review
- Papers containing at least three keywords in the title, abstract, keywords

Exclusion Criteria. Sources are excluded from the research for the following reasons:

- Papers with no explicit discussion about DevOps
- Papers that do not provide information on factors contributing to DevOps success
- The full text of a paper is not available
- The paper is not written in English
- Duplicate papers (the same paper taken from different databases)
- Duplicate reports of the same study (only the most complete version is included)

The selection criteria do not include a date range. However, mid-year 2021 was the period when database searching concluded.

2.4 Source Selection

The following data sources were selected to perform the search:

a) IEEE Xplore Digital Library
b) Science Direct
c) ACM Digital Library
d) Springer Link

These databases are all recognized research databases in the information technology field. In addition, Google Scholar was used for citation searches to find additional sources missing in the abovementioned databases. This study uses the words article and paper interchangeably to reference the data source. However, the data source is either a journal article, conference paper, book chapter, dissertation, or thesis.

Figure 1 displays a Prisma Flowchart with four steps, Identification, Screening, Eligibility, and Included. Sources are excluded during each step, as depicted in the flowchart, based on the predefined selection criteria detailed in Sect. 2.3.

A search using the search string mentioned in Sect. 2.2 was performed on the selected databases, returning 2179 articles. The Google Scholar citation search found an additional 10 records. After that, 291 duplicate papers were removed. Screening by the title was conducted, leaving 262 full-text articles. An abstract screening was then performed on these articles, leaving 108. After the full-text assessment for eligibility, the 33 remaining articles were used for data extraction and synthesis.

2.5 Prisma Flowchart

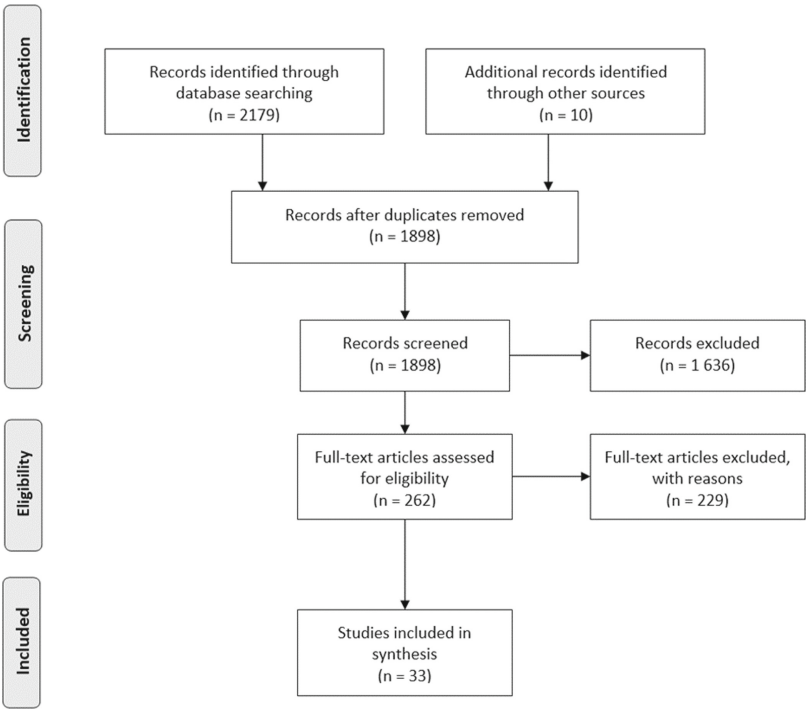


Fig. 1. Prisma flowchart

2.6 Quality Assessment

The included papers were assessed using four quality assessment questions. The questions aimed to evaluate the quality aspects mentioned by [15]. These aspects are characterised as **objectivity** – if the research is free of bias; **reliability** – the accuracy and reliability of the research instruments used; **internal validity** – whether the research was well structured, so data was collected from suitable sources, and **external validity** – determines if the findings can be predicted for subsequent occasions.

Therefore, the following questions were devised to assess the quality of the selected literature:

Q1. Are DevOps and the DevOps environments the core of the discussion?
Q2. Does the research have a well-defined aim?
Q3. Does the article follow a clear research process with a clear description of the methods used to analyze data?
Q4. Does the article report its findings based on evidence and argument?

The response options to these questions were Yes, No and Partially. The weighting for each response is as follows: Yes = 1, Partially = 0.5, No = 0. The total score was then recorded and used as a guide to indicate the quality of the selected literature ranging from 0 to 4. The quality evaluation of the articles is described in Sect. 3.2.

2.7 Data Extraction

The data extraction was carried out on the 33 papers included in the SLR. After that, a qualitative thematic analysis was conducted to synthesize the extracted data. While reading each paper, sections of its text were highlighted to identify the concepts. These concepts, also known as codes, were closely examined to categorize them into common themes. All the relevant information that helped answer the research question was extracted, including the citation, the research title, the source database, the year published and study type, article sub-concepts, and the central concept. Microsoft Excel was used for data extraction and synthesis.

3 Results

An analysis of the extracted data was conducted, which includes information on the search results, quality evaluation of the included papers, and the synthesis of identified factors.

3.1 Search Results

According to [23], the first publications devoted to DevOps originated in 2011. However, in this study, the earliest publications on the factors that influence a DevOps environment were published in 2015. Figure 2 depicts the distribution by year of the included papers that satisfied the selection criteria of this study.

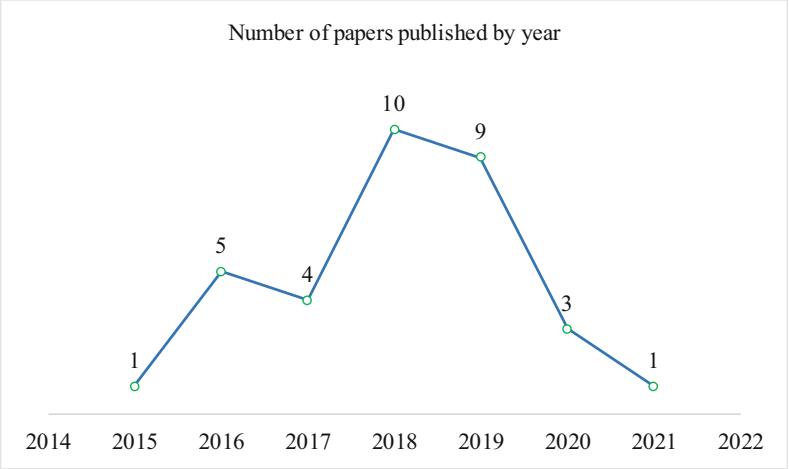


Fig. 2. Distribution of selected literature by publication year

Four databases were used to search for the papers as mentioned in the source selection section, and additional papers were found through Google Scholar. The percentage distribution is shown in the graph above (Fig. 3). The highest number of papers were from the ACM digital library (14), regarded as an extremely comprehensive database for literature on computing and information technology.

IEEE Xplore had 7 included papers, Springer Link (3) and Science Direct (3) yielded the least number of papers as they have a broad focus ranging from scientific to technological to medical. Google Scholar (7) was used for citation searches for papers relating to the research question not found in the selected databases but referenced by other papers found.

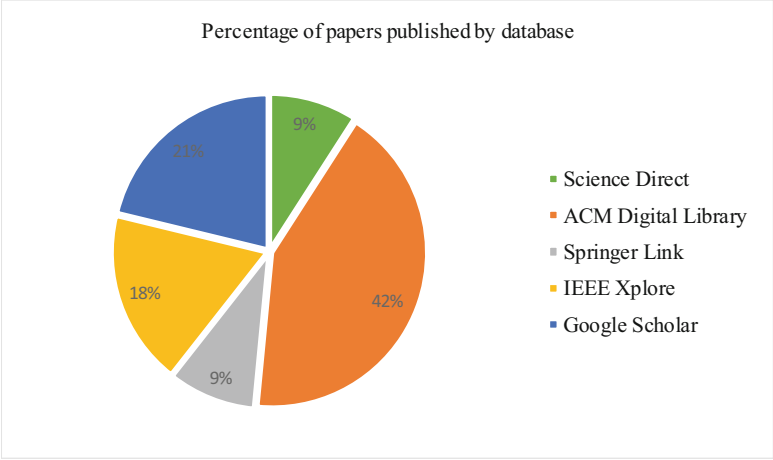


Fig. 3. Distribution of selected literature per database

Figure 4 depicts the study type distribution for the selected literature. These study types are defined in Table 1. The results show that the most used method was an SLR (45%), followed by a case study (33%). The reason could be because, as [36] stated, a case study is best for situations where the boundaries between context and phenomenon are unclear. Similarly, DevOps is a new concept. Additionally, SLRs can combine findings from different studies and identify critical areas that need to be sufficiently addressed with past empirical studies [36].

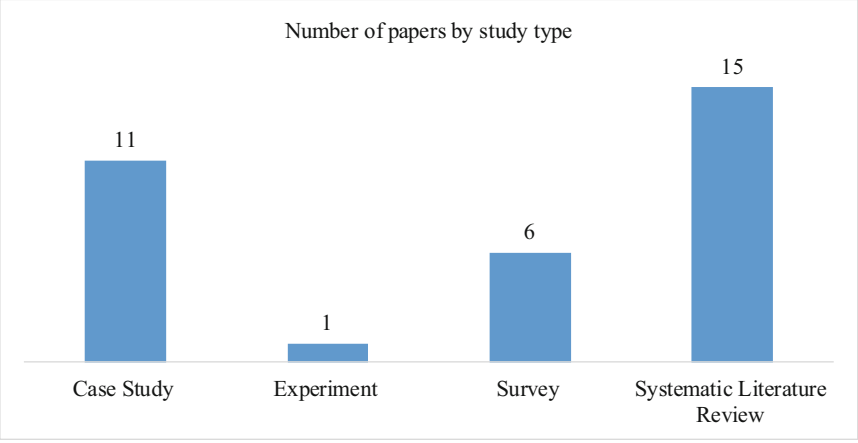


Fig. 4. Distribution of selected literature by study type

Table 1. Study type descriptions

Study type	Definition
Case study	A study that investigates a research problem within a real-life context and aims to uncover the contextual conditions [36]
Experiment	A study that investigates the causal relations and process of a research problem [36]
Survey	A retrospective study of a research problem that examines relationships and outcomes using standard operating procedures available in the guidelines [36]
Systematic literature review	Research that uses information from previously published studies for research synthesis [36]

3.2 Quality Evaluation of Articles

Four questions were used to assess the quality of the selected literature, as mentioned in the quality assessment section. The majority of the included papers had a quality assessment score of 4 (85%), with the remaining papers having a score between 2.5 and 3.5 (15%). Figure 5 displays the distribution of the quality assessment scores.

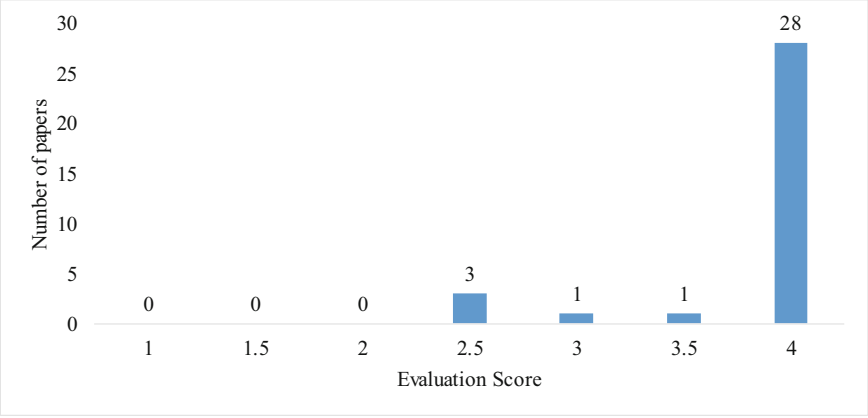


Fig. 5. Quality assessment scores distribution

3.3 Synthesis of Identified Factors

A thematic analysis was conducted to identify the core themes and factors within the selected literature. The themes corresponded with the constructs that held the factors identified. Initially, 16 factors were identified as contributing to the success of the DevOps environment. However, three factors merged into others, leaving 13 factors. The 13 factors result from re-examining each of these factors' definitions and the references made by the authors of the selected literature. The factors that were merged into others were Cooperation, Coordination and Knowledge sharing. Although none of the authors clearly defined Cooperation and Coordination, they were referenced in relation to collaboration by [29] and sometimes used in place of the word collaboration [22]. Therefore, Cooperation and Coordination were merged into the Collaboration factor. Knowledge management was found in only two papers and was defined by [5] as collection, visualisation and sharing of information and knowledge to support collaboration. Thus, this factor was then merged into the Sharing factor, which includes knowledge sharing and had more existing literature supporting it.

The remaining 13 factors were then analyzed to identify any relationships that could help categorize them. Categorizing the factors would help to understand the more significant themes and give deeper insight into the factors. Four common themes were identified: Collaborative Culture, Organizational Aspects, Tooling and Technology, and Continuous Practices, as illustrated in Table 2.

In addition, a qualitative data analysis tool called NVivo was used to refine the grouping of the factors and assess the validity of the selected themes. All the data extracted was processed in NVivo, and it generated 19 codes. Several of these codes corresponded with the initially identified thematic analysis factors. This tool grouped the codes into categories, and the result was that NVivo also grouped the factors as they were grouped in the thematic analysis. Albeit, NVivo gave a few more codes than initially identified in the first iteration of the data syntheses. After analyzing and considering each of these codes as possible factors, two additional codes were included as part of the synthesized data, resulting in a total of 15 factors (see Table 2). These two additional

codes are Continuous Deployment and Continuous Delivery. The rest of the codes were excluded as they were ambiguous or it did not contribute to answering the research question. The DevOps success factors conceptual model is shown in Fig. 6, depicting the core themes and factors.

Table 2. Identified DevOps environment success factors summary

Success factor	Sources
Theme: Collaborative Culture	
1. Communication	[2, 16, 18, 24, 25, 31]
2. Collaboration	[2, 6, 9, 10, 12, 14, 17, 18, 20, 23, 28, 29, 32, 33]
3. Sharing	[1, 4–9, 11, 17, 18, 22, 28]
Theme: Organizational Aspects	
4. Team roles setting	[6, 10, 19, 21, 26]
5. Organizational culture	[1, 4, 5, 7, 9, 12, 13, 17, 19, 25, 31]
6. Training	[2, 16, 31]
7. Customer involvement	[1, 5, 7, 17, 24]
Theme: Tooling and Technology	
8. Automation	[1, 3–9, 11, 17, 18, 19, 22, 28, 32]
9. Measurement	[1, 3–5, 9, 17, 19, 23, 28]
10. Tooling	[7, 13, 16, 19, 22]
Theme: Continuous Practices	
11. Continuous integration	[8, 14, 19, 27, 29, 30]
12. Continuous experimentation	[3, 11, 14, 17, 19, 22, 27, 30]
13. Continuous delivery	[14, 17, 30]
14. Continuous deployment	[3, 11, 14, 19, 27, 30, 32]
15. Continuous monitoring	[8, 11, 19, 22]

4 Discussion

This section aims to answer the research question. The core themes identified are 1) Collaborative Culture; 2) Organizational Aspects; 3) Tooling and Technology, and 4) Continuous Practices. Each of these themes has three to five factors. Therefore, the discussion will first explain how these themes were formed and then discuss the factors under each theme.

4.1 Collaborative Culture

A collaborative culture is unequivocally a core theme regarding the success of the DevOps environment [17, 18]. Three factors that make up this Collaborative Culture are Collaboration, Communication and Sharing.

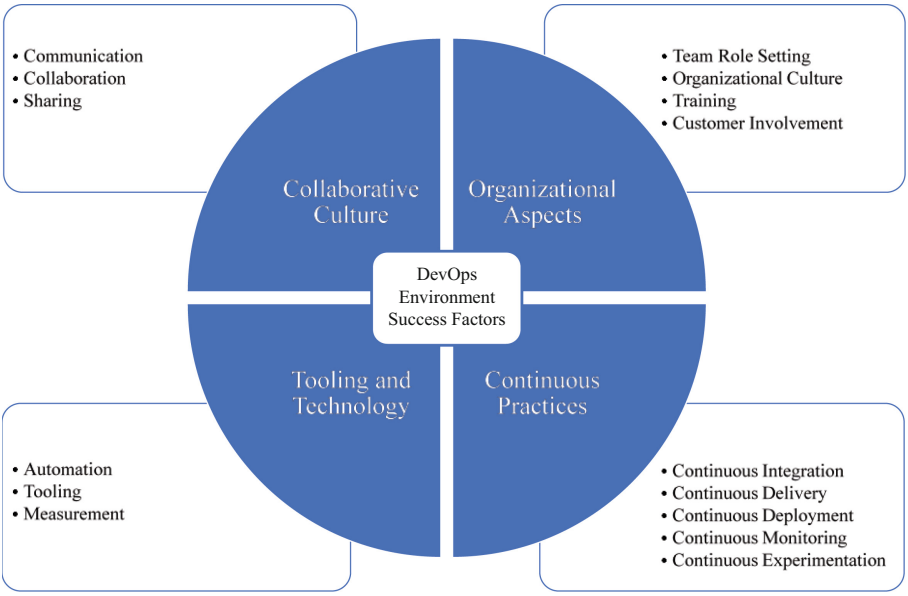


Fig. 6. DevOps environment success factors model

Success Factor 1: Collaboration. Collaboration is critical in a DevOps environment because it is the essence of DevOps itself as it focuses on working together, and the exchange of knowledge and experience between teams. DevOps is all about connecting Development and Operations, which entails collaboration [23]. The following 13 authors identified collaboration in this SLR as one of the essential aspects of DevOps environments [2, 6, 9, 10, 12, 14, 17, 18, 20, 23, 28, 29, 32, 33]. This is because numerous benefits have been seen when a DevOps environment has a strong culture of collaboration. These advantages include enhanced development and operations processes, as well as improved quality of software services and products [9].

[32] concluded that their findings on IT governance in DevOps teams support the concept that cross-functional collaboration can result in competitive advantages. Furthermore, collaboration leads to significant business improvements, enhanced operation software quality, and better-quality IT services ([29]). It is posited that collaboration within a DevOps Environment is one of the contributors to success.

Success Factor 2: Communication. Communication between the operations and development teams is essential for DevOps to succeed [2, 18]. A DevOps environment

is essentially made up of a cross-functional team. Thus, for the teams to work together effectively, information within and between the teams needs to be shared effectively and frequently. The presence of quality communication helps establish a collaborative culture which contributes to a successful DevOps environment. [2] emphasize that communication is crucial for DevOps to succeed because not sharing important information may have a significant negative impact, therefore, managers must impose regular communication. [21] cited communication as a critical success factor for DevOps improvement. [21] further posits that effective communication is the most important success factor during DevOps implementation in a software organization.

Success Factor 3: Sharing. Sharing is prevalent in a thriving DevOps environment because it has a tremendous positive impact on its process and output. Sharing in DevOps means the joint use of tools, practices, processes, and knowledge [33]. [1, 4–9, 11, 17, 18, 22, 28] states how sharing tools, practices, and processes impact the DevOps environment. [29] identified sharing resources as one of the success factors of collaboration in a DevOps context. [33] point out how sharing failures and successes in DevOps practices helps team members improve quicker by learning from others.

While [3] mentions sharing tools, experiences and findings allow individuals with similar interests to meet and interact. The author also states how this brings a sense of commitment to individuals and reduces duplicates in work. Sharing facilitates successful DevOps teams and helps magnify their success [32].

4.2 Organizational Aspects

Certain aspects within a software organization influence the success of the DevOps environment. These aspects include Team Role Setting, Culture, Training and Customer Involvement.

Success Factor 4: Team Role Setting. Team role setting is one of the factors that are still blurry in many DevOps environments, but when team roles are set, it impacts the success of these environments. This factor has to do with the precise definition of roles and responsibilities within a DevOps team. This definition of roles is crucial because the observed trend in DevOps is the reorientation of roles and responsibilities between software development and operations teams [19]. Clear team roles have several benefits. Some of the benefits are; positive effects on team members' performance [20], helping practitioners in recognizing their crucial role in DevOps processes [2] and assisting in aligning the motivations of the key roles involved in software delivery [20]. [30] also identified clear role setting as one of the success factors of collaboration in the DevOps context.

Success Factor 5: Organizational Culture. Organizational culture is an integral part of the DevOps environment, and changing the organizational culture to have the attributes of DevOps culture is a decisive success factor. DevOps features culture as one of the vital aspects to safeguard the success of DevOps adoption [28]. Even though DevOps have varying definitions, most authors agree that culture is the epitome of DevOps. The popular interpretation is that DevOps is about culture [24], and culture change is

seen as the beginning of DevOps [23]. In an organization, you will find team culture and organizational culture. Team culture flows from the organizational culture [24]. Therefore, it is integral that the organizational culture represents a DevOps culture.

Success Factor 6: Training. Training is one of the more overlooked success factors for multiple reasons. However, although it was cited the least in this SLR (tied with continuous delivery), its impact is undoubtedly noteworthy. Appropriate training is required to grasp the DevOps and DevOps environment concepts well [33]. DevOps requires supplementary technical skills from both software developers and operators since operators must pick up some skills from developers and vice versa [10].

The training given to teams should be all-encompassing, including the use of DevOps concepts [16], new processes and tools [10], and DevOps culture [2]. Well-rounded training helps keep pace with DevOps practices and tooling evolution [16]. Training contributes to a DevOps environment being successful through several realized benefits. These benefits include combating the lack of awareness challenge in an organization [10] and helping team members know how their work will interact with and impact other areas [33]. Training and education help organizations successfully work on DevOps activities [21], contributing to a thriving DevOps environment. Therefore, training is imperative for DevOps success and must be considered a priority for organizations.

Success Factor 7: Customer Involvement. Customer involvement was identified as a success factor for software organizations' DevOps environment due to the customer feedback's impact on software product improvement. The term customer involvement is extensive and has several definitions. [31] describes customer involvement as "*preparing and receiving customer input, establishing a customer sample group, and delivering feature growth.*" Through verified performance improvement, [31] identified customer involvement as a critical success factor during the adoption of continuous deployment and continuous delivery in the context of DevOps. In a study by [21], 87% of the respondents considered 'Customer feedback to improve development' the second most crucial success factor in DevOps adoption. [21] express how customer involvement positively impacts different areas in a software organization's DevOps environment.

Interacting with their customers allows organizations to understand their customer's preferences and modify their software products to meet market demands [22]. Other benefits include the quicker discovery of bugs and defects in the software [3] and saving time in market development [28]. In addition, customer involvement plays a crucial role in validating the continuously released features; it keeps the feedback loop tight and helps the teams be more focused [10]. [3] emphasizes that transparency between teams and customers helps the developers assess customers' reactions to the software products allowing the team to make the necessary improvements. This factor has several reported and proven benefits in the existing literature and is posited to contribute to a DevOps environment's success.

4.3 Tooling and Technology

The selection of tools and use of technology in DevOps significantly affects the success of a DevOps environment [6]. This section discusses how three factors, Automation, Measurement, and Tooling, contribute to a DevOps environment's success.

Success Factor 8: Automation. Automation is referred to as one of the key aspects of DevOps because of the advantages it brings to a DevOps environment. Making repetitive manual tasks automated has contributed to success in DevOps environments within a software organization. [18, 23] identified six concepts of automation in DevOps. 1) Build automation, 2) recovery automation, 3) deployment automation, 4) monitoring automation, 5) test automation and 6) infrastructure automation (also referred to as infrastructure as code). However, three concepts of automation that are widely agreed upon in our SLR as having benefits that contribute to success in a DevOps environment are test automation, deployment automation, and infrastructure automation.

Deployment automation automates software deployments and rollbacks of software build to be transferred to development, testing, acceptance, and production environments quicker [7]. Deployment automation reduces user errors and increases deployment speed [10].

Test automation is executing tests and repeating activities using software [33]. Test automation reduces manual testing and accelerates the testing process significantly [10]. In an experiment conducted by [26], they observed that test automation strongly affected deployment speed.

Infrastructure automation provides infrastructure elements as code and configuration of application servers using pre-set automation logic [10]. Automation improves the consistency of infrastructure, team productivity, and repeatability of activities [9].

Each of these forms of automation has its benefits. When all these areas within a DevOps environment are automated, it increases the benefits realized, contributing to success. Some significant positive impacts of automation include increased agility and reliability [17] and improved quality assurance [22].

Furthermore, automation decreases the effort needed to set up releases, making it possible for organizations to release software frequently [22]. [28] stated that automation is seen as the most significant characteristic in DevOps practice that is advantageous to software production.

In addition, deployment, test, and infrastructure automation form part of the DevOps continuous practices, discussed in Sect. 4.4.

Success Factor 9: Tooling. Tooling in DevOps is posited to affect the success of a DevOps environment because it is not just about using tools but using the right tools. Selecting the wrong tools might negatively influence the productivity of work [3].

The term 'DevOps tools' was described by [16] as a set of tools that achieve one of the following goals: 1) supporting collaboration throughout different divisions, 2) sustaining software reliability, or 3) aiding continuous delivery. The tools need to be appropriate to a specific DevOps environment because, as [10] explained, DevOps is not standardized to any prescribed toolset; thus, individual setups will differ across organizations. In

addition, DevOps tools need to support DevOps practices as this is the primary purpose of the tools. Tools are used to build DevOps capabilities and collaboration [14]. DevOps tools assist in monitoring, implementing, deploying, automating, and analyzing each process in the DevOps journey [3]. Therefore, selecting the correct tools for your DevOps environment is essential [6].

Success Factor 10: Measurement. Measurement in DevOps environments is important because it influences the software development process. Measurement focuses on specifying production metrics to monitor and evaluate the performance of a process in software development and operations [19]. As indicated, this factor's significance stems from the impact of having metrics and a baseline to assess performance in a software organization's DevOps environment. The sooner an organization applies measurement, the sooner they establish a baseline to evaluate improvement [8].

All processes must be quantified for an organization to recognize its delivery capability [5]. Measurement affects how team members work because visualization of metrics helps team members have a realistic view of their team's performance [10]. In addition, the aggregation of efficient metrics supports the decision-making process during development and operations [9]. Observable and easy-to-read data is crucial to making the correct choices [3]. However, the impact of measurement goes beyond decision-making when it is applied to software systems. Software measurement aids in identifying vulnerability, quality assessment, productivity improvement, compliance management, and enhancement of development practices [3]. Therefore, as stated by [8], measurement is paramount to forming an effective and valuable software development stream. Measurement is posited to contribute to the success of a DevOps environment.

4.4 Continuous Practices

The practices done constantly in a DevOps environment are considered vital to the environment's success; these are Continuous Practices. Continuous Practices are Continuous Integration, Continuous Delivery, Continuous Deployment, Continuous Monitoring, and Continuous Experimentation.

Success Factor 11: Continuous Integration. Continuous integration is a straightforward success factor in a DevOps environment, as DevOps demands a need for this [27]. It is defined by [31] as the part that deals with code compilation, acceptance testing, unit testing, confirming code coverage, examining compliance to coding standards, and deployment package building. Continuous integration is an incremental process that prevents last-minute integration issues and is about achieving one of the essential values of DevOps, which is speed [14]. In addition, continuous integration reduces the risk and cost associated with software releases [33]. Therefore, partaking in continuous integration activities contributes to the success of a DevOps environment within a software organization.

Success Factor 12: Continuous Delivery. Continuous delivery is an integral part of the DevOps environment, as it allows DevOps teams to automate their software delivery process up to the point just before it is deployed [34]. Continuous delivery as a

DevOps practice correlates to performance [14], and it has been argued that measurement contributes to continuous delivery [17, 30]. An essential benefit of this practice, as mentioned by [14, 17, 30], is how it creates an environment for quality outcomes. In other words, by implementing the continuous delivery practice, quality can be seen as a by-product.

Success Factor 13: Continuous Deployment. Continuous deployment allows DevOps teams to automatically deploy new software features onto a production environment [19, 35]. No manual human intervention is needed to enable the software to be deployed. [19] states that the popularity of DevOps results from the emergence of continuous deployment, and [30] suggests this practice reduces the costs of detecting defects due to faster feedback from users. Therefore, it is no surprise that seven of the included studies mentioned it as a factor. Continuous development is posited to contribute to DevOps success within software organizations.

Success Factor 14: Continuous Monitoring. Continuous monitoring is a factor whose contribution to the DevOps environment hinges on the presence of another factor, namely, measurement. However, monitoring is a continuous and ongoing activity, and therefore it is classified as a continuous practice. Once metrics in measurement have been set, continuous monitoring commences. Continuous improvement in DevOps is enabled by constantly monitoring processes and measuring them against metrics [7].

[14] state that DevOps success possibly hinges on how a system is monitored, as monitoring aids in understanding a system inside-out. [22] identified that developers emphasized real-time monitoring as a factor that helped create fault-aware systems and that well-timed feedback from monitoring enabled continuous delivery of software. [7] outline how monitoring for incidents aids in continually improving a product inherent to continuous improvement. Similarly, [21] cited that 'Monitoring automation for continuous delivery' was an essential factor in DevOps and a suitable solution to delivering products on time. Therefore, continuous monitoring contributes to a software organization's DevOps environment by enabling continuous improvement.

Success Factor 15: Continuous Experimentation. Continuous experimentation is an infrequently used term but was identified in eight of the selected literature as having a positive impact on a DevOps environment that could contribute to success. Continuous experimentation is defined by [22] as the continual testing of propositions to determine the benefits for the customer and the software organization. An in-depth analysis by [19] of their primary studies revealed continuous and rapid experimentation as one of the recurring factors in continuous deployment. [24] identified experimentation as one of the characteristics of a healthy DevOps environment.

Experimentation is one of the attributes DevOps teams must exhibit to support DevOps capabilities [25]. Another aspect of continuous experimentation that positively impacts a DevOps environment is giving developers time to experiment with new ideas, driving the discovery process, and inspiring team members to share their discoveries with the rest of the organization [10]. In addition, an experimental culture involves more

customer interactions [22]. These findings are of great importance because continuous experimentation is one of the essential practices in a DevOps environment.

The systematic literature review identified factors that are posited to contribute to the success of DevOps environments. Collaboration and organizational culture were the most cited factors, which can infer that they are the most important. However, this research did not prioritise the factors because it is beyond the scope of this study.

It was observed that the influence of some factors was at distinct phases of software development, while others influenced the entire software development process. The factors associated with Organization Aspects and Collaborative Culture influence the DevOps environment at every phase of software development, from initial team setup to product delivery. In contrast, factors related to Tooling and Technology and Continues Practices have more impact once software development begins.

5 Conclusion

A systematic literature review (SLR) was conducted to identify the factors that contribute to the success of a software organization's DevOps environment. Thirty-three articles were selected from five data sources for data extraction and synthesis. The data synthesis was conducted using a thematic analysis following the data extraction. For the first iteration of data synthesis, 16 factors were identified. After the second iteration of data synthesis, the factors were refined to 13. After the final iteration of data synthesis with the help of the NVivo tool, two additional factors were included, settling on 15 factors posited to contribute to the success of a DevOps environment within software organizations. As a result of this study, an initial conceptual model was developed. The model developed provides the mapping of the identified factors to constructed themes. The factors were classified into four themes to answer the research question – Collaborative Culture (3 factors), Organizational Aspects (4 factors), Tooling and Technology (3 factors), and Continues Practices (5 factors). The research aims were met, aiding in answering the research question.

This study's limitations are that English papers were only considered for this SLR; thus, any information that might have been relevant to the research question but was written in another language is excluded. Secondly, only five data sources were used to search for the literature for this SLR; therefore, relevant literature from other databases might have been missed. Thirdly, some factors were excluded or merged into others due to a lack of a clear definition or information indicating how they contribute to success. The implication might be that some standalone factors in a DevOps environment were merged with others in this study. Fourth, the search terms used in the search string may not be of adequate rigour, which could have resulted in missed factors and themes.

Researchers and practitioners could use the model derived from this SLR to discover the critical success factors in a DevOps environment. Furthermore, another study could investigate the theme's and factor's level of contribution towards DevOps environment success. In addition, a similar study could be undertaken to confirm or disprove this study's success factors and themes.

Acknowledgements. This work is based on the research supported by the National Research Foundation of South Africa (Grant Number 138325).

References

1. Agarwal, A., Gupta, S., Choudhury, T.: Continuous and integrated software development using DevOps. In: 2018 International Conference on Advances in Computing and Communication Engineering (ICACCE), pp. 290–293. IEEE (2018)
2. Akbar, M.A., Mahmood, S., Shafiq, M., Alsanad, A., Alsanad, A.A.A., Gumaei, A.: Identification and prioritization of DevOps success factors using fuzzy-AHP approach. Soft Comput. **27**, 1–25 (2020). https://doi.org/10.1007/s00500-020-05150-w
3. Aljundi, M.K.: Tools and practices to enhance DevOps core values. LUTPub (2018). http://urn.fi/URN:NBN:fi-fe201801121311
4. Bass, L.: The software architect and DevOps. IEEE Softw. **35**(1), 8–10 (2018). https://doi.org/10.1109/MS.2017.4541051
5. Colomo-Palacios, R., Fernandes, E., Soto-Acosta, P., Larrucea, X.: A case analysis of enabling continuous software deployment through knowledge management. Int. J. Inf. Manage. **40**, 186–189 (2018)
6. Ebert, C., Gallardo, G., Hernantes, J., Serrano, N.: DevOps. IEEE Softw. **33**(3), 94–100 (2016). https://doi.org/10.1109/MS.2016.68
7. Feijter, R., Vliet, R., Jagroep, E., Overbeek, S., Brinkkemper, S.: Towards the adoption of DevOps in software product organizations: a maturity model approach. In: UU BETA ICS Departement Informatica (2017)
8. Forsgren, N., Kersten, M.: DevOps metrics. Commun. ACM **61**(4), 44–48 (2018). https://doi.org/10.1145/3159169
9. França, B.B.N.d., Jeronimo, H., Travassos, G.H.: Characterizing DevOps by hearing multiple voices. In: Proceedings of the 30th Brazilian Symposium on Software Engineering, Maringá, Brazil (2016). https://doi.org/10.1145/2973839.2973845
10. Hamunen, J.: Challenges in adopting a Devops approach to software development and operations (2016)
11. Hussaini, S.W.: A systemic approach to re-inforce development and operations functions in delivering an organizational program. Procedia Comput. Sci. **61**, 261–266 (2015). https://doi.org/10.1016/j.procs.2015.09.209
12. Jabbari, R., Ali, N.B., Petersen, K., Tanveer, B.: What is DevOps? A systematic mapping study on definitions and practices. In: Proceedings of the Scientific Workshop Proceedings of XP2016, Edinburgh, Scotland, UK (2016). https://doi.org/10.1145/2962695.2962707
13. Kamuto, M.B., Langerman, J.J.: Factors inhibiting the adoption of DevOps in large organizations: South African context. In: 2017 2nd IEEE International Conference on Recent Trends in Electronics, Information and Communication Technology (RTEICT), pp. 48–51. IEEE (2017)
14. Katal, A., Bajoria, V., Dahiya, S.: DevOps: bridging the gap between development and operations. In: 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC), pp. 1–7. IEEE (2019)
15. Kitchenham, B.: Procedures for performing systematic reviews. Keele, UK, Keele Univ. **33**(2004), 1–26 (2004)
16. Leite, L., Rocha, C., Kon, F., Milojicic, D., Meirelles, P.: A survey of DevOps concepts and challenges. ACM Comput. Surveys (CSUR) **52**(6), 1–35 (2019)
17. Luz, W.P., Pinto, G., Bonifácio, R.: Building a collaborative culture: a grounded theory of well succeeded DevOps adoption in practice. In: Proceedings of the 12th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement, Oulu, Finland (2018). https://doi.org/10.1145/3239235.3240299
18. Luz, W.P., Pinto, G., Bonifácio, R.: Adopting DevOps in the real world: a theory, a model, and a case study. J. Syst. Softw. **157**, 110384 (2019)

19. Lwakatare, L.: DevOps adoption and implementation in software development practice. Concept, practices, benefits and challenges. Ph.D. Dissertation (2017)
20. Masombuka, T., Mnkandla, E.: A DevOps collaboration culture acceptance model. In: Proceedings of the Annual Conference of the South African Institute of Computer Scientists and Information Technologists, Port Elizabeth, South Africa (2018). https://doi.org/10.1145/3278681.3278714
21. Rafi, S., Yu, W., Akbar, M.A.: RMDevOps: a road map for improvement in DevOps activities in context of software organizations. In: Proceedings of the Evaluation and Assessment in Software Engineering, pp. 413–418 (2020)
22. Riungu-Kalliosaari, L., Mäkinen, S., Lwakatare, L.E., Tiihonen, J., Männistö, T.: DevOps adoption benefits and challenges in practice: a case study. In: Abrahamsson, P., Jedlitschka, A., Duc, A.N., Felderer, M., Amasaki, S., Mikkonen, T. (eds.) Product-Focused Software Process Improvement, pp. 590–597. Springer International Publishing, Cham (2016). https://doi.org/10.1007/978-3-319-49094-6_44
23. Rütz, M.: DevOps: A Systematic Literature Review No. August, 23, 25 (2019)
24. Sánchez-Gordón, M., Colomo-Palacios, R.: Characterizing DevOps Culture: A Systematic Literature Review: Software Process Improvement and Capability Determination. Springer, Cham (2018)
25. Senapathi, M., Buchan, J., Osman, H.: DevOps capabilities, practices, and challenges: insights from a case study. In: Proceedings of the 22nd International Conference on Evaluation and Assessment in Software Engineering 2018, Christchurch, New Zealand (2018). https://doi.org/10.1145/3210459.3210465
26. Shahin, M., Babar, M.A.: On the role of software architecture in DevOps transformation: an industrial case study. In: Proceedings of the International Conference on Software and System Processes, Seoul, Republic of Korea (2020). https://doi.org/10.1145/3379177.3388891
27. Stahl, D., Martensson, T., Bosch, J.: Continuous practices and devops: beyond the buzz, what does it all mean? In: 2017 43rd Euromicro Conference on Software Engineering and Advanced Applications (SEAA), pp. 440–448. IEEE (2017)
28. Toh, M.Z., Sahibuddin, S., Mahrin, M.N.R.: Adoption issues in DevOps from the perspective of continuous delivery pipeline. In: Proceedings of the 2019 8th International Conference on Software and Computer Applications, Penang, Malaysia (2019). https://doi.org/10.1145/3316615.3316619
29. Van Belzen, M., DeKruiff, D., Trienekens, J.J.: Success factors of collaboration in the context of devops. In: Proceedings of the 12th IADIS International Conference Information Systems 2019, IS 2019, pp. 26–34 (2019)
30. Van Belzen, M., Trienekens, J., Kusters, R.: Critical success factors of continuous practices in a DevOps context (2019)
31. Vonk, R., Trienekens, J.J., van Belzen, M.: A study into critical success factors during the adoption and implementation of continuous delivery and continuous deployment in a DevOps context. ACM (2021)
32. Wiedemann, A.: IT governance mechanisms for DevOps teams – how incumbent companies achieve competitive advantages. In: Proceedings of the 51st Hawaii International Conference on System Sciences (2018)
33. Wiedemann, A., Forsgren, N., Wiesche, M., Gewald, H., Krcmar, H.: Research for practice: the DevOps phenomenon. Commun. ACM **62**(8), 44–49 (2019)
34. Chen, L.: Towards architecting for continuous delivery. In: 2015 12th Working IEEE/IFIP Conference on Software Architecture, pp. 131–134. IEEE (2015)

35. Rodríguez, P., et al.: Continuous deployment of software intensive products and services: a systematic mapping study. J. Syst. Softw. **123**, 263–291 (2017)
36. Sjoberg, D.I., Dyba, T., Jorgensen, M.: The future of empirical methods in software engineering research. In: Future of Software Engineering (FOSE 2007), pp. 358–378. IEEE (2007)

Exploring the Business Value that South African Services-Oriented Micro-businesses Derive from Mobile Applications

Anele Cele and Jean-Paul Van Belle(✉)

University of Cape Town, P Bag, Rondebosch 7701, South Africa
jean-paul.vanbelle@uct.ac.za

Abstract. Small and Micro-Sized Enterprises (SMEs) play a significant role in economic development, job creation, and poverty alleviation. Despite the many advantages SMEs provide for an economy, they still face a range of challenges and have an alarming survival rate. This study examines the business value that organisations derive from the adoption of mobile applications, considering the positive impact that digital technologies can have on an organization's success. This research focuses on SMEs which are classified as personal service establishments in the Western Cape. The study was guided by a conceptual framework that combines the Technology-Organization-Environment framework, the Technology Acceptance Model, and the M-Business Value model. A thematic analysis approach was used to analyze the data collected through semi-structured interviews. The research shows that the value SMEs derive from mobile app users is a result of the apps being used to drive both sales and marketing and streamline internal operations. Based on these findings, SMEs found that they could derive business value from mobile applications when it was aligned with the task required.

Keywords: Micro-business · Mobile applications · Business value · M-Commerce · Mobile adoption · South Africa

1 Introduction

Small-medium enterprises (SMEs) cannot compete competitively without properly adopting technology at the correct market levels [1]. "Performance" can refer to both intermediate process-level measures and organizational measures and revealed two contributors to performance being both efficiency and effectiveness [2]. Business process performance entails the operational efficiency of specific business processes within an entity and their measures, including information sharing, flexibility, and inventory management [2].

However, South African SMEs experience several barriers in adopting technology such as lack of IT literacy, lack of internet access, and software [3]. Mobile technology could alleviate some of the traditional IT barriers. This study, therefore, explores the impact that mobile applications have on SMEs in South Africa by looking at business value derived and how it is derived.

P. Ndayizigamiye et al. (Eds.): IDIA 2022, CCIS 1774, pp. 89–100, 2023.
https://doi.org/10.1007/978-3-031-28472-4_6

The main goal of this study is to explore what business value SMEs in South Africa derive from the use of mobile applications. Current literature identifies a growing number of SMEs adopting mobile technologies in South Africa to solve their problems; however, there is a lack of research outlining what value is derived from mobile applications by SMEs and what influences its adoption [4]. This study's objective is to understand the experiences, understanding and meanings that respondents may have concerning the mobile applications in South African service-oriented micro-business. The questions this research aims to answer are: *"How do they use mobile applications to drive value?"* and "*What value do these micro-businesses derive from the use of mobile applications?"*.

2 Literature Review and Theoretical Framework

2.1 Definitions

For the context of this study, a small business is a "distinct business entity managed by one owner or more" [5]. The act categorizes SMEs into four different categories being micro, very small, small and medium enterprises. Micro-businesses are classified as those with less than 10 employees, a turnover below ZAR200,000 or an asset base of less than ZAR500,000. For ease of reading, the *micro-businesses in this study are referred to by the more universal "SME" acronym.*

Mobile apps are applications that are installed or are accessible on mobile phones or smartphones, and These applications offer up a range of functionalities by utilising the hardware of the device it is installed to collect, retrieve or store data on the device or a virtual cloud [4]. Apps can be either native, web apps or hybrid applications. Statistics relating to the number of application downloads in South Africa could not be obtained. However, internet penetration in South Africa has grown by 4.5% between 2020 and 2021 and now stands at 64%, with 38.19 million internet users [6].

2.2 Application of Mobile Apps in SMEs

Mobile applications impact a business's success in two ways: streamlining communication within an organisation and external communications makes communication more accessible and affordable [7]. It can access the internet through cheaper platforms such as social media and allow businesses to have a global reach cost-effectively through these social media platforms. The second way is by streamlining admin tasks in a business, mobile devices provide convenience as they simplify access to business records and documentation when one needs them.

With the growing adoption of mobile technologies in Africa, SMEs are turning to them to find solutions to their problems [4]. However, SMEs face barriers in adopting mobile apps, such as a lack of ICT skills, financial resources, and high dependence on business partners. There is also a lack of awareness of ICT benefits and the perception that adoption is costly [8]. ICT adoption in developing countries is hindered by many obstacles, including and not limited to lack of facilities, technological capabilities and lack of legal determinants. Most African developing countries struggle with finances which impacts infrastructure development and a poorly skilled population to support development [9].

SME owners were asked whether technological limitations are an obstacle for them to grow their business. Fifty percent of the participants responded that technological limitations pose an obstacle. Unstable internet access was the main obstacle, as depicted in Fig. 1 [3]. The reasons cited by these SMEs as obstacles to their adoption also include software, tech skills and hardware. The cost of data is a significant barrier to accessing stable internet access, mobile data is the most pervasive method of connectivity, with 71 percent of SMEs reporting to use mobile LTE networks (SME South Africa, 2020). The average cost of 1 Gb of data in South Africa is $7.50 compared to the average cost of $4.90 in Kenya or $2.80 in Uganda [10].

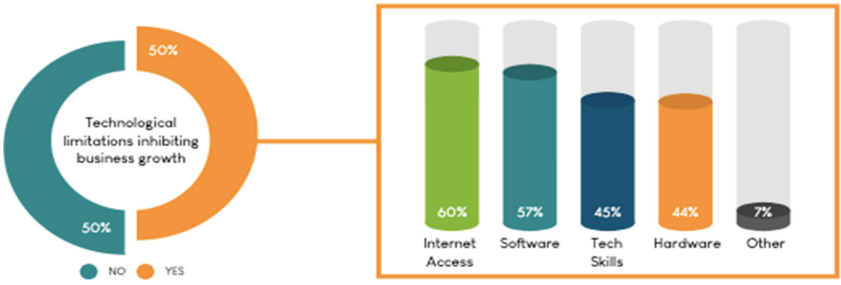


Fig. 1. Tech limitations inhibiting business growth [3].

2.3 Business Benefits of Mobile App Adoption

Mobile technology opens up enormous benefits for SMEs. Technology has enabled SMEs to access new markets, reduce costs, and improve efficiency and competitiveness [3]. Businesses need to adopt processes and technology that will allow them to bring about services that will give them a competitive advantage [11].

Technology has proven to positively influence small businesses within the Gauteng province, affecting in two ways via their productivity and through their competitiveness. Mobile applications impact a business's success in two ways: streamlining communication within an organisation and external communications, making communication more accessible and more affordable [7]. It can access the internet through cheaper platforms such as social media and allow businesses to have a global reach cost-effectively through these social media platforms. The second way is by streamlining admin tasks in a business, mobile devices provide convenience as they simplify access to business records and documentation when one needs them [7]. They enable businesses to improve their productivity and streamline their communications which therefore gives them a competitive edge. A business's growth is a function of mobile tech for its business processes and a founders' intent to grow it and argues that it is also dependant on the nature of the business, its industry and strategy [12]. [12] also contends that mobile phones in business activities had the most significant effect on business growth when coupled with a robust entrepreneurial spirit.

Technology adoption is becoming more affordable and necessary to survive in today's ultra-competitive business environment [13]. Empirical evidence shows that small and

medium-sized enterprises cannot achieve competitiveness and remain profitable without proper technology adoption [1]. Mobile technology adoption has improved business productivity which can be attributed to communications, market solicitation, social interaction and operational cost reduction [11].

2.4 The M-Business Value Model to Understanding Mobile App Adoption by SMEs

The M-Business Value Model (Fig. 2) is used to help understand the impact of mobile applications on an organisation's value chain activities that bring about business value. In this model, the business value is the impact on value chain activities such as sales and marketing and internal operations. With this model, mobile app usage drives business value and task requirements influence the usage of mobile applications [14]. This conceptual framework being utilized to add value to the understanding of what value businesses derive from mobile applications and how a business uses mobile applications drive value in certain business areas.

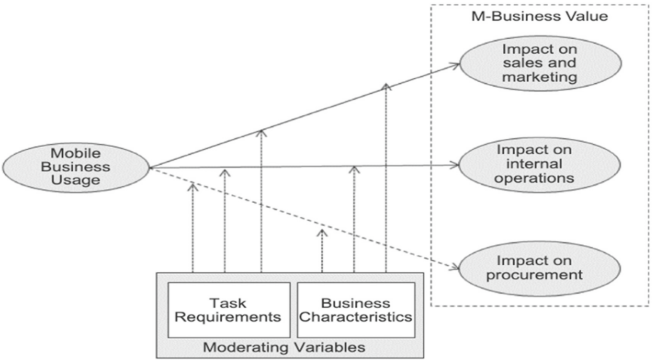


Fig. 2. M-Business Value Conceptual Model [14].

3 Research Methodology

This study takes a qualitative approach to investigate SME respondents' experiences, behaviours, organisational interactions, and contextual phenomena around mobile applications to unpack their impact on their business [15]. Maximizing efficiency and validity is the goal of sampling methods. Qualitative research allows for the research participants' subjective meanings, actions, and social contexts to be illuminated [16]. The use of qualitative research is beneficial when the research focuses on complex issues, such as human behavior and felt needs. Consequently, qualitative research aims to enhance our understanding of social phenomena by drawing on the views and experiences of all participants.

Purposive sampling was adopted as a technique to select potential candidates. Purposive sampling seeks to select individuals by virtue of their knowledge and experience

[17]. Identifying and selecting information-rich cases relating to the phenomenon of interest is a widely used method of qualitative research (Patton, 2002). This approach works well with the qualitative approach when taking into consideration the importance of availability and willingness to participate, as well as the ability of participants to articulate, express, and reflect on experiences and opinions. A list of potential candidates was drawn up by obtaining a list of laundry services, beauty salons and accommodation establishments in the Western Cape through Google search and neighbourhood scoping. Potential candidates were then contacted via phone and email. Participants were given a list of mobile apps based on [4] and 5 common barriers SMEs face when adopting mobile applications extracted from the SME South Africa landscape report [3]. This was followed by a semi-structured interviews which asked demographic information and questions around what motivates them to select particular mobile applications, the business value derived from using the apps, and any barriers they face in adopting mobile apps. Thematic analysis was used to analyse the interview transcripts. The research was subject to the research ethics guidelines of the University of Cape Town and informed consent was obtained from all respondents.

4 Findings

4.1 Demographics of Respondents

This study examined ten micro-businesses in the service industry, with 1 to 7 employees each (Table 1). The majority of respondents were in management positions within the organization, with 60% of respondents stating that they held an ownership position and 20% holding a management position. Five SME owners said they hired part-time staff when there were new projects, so the number of employees fluctuated. The SMEs have been in business from one to twelve years.

Table 1. Micro-businesses sampled.

Number	Business type	Professional position	No. of employees	Business age (Yrs)
DC1	Hair salon	Owner	2	4
DC2	Nail Parlour	Senior Employee	1	1
DC3	Hair salon	Owner	3	5
DC4	Laundromat	Senior Employee	2	2
DC5	Laundromat	Manager	5	6
DC6	Hair salon	Owner	2	12
DC7	Laundromat	Owner	7	5
DC8	Nail Salon	Owner	3	3
DC9	Laundromat	Manager	3	2
DC010	Laundromat	Owner	2	1

4.2 Mobile Application Adoption

Mobile App adoption was driven by four major themes: namely ease of use, accessibility, sales/promotion and speed.

Ease of Use
SMEs stated that what largely influenced their adoption of mobile applications was their ease of use. This was reported by respondent DC6 who said: "Yah like apps that are easy to use, that I can just find my way around without needing my son to explain how to use it." The ability of an app to be simple to use was perceived as a key criterion influencing the adoption of mobile applications. Respondent DC3 confirmed the need for mobile applications to be easy to use by stating: "These apps I use like WhatsApp make it easy to talk to my customers. A lot of my customers find me through these apps and talk to me there." The ability to facilitate bookings was identified as a factor influencing the adoption of mobile applications, as it meant that they'd be able to conduct their services.

Accessibility
All SMEs stated that the ability to access and reach more customers played a large role when deciding which application to adopt: "For the current market, I'd say like, if you're looking at the youth of today, and everyone's on apps. So with that regard, I'd say that has helped a lot getting more people when they pick up the phone." [DC5]. The ability to reach more customers is also influenced by the popularity of the applications and what most customers are using. "So the main reason behind, like, using certain apps is just because to keep up with what customers are using, and what the world is doing." [DC6]. Using applications that allow businesses to access their customers enables them to also engage with new customers that find them via word of mouth. " like after I did a job nice you go you send me a friend, that happens via WhatsApp then I talk to them and plan when they come to me." [DC3]. In turn, accessibility brings them more clients thus more sales: "The main benefits? I just mentioned that you get a lot of customers, which brings me more money" [3].

Business Operations
SMEs noted that their adoption of mobile applications was significantly influenced by the tasks that they performed. These tasks were common tasks that entailed communications, sales and promotion. This was supported by DC6: "They help with saving time, you know, instead of waiting and do other things, with this kind of apps that is available you get things quicker, you know, the times that you waste waiting for people, you can use the apps, you can reach people in time, and you can do banking in time and again, you know." The ability to perform business tasks quickly and effectively was perceived as an important feature offered by mobile applications as SMEs could improve on their marketing strategies and reach more customers in one go. This is highlighted by DC5: "Process for deciding was speed because if you're trying to get a hold of, say 100 people in a short amount of time, then the WhatsApp application or messenger application on Facebook or SMS system will work a lot faster than having to pick up the phone and phone. Yeah, generally just speed." In addition two respondents, although less common, stated that they adopted mobile applications to facilitate payments, stating that: "Okay,

we mostly dealing with students, so they don't go out with cash most of the time, so they got to swipe and also to avoid being robbed. We don't like to keep to keep cash on us."

Competitive Pressure

Five SMEs implied that they adopted applications due to competitive pressure. They used apps to reach customers in the same way as their competitors: "Yeah there're a lot of laundry stores and we have similar customers, so we had no choice but to also be on socials just to keep up" [7]. This sentiment was also emphasized by DC8: "Everyone is on socials; we have to be there so we don't lose out customers to other stores."

4.3 Mobile Application Usage

Figure 3. shows the application usage by function. Social media use ranks highest although payments/collections and selling products/services rate relatively highly too. Organisation/scheduling, job management and information analysis is hardly used.

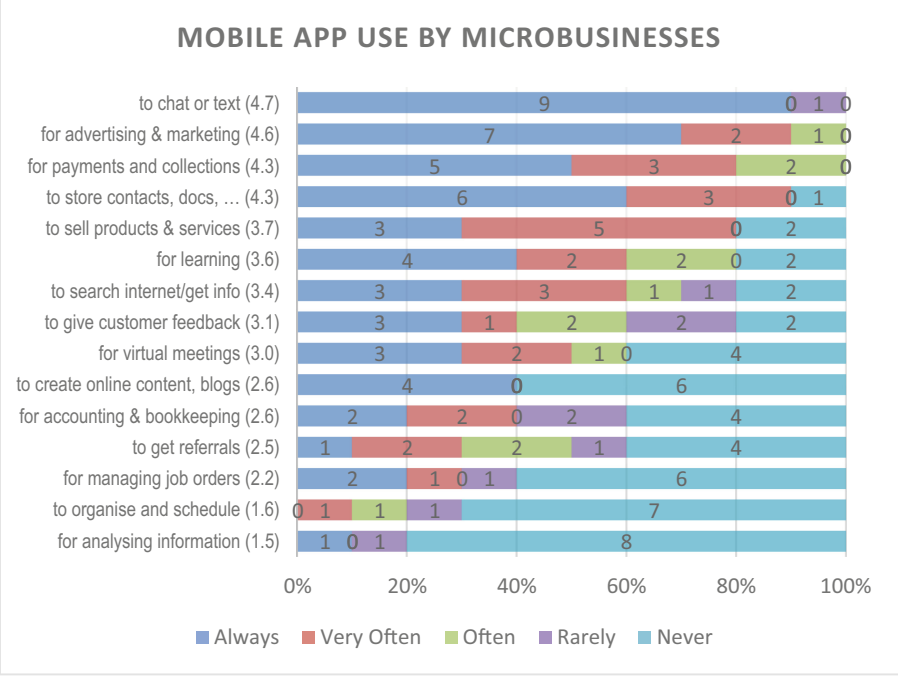


Fig. 3. Mobile app usage (data from fieldwork; mobile apps from [4].

The underutilisation may be due to a lack of contextually appropriate applications or a knowledge gap. It is possible to use the Excel mobile app for bookkeeping as a temporary fix, but it is not necessarily appropriate. However, that many SMEs reported using manual means to plan their schedules and conduct their bookkeeping: "I just like write it down. So that I know when the other a pick up tomorrow they know who's

coming in and who's paid." [DC4] and "No, no, I write it down or keep it in my head; when I use the phone, I can write it in my notes app." [DC3].

4.4 Business Value from Using Mobile Applications

This section presents and discusses findings related to the Business Value theme and its two sub-themes as presented in the M-Business Value component of the conceptual framework. The sub-themes are namely Impact on sales and internal operations.

Impact on Marketing and Sales.
Firms can benefit significantly from using mobile technologies and applications to enhance their sales- and marketing efforts marketing-related activities. Since more and more clients use their mobile devices, phones or other mobile devices to perform different types of tasks, using mobile applications has been recognized as a key channel in the interviews. All respondents emphasized the importance of mobile applications that they could have a positive impact on this organizational dimension. In fact, applications that are relevant to sales forces could increase sales and allow the products to be positioned more effectively on the shelf. For DC3, for example, Despite the fact that mobile applications are used by a variety of users (e.g. senior employees, managers and owners), they strive for the same goal: the improvement of client services through sales, marketing, and after-sales customer support. As our interviewee stated: "Even more people come in when I post online, […] more than when I didn't post".

Because the business uses online marketing, a broader audience can search for the services they need using their phones. Mobile apps can also improve customer service. Some respondents noted that mobile apps allow companies to: "respond faster to customers and stand out from other businesses that use the same services". Even if it does not provide a competitive advantage, some companies must offer some type of mobile functionality in certain contexts, such as banking.

Clearly, there are many benefits for clients: improved client experiences, motivated employees who are more client focused and the business is able to respond to client demands, enhancing client service e.g. mobile banking, for example. DC5 supports this capability: "We kind of had to use the Snapscans and Yocos, so we don't turn away customers who don't have cash and find it easier to do things electronically."

The use of mobile applications also allows for faster responses to client requests, which has significantly increased customer satisfaction. One interviewee cites texting as a useful tool in facilitating communication with clients. Another example is when the client makes an inquiry about availability in hopes to make a booking, that client can quickly receive a response. One of the informants reinforced this finding revealing that they send bulk messages to their clients if they have a promotion and reach 100 people in one go.

Impact on Internal Operations.
The internal dimension of business value derived from mobile applications indicates that businesses can reap substantial benefits from using mobile technologies and applications to execute internal business processes and support employee tasks. SMEs recognize that mobile applications are useful for managing internal operations. They found that

they were useful in improving internal operations, such as making client responses and updates timelier. In one instant, the respondent explained there was an improvement in communications and productivity service turnaround time, as the time to contact other branches or clients was cut down and time to troubleshoot cases of special laundry items being brought in, "Like when I need to refer to them sometimes call me or ask me, as far as in time need to come all the way to Observatory when there's another branch? Oh yeah, yes, I just refer them to the nearest branch, which is the one in town. And whatever they want to know, this saves us and client time and makes them happy". The increase in productivity increases employee satisfaction as they are able to respond more rapidly to clients.

The comments of respondents also indicate that mobile applications play an important role in facilitating communication among employees. They recognize that it is a good way to pass information among mobile workers, allowing workers to be more flexible at work, which is associated with their increased motivation. One of the interviewees (DC1) explained: "part-time worker comes in when she needs to, I message her when I need her help so she doesn't come in when she won't work".

SMEs interviewed agreed that mobile apps could make information more accessible and easier to analyse, especially when urgent and critical decisions need to be taken. DC1 supports this by saying: "Trends are always changing, and I always get clients wanting something I never heard of so I go to Pinterest or YouTube to learn it so I don't lose customer."

5 Discussion

The findings from this study have addressed the value that micro-businesses derive from the use of mobile applications by investigating the adoption of mobile apps and the factors that influenced the adoption as well as the reasons why mobile applications are being utilized. It was found that the main reasons behind utilizing applications Is to chat/text as well as marketing related activities, this pointed to the greatest value of mobile applications is derived in marketing and sales activities as well as for internal operations. The impact of mobile applications on marketing and sales activities and on internal operations, this aligns with the Business Value theme presented in M-Business Value component of the conceptual framework which speaks on the impact of value chain activities. The sections below will discuss findings further.

5.1 Mobile Application Adoption

The **ease of use** and **accessibility** of an application is essential in the adoption of an application to ensure that value is derived. This study recognises 2 types of business value, being impact on sales & operations and impact on internal operations as observed by [14]. [3] says technical skills are an issue amongst SME founders; hence applications that are easy to use and understand will be used by SMEs. As observed by [4] participants highlighted that that the adoption of mobile applications was heavily influenced by its ability to access their customers through namely chatting, customer feedback and referrals. This is due to respondents noting that the accessibility of customers via mobile

devices is crucial for the promotion of their business and communication flow so that information can be shared and feedback provided [1].

Participants also adopted mobile applications in order to remain competitive, they noted their competitors used them; thus, they also had to increase their customer base and sales. This finding is in accordance with [13] who noted in their findings that technology adoption is becoming more affordable as well as a necessity to survive in today's ultra-competitive business environment and [1] who said "small and medium scale enterprises cannot achieve competitiveness and remain profitable without proper adoption of technology at the right market levels".

5.2 Mobile App Usage

Benefits derived from mobile app adoption in SMEs which are communications, market solicitation and social cohesion according to [11]. It was found that mobile applications are frequently used for customer interactions for feedback and bookings as well promotions that's are being run via text/SMS or being run on social media applications due to their speed and efficiency. Applications that were used most frequently amongst SMEs were found to be Whatsapp, phonebook and Google. Whatsapp is used by all SMEs to communicate with customers and employees due to speed and cost-efficiency, it was the most popular application used and used across functions across customer feedback, referrals, chatting and promotions, this was anticipated given that 89% of South African internet users using it per month [6]. Social Media applications scored highly via their adoption and this can be explained by how a majority of respondents use mobile applications for their sales and marketing activities.

5.3 Business Value of Mobile Apps

The business value derived from mobile applications which stem from the m-business conceptual framework discussed above. The findings from the study acknowledge that mobile app usage plays an integral role in driving business value for SMEs as identified by [14]. Participants identified that **sales and marketing** were promoted through mobile app usage which afforded SMEs the ability to respond quicker to their clients demands which increase customer satisfaction and the ability to communicate and interact with clients through various channels which aided in relationship building. It was also noted that respondents felt that the impact on sales and marketing allowed for to remain competitive in their respective environments. Participants also identified the impact on internal operations was promoted through mobile app adoption, with which they identified to be client communications which entail bookings and feedback, service turnaround time and as well as payments through the use of banking applications. Respondents felt the impact of mobile application on internal operations improved their productivity and efficiency by simplifying their internal processes and increasing organisational flexibility which supports the findings of [2].

6 Conclusion

The first research question presented was: *What value do micro-businesses derive from the use of mobile applications?* All respondents stated that they adopted mobile applications to communicate more easily with the customer and to conduct sales and or promotions. Almost all the respondents always use their applications for chatting and texting, and most always use mobile applications for marketing and promotions. So they derive business value from the impact of the mobile application on sales and marketing as well as from internal operations. In addition, through the use of these applications, they were able to reach a broader audience and improve their customer services as they were able to respond faster to customers and in turn increase their sales. They were also able to improve productivity and efficiency by simplifying their internal processes such as bookings and internal communications and increase organisational flexibility by keeping up with trends.

The second research question was: *ow do South African service-oriented micro businesses use mobile applications to drive value?* Mobile applications promoted various value drivers and presented the usages that promoted business values. SMEs primarily derived business value from sales and marketing as well as from internal operations. The ability to derive business value was encouraged by two factors, namely the ability of an organisation to adopt a mobile application and to use an application to promote accessibility and business operations. Common apps to drive business value included chatting and texting, marketing and sales as well as payments and collections.

As a limitation, this study may not accurately represent the entire population of SMEs in South Africa due to a limited sample of 10 respondents being used for the study. Language was another limiting factor as some respondents did not have English as their native language, thus may have omitted information when answering questions and may have misinterpreted the question being posed.

This study found that SMEs were able to utilize mobile applications, realise critical value drivers, and derive business value from mobile applications. Future studies could explore what capabilities SME owners are not utilizing from their mobile applications and what the infrastructure gap is in order for them to effectively adopt mobile applications.

References

1. Eze, S.C., Chinedu-Eze, V.C., Bello, A.O., Inegbedion, H., Nwanji, T., Asamu, F.: Mobile marketing technology adoption in service SMEs: a multi-perspective framework. J. Sci. Technol. Policy Manage. **10**(3) (2019)
2. Barua, A., Kriebel, C.H., Mukhopadhyay, T.: Information technologies and business value: An analytic and empirical investigation. Info. Sys. Res. **6**(1), 3–23 (1995)
3. SME South Africa: SME Landscape Report (2020). Available: https://smesouthafrica.co.za/south-africas-smes-landscape-report. Last accessed 20 September 2022
4. Owoseni, A., Twinomurinzi, H.: Evaluating mobile app usage by service sector micro and small enterprises in Nigeria: an abductive approach. Info. Technol. Develop. **26**(4), 762–772 (2020)
5. Republic of South Africa: National Small Business Act 102. Government Gazette 27 (Nov 1996)

6. Kemp, S.: Digital 2021: South Africa. https://datareportal.com/reports/digital-2021-south-africa. Last accessed 20 September 2022
7. Agwa-Ejon, J.O.H.N., Mbohwa, C.H.: The impact of technology on small, micro and medium scale enterprises in Gauteng South Africa (2006)
8. Chaua, N.T., Denga, H.: Critical Determinants for Mobile Commerce Adoption in Vietnamese SMEs: A Conceptual Framework. Procedia: Computer Science **138**, 433–440 (2018). https://doi.org/10.1016/j.procs.2018.10.061
9. Kossai, M., Piget, P.: Adoption of information and communication technology and firm profitability: Empirical evidence from Tunisian SMEs. The J. High Technol. Manage. Res. **25**(1), (2014). https://doi.org/10.1016/j.hitech.2013.12.003
10. Research ICT Africa: Cheapest price for 1GB basket in Africa by country. ResearchICTAfrica.net. Last accessed 20 September 2022
11. Melchioly, S.R., Sæbø, Ø.: ICTs and development: nature of mobile phones usage for SMEs economic development - an exploratory study in morogoro, tanzania. Int. Feder. Info. Proc. (IFIP) Techni. Comm. **9**(4) (2015)
12. Chew, H.E.: The Effects of Mobile Phones in Social and Economic Development: The Case of Female Micro entrepreneurs in Chennai. Michigan State University (2012)
13. Khaskheli, A., Jun, Y., Bhuiyan, M.A.: M-commerce and mobile apps: opportunities for SMEs inDeveloping Countries. J. Int. Bus. Res. Mark. **2**(2) (2017)
14. Picoto, W.N., Bélanger, F., Palma-dos-Reis, A.: M-business organizational benefits and value: a qualitative study. J. Organi. Comp. Elec. Commer. **23**(4), 287–324 (2013). https://doi.org/10.1080/10919392.2013.837789
15. Rahman, M.S.: The advantages and disadvantages of using qualitative and quantitative approaches and methods in language "Testing and Assessment" research: a literature review. J. Edu. Learn. **6**(1), 102 (2016). https://doi.org/10.5539/jel.v6n1p102
16. Fossey, E., Harvey, C., Mcdermott, F., Davidson, L.: Understanding and evaluating qualitative research. Australian and New Zealand Journal of Psychiatry **36**(6), 717–732 (2002). https://doi.org/10.1046/j.1440-1614.2002.01100.x
17. Ritchie, J., Lewis, J., Ormston, R., Nicholls, C.M.: Qualitative Research Practice: A Guide for Social Science Students and Researchers. Sage Publications, London (2013)

The Importance of Knowledge for Influencing Citizens' Attitudes and Practices of Water Sustainability: The Case of the Eastern Cape

Dumani Kunjuzwa, Brenda M. Scholtz(✉), and Ifeoluwapo Fashoro

Nelson Mandela University, Gqeberha, South Africa
{dumani.kunjuzwa,Brenda.Scholtz,Ife.Fashoro}@mandela.ac.za

Abstract. Water shortages in developing economies exacerbate the challenges of socio-economic growth and the development of Smart Cities. Information flow through intelligent technologies is a crucial element of Smart Cities that has the potential to improve the knowledge of citizens regarding water sustainability. The role of knowledge in citizens' attitudes and practices of water sustainability is under-researched. This study used a positivist paradigm and collected quantitative data using an online survey to examine the relationships between knowledge, attitudes, and practices of citizens regarding water sustainability. The respondents were citizens in the Nelson Mandela Bay and Buffalo City Metro municipalities in the Eastern Cape Province of South Africa. The findings revealed that knowledge is closely correlated to the attitudes and practices of citizens regarding water sustainability. The contribution is the Knowledge, Attitude, and Practices (KAP) Model for Water Sustainability, which provides a deeper understanding of the KAP profile of citizens. It can be used by both researchers and practitioners for designing campaigns and technological solutions that aim to improve citizens' knowledge, attitudes, and practices toward water sustainability.

Keywords: Attitudes · Awareness · Knowledge · Practices · Water sustainability

1 Introduction

The rate at which water demand has increased globally in the past century is more than twice that of population growth, and more and more regions, particularly arid ones, are approaching the point at which water services cannot be provided in a sustainable manner [1]. South Africa has been struggling with limited water resources and a rapidly worsening drought situation over the past four to six years, especially in the provinces of the Eastern Cape, Northern Cape, and the Western Cape [2, 3]. Both the City of Cape Town and the Nelson Mandela Bay Municipality (NMBM) have been tackling severe water shortages and the arrival of "*Day Zero*" has been threatened. Although it is normal for water consumption to rise as the population and economy expand, the rate at which it is doing so is much faster, with a negative impact on water sustainability and the South African economy [4]. There is a need for research addressing the relative lack of knowledge on water resource issues, and thus the low level of citizens' awareness of

P. Ndayizigamiye et al. (Eds.): IDIA 2022, CCIS 1774, pp. 101–117, 2023.
https://doi.org/10.1007/978-3-031-28472-4_7

water shortages [5]. Despite the research conducted in the field of solutions for water sustainability and water knowledge, the majority of studies adopt a Euro-centric view and do not consider indigenous knowledge, which is unique, trustworthy, and confined to a particular culture. There is also a gap in research that provides a detailed understanding of the various sub-factors of knowledge in the context of developing economies, or the relationships between the sub-factors. This paper addresses these gaps and provides deeper insight into the relationships between all three factors of knowledge, attitude and practices as well as between the various sub-factors of each of these. The sub-factors of knowledge investigated include those such as that related to the local dams, water filtering approaches, conservation methods, indigenous knowledge, sources of knowledge and knowledge sharing methods used.

The main research question addressed in this paper is: *"What is the relationship between citizens' knowledge of and their attitudes and practices toward water sustainability?"* The two sites of the Buffalo City Municipality (BCM) and the NMBM in the Eastern Cape were used as cases to investigate the phenomenon of the water shortage problem, particularly to understand the knowledge, attitude and practices of citizens of these municipalities. An online survey approach was adopted for data collection in both municipalities. The contribution of the paper is a model of the relationships between the explicit knowledge of water sustainability of citizens and their attitudes and practices of water sustainability. An understanding of these factors can assist the designers of initiatives and technological solutions in developing awareness and knowledge-sharing campaigns that can ultimately add value to the social and economic growth of a country. The structure of the paper is as follows: the next section presents the theoretical context and the research design. In Sect. 3 the findings are presented and then these are further discussed in Sect. 4. Section 5 concludes this paper and proposes recommendations for future research.

2 Theoretical Context and Research Design

2.1 Literature on Solutions for Water Sustainability

Increasing urbanisation has been widely acknowledged as one of the factors that pose a threat to urban water sustainability [6]. Water sustainability is an issue highlighted in Goal 6 of the United Nations' Sustainable Development Goals (SDGs), which aims to ensure access to water and sanitation for all. Urbanisation on the other hand has significant effects on production and consumption, making it a crucial component of efforts to advance sustainability and a greener economy [7]. Urban planning needs to improve human capital development, enlightening inhabitants with pertinent information about the city's resources, such as water, healthy living standards, and environmental concerns, to achieve sustainability and reach new frontiers [8, 9]. Water's importance in all its functions should therefore be considered in this planning. One response to solve the urbanisation problems and address the challenge of limited resources such as those of water and energy has been the development of sustainable or Smart City solutions [10]. Sustainability, technology, and information flow are three important characteristics of these solutions [11, 12]. Social computing and other technologies can be used to instrumentally increase the awareness of citizens on the importance of water sustainability

[13]. This argument is supported in [14] and [15], where it was reported that the use of technologies such as social platforms can be used for communication and information sharing to make communities smart, in terms of empowering citizens' awareness with relevant knowledge and information about limited natural resource issues such as water. The findings in [13] revealed that increasing awareness of water issues can positively influence the attitudes of citizens towards water sustainability. For a city to be considered sustainable, certain key elements must be present. These elements include renewable energy, smart water management, sustainable transportation, sustainable buildings, waste management, and sustainable education [16]. Some research has addressed water sustainability by providing smart metering and monitoring systems for water usage and leakage detection [17], or by improving water infrastructure coupled with intelligent systems [16]. Sustainable education and water management with micro-education systems were successfully proposed in [18]. However, these solutions may not be feasible in developing economies due to the associated costs. Lower-cost solutions can be provided by using the opportunities provided by the information explosion and Big Data phenomena [19]. These solutions can assist governments in developing economies to promote effective service delivery to their citizens. One example solution is the use of social media by citizens to express their opinions and to keep governments accountable [20]. Social media can also overcome challenges related to e-government such as low accessibility and the lack of citizen e-participation [21]. Knowledge sharing on social media has also been reported by [22] for improving people's awareness and education on environmental issues. However, most studies on knowledge have a westernised view and do not consider indigenous knowledge, which is unique, trustworthy, and confined to a particular culture [23]. In the context of South Africa, which is a developing economy in Africa, the study of [24] examined urban households' attitudes and perceptions for crafting potential intervention strategies for dealing with future drought impacts in the Makhanda region of the Eastern Cape Province. In their study it was revealed that technical solutions are important but are not sufficient for addressing wasteful practices. The focus in [24] was on practices, barriers and challenges of water resources in one city. It did not investigate the knowledge levels of respondents and its influence on practices or attitudes. In [25], water supply challenges in urban and rural areas of the Eastern Cape were highlighted, whilst in the same province water scarcity and food security challenges in the Ngqeleni locality were reported in [26].

2.2 Theoretical Framework

The research was undergirded by the theory of Knowledge, Attitude, and Practices (KAP), which is a theory proposed for health behaviour change [27], where human behaviour change is classified into three successive events: the acquisition of knowledge, attitude generation, and the formation of behaviour or practices. Knowledge has become a crucial asset to organisations and it provides them with a sustainable competitive advantage [28]. Knowledge sharing is a key element of knowledge management and it includes activities for information dissemination, collaboration with others, problem solving, innovation, and policy implementation. The benefits of knowledge and knowledge sharing can be realised by governments as well as corporate organisations. For

example, in [29] it was revealed that knowledge sharing using ICTs and smart technologies can contribute to the growing self-awareness of citizens. It has also been argued that, conversely, a lack of knowledge on water resource issues contributes to the low level of citizens' awareness of water shortages [5]. There are two main types of knowledge, i.e. explicit and tacit knowledge [30]. Explicit knowledge is systematic knowledge and is mostly in written form, for example books and reports. Knowledge that is explicit can easily be codified, stored, and transferred across time and space independent of individuals. On the other hand, tacit knowledge is hidden, unspoken and resides in the minds of human beings. It is often deeply rooted in actions, procedures, routines, commitment, ideals, values, and emotions. Knowledge that has not been considered in the management of information in Africa is indigenous knowledge, in spite of the fact that we live in a knowledge management era and that it can play a crucial role in the development of a sustainable society [31]. This knowledge is orally transmitted from generation to generation.

Environmental knowledge sharing has been shown to improve the awareness of citizens [15]. In [32] it was found that citizens' lack of awareness of water shortages contributed to irresponsible water usage practices and behaviour, thus worsening the shortage situation. However, to encourage participation in virtual knowledge-sharing communities of practice, citizens need to be motivated and barriers need to be overcome [33]. In [34] it was reported that perceptions are the way in which water is perceived by citizens. For example, that it is a human right or a gift from God, a public resource, and/or a natural resource. These perceptions can influence a citizen's attitude towards water resource issues. Studies such as [35] and [36] confirm this and showed that there is a relationship between water use behaviour or practices and attitudes and perceptions. An example of attitude is the willingness to pay for water. The early study by [37] argues that poor urban residential communities are willing to pay a reasonable price for water usage. However, in both [38] and in [39] it was reported that poor urban residential communities are unwilling to pay.

Socio-economic factors may also influence knowledge and practices of water sustainability. For example, in [40] it is argued that people with a high level of education tend to know more about environmental issues. In [41] it is reported that households with lower incomes are more engaged in water conservation and use less water than those with a high-income. Similarly in [42] the findings revealed that financial status is influential on water usage within the household since "*as a household's income rises members of the household become less sensitive to water billing*".

2.3 Context and Design of Survey

The Eastern Cape province in South Africa has been battling with water shortages since 2015 and is one of the poorest provinces in the country with the highest unemployment rate of 44% [43]. In this province, the NMB municipality saw significant drops in dam levels during 2022. The municipality has now reached the situation where the main water supply dam, the Impofu Dam, has dropped too low for any water to be extracted. A nearby municipality, the BCM, is also experiencing severe water shortages. Citizens from these two municipalities (NMB and BCM) were used as the population researched in this paper. A positivist paradigm was adopted and quantitative data was collected using

an online questionnaire. The required ethics approval was received from the university. The questionnaire was assessed by the university's statistical consultant to get expert feedback and based on this feedback the questionnaire was subsequently improved. A pilot test with five respondents was conducted to get feedback regarding the design of the questionnaire, the clarity of the questions, and the relevance of questions included. The first section of the questionnaire related to social demographics such as gender, age, and education. The remaining three sections were designed based on the three factors of KAP theory, Knowledge (K), Attitude (A) and Practices (P). The Knowledge (K) factor is related to explicit knowledge of water obtained from reputable websites such as [44] and [45], and literature on indigenous knowledge (for example, in [46]), amongst others. This factor was further classified into the following six sub-factors (Fig. 1):

- Knowledge Dams (K-Dams) – related to dam supply of water in their respective municipalities.
- Knowledge Water General (K-General) – includes general knowledge of water; for example: a) how long can a healthy person live without water? and b) what portion of the world's population may face water shortages by the year 2025?
- Knowledge Water Filtering (K-Filtering) – consisted of knowledge of thermal distillation, water filtration, solar disinfection, and chlorination.
- Knowledge Water Conservation (K-Conservation) – included knowledge of water harvesting, greywater systems and knowledge of planting a water-wise garden.
- Knowledge Information Sources (K-Sources) – related to where knowledge is obtained from, e.g., social media, friends and family, and other media.
- Knowledge Indigenous Indicators (K-Indigenous) – considers knowledge from generation to generation for sustainable living under ecosystem conditions that include weather conditions. For example, some cultures use cloud colour, river flow or the moon shape as indicators for predicting rainy weather [46].

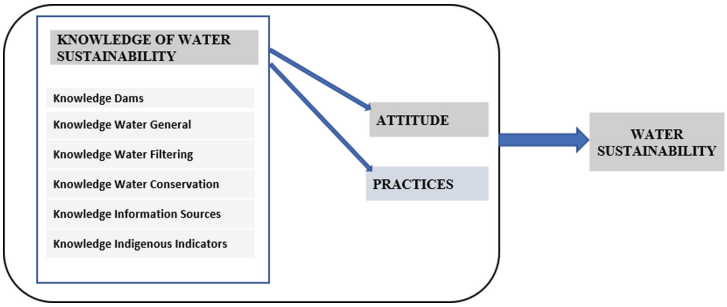


Fig. 1. KAP for water sustainability model

Each of the six Knowledge sub-factors had several related items where respondents were required to select the correct options associated with that aspect of knowledge of water sustainability. The Attitude factor had eight items related to attitude that respondents had to rank on a scale from Strongly Disagree to Strongly Agree. For example,

"To save water, a person should shower instead of taking a bath". Other examples of attitude related to buying water-saving appliances and reporting leakages and burst pipes. In the Practices section, there were 10 items that respondents had to rate based on the frequency with which they did the water sustainability practice. The five possible responses were Never, Sometimes, Regularly, Often or Always. One example was *"Turn off the tap when brushing your teeth"*.

Proportionate stratified random sampling was adopted to ensure that the sample represented specific sub-groups, in this case, the citizens from the NMB municipality and the BCM. The inclusion criteria for these participants were therefore: a) they are citizens of either the NMB or BCM b) they are from a South African ethnic groups and c) they had access to municipal water or through shared taps. An advert for the survey was compiled that included a short overview of the objective of the survey and a link to the online questionnaire. The adverts were placed on two Facebook pages, the researchers' university department's Facebook page and the Project ArCc's Facebook page (using #ArCcSavesWater). The advert was also placed on the university's environmental awareness webpage, which used the #savewater. The Facebook accounts of the participants were not checked. Data collection took place between September and November 2021. The collected data were anonymised and statistically analysed using descriptive statistics, inferential statistics, and Analysis of Variance (ANOVA) calculations.

3 Results

3.1 Socio-economic Profile of the Sample

The profile of the respondents is summarised in Table 1. A large portion (59%) of the sample were young (between 18 and 30 years of age) whilst 24% of the respondents were between the ages of 30 and 39 years. The majority (94%) of the sample were not members of any water awareness groups. Nearly three-quarters (73%) of were from the NMB municipality (n = 125) and the remaining 27% were from the BCM (n = 47). There was a fairly even split between males and females. The sample was therefore slightly skewed to the NMB municipality and younger respondents. The largest group (47%) of the sample were employed and self-employed respondents, 12% were unemployed, retired or pensioners and 42% were students. The largest group (24%) of the sample respondents' household monthly income was between R20 000 and R39 999.

3.2 Descriptive Statistics – Knowledge, Attitudes, and Practices Items

For Knowledge Dams, the sample from the NMB municipality scored higher (with an average correctness score of 59%) compared to the BCM sample (with an average correctness score of 43%). For Knowledge Water General, for example, less than half of the total (37%) of the sample knew how long a healthy person can live without water, which is three to seven days. Less than a third (28%) of the respondents gave the correct answer that two-thirds of the world's population may face water shortages by the year 2025. The frequency distributions of sources for acquiring information about water where classified into three categories: Low – which relates to never and rarely,

Table 1. Participant profile

Gender	**n**	**%**	**Education Level**	**n**	**%**
Male	75	44%	Grade 12 or less	63	37%
Female	95	55%	Certificate/Diploma	30	17%
Other	2	1%	Degree	44	26%
Total	**170**	**100%**	Post-Graduate Degree	35	20%
Municipality	**n**	**%**	**Total**	**172**	**100%**
NMB	125	73%	**Employment Status**	**n**	**%**
BCM	47	27%	Employed/Self-Employed	80	47%
Total	**172**	**100%**	Unemployed/Retired/Pensioner	20	12%
Age	**n**	**%**	Student	72	42%
18 - 29	101	59%	**Total**	**172**	**100%**
30 - 39	41	24%	**Household Income (Month)**	**n**	**%**
40+	30	17%	R 0 - R 1 999	30	17%
Total	**172**	**100%**	R 2 000 - R 9 999	34	20%
Member - Group	n	%	R 10 000 - R 19 999	36	21%
Yes	11	6%	R 20 000 - R 39 999	41	24%
No	161	94%	R 40 000+	31	18%
Total	**172**	**100%**	**Total**	**172**	**100%**

Medium – which relates to sometimes, and High – which relates to often and always. The three sources that were rated as High were other media (43%), family and friends (42%), and social media (40%). None were rated Medium and the rest were rated Low. Seven of the eight items for Attitude were scored positively by respondents as Agree or Strongly Agree. The only item that was overall ranked as Strongly Disagree or Disagree was *"Water is a free natural resource, no payment should be associated with it."*, which implies that the majority of the sample believed they should be required to pay for water usage. The positive attitude of the sample was also indicated in the frequency distribution analysis for the combined sample of both BCM and NBMM. Out of the 10 practice items only four were reported to be good practices, and the rest were deemed poor practices.

The frequency distributions are listed in Table 2, which were examined using three rankings where a frequency count (n) of 0 to 39 was categorised as Lower, 40 to 60 was ranked as Middle and 61 to 100 was ranked as Higher. The largest percentage for each of the six knowledge sub-factors and for each of the three KAP factors are shaded in grey. Attitude was the only factor to have the largest percentage of respondents scoring in the Higher category (with 81%). None of the Knowledge scores were in the Higher category. Only two Knowledge scores were in the Middle category, namely Knowledge Dams (59%) and Information Sources (44%). The lowest Knowledge sub-factor that had the highest percentage in the Lower category was Water General with 66%; thus, it was the lowest scoring factor. Following closely behind was the sub-factor of Indigenous Knowledge with 65% of the sample in the Lower category. The largest samples for both overall Knowledge (48%) and Practices (44%) were in the Lower category.

Table 2. Frequency distributions: KAP scores (n = 172)

	Lower 0 - 39		Middle 40 - 60		Higher 61 - 100	
Knowledge Dams	22	13%	101	59%	49	28%
Knowledge Water General	113	66%	53	31%	6	3%
Knowledge Water Filtering	98	57%	54	31%	20	12%
Knowledge Water Conservation	72	42%	56	33%	44	26%
Knowledge Information Sources	68	40%	76	44%	28	16%
Knowledge Indigenous	112	65%	37	22%	23	13%
Knowledge	83	48%	75	44%	14	8%
Attitude	5	3%	27	16%	140	81%
Practices	75	44%	49	28%	48	28%

3.3 One-Sample t-tests and Inferential Ranking for the KAP Scores

To compare and rank the nine KAP factors, inferential ranking was conducted where variables were ranked using matched-pair t-tests (statistical significance) and Cohen's d coefficient (practical significance) as follows: a) Cohen's *d* coefficient is used for practical significance in a one-sample t-test and b) it is used to represent the extent of differences between two or more groups [47]. The details of the statistics for obtaining the inferential ranking of the KAP scores are provided in Appendix A, Table A1. The comparison of scores revealed that the following differences were significant (statistically and practically):

- Attitude and Knowledge Dams (M = 17.73 and d = 0.80),
- Knowledge Dams and Practices (M = 8.55 and d = 8.55),
- Practices and Knowledge (M = 5.87 and d = 0.28), and
- Knowledge and Knowledge Water Filtering (M = 6.82 and d = 0.41).

The one-sample t-test classification and inferential rankings of the KAP scores are reported in Table 3, where the intervals for the Cohen's d value are ranked as: Very Low: < 20.00; Lower: 20.00 to 39.99; Middle: 40.00 to 60.00; Higher: 60.01 to 80.00 and Very High: > 80.00. Shaded rows are both statistically and practically significant. According to these rankings Attitude (M = 72.37) was ranked the highest of all the factors (Higher category). It had a significantly larger population mean than Knowledge Dams (M = 54.65), which ranked second (Middle category).

Practices (M = 46.09), Knowledge Water Conservation (M = 45.94), and Knowledge Sources of Information (M = 43.10) all occupied the third inferential ranking group with a Middle category. They were in the same group since there was no significant differences between their sample means. Knowledge (M = 40.22) occupied the fourth group position with a Lower to Middle category. Knowledge Water Filtering (M = 33.40), Water Indigenous (M = 32.78) and Knowledge Water General (M = 31.05) occupied the fifth and lowest position of the inferential ranking groups in the Lower category.

Table 3. Inferential Ranking of KAP Scores (n = 172)

Factors	Descriptive Statistics		One-Sample t-test Classification					Inferential Ranking	
	Mean (M)	S.D.	H_1	t-value	p-value	Cohen's d	Category	Rank	Sig. Group
Attitude	72.37	14.70	M ≠ 80.00	-6.81	<0.0005	0.52	Higher	1	1
K-Dams	54.65	18.90	M ≠ 60.00	-3.72	<0.0005	0.28	Middle	2	2
Practices	46.09	25.38	M ≠ 40.00	3.15	0.002	0.24	Middle	3	3
K-Conservation	45.94	26.17	M ≠ 40.00	2.97	0.003	0.23	Middle	4	3
K-Sources	43.10	19.17	M ≠ 40.00	2.12	0.035	0.16	Middle	5	3
Knowledge	40.22	13.53	M ≠ 40.00	0.21	0.831	n/a	Lower to Middle	6	4
K-Filtering	33.40	25.01	M ≠ 40.00	-3.46	0.001	0.26	Lower	7	5
K-Indigenous	32.78	22.60	M ≠ 40.00	-4.19	<0.0005	0.32	Lower	8	5
K-General	31.05	16.33	M ≠ 40.00	-7.19	<0.0005	0.55	Lower	9	5

3.4 Relationships Between KAP Scores

To determine the correlations between the nine factors, the Pearson correlation coefficient r was calculated; this was considered statistically significant at the 0.05 level for n = 172 if $|r| >= 0.150$, and practically significant, regardless of the sample size, if $|r| >= 0.300$. It was thus significant (both statistically and practically) if $|r| >= 0.300$ [48]. Statistical but not practically significant correlations are shown in black bold in Table 4. Significant correlations, both statistically and practically, ($r > 0.300$) are shown in red and existed for all three main factors as follows: Knowledge and Attitude (r = 0.366), Knowledge and Practices (r = 0.571), and Attitude and Practices (r = 0.385).

The sub-factors of Knowledge were mostly correlated with each other and with Knowledge itself, which is to be expected as they were designed as sub-factors of knowledges. This finding also provides additional validity to the research design. For brevity, in this paper we only highlight the significant correlations between the three KAP factors and between the sub-factors of Knowledge, Attitude and Practices. These key significant correlations were between: Knowledge General and Attitude (r = 0.310); Attitude and Knowledge Conservation (0.317); Knowledge Filtering and Practices (r = 0.366); Practices and Knowledge Conservation (r = 0.366) and Practices and Knowledge Indigenous (r = 0.489).

3.5 Analysis of Variance (ANOVA) Results

ANOVA is used to determine whether there are any statistically significant differences between the means of two or more independent and unrelated groups or factors [49]. For each significant result, a post-hoc test was conducted to determine which specific group differed from the others. Only significant results are reported in Table A2 (Appendix A), which shows a statistical and practical significance for Knowledge Dams in: Age

Table 4. Correlations between KAP Scores (n = 172)

	K - Dams	K - General	K-Filtering	K-Conservation	K-Sources	K-Indigenous	Knowledge	Attitude	Practices
K-Dams	-	***0.375***	**-0.159**	0.040	0.040	**-0.168**	**0.235**	**0.152**	**0.160**
K-General	***0.375***	-	0.076	**0.207**	0.078	0.069	***0.417***	***0.310***	**0.189**
K-Filtering	**-0.159**	0.076	-	***0.666***	***0.455***	***0.607***	***0.779***	**0.169**	***0.366***
K-Conservation	0.040	**0.207**	***0.666***	-	***0.396***	***0.535***	***0.823***	***0.317***	***0.566***
K-Sources	0.040	0.078	***0.455***	***0.396***	-	***0.445***	***0.652***	**0.176**	**0.271**
K-Indigenous	**-0.168**	0.069	***0.607***	***0.535***	***0.445***	-	***0.719***	**0.260**	***0.489***
Knowledge	**0.235**	***0.417***	***0.779***	***0.823***	***0.652***	***0.719***	-	***0.366***	***0.571***
Attitude	**0.152**	***0.310***	**0.169**	***0.317***	**0.176**	**0.260**	***0.366***	-	***0.385***
Practices	**0.160**	**0.189**	***0.366***	***0.566***	**0.271**	***0.489***	***0.571***	***0.385***	-

($p = 0.001$); Member of Water Awareness Group ($p = 0.050$); and Municipality ($p < 0.0005$).

The post-hoc test for Knowledge Dams revealed a statistical and practical significance ($p = 0.000$; $d = 1.09$) is evident in the age group of 18 to 29 years ($M = 50.11$) compared to the age group of 40 + years ($M = 69.11$), which was higher than the younger age group. The age group of 40 + years was practically and statistically significant ($p = 0.001$; $d = 0.78$) when compared with the age group of 30 to 39 years. The age group of 40 + years had a higher mean ($M = 69.11$), suggesting that the age group of 40 + years had a higher knowledge of dams than all other age groups. Member of Water Awareness Group was statistically and practically significant ($p = 0.050$; $d = 0.084$) among members who belonged to awareness groups and those who did not belong to any awareness groups. Members of awareness groups had a higher mean ($M = 68.80$) compared to those who were not members of any awareness groups ($M = 53.40$). The respondents from the NMB municipality showed a significantly higher knowledge of dams ($M = 58.82$) compared to the BCM sample ($M = 42.49$).

For Knowledge Water General, there was a statistical and practical significance for Age ($p = 0.038$) and Education Level ($p = 0.009$). The post-hoc results highlighted that the age group of 30 to 39 years had a higher mean ($M = 25.85$) than the older age group ($M = 32.14$) for Knowledge Water General. Those who had an education of Grade 12 or less had a lower mean ($M = 28.39$) compared to those with higher education (Degrees) ($M = 38.14$). For Knowledge Water Filtering, there was a statistical and practical significance in Gender ($p = 0.049$; $d = 0.22$) and Member of Water Awareness Group ($p = 0.008$; d

= 0.58). Males had more knowledge related to Water Filtering (M = 36.47) compared to Females (M = 30.95). Members of awareness groups (M = 47.00) scored higher than those respondents who were not members of any awareness groups (M = 32.53). For Knowledge Water Conservation, there was a statistical and practical significance in Age (p = 0.039) and in Members of Water Awareness Group (p = 0.039; d = 0.64). The post-hoc results showed that the older age group of 40 + years had a higher knowledge related to Water Conservation (M = 51.39) compared to the young age group of 18 to 29 years (M = 46.27) and the middle age group of 30 to 39 years (M = 42.24). The results showed that the members of the awareness groups were mostly in the younger age group of 18 to 29 years (M = 61.90) compared to the older age group of 40 + years (M = 45.16). This result indicated that the members of awareness groups had more knowledge related to water conservation.

The ANOVA results for Attitude showed that of all the factors only age had a statistical and practical significance (p = 0.008). The post-hoc results showed that the age group of 40 + years had a higher significance (M = 79.79) compared to the young age group of 18 to 29 years (M = 71.69) and the middle age group of 30 to 39 years (M = 69.22). This result confirms that the older age group had a more positive attitude towards water than the other age groups.

For Practices (related to water sustainability), the three factors that were statistically significance were Age (p = 0.001), Member of Awareness Group (p = 0.047; d = 0.66), and Municipality (p = 0.038; d = 0.30). The older age group (40 + years) had good practices (M = 59.89) towards water (e.g. usage and/or conservation) compared to the young age groups of 18 to 29 years (M = 43.24) and the middle group of 30 to 39 years (M = 44.54). The results highlighted that the respondents who were members of the awareness groups also had better water practices (M = 62.00) compared to those respondents who were not members of any water awareness groups (M = 45.31). The respondents from the BCM had better water practices (M = 51.70) compared to the NMB municipality respondents (M = 44.23).

4 Discussion of Findings

The largest frequency scores for both overall Knowledge and Practices were in the low category while Attitude was the only factor that scored highly according to frequency scores. For the inferential rankings, Attitude was also ranked the highest of all nine factors, with a significantly larger population mean than Knowledge Dams (the second highest) whereas Knowledge of water in general was the lowest ranked of all the factors. Most of the respondents (NMB municipality and BCM combined) had a reasonable (middle) knowledge of dams as their suppliers for water within their municipalities as well as in water conservation mechanisms. However, the general water knowledge of these groups was low and even lower for water filtration knowledge. The lack of knowledge in general water knowledge and filtering could suggest that an intervention strategy is needed to address these areas of knowledge of citizens in the two municipalities in the Eastern Cape. The correlation results showed significant correlations between all three of the main factors Knowledge, Attitude and Practices, thereby confirming the adoption of the KAP model to the context of water sustainability, as shown in Fig. 2.

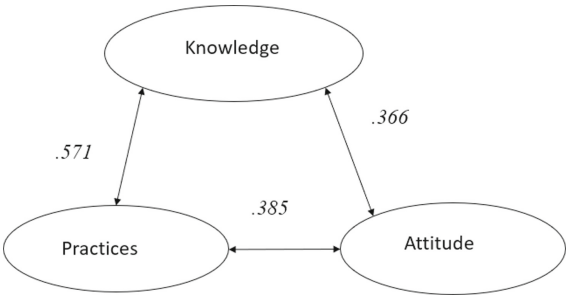


Fig. 2. Model of KAP for water sustainability

The significant social demographics were age, education level, gender, municipality, and member (or not) of water awareness groups. Other demographics were not found to be statistically significant, namely employment status and household monthly income. The income finding contrasts with that of [42] and [41]. Although financial status was not significant in the survey findings, it could be due to the limitation of the context of the sample being from only two municipalities. Since South Africa is a growing economy and the Eastern Cape is a poor province, the sample can be referred to as the urban poor.

The older age group (40 + years) was the most well-informed with their knowledge of dams (as sources of water supply) in their municipalities. Members of awareness groups had more knowledge of dams than those respondents who were not. This evidence could be due to the fact that the knowledge related to water resources is shared within their awareness groups (social connections). Respondents with degrees had a higher general knowledge of water than those with certificates or Grade 12 or less. This finding therefore supports that of [40] highlighting that more educated citizens have a higher knowledge of water sustainability. Male respondents were found to have more knowledge of water filtering methods than the female respondents. This correlation was not mentioned in any of the reviewed literature.

The findings for Attitude revealed a positive attitude for all items used as measurements thereof. Respondents said they were willing to pay for water usage to enhance socio-economic development, which agrees with the results of [37], but contradicts the findings of [38]. One possible reason for this unexpected result could due to the sample's demographics. The older age group (40 + years) had a more positive attitude towards water than the younger respondents. It could be argued, with the supporting evidence in [50], that older citizens value water as a natural resource that is embedded within the cultural and religious values of societies.

Practices ranked the lowest of all three main factors, implying that this area needs the most attention. Since knowledge and attitude are correlated to practices stakeholders who are planning projects to improve water sustainability through awareness or knowledge campaigns should focus on encouraging good practices. The respondents of the younger age groups had lower levels or practices of water than the older respondents. Some good practices were also reported in the survey findings; however, they were outweighed by bad practices. The results reported that most respondents do not share water-saving tips with others within their communities, which could be due to the fact that there is a lack of

social community networks, where communities can share knowledge relevant to their community issues.

5 Conclusions and Recommendations

This paper answered the research question and statistically confirmed relationships between knowledge, attitude and practices by citizens for water sustainability, thus validating the KAP for Water Sustainability model. An investigation of the level of knowledge, attitudes, and practices on water resource issues resulted in the conclusion that the citizens in the two municipalities of the Eastern Cape have a lower level of knowledge than they do for practices and attitude. Since the sample had a low level of knowledge about their water resource issues, it can be argued that knowledge sharing campaigns are required in these areas. On the other hand, the sample had a fairly positive attitude towards water resource issues. This finding is slightly unexpected since other literature and the correlation findings in this sample imply that a higher attitude of citizens results in higher levels of knowledge and vice versa. There could be a number of reasons for the lower level of knowledge and an investigation of the detailed correlation results of the different types or sub-factors of knowledge could provide more understanding of this. For example, some of the sub-factors of knowledge rated much higher than others. Another reason could be because of the socio-economic factors, where for example the older citizens in the sample had higher levels of knowledge and attitude than the younger citizens. For practices, there were mixed responses, with several aspects of practices that should be improved.

The significant socio-economic demographics were age, education level, gender, municipality, and member (or not) of water awareness groups. Other social demographics were not found to be statistically significant, namely employment status and household monthly income. The income finding contrasts with that of [42] and [41]. Although financial status was not significant in the survey findings, it could be due to the limitation of the context of the sample being from only two municipalities in a poor province. Another limitation of the study was that it was scoped to only explicit knowledge and did not investigate tacit knowledge factors.

The contribution of the research is a deeper understanding of citizens and their knowledge, attitude, and practices toward water sustainability. The proposed model can be used by designers of awareness campaigns and by those designing technological solutions to improve the KAP areas of water sustainability. SDG 6 acknowledges that South Africa is a water-scarce country and that water scarcity is likely to increase as a consequence of climate change. This paper contributes to addressing water scarcity (water shortages) by presenting a model of relationships between attitude, practices, and knowledge of water sustainability ahead of 2030. The opportunity cost of this research is that a lack of awareness of water sustainability can lead to reduced access to clean/fresh water, which will mean a poorer social standard of living for citizens that can lead to poor economic growth in the province of the Eastern Cape. The study was limited to the context of only two municipalities in the Eastern Cape. Future research should investigate the adoption of the KAP model in other municipalities or cities in the Eastern Cape and in other South African provinces and African countries. In addition, the relationship between tacit knowledge factors and practices should also be investigated in future studies.

Acknowledgements. Acknowledgements are due to the National Research Fund (NRF) and the Nelson Mandela University Research department for the funding of this research.

Appendix A

Table A1. Inferential Ranking Statistics – KAP Scores (n = 172)

	Difference		Inferential Statistics		
Scores Compared	**Mean (M)**	**S.D**	**t-value**	**p-value**	**Cohen's d**
Attitude & K-Dams	17.73	22.11	10.52	<0.0005	0.80
K-Dams & Practices	8.55	29.13	3.85	<0.0005	0.29
Practices & K-Conservation	0.16	24.03	0.09	n/a	0.01
Practices & K-Sources	2.99	27.35	1.44	n/a	0.11
Practices & Knowledge	5.87	20.87	3.69	<0.0005	0.28
Knowledge & K- Filtering	6.82	16.78	5.33	<0.0005	0.41
K-Filtering & K-Indigenous	0.62	21.22	0.38	n/a	0.03
K-Filtering & K-General	2.35	28.82	1.07	n/a	0.08

Table A2. Univariate ANOVA results

Knowledge Dams (K-Dams)				
Effect	***F-value***	***D.F.***	***p***	***Cohen's d***
Age	7.55	2; 150	0.001	n/a
Member of Awareness Group	3.90	1; 150	0.050	0.84
Municipality	14.01	1; 150	<0.0005	0.94
Knowledge Water General (K-General)				
Effect	***F-value***	***D.F.***	***p***	***Cohen's d***
Age	3.33	2; 150	0.038	n/a
Education Level	4.01	3; 150	0.009	n/a
Knowledge Water Filtering (K-Filtering)				
Effect	***F-value***	***D.F.***	***p***	***Cohen's d***
Gender	3.93	1; 150	0.049	0.22
Water Awareness Group	7.25	1; 150	0.008	0.58
Knowledge Water Conservation (K-Conservation)				
Effect	***F-value***	***D.F.***	***p***	***Cohen's d***
Age	3.33	2; 150	0.039	n/a
Member of Awareness Group	4.32	1; 150	0.039	0.64

References

1. Guterres, A.: The sustainable development goals report 2020. United Nations Publ. issued by Dep. Econ. Soc. Aff., 1–64 (2020)
2. Rawlins, J.: Political economy of water reallocation in South Africa: Insights from the Western Cape water crisis. Water Secur. **6**, 100029 (2019). https://doi.org/10.1016/J.WASEC.2019.100029
3. Mutamba, J.: Water security: Is South Africa optimally pursuing its options? WIT Trans. Ecol. Environ. **239**, 47–54 (2019). https://doi.org/10.2495/WS190051
4. Blignaut, J., Van Heerden, J.: The impact of water scarcity on economic development initiatives. Water SA **35**(4), 415–420 (2009)
5. Clifford, E., et al.: Interactive water services: The WATERNOMICS approach. Procedia Eng. **89**, 1058–1065 (2014). https://doi.org/10.1016/j.proeng.2014.11.225
6. Enqvist, J.P., Ziervogel, G.: Water governance and justice in Cape Town: An overview. Wiley Interdiscip. Rev. Water **6**(4) (2019). https://doi.org/10.1002/wat2.1354
7. Turok, I., Borel-Saladin, J.: Is urbanisation in South Africa on a sustainable trajectory? Dev. South. Afr. **31**(5), 675–691 (2014). https://doi.org/10.1080/0376835X.2014.937524
8. Aquilani, B., Silvestri, C., Ioppolo, G., Ruggieri, A.: The challenging transition to bio-economies: Towards a new framework integrating corporate sustainability and value co-creation. J. Clean. Prod. **172**, 4001–4009 (2018). https://doi.org/10.1016/J.JCLEPRO.2017.03.153
9. Szopik-Depczyńska, K., Kędzierska-Szczepaniak, A., Szczepaniak, K., Cheba, K., Gajda, W., Ioppolo, G.: Innovation in sustainable development: An investigation of the EU context using 2030 agenda indicators. Land use policy **79**, 251–262 (2018). https://doi.org/10.1016/J.LANDUSEPOL.2018.08.004
10. Balkaran, S.: Smart cities as misplaced priorities in South Africa: A complex balance of conflicting societal needs. J. Manag. Adm. **2**(1), 1–30 (2019)
11. Malik, B.H., et al.: Investigating technologies in decision based internet of things, internet of everythings and cloud computing for smart city. Int. J. Adv. Comput. Sci. Appl. **10**(1), 580–587 (2019). https://doi.org/10.14569/IJACSA.2019.0100173

12. El-Kholei, A.O., Yassine, G.: The mirage of smart sustainable cities in the Arab region. Open House Int. **44**(1), 8–16 (2019). https://doi.org/10.1108/ohi-01-2019-b0002
13. Hamdy, A.: Information and communication technologies in smart water management. In: Sixth Int. Sci. Agric. Symp. (2015). https://doi.org/10.7251/AGSY15051530H
14. Zhao, B., Guo, X., Marinova, D.: To pay more or use less: China's urban water pricing regime and water awareness. In: MODSIM 2011 - 19th Int. Congr. Model. Simul. - Sustain. Our Futur. Underst. Living with Uncertain, pp. 1724–1730 (2011). https://doi.org/10.36334/modsim.2011.d12.zhao
15. Kunjuzwa, D., Scholtz, B., Fashoro, I.: Incorporating indigenous knowledge within appropriate technologies for promoting awareness of water resource issues. In: 2020 IST-Africa Conf., pp. 1–9 (2020)
16. Bai, Z., Zhao, J., Wei, Z., Jin, X., Ma, L.: Socio-economic drivers of pig production and their effects on achieving sustainable development goals in China. J. Integr. Environ. Sci. **16**(1), 141–155 (2019). https://doi.org/10.1080/1943815X.2019.1671463
17. Aivazidou, E., et al.: Smart technologies for sustainable water management: An urban analysis. Sustain. **13**(24), (2021). https://doi.org/10.3390/su132413940
18. Van Den Branden, K.: Sustainable education: Exploiting students' energy for learning as a renewable resource. Sustain. **7**(5), 5471–5487 (2015). https://doi.org/10.3390/su7055471
19. Al-Sai, Z. A., Abualigah, L. M.: Big data and E-government: A review. In: ICIT 2017 - 8th Int. Conf. Inf. Technol. Proc., pp. 580–587 (2017). https://doi.org/10.1109/ICITECH.2017.8080062
20. Belkahla Driss, O., Mellouli, S. Trabelsi, Z.: From citizens to government policy-makers: Social media data analysis. Gov. Inf. Q. **36**(3), 560–570 (2019). https://doi.org/10.1016/J.GIQ.2019.05.002
21. Singh, R.K., Verma, H.K.: Effective parallel processing Social Media Analytics Framework. J. King Saud Univ. - Comput. Inf. Sci. **34**(6), 2860–2870 (2022). https://doi.org/10.1016/J.JKSUCI.2020.04.019
22. Okuah, O., Scholtz, B. M., Snow, B.: A grounded theory analysis of the techniques used by social media influencers and their potential for influencing the public regarding environmental awareness. In: ACM Int. Conf. Proceeding Ser. (2019). https://doi.org/10.1145/3351108.3351145
23. Senanayake, S.: Indigenous knowledge as a key to sustainable development. Journal of Agricultural Sciences. **2**, 87 (2006)
24. Pamla, A., Thondhlana, G. Ruwanza, S.: Persistent droughts and water scarcity: Households' perceptions and practices in Makhanda, South Africa. Land **10**(6), (2021). https://doi.org/10.3390/land10060593
25. Dolo, M.: Water supply challenges in urban and rural areas of Eastern Cape, Masters Dissertation, p. 165. Durban University of Technology (2019)
26. Amoah, L.N.A.: Water scarcity and food security in Ngqeleni Locality in the Eastern Cape Province - South Africa. African J. Hosp. Tour. Leis. **10**(1), 40–53 (2021). https://doi.org/10.46222/AJHTL.19770720-85
27. Ross, J., Smith, D.: Korea-trends in 4 national KAP surveys. Stud Fam Plan **43**, 6–11 (1964)
28. Choudhary, P., Sarikwal, L.: Conceptual framework of organisational factors on explicit and tacit knowledge sharing. IOSR J. Bus. Manag. **19**(01), 34–38 (2017). https://doi.org/10.9790/487x-1901023438
29. Brouwer, S.: Public participation in science: The future and value of citizen science in the drinking water research. Water Journal. 1–15 (2018). https://doi.org/10.3390/w10030284
30. Olomolaiye, A., Egbu, C.: The impact of human resource policies and practices on knowledge management in the construction industry. In: Assoc. Res. Constr. Manag. ARCOM 2005 - Proc. 21st Annu. Conf., pp. 125–135 (2005)

31. Ngulube, P.: Managing and preserving indigenous knowledge in the knowledge management era: Challenges and opportunities for information professionals. **18**(2), 95–102 (2016). https://doi.org/10.1177/026666602400842486
32. Van Der Vyver, C.: Improving awareness of the water crisis in South Africa. Int. J. Appl. Bus. Econ. Res. **14**(2), 577–590 (2016)
33. Javadpour, A., Samiei, S.: Motivation and barriers to participation in virtual knowledge-sharing communities of practice. Manag. Sci. Lett. **7**(2), (2017). https://doi.org/10.5267/j.msl.2016.11.011
34. Noga, J., Wolbring, G.: Perceptions of water ownership, water management, and the responsibility of providing clean water. Water (Switzerland) **5**(4), 1865–1889 (2013). https://doi.org/10.3390/w5041865
35. Jorgensen, B., Graymore, M., O'Toole, K.: Household water use behavior: An integrated model. J. Environ. Manage. **91**(1), 227–236 (2009). https://doi.org/10.1016/J.JENVMAN.2009.08.009
36. Willis, R.M., Stewart, R.A., Panuwatwanich, K., Williams, P.R., Hollingsworth, A.L.: Quantifying the influence of environmental and water conservation attitudes on household end use water consumption. J. Environ. Manage. **92**(8), 1996–2009 (2011). https://doi.org/10.1016/j.jenvman.2011.03.023
37. Addo, I.A.: Assessing residential satisfaction among low income households in multi-habited dwellings in selected low income communities in Accra. Urban Stud. J. Ltd. **53**(4), 631–650 (2016). https://doi.org/10.1177/0042098015571055
38. Adams, E.A., Vásquez, W.F.: Do the urban poor want household taps? Community preferences and willingness to pay for household taps in Accra. Ghana. J. Environ. Manage. **247**, 570–579 (2019). https://doi.org/10.1016/J.JENVMAN.2019.06.113
39. Price, H.D., Adams, E.A., Nkwanda, P.D., Mkandawire, T.W., Quilliam, R.S.: Daily changes in household water access and quality in urban slums undermine global safe water monitoring programmes. Int. J. Hyg. Environ. Health **231**, 113632 (2021). https://doi.org/10.1016/J.IJHEH.2020.113632
40. Mondéjar-Jiménez, J.A., Cordente-Rodríguez, M., Meseguer-Santamaría, M.L.: Gázquez-Abad: Environmental behavior and water saving in Spanish housing. Int. J. Environ. Res **5**(1), 1–10 (2011)
41. Gregory, G.D., Di Leo, M.: Repeated behavior and environmental psychology: The role of personal involvement and habit formation in explaining water consumption. J. Appl. Soc. Psychol. **33**(6), 1261–1296 (2003). https://doi.org/10.1111/j.1559-1816.2003.tb01949.x
42. Viljoen, N.: City of Cape Town residential water consumption trend analysis 2014/2015 (2015)
43. Alexander, M.: Mapping poverty in South Africa | South Africa Gateway. South Africa Gateway (2021)
44. Colgate: Five Water Conservation | Colgate®, https://www.colgate.com/en-us/oral-health/brushing-and-flossing/five-water-conservation-tips. Last accessed 26 August 2022
45. MyGqeberha: Dam levels, https://mygqeberha.com/dam-levels/. Last accessed 2022
46. Mapfumo, R., Mtambanengwe, P., Chikowo, F.: Building on indigenous knowledge to strengthen the capacity of smallholder farming communities to adapt to climate change and variability in southern Africa. Clim. Dev. **8**(1), 72–82 (2016)
47. Salkind, N.J.: Encyclopedia of research design, 1st edn. Sage (2010)
48. Gravetter, F.J., Wallnau, L.B.: Statistics for the Behavioural Sciences. Cengage Learning, Wadsworth, Belmont, CA (2009)
49. George, D., Mallory, P.: IBM SPSS statistics 23 step by step: A simple guide and reference, 14th edn. Routledge, New York (2016)
50. Cosgrove, W.J., Loucks, D.P.: Water Resources Research Institute. Wyo Univ Water Resour Res Inst Water Resour Ser **44**, 4823–4839 (2015). https://doi.org/10.1002/2014WR016869

Open Data Accessibility Mechanisms for Tourism Development in South Africa

Siviwe Bala[1(✉)], Judy van Biljon[1], and Marlien Herselman[1,2]

[1] University of South Africa, Pretoria, South Africa
64074633@mylife.unisa.ac.za
[2] Council for Scientific and Industrial Research, Pretoria, South Africa

Abstract. Employment is a means of addressing the socio-economic developmental challenges of inequality and poverty. In South Africa, tourism has been identified as an important sector for creating employment and various initiatives are being investigated to support the tourism sector. Open Data has been proposed as a mechanism to promote tourism development but how that can be done has not been theorised. Applying Critical Realism as a research philosophy, this study seeks to identify potential causal mechanisms underlying open data accessibility and use towards sustainable tourism development. An exploratory qualitative case study and semi-structured interviews with 25 tourism information managers from the local government and private sector was used to collect data. Thematic Analysis was conducted in phase one of this study's qualitative analysis. The 15 themes identified were then subjected to affordance-based causal analysis in phase two of the analysis. Three mechanisms (innovation, efficiency and marketing) linked to touristic open data access and use for digital platforms development were identified. This study contributes to the limited body of open data literature in the South African tourism context. The contribution to theory building lies in the causal mechanisms proposed as underlying open data accessibility and use in tourism.

Keywords: Open data · Critical realism · Affordances · Socio-economic development · Tourism

1 Introduction

In South Africa, the tourism industry has been identified as a sector that provides better job opportunities than most other growth sectors with a 2,7% GDP contribution and employs 4,5% of the total South African workforce [1]. Therefore, the South African President announced support for destination marketing and ways to reduce bureaucratic red tape and prioritise support for developing tourism businesses [2]. Stakeholders in the tourism sector, including investors, rely on access to the relevant information that is up-to-date for planning and decision making [3] increasing the demand for tourism related open data [4]. The World Bank Group [5] asserts that open data provides the potential for poverty reduction and employment creation while Bonina and Scrollini [6] highlight the importance of understanding and improving the provision of open

P. Ndayizigamiye et al. (Eds.): IDIA 2022, CCIS 1774, pp. 118–132, 2023.
https://doi.org/10.1007/978-3-031-28472-4_8

data in its adoption for promoting the uptake toward solving real-world problems. Most open data studies focused more on unravelling conceptual and theoretical dimensions of open data while the empirical investigation of open data lags, especially in the African context [7]. As such Žebrytė et al. [8] called on researchers to conduct empirical studies that will investigate the area of access to and use of open data platforms. Despite the growing adoption and importance of open data in influencing economic growth and socio-economic development a grounded understanding of what supports the access and use of open data in the South African Tourism Sector is lacking. Against this background, the research was guided by the question: *What are the potential causal mechanisms for open data driven digital platforms developed in the South African tourism sector?*

The remainder of the paper is structured as follows; in Sect. 2 the review of open data literature is discussed. Section 3 presents the core concepts of critical realism and associated methodological implications. Key procedures for data collection and analysis are described in Sect. 4. The thematic analysis and the six stepwise-framework for identifying mechanisms were operationalised in Sect. 5 to present the findings. The discussion of the findings is presented and the paper concludes by discussing the contribution and key takeaways of the study in Sect. 6.

2 Review of Literature

In the tourism field, the focus is on open data reuse for the development of digital solutions such as Mobile Apps, APIs, Maps, Websites, Mobile Augmented Reality etc.[9]. The review in the following sub-sections covers key terminology and concepts related to open data including the stakeholders, accessibility and utilisation. Open Data (OD) relates to data and information that can be freely utilised, modified, and shared by anyone for any purpose [10]. In a report produced by The World Bank [11], typical stakeholders of open data can be grouped into six main archetypes namely; Suppliers, Aggregators, Developers, Enrichers, Enablers and ICT Sector. These stakeholder categories are described in Table 1 for a clear distinction. Governments, civil society, researchers, activists, the private sector and individuals can use open data for socio-economic and environmental gains [12].

Synthesising the transformative potential of open data Davies et al. [13] views open data as a "digital fuel of the twenty-first century", a promising resource that can support new economic activity and novel ideas. In its existence open data overlaps and sometimes cross pollinates with various other terms in the Information System domain. Figure 1 illustrates a nexus of open data related term and conceptual linkages between them.

There are eight principles that should be considered by governments and public funded organisations to ensure that the data is open when it is released [15]. The principles include data being timely, easily accessible, available in different data formats, openly licensed, complete, raw as possible, machine readable, free, and non-discriminatory. For this study the researcher examined "accessibility" principle and data disclosure characteristic due to their relevance to the broader aim of this study of providing causal mechanism underlying open data accessibility in support of South African tourism development.

Table 1. Open data stakeholder types [11].

Stakeholders	Description
Suppliers	These are organisations that publish their data as Open Data to allow others to use and reuse it. They include not only public sector bodies but also some private sector companies
Aggregators	These are organisations that collect and aggregate open data and, sometimes, other proprietary data. This is typically on a sectorial or geographic theme
Developers	These are organisations and individual developers who design, build, and sell web or smartphone applications to deliver government open data to customers (normally in the personal sector) in attractive and informative ways, sometimes in competition with "official" applications
Enrichers	These are organisations which use open data to gain new or better insights that they can deliver in services or products to their customers - often completely new services which could not exist before Open Data
Enablers	These are organisations which provide platform and technologies that other businesses and individuals use. They are a vital part of the Open Data "ecosystem" - while being revenue generating themselves, they also provide cost-effective and easy-to-access services for both data suppliers and data consumers
ICT Sector	These are companies in the "traditional" ICT sector of hardware supply, software development and systems integration that are providing services to companies working to create data-rich services. Some are also providing enabling services

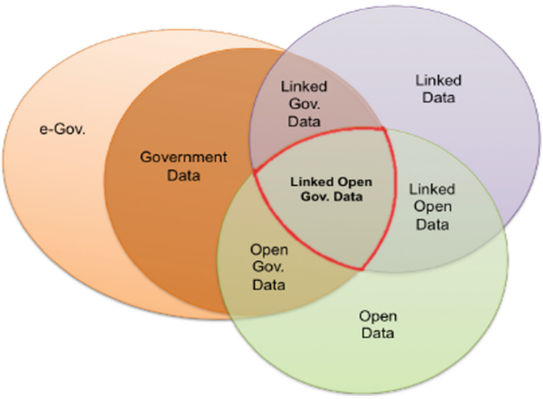


Fig. 1. Relationship Between Open, Government and Linked Data [14].

2.1 Open Data Accessibility

Open Data accessibility refers to the availability and access to a comprehensive content and information at a reasonable fee, obtainable through the use of internet, in an appropriate and easy to amend manner [10]. Factors such as context, manner of publishing, support provided to the users, open data tools, all these factors have an impact on how far the open data will be accessible. In the African Data Revolution report Van Belle [16] notes that in the African context open data is usually not available in most public sector organisation's online platforms. As a result, it is difficult to establish which government entity has and or is publishing open data since most information and data desired by the public ought to be derived from institutional documents in order to be accessible. Sapkota [17] maintains that a commitment to open data means making information and data resources accessible to all without discrimination; and actively engaging [users and other stakeholders] to ensure that information and data can be used in a wide range of ways. Davies et al. [13] argue that putting data online under open license does not equate to it being accessible and usable. With accessibility perceived as an impediment, Zuiderwijk et al. [18] proposed that more research should be conducted on open data accessibility in order to uncover other possible impediments that inhibit and constrain open data accessibility. Lastly but not least, the World Bank Group [5] suggested that in order to effectively open up government data, a constant interaction and collaboration between users of open data, government and ICT sectors is imperative so as to not only leave the function of opening the data to the public sector organisation alone.

2.2 Open Data Utilisation

Enormous quantities of open data are associated with the tourism industry [19]; particularly, the travel data in the form of transportation services (e.g., the bus and train schedules, bikes lanes and walkways information) needed for developing digital platforms such as mobile applications [20]. Table 2 depicts the categorisation of open tourism data and its prevalent utilisation within the tourism sector. However, in spite of tourism widely known as an industry concerned with open data, Pantano [21] underscored the limited application of open data in the tourism industry sector. This is supported by Pesonon and Lampi [19] and Hassan and Twinomurinzi [22], calling for more studies to investigate the use of open data, particularly, examining open data benefits for tourism businesses and destinations in developing countries.

3 Critical Realism

Critical Realism (CR) is a new direction in the philosophy of science and social science [23], which provides an alternative to the positivist and interpretivist paradigms that have shaped our view of science and causal laws [24]. There are two critical realism enigmas asserting that i) knowledge is fallible (transitive dimension) – arguing that the knowledge about the real-world is socially constructed; and that ii) there is a real-world (intransitive dimension) – the world that exists independent of our knowledge and awareness about it [25]. Such that one of the vital tenants of critical realism is that knowledge is distinguished from existence. The central idea of critical realism is the stratification of social

Table 2. Categorisation of open tourism data [19].

Open data type	Description	Where data has been used
Geographic data	GPS Location	Mobile applications, Websites
Event data	Description of events, bands playing timetable, event type	Mobile applications, Websites
Visitor statistics	Number of overnights	Mobile applications, Websites
Supply statistics	Number of businesses, types of businesses, number and information on attractions and museums	Mobile applications, Websites
Survey data	Data from survey studies	Mobile applications, Websites, academic and business research
Supply information	Information on travel destinations, attractions, restaurants, and happenings	Mobile applications, Websites, academic and business research
Transit data	Timetables	Mobile applications, Websites
Governmental data	Tax distribution and collection	Mobile applications, Websites, academic and business research
All the above		Smart Tourism City, augmented reality applications, services that combine data from several sources

reality into three nested domains, namely, real, actual and empirical domain (as illustrated in Fig. 2) [26]. Critical realism literature further asserts that reality has an ontology differentiated by structures, mechanisms, and events operating in an open system [27]. In terms of practice and methodology, critical realism is not committed to the qualitative and quantitative forms of research. CR goes beyond the "what" question that the qualitative and quantitative usually describes, to the "why" and "how" questions in order to understand and explain phenomenon such as why, i.e., Information Communication Technologies lead to Development in the Information Communication Technologies for Development (ICT4D) research field and hypothesise the causal structures that might have caused the observed phenomenon or outcomes [28].

One distinguishing critical realism methodological feature is the process of retroduction. Bygstad, Munkvold and Volkoff [29] draws on Weick's description that the process of retroduction is a "thought trial" [29] employed towards identify the mechanisms underpinning empirically observed events. In his study Mungai [30], drawing from Easton [31], explains that retroduction involves tracing backwards in retrospect to identifying circumstances and setting that led to the observed events. The identified mechanism is usually a logical explanation of the contingent reality through an iterative process corroborated by the collected empirical data. According to Raduescu and Vessey [32], "causality is central to the world of scientific investigation and is generally understood to be the relationship between cause and effect, while causal mechanisms are the process or pathways through which an outcome is brought about". The explanation

of causal powers that were activated at a given place and time is fundamental to critical realism's approach of understanding the ontology of social reality.

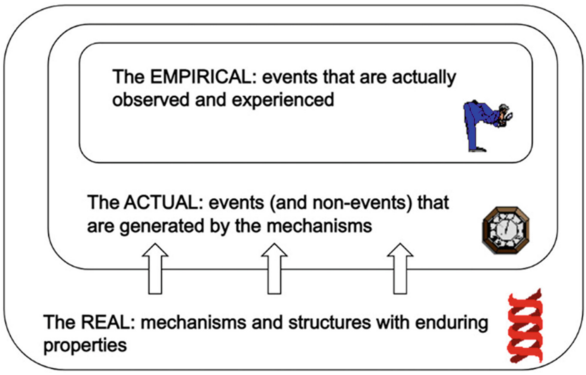


Fig. 2. Stratified ontology of critical realism [26].

Critical realism's view on causality is such that mechanisms produce certain outcomes in one context and different outcomes under different context and circumstance as a result of contingent causality [27]. Critical realism's notion of causality is illustrated in Fig. 3. Mechanisms also occur as a result of an interplay between human, social and technical elements that produce contingent events. However, those events or outcomes cannot be separated to the elements that caused them [27]. Simultaneously, the results of the abstract or high-level mechanisms may arise due to the interplay of various affordances at a granular level [29].

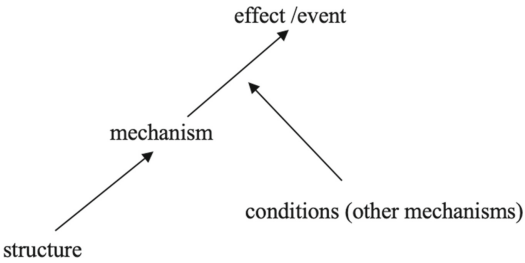


Fig. 3. Critical realism view on causality [33].

Critical Realism was considered appropriate for investigating this practice-based research domain of tourism information access as CR's search for causation is useful to researchers who aim to explain social events towards suggesting practical policy recommendations that could address social problems [34]. This also included CR's unique features of reflexivity, pluralism and iterative retroduction, features that are not supported by interpretivism and positivism [35].

3.1 Affordances as Generative Mechanisms

The original tenant and basic assumption of Affordance theory is 'action possibility' provided by the environment [36] to an actor that is a result of the interaction between the environment and the actor. This research project adopted the affordance concept application akin to Bygstad et al. [29] also observed in the IS literature where affordance was used as a relational term and "the potential for behaviours associated with achieving an immediate concrete outcome" [37] emanating from interaction between an object (such as an IT artefact) and goal-oriented actor or actors [38]. The relational notion of affordance includes viewing affordances as neither exiting in the artefact nor the actor but as a result of the interaction between the two objects or structures [37]. Affordance potentialities associated with IT artefacts are methodologically easier to identify than whole mechanisms [37]. In this instance affordances are viewed as elementary units or "building blocks" of more complex mechanisms [37, 38]. This study employed affordance-based causal analysis to examine the contextual conditions and factors influencing the accessibility and utilisation of touristic open data for the development of travel and tourism mobile applications.

4 Research Methodology

An exploratory qualitative research design was employed using purposive sampling, which involves "the deliberate choice of an informant due to the qualities the informant possesses" [39]. Greater interest in reusing open data usually comes from visionaries, entrepreneurs, and small businesses who seek to make new products and services that are data-driven [5]. Therefore, the study was conducted with fifteen private sector participants from mobile application developing companies (Start-ups) that have South African tourism open data driven mobile applications on Google Play Store. The public sectors participants were eight Gauteng municipalities, including Gauteng Tourism Authority and Gauteng Department of Economic Development. Gauteng province as an economic hub with the largest Internet penetration (85.2%) was selected for this study. This cross-sectional study used semi-structured interviews for qualitative data collection where participants were asked questions about their experiences concerning accessing, using, and their interaction with tourism-related open data. The questions asked were geared towards participant thoughts, perception and experiences about tourism open data accessibility and use underpinned by Open Data in Developing Countries (ODDC) framework to inform the interview questions available in this link: https://bit.ly/3L8CnRS. The duration to complete one semi-structured interview was approximately 30 to 40 min each. The researcher recorded the interview sessions to ensure comments are not missed, although some notes were also taken during the interview session. In total, 25 interviews were conducted in which 15 of the interviews were with the participants from private companies while ten were with public sector participants. The researcher was granted ethical clearance with reference number – Ref #:074/SGB/2029/CSET_SOC for the purposes of this study. This study's qualitative data analysis process for identifying mechanisms began in phase one with the Thematic Analysis (TA) as advocated by Braun and Clarke [40] was conducted as a useful and flexible method for qualitative

research, followed by phase two of stepwise framework for critical realist data analysis as advocated by Bygstad et al. [29].

5 The Findings

5.1 Thematic Analysis

This study employed thematic analysis approach for qualitative analysis of the collected research data using Atlas.ti software package, which entailed coding, identifying and analysing patterns or themes emerging from the data and reporting on them. The thematic analysis (phase one) followed six stages as advocated by Braun and Clarke [40]. A total of 15 themes were identified and their descriptions, including summaries of raw data extracts from individual transcripts that relate to the themes, were created as part of thematic analysis, as illustrated in Table 3. The full list of themes is available in this link: https://bit.ly/3P2J2xR with theme descriptions and sample quote.

Table 3. Themes derived from transcribed data.

Theme	Theme	Theme
1) Government support for open data initiatives is lacking or limited	6) Poor open data platforms formats and standards	11) Open Data utilised to provide tourist convenience and efficiency
2) Unavailability and challenges in accessing open data	7) Innovation led to open data application development	12) Open Data use required compliance with GDPR and POPI Act
3) Open data quality	8) Non-existent Open Data policies	13) Touristic open data and information easily obtainable
4) Cost and competitiveness of Open Data driven mobile app development	9) Entrepreneurship informed the use of open data for application development	14) Users engaged on published tourism data and information
5) Marketing and information provision for tourist decision making informed the development of open data driven mobile app	10) Collaboration with relevant stakeholders crucial for open data supply and publication	15) Private company collaborations crucial for open data sourcing, access and sharing

In line with CR ontology, critical realist data analysis begin with the search for 'demi-regularities' at the empirical level of reality [34]. CR searches for tendencies not laws, as predictions that suggest that 'whenever event x, then y' are not possible in the social world that consist of open system [41]. The tendencies that CR look for may be recognised from the rough trends or broken patterns in empirical data that are known as themes from phase one of thematic analysis. Critical realists call such tendencies demi-regularities and can be identified through qualitative data coding [34]. In CR the

identification of themes in the form of demi-regularities is crucial as it also represents the beginning of abduction and retroduction [34]. The themes and codes that were identified as part of thematic analysis were adopted for abstraction and retroduction in a stepwise framework for critical realist data analysis for uncovering affordances and mechanisms underlying open data accessibly and use in support of the South African tourism sector as explained in the next section. The stepwise framework was operationalised using the concept of affordances as advocated by Bygstad et al. [29].

5.2 Stepwise Framework for Critical Realist Data Analysis

Step 1: Description of events and issues that constitute the phenomenon of interest. Events are basically the clustering of themes derived from the empirical evidence, so that the identification of events can be corroborated by direct observations and interview quotes from the empirical data[42]. The researcher continued with the process of identifying events until no new information on the events was discovered. The consolidation of themes that were derived from this study's phase one of thematic analysis served as events. Crucial events that were identified during data collection and analysis were: (1) the desire to innovate through open data (2) open data driven app developed for marketing and information provision to support tourist decision making (3) touristic open data used for generating income (4) open data used for user convenience and efficiency.

Step 2: Identification of Key Entities. The second step focused on identifying associated entities and objects of the case that characterise the phenomenon being studied and collecting data about these entities. These entities, for example, could be organisations, technology artefacts, individuals and interactions between them with emergent causal powers [29]. The combination of these entities may constitute physical or social structures that reside in the real domain. Key entities that were distinguished concerning events under analysis in this study include among others developers, suppliers, aggregators, enablers and the general ICT sector. Seven other technical, physical or social structures that were identified include: 1) Internet connection 2) Websites 3) Technical expertise 4) Smartphones, Open datasets, Open Data APIs 5) Voluntary involvement of individual innovators and entrepreneurs, tourism practitioners 6) Collaboration between companies and online user communities 7) Private funding for mobile application development (personal savings, financial support from family members, private investment, i.e. sweat equity funding). The entities may sometimes form a collection called an assemblage which consist of interconnected entities that act, enact, and interact to generate events [43] as it was observed in this study from the collaboration of various private companies who collected and shared open data among themselves for the development of mobile applications.

Step 3: Theoretical Re-description (Abduction). The third step of this study's critical realist data analysis was the interpretation of the empirical data using theoretical concepts called theoretical re-description or abduction. In this study, the researcher applied Affordance Theory as a relevant IS theory and lens to enhance understanding of actions and events through redescription and recontextualisation. According to Danermark, Ekström and Karlsson [41], "abduction is to move from a conception of something to a different,

possibly more developed or deeper conception of it". This abduction process is a form of inference different from induction and deductions in that the case under investigation is abstracted, described in general patterns and concepts, while conclusions are not part of the claims of given statements or ideas. For theoretical re-description and abduction, the concept of affordance was adopted towards identifying and analysing potential causal mechanisms underlying open data accessibility in support of South African tourism development.

Step 4: Retroduction: Identification of Candidate Mechanisms. Retroduction was the fourth step in the critical realist data analysis as a key epistemological process. This step focused on identifying underlying causal mechanisms and conditions that may have caused patterns of events under review. Retroduction is the mode of inference where the researcher postulates the continuous process by which observed events were generated [44]. In practical terms, this process was conducted iteratively and creatively by the researcher as a meta-process whose outcome was identifying mechanisms that explain what caused the observed events [31]. Through the lens of the affordance concept four candidate causal mechanisms underlying open data accessibility and use in support of South African tourism development were identified. The affordances are modelled on the structures proposed by Bygstad et al. [29] as described below:

The Innovation Mechanism. The innovation mechanism is based on two affordances, the *reusability* and *redistributability* affordances of open data. Figure 4 the reusability affordance actualisation, was supported by the collaboration with user communities and understanding of user needs and requirements as an enabling condition for the use of open data.

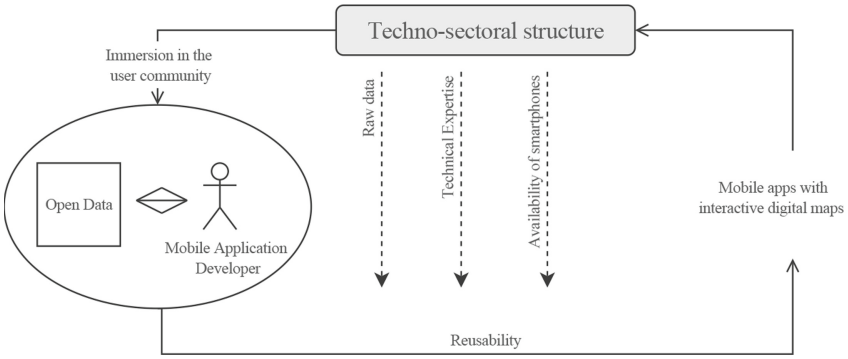


Fig. 4. Reusability affordance.

The immediate concrete outcomes are a result of goal oriented user interacting with an artefact (i.e. open data dataset) [29]. The mobile applications with interactive digital map were identified as a concrete outcome by the mobile application developers. Participant was quoted stating the following concerning the open data reuse linked

to the identified reusability affordance outcomes i.e., *"open data enabled the mobile application development with digital maps" C14.*

Davies et al. [13] maintains that open data enables private innovators to develop solutions for improvement of public service and build unique, new services and products with economic and social value, thus contributing to economic development and innovation. Furthermore, innovation mechanism as a source of value creation indeed supports tourism development [45]. This is through the use and application of touristic open data for the development of new and innovative digital products and services that transform tourism industry.

The Marketing Mechanism. The *boosterism (marketability)* affordance that was actualised contributed to the activation of the marketing mechanism. Boosterism is an approach on tourism initiatives such as stimulating tourism demand, planning, analysis and goal setting purely within marketing context for tourism growth [46].

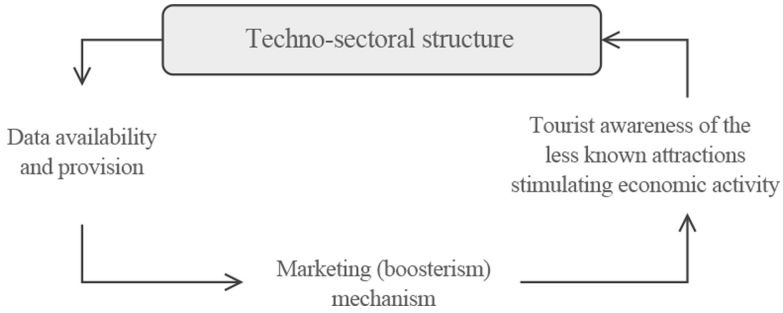


Fig. 5. Marketing mechanism.

The availability of data provided by public entities also shared by private companies among each other as data aggregators, enablers and suppliers lead to the enactment of the marketing mechanism illustrated in Fig. 5. The observed concrete outcomes from the activation of the marketing mechanism were among others the ability to draw tourist to less known tourist attractions and the increased economic activity in the tourist destinations and attractions. In tourism as a data driven industry, the marketing and management of tourism businesses and destinations depend on the collection and analysis of data such as tourist arrivals, tourism spent, bed nights etc. The effective use of this data can result into new products and business models that may lead into new employment opportunities that never existed before and increased economic activity in tourist attraction and destinations.

The Efficiency Mechanism. The *efficiency* affordance was actualised as a subset of generative mechanism through the interaction between the goal driven users and the open data datasets as an ICT artefact. A participant stated the following regarding the access and use of open data for achieving efficiency affordances: *"It was the only one[app] that was allowing people to use an app to find wildlife sightings, and it helps people to get*

more out of their time in South Africa". The actualisation of efficiency affordance as a subset of generative mechanism resulted in the enactment of the efficiency mechanism in Fig. 6. Conradie and Choenni [47] argue that the benefits associated with open data include efficiency and effectiveness both in government and in social setting.

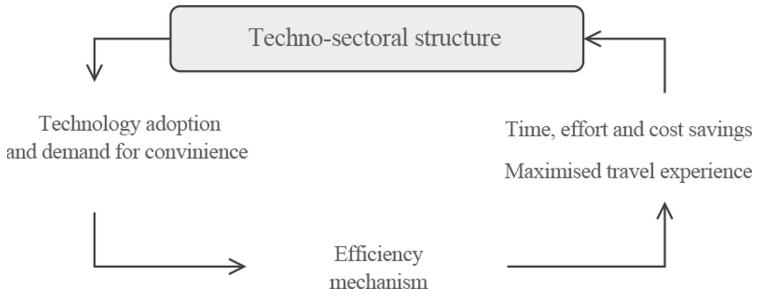


Fig. 6. The efficiency mechanism.

The Entrepreneurial Drive Mechanism. The entrepreneurial drive mechanism is based on the actualisation of the *earnability* affordance through the access and use of open data. Revenue and income generation from the use of open data was specified as one of the perceived goals that was indicated by this study's participants as open data users. The users of touristic open data pursued the goal of revenue generation through start-ups companies for mobile application development and as entrepreneurs. The enabling conditions for the actualisation of the earnability affordance were the opportunity to create a touristic open data driven mobile applications in order to fulfil user needs gap and space of possibilities provided by the tourism sector. The data availability, technical expertise and internet connectivity provided stimulating conditions for the earnability affordance. Two concrete outcomes were indicated by the participants namely; the revenue generation and job creation with the possibility of earning income. The touristic open data, as an ICT artefact, was engaged by goal-oriented mobile application developers to generate economic value through the activation of the entrepreneurial drive mechanism. Furthermore, Carrara et al.[48] confirms that even though there is a limited number of studies that measure the number of jobs created in the Open Data field of practice, there is evidence that supports causal relationship between the release of Open Data and subsequent increase in the number of jobs created by small-medium sized enterprises.

Step 5: Analysis of the Set of Affordances and Associated Mechanisms. In this step several avenues of analysis are proposed in line to this study's research questions [29]. The iterations suggested as approaches that may be explored to examine the explanatory power of the initial results are: i) Analysing the dependencies between affordances ii) Grouping affordances iii) Identifying focal affordances, and their relationships to other mechanisms iv) Abstracting affordance into height-level mechanisms. The five identified affordances (reusability, redistributability, boosterism, efficiency, and earnability) were abstracted into four mechanisms of innovation, marketing, efficiency and entrepreneurial drive. The innovation and entrepreneurial drive were combined to form one mechanism

"the innovation mechanism". What enabled the combination of the two mechanisms (innovation and entrepreneurial drive) was the enabling conditions that were relatively similar, the affordances that are of the same nature and the functional outcomes of the mechanisms were also relatively similar.

Step 6: Assessment of Explanatory Power. The results of this study's stepwise critical realism based methodological analysis provided three mechanisms namely innovation, efficiency and marketing. Although other mechanism could also have been at play, the three final mechanisms were identified for being consistent with the data material and feedback that was received from the study's participants.

6 Discussion and Conclusion

This study is a first step towards enhancing our understanding of open data, its use and accessibility in the South African tourism sector context. Open data as an approach in supporting the accessibility of tourism related information in South Africa has a potential for supporting socio-economic development efforts as shown by the three identified mechanisms namely; *innovation, marketing, efficiency*. Critical Realism was found appropriate as a research philosophy for guiding the non-deterministic view on causality where an exploratory case study was used to collect data for identifying potential causal mechanisms underlying open data driven digital platforms developed in the South African tourism sector. The findings provide various opportunities for further research. The literature on open data in relation to the South African context was found lacking in terms of strategies and relevant policies. Therefore, this study does not only contribute to the body of open data literature in South African context particularly in the tourism domain, but it could also be used to inform policy development on open data. The study also demonstrates theory building in ICT4D research based on the abstraction of causal mechanism-based explanations underpinned by Critical Realism philosophy and the theory of Affordances.

Acknowledgements. This paper is based on the research supported by the South African Research Chairs Initiative of the Department of Science and Technology and National Research Foundation of South Africa (Grant No. 98564).

References

1. Statistics South Africa: Tourism Satellite Account for South Africa, final 2016 and provisional 2017 and 2018 (2019)
2. Ramaphosa, C.: State of the Nation Address. https://bit.ly/3Eov6Mu
3. Department of Tourism: National Tourism Sector Strategy Executive Summary. South Africa (2011). https://bit.ly/3rF9HY0
4. Wu, C.. Te., Liu, S.C., Chu, C.F., Chu, Y.P., Yu, S.S.: A study of open data for tourism service. Int. J. Electron. Bus. Manag. **12**, 214–221 (2014)
5. World Bank Group: Open Data for Sustainable Development. Policy Note ICT01. 33 (2015)

6. Bonina, C., Scrollini, F.: 11 Governing Open Health Data in Latin America. Mak. Open Dev. Incl. Lessons from IDRC Res. (2020)
7. Saxena, D.: The search for mechanisms in business research: Reflections on retroductive analysis in a multilevel critical realist case study. Electron. J. Bus. Res. Methods. **17**, 17–27 (2019)
8. Žebrytė, I., Fonseca-Vasquez, F., Hartley, R.: Emerging economy entrepreneurs and open data: Decision-making for natural disaster resilience. J. Small Bus. Strateg. **29**, 43–54 (2019)
9. Pereira, R.L., Sousa, P.C., Barata, R., Oliveira, A., Monsieur, G.: CitySDK Tourism API - building value around open data. J. Internet Serv. Appl. (2015). https://doi.org/10.1186/s13174-015-0039-z
10. Open Knowledge Foundation: The Open Definition. http://opendefinition.org/
11. The World Bank: Open data for economic growth. https://bit.ly/3rF7nQU
12. Open Knowledge Foundation: Open Data Handbook. https://bit.ly/3CEqMaG
13. Davies, T., Perini, F., Alonso, J.: Researching the emerging impacts of open data ODDC conceptual framework. J. Community Informatics. **12**, 148–178 (2016)
14. Attard, J., Orlandi, F., Scerri, S., Auer, S.: A systematic review of open government data initiatives. Gov. Inf. Q. **32** (2015). https://doi.org/10.1016/j.giq.2015.07.006
15. OpenGovData.org: The 8 Principles of Open Government Data. 2007, 1–8 (2007)
16. Van Belle, J.-P.: African Data Revolution Report - The Status and Emerging Impact of Open Data in Africa. 1–111 (2018)
17. Sapkota, K.: Exploring the emerging impacts of open aid data and budget data in Nepal. Free. Forum. 1–67 (2014)
18. Zuiderwijk, A., Janssen, M., Choenni, S., Meijer, R., Alibaks, R.S.: Socio-technical Impediments of Open Data. Electron. J. e-Government. **10**, 156–172 (2012). https://doi.org/10.1641/b570402?ref=search-gateway:885882d1830675b0f27af0760faeaef8
19. Pesonen, J., Lampi, M.: Utilizing open data in tourism. Conf. Enter 2016 (2016)
20. Longhi, C., Viallis, L.: Open Data: Challenges and Opportunities for The Tourism Industry Data: Challenges and Opportunities for The Tourism Industry 13 (2013)
21. Pantano, E., Priporas, C., Stylos, N.: You will like it !'using open data to predict tourists' response to a tourist attraction. Tour. Manag. **60**, 430–438 (2017). https://doi.org/10.1016/j.tourman.2016.12.020
22. Hassan, M.I.A., Twinomurinzi, H.: A systematic literature review of open government data research: challenges, opportunities and gaps. In: 2018 Open Innovations Conference, OI 2018, pp. 299–304. Institute of Electrical and Electronics Engineers Inc. (2018)
23. Archer, M., Bhaskar, R., Collier, A., Lawson, T., Norrie, A.: Critical realism: essential readings. Crit. Realis. Essent. Readings. 1–756 (2013). https://doi.org/10.4324/9781315008592
24. Bhaskar, R.: A Realistic Theory of Science. Routledge (2013)
25. Zachariadis, M., Scott, S., Barrett, M.: Methodological implications of critical realism for mixed-methods research. Manag. Inf. Syst. Q. **37**, 855–879 (2013)
26. Heeks, R., Wall, P.J.: Critical realism and ICT4D research. Electron. J. Inf. Syst. Dev. Ctries. **84**, 1–11 (2018). https://doi.org/10.1002/isd2.12051
27. Bygstad, B., Munkvold, B.E.: In Search of Mechanisms Conducting a Critical Realist Data Analysis. 1–15 (2011)
28. Omland, H.O., Thapa, D.: Methodological approach for identifying mechanisms in ICT4D: A critical realism perspective. In: IFIP Advances in Information and Communication Technology (2017)
29. Bygstad, B., Munkvold, B.E., Volkoff, O.: Identifying generative mechanisms through affordances: a framework for critical realist data analysis. J. Inf. Technol. **31**, 83–96 (2016). https://doi.org/10.1057/jit.2015.13

30. Mungai, P.W.: Casual Mechanisms that enabale Institutionalisation of open Government Data in Kenya (2017)
31. Easton, G.: Critical realism in case study research. Ind. Mark. Manag. **39**, 118–128 (2010). https://doi.org/10.1016/j.indmarman.2008.06.004
32. Raduescu, C., Vessey, I.: Causality in Critical Realist Research : An Analysis of Three Explanatory Frameworks (2008)
33. Sayer, A.: Realism and Social Science. Sage, London, UK (2000)
34. Fletcher, A.J.: Applying critical realism in qualitative research: methodology meets method. Int. J. Soc. Res. Methodol. **20**, 181–194 (2017). https://doi.org/10.1080/13645579.2016.1144401
35. Heeks, R., Wall, P.J.: Critical realism and ICT4D research. In: IFIP Advances in Information and Communication Technology (2017)
36. Gibson, J.J.: The theory of affordances. Hilldale, USA **1**, 67–82 (1977)
37. Volkoff, O., Strong, D.: Affordance Theory and How to Use it in IS Research Nothing is so practical as a good theory. routledge companion to Manag. Inf. Syst. 232–245 (2017)
38. Volkoff, O., Strong, D.M.: Critical Realism and Affordances: Theorizing IT-Associated Organizational Change Processes **37**, 819–834 (2013)
39. Tongco, M.D.C.: Purposive sampling as a tool for informant selection. Ethnobot. Res. Appl. **5**, 147–158 (2007)
40. Braun, V., Clarke, V.: Using thematic analysis in psychology. Qual. Res. Psychol. **3**, 77–101 (2006)
41. Danermark, B., Ekström, M., Karlsson, J.C.: Explaining Society: Critical Realism in the Social Sciences (2002)
42. Thapa, D., Sein, M.K.: Trajectory of Affordances: Insights from a case of telemedicine in Nepal. Inf. Syst. J. **28**, 796–817 (2018). https://doi.org/10.1111/isj.12160
43. Thapa, D., Omland, H.O.: Four steps to identify mechanisms of ICT4D: A critical realism-based methodology. Electron. J. Inf. Syst. Dev. Ctries. **84**, 1–10 (2018). https://doi.org/10.1002/isd2.12054
44. Sayer, A.: Method in social science: A realist approach. Psychology Press (1992)
45. Avgerou, C.: Discourses on ICT and Development. Inf. Technol. Int. Dev. **6**, 1–18 (2010)
46. Marcouiller, D.: "Boosting" tourism as rural public policy: Panacea or Pandora's box? J. Reg. Anal. Policy. **37**, 28–31 (2007)
47. Conradie, P., Choenni, S.: On the barriers for local government releasing open data. Gov. Inf. Q. **31** (2014). https://doi.org/10.1016/j.giq.2014.01.003
48. Carrara, W., Chan, W.S., Fische, S., Steenbergen, E. van: Creating value through open data: Study on the impact of re-use of public data resources. Eur. Comm. (2015)

Emerging Technologies for Transformation, Inclusion and Sustainable Development

Social Media Affordances for Disaster Management

Omar Safianu and Jean-Paul Van Belle(✉)

Department of Information Systems, University of Cape Town, Cape Town 7700, Rondebosch, South Africa
omrsaf003@myuct.ac.za

Abstract. Social media has seen a rise in usage especially for disaster communication. Existing research that investigated the use of social media for disaster management have focused on information dissemination affordance of social media especially at the initial response phase, hence, the full potential of social media is not being realized. Social media offers many affordances in all phases of disaster management and holistic integration of the affordances may lead to better communication outcomes. This paper studied the extent of the knowledge gap through systematic literature review using the preferred reporting items for systematic reviews and meta-analysis (PRISMA) technique. Open, axial and selective coding strategies were used for the data analysis which also guided the study in defining a holistic framework for social media affordances for disaster management. The study provides disaster responders with an understanding of the social media functionalities that are best suited for each phase of disaster management. Essentially, the holistic understanding of social media functionalities in all disaster management phases can assist in minimizing the impact of disasters and help in managing them better.

Keywords: Social media affordances · Technology affordances · Disaster management · Disaster phases · Affordances theory

1 Introduction

Disaster events are threatening conditions that demand urgent action. In recent times, the magnitude and the number of disasters has increased with 2021 recording 432 disaster events which is above the 2001–2020 average of 396 cases and has impacted about 101.8 million people causing approximately 252.1 billion USD loss of properties [1]. Disaster losses are growing globally due to many reasons which include climate change and population growth with its concomitant increasing demand for agriculture products, unplanned urbanization, and growing pressure on natural resources [2]. While disaster response systems would save lives and properties, a number of these losses can be curbed if proper management strategies are implemented to tackle the root causes.

In recent times, social media has seen a surge in its use for disaster management especially as a communication channel during crisis events. The general public utilize

P. Ndayizigamiye et al. (Eds.): IDIA 2022, CCIS 1774, pp. 135–153, 2023.
https://doi.org/10.1007/978-3-031-28472-4_9

social media for disaster information sharing [3] or for seeking information [4]. Consequently, disaster management organizations (DMOs) are using social media as one of the key communication platforms and have integrated it into their disaster management efforts [5, 6].

Social media has the ability to establish extensive and far-reaching communication channel and can support information flow while being flexible to the communication needs of DMOs [6, 7]. These features of social media make it an invaluable tool that can support communication needs in times of disasters. Although existing literature has shown strong evidence of the value of social media to disaster communication management, very little is known about how this innovative communication tool could be harnessed and used more effectively in different phases of disaster management [7, 8]. Social media offers many affordances and functionalities and its usage scope for disasters varies significantly [8–10]. Though social media is a powerful communication channel, there is little research on how effective strategies can be deployed by DMOs to fully incorporate social media affordances for disaster management [6, 7, 11] and this may be due to lack of understanding of different affordances and functionalities of social media [7, 8]. The extensive use of social media in the response phase of disaster management is promising, however, its minimal use in other phases is cause for concern. Therefore, the objective of this paper is to provide insights into the following research questions:

1. What social media affordances exist in the different phases of disaster management?
2. What holistic framework can be deployed for disaster management using social media?

To highlight these insights, we examined the literature published from 2010 to 2021 using the preferred reporting items for systematic reviews and meta-analyses (PRISMA) then we utilized open, axial and selective coding techniques for data analysis and finally, a framework was developed to address the gap.

The paper is organized as follows. The first section presents an overview of the present state of social media use for disaster management and why further studies need to be carried in that area. Theoretical background is presented next where disaster management phases, social media use for disaster management, and affordances theory are discussed. We then presented our research methods, materials used, data analysis, and the results from the analysis. Afterwards, social media affordances for disaster management framework is presented followed by study conclusion, implications and directions for future research.

2 Theoretical Background

2.1 Disaster Management Phases

A disaster is "a serious disruption of the functioning of a community or a society involving widespread human, material, economic or environmental losses and impacts, which exceeds the ability of the affected community or society to cope using its own resources" [12] (p. 9). As presented in Fig. 1, disaster management has been generally classified into

four categories (often occurs in a cycle or phases) which are *mitigation, preparedness, response, and recovery*. This classification was first introduced by the FEMA [13] and is consistently used in the literature (e.g., Ngamassi, Ramakrishnan, Rahman [8], Chikoto, Sadiq, Fordyce [14], Islam, Chik, [15]).

The first phase of the disaster management is mitigation. This phase looks into possible measures to be taken to lessen long-term risks to people and properties in times of disasters. The actions and processes that would be needed are taking into consideration long before the disaster strikes. The purpose of the activities and processes taken in this phase is to protect lives and properties and minimize the cost of activities in the other phases [8, 13]. Some of the actions taken in this phase include public education and vulnerability analysis.

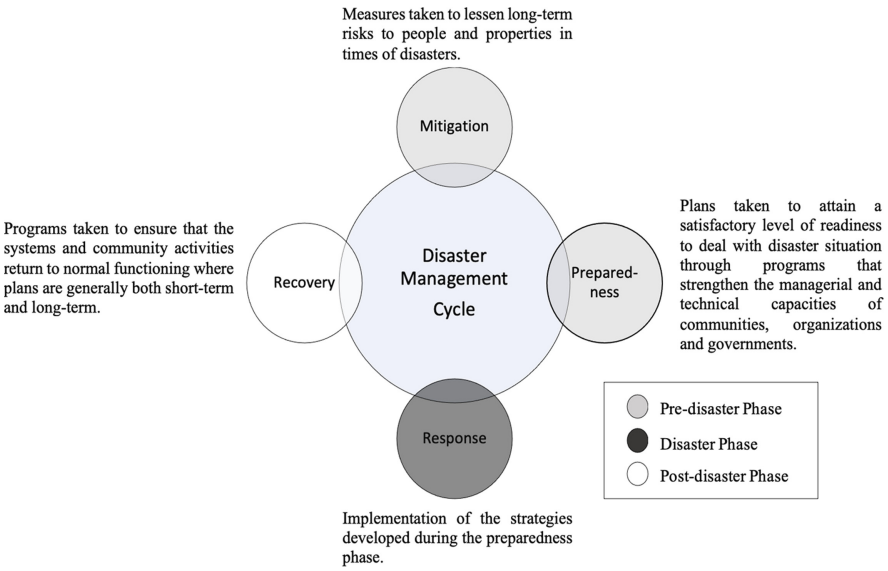


Fig. 1. Phases of disaster management (our elaboration)

Preparedness is the second phase of disaster management. It is almost impossible to mitigate all potential casualties and risks during disasters, therefore, it is essential to have some level of preparedness to reduce the impact of disaster [12]. The goal of disaster preparedness plans is to attain an adequate level of readiness to deal with any disaster situation through programs that strengthen the managerial and technical capacities of communities, organizations and governments [16]. This is often use to ascertain the logistical readiness to respond to disasters which can be achieved by putting in place good procedure and response mechanisms, public education/information, developing short-term and long-term strategies as well as early warning systems. The effectiveness of the preparedness phase relies on information availability on hazards, emergency risks and the countermeasures to adopt, and the degree to which the general public, government agencies and private organizations are able to make use of this information [16].

The third phase of disaster management is the response phase. The response phase starts when a disaster is looming or immediately after the disaster strikes [13]. This phase mainly entails implementation of the strategies developed during the preparedness phase [8]. It begins with an assessment of the situation and then providing assistance to victims while restoring important services like communications, transportations, water supply and ensuring continuity of vital services such as fire and police services [8, 13, 17]. The main focus of this phase is to provide immediate assistance to sustain life and support the morale of affected communities.

The fourth and final phase is the recovery phase. The goal of this phase is to ensure that the systems and community activities return to normal functioning where plans are generally both short-term and long-term [8]. Short-term and long-term recovery measures include returning support systems to minimum operating standards, improving quality of life, reconstruction, improving public information, safety and health education, and economic impact analysis [16]. Information services and resources may include data gathering related to rebuilding and documenting and documentation of lessons learned. Generally, the recovery phase provides information for effective mitigation and preparedness thereby reducing vulnerability in future disasters [16].

2.2 Social Media Use for Disaster Management

The low cost of social media for communication and its enhancement of management functions have made several organizations including DMOs to adopt it in their daily activities [18]. When information is not available during disasters, people use other means to satisfy their communication needs which may give rise to misinformation. During such times, DMOs are seen as reliable source of information because they are able to moderate and share information on disaster events [5]. Thus, DMOs are responsible for informing the public throughout disaster events and social media provides a platform for such services. DMOs therefore need to adjust their communication strategies on social media in order to be effective during disaster events. Artman et al. [18] propose two concepts in making effective communication during disaster events: dialogical disaster management, which refers to active monitoring of posts on social media in order to adjust the communication strategy based on the posts; and strategic awareness which refers to being aware of the understanding of the receivers of disaster information as well as the disaster incidence itself. Thus, DMOs would need to see the general public as co-creators of disaster information and not just consumers of information.

A number of researchers have studied how DMOs use social media by investigating their social media communication strategies during disasters [6, 19–22]. Generally, DMOs use social media services for information sharing, issuing warning signals, requesting specific information from the general public and fighting misinformation and rumors [5, 20] and influencing the behaviour of the public and media during disaster events [19, 22]. San et al. [23] studied the usage of social media by DMOs during disaster events and found out that they had a social media presence though it was built in an ad-hoc manner and they had only basic communication capabilities [8]. This finding was corroborated by Potter [24] and Ehnis and Bunker [20] who also found out that DMOs mainly use social media at the response phase of disaster management and at

both response and recovery phase [25, 26]. Even then, they use social media for information dissemination only [5, 27] while social media can be deployed for both disaster information dissemination and gathering [28, 29]. The challenge encountered by DMOs is to understand the affordances and capabilities of social and how it can be used for better communication outcomes in all phases of disaster management, i.e., *mitigation, preparation, response and recovery* phases.

2.3 Affordances Theory

The origin of affordances theory can be traced in ecological psychology and was first propounded by Gibson [30] to explain the perception of actors (human beings and animals) to the properties or objects of the environment to perform certain actions. Despite affordances being inherent in artefacts or objects, they are not their properties and are needed to produce intended actions [30]. Even though Gibson's [30] explanation of affordances refers to man-made tools or objects, Norman [31] further expands the theory to technology design to include everyday objects. Norman [32] later crystalised his views to explain affordances as "the relationship between properties of the artefacts and capabilities of the users that establishes the way that the artefact would be used" (p. 11). In other words, technology affordances refers to the arrays of potential uses enabled by ICT artifacts [33] and permeates the relationship between technology and organization [34]. Although affordances emerged from the relationship between the characteristics of the actor(s) and the technology, in this paper we explored the relationship at organizational level and not individual level. Majchrzak and Markus [35] suggest that "uses and outcomes of information systems and technology are best understood in terms of relationships between individuals or organizations and technology features." This suggestion makes the technology affordances theory appropriate for this study which attempts to examine social media use for disaster management based on literature review. Gaver [36] argues that technology affordances are independent of individual perception and/or appropriation which means that technologies are designed with affordances and whether users are aware of this or not does not make it invalid. Applied in this paper's context, it means whether DMOs appropriated social media affordances or not, does not mean social media lacks the affordances.

The concept of affordances has attracted the attention of information systems researchers [37] and has been proposed as a powerful concept in understanding the dynamics of social and material in organizational certain [38]. Strong et al. [39] expatiate the nature of affordances in organizations to mean "the potential for behaviors associated with achieving an immediate concrete outcome and arising from the relation between an artifact and a goal-oriented actor or actors." (p. 69). In the context of social media, several scholars have used the concept to explain the role of social media in organisations (e.g., Ellison [40], Mohajerni et al. [41], Majchrzak et al. [42], Zheng, Yu [43]). Chen et al. [44] have recorded ten technology affordances as the common social media affordances which are *visibility, persistence, editability, association, metavoicing, generativity, communicability, interactivity, collaboration*, and *anonymity*. These affordances are explained in Table 1. Some affordances are applicable to certain actions while others are not. For instance, in presenting the affordances, Chen et al. [44] distinguish between affordances which are related to government services innovation and those related to

government openness and accountability. In this work, we selected eight affordances to frame the study and are considered relevant to the study's objective which are *visibility, persistence, association, metavoicing, generativity, communicability, interactivity,* and *collaboration.*

Table 1. Technology affordances [44]

Affordance	Definitions
Visibility	The potential of social media which enables users' actions, emotions and connections to be visible to others
Persistence	The potential of social media which enables users to retrieve relevant actions beyond the time boundary
Editability	The potential of social media which enables users to craft and recraft their communicative actions freely before presented to others
Association	The ability of social media which enables users to establish connection among individual, organizations and actions
Metavoicing	The potential of social media to enable users to react to content
Generability	The potential of social media which enable users to engage and innovate by expressing ideas and experiences of public service
Communicability	The potential of social media which enables users to communicate with each other with common languages and syntaxes
Interactivity	The potential of social media which enables users to interact with others (especially government-citizen interaction)
Collaboration	The potential of social media to promote government-citizen collaboration that can improve decision making
Anonymity	The potential of social media which enable citizens to express their opinion anonymously

2.4 Data Collection and Methods

To realize the research objectives, we used a systematic literature review (SLR) based on PRISMA which we adapted from Moher et al. [45]. An SLR is a systematic process used to identify, evaluate, and summarize themes in the literature. It allows the restrictive collection of data from literature through a rigorous methodological analysis with minimum biases as compared to traditional reviews [46, 47]. In an SLR, the aim is to develop a general idea of a specific question and make it a reasonable summary of the literature [47] through quantitative and/ or qualitative approach, reporting evidence to crystalise implicit and explicit knowledge in relation to the objective of investigation [48].

To conduct a replicable and transparent analysis, we adopted an iterative process which is made up of the following stages [49, 50]:

- Formulating research questions
- Framing the research protocol
- Text sampling and selection of relevant text for analysis
- Developing the coding framework; and
- In-depth analysis of selected texts

The first stage entailed definition of the research questions which focused on understanding how social media enables disaster management and what affordances emerged during the management phases. The research questions were summarized in the following form:

1. What social media affordances exist in the different phases of disaster management?
2. What holistic framework can be deployed for disaster management using social media?

In the second stage, we defined the research protocol in order to ensure objectivity and support evidence-based practice [51]. The focus in this stage was to execute the research strategy which is implementing the inclusion and exclusion criteria of data sources in accordance with the research question [52]. We followed a detailed analysis perspective, selecting the most relevant articles from academic databases in social media disaster management. The literature selection process was made up of four steps as follows:

1. Identification: this step comprised of applying the search strings in Table 2. Two hundred and thirteen (213) papers were retrieved from this step.
2. Screening: In this step, all papers whose titles and abstracts were not directly related to social media disaster management were removed. The total number of papers excluded in this step was 46.
3. Eligibility: The number of works assessed for eligibility was 156. Using the following inclusion and exclusion criteria, the final list of papers was used for the review.

   Studies published from 2010 and 2021.
   Studies published in the English language.
   Only studies related to social media disaster management were considered.
   Secondary and tertiary studies were discarded (e.g., systematic literature reviews, surveys, etc.).
   Studies published in Scopus and Web of Science platforms.

4. Included: Twenty-three (23) works were selected for full text review after the eligibility assessment.

The aim of the third stage was to identify papers for the analysis by defining a research string to use for searching the databases. The articles were collected via keyword-title-abstract field codes of the individual databases. Using Boolean operators, “AND” and ‘OR”, as connectors, the search parameters were applied in the social media disaster management area as described in Table 2.

Table 2. Databases and key terms used for searching

Database	Search String
ScienceDirect EBSCO Host Google scholar	((“disaster management” OR “emergency management “OR “disaster phases”) AND (“social media” OR “social network))

We entered the search query in ScienceDirect, EBSCO Host, and Google Scholar databases which made it possible to gather a total of 81, 93, and 39 articles respectively, for a total of 213 articles published from 2010 to 2021. The three databases were selected because of their efficient algorithms for federated searching and retrieval of scientific and scholarly articles and the choice of the time frame was made in light of the fact that the use of social media in disaster management became popular in the last decade and its usage is regarded by various DMOs as experiment stages and therefore the best practices are developing [53]. Duplicate and non-English articles were eliminated which were insignificant with respect to the research question. Also, we restricted data gathering to articles in the Scopus and Web of Science lists of journals and scientific articles because of their high standards and quality of knowledge they provided. By carefully reading the abstracts, secondary and tertiary studies such as survey papers and literature reviews, and studies not related to social media disaster management were removed. Furthermore, specific technical articles in which the connection between social media and disaster management was only mentioned but not developed were discarded. Following these processes, we obtained a valid sampled articles then we checked the full text articles to further assess the eligibility and quality of the articles. Making a thorough reading, we selected 23 papers that were relevant to the research questions for the study. Figure 2 presents the process adopted for the study.

In the fourth stage, we defined the coding framework. We used open coding strategy to break the data into the four disaster management phases which are *mitigation, preparedness, response, and recovery.* The papers did not have to categorically mention a particular phase by name but certain actions and processes as listed by FEMA [13] and Srinivas [16] to belong to a particular phase were used as the code categories. Next, we used axial coding to connect social media affordances to the various disaster management phases. We categorized all the social media affordances into eight classes using Chen et al.’s [44] technology (social media) affordances. The eight classes of affordances are *visibility, persistence, association, metavoicing, generability, communicability, interactivity,* and *collaboration.* Then we used selective coding to define a holistic framework for social media affordances for disaster management. All the articles were coded using NVivo software package because of its coding capabilities which significantly aided analysing process of this study. The last stage was to do a comprehensive analysis of the selected articles. The discussion of results is presented in the next section.

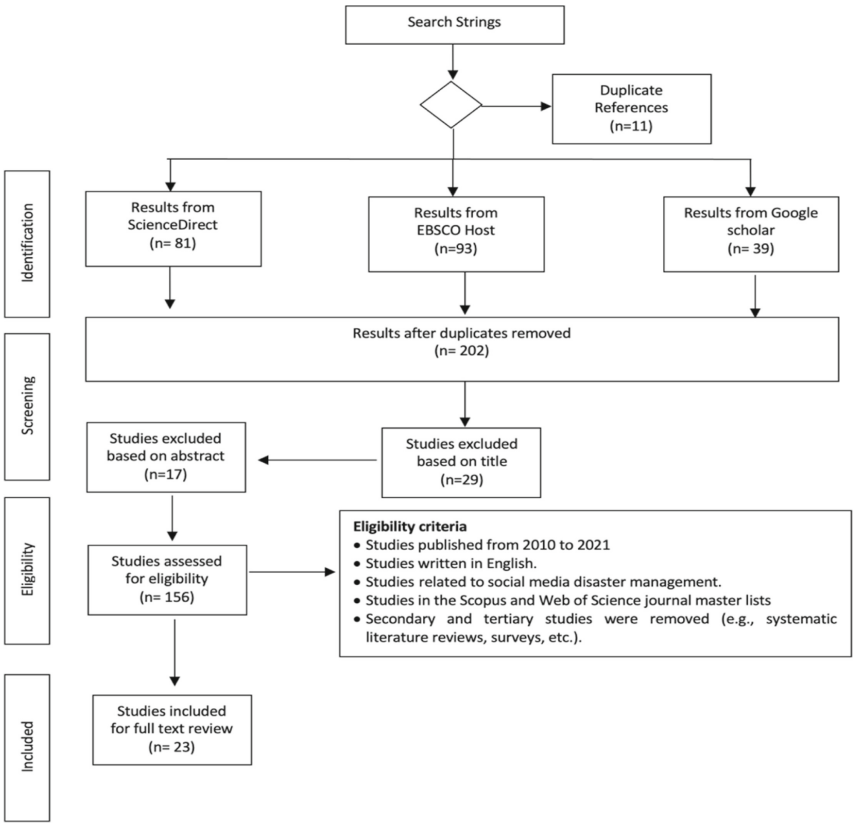


Fig. 2. Search and selection process for articles

3 Data Analysis and Results

In Table 3, the distribution of the analyzed works is presented which clearly shows a majority of the studies (82%) researched on the use of social media in the response phase of disaster management while 30.1% of the studies were done at the recovery phase. This shows that social media was largely used for disaster management at the response and recovery stage of disasters which leaves a gap on social media use at the mitigation and preparation phases.

We further analyzed the data based on the technology affordances theory (Chen et al. [44]). Eight of the social media affordances are relevant to this study which are *visibility, persistence, metavoicing, generability, association, communicability, interactivity, and collaboration.* Apart from opting for affordances which offered relevant insights to the research questions, we selected affordances which are more appropriate to disaster management and are applicable to organizations over those which are more applicable to other domains and individuals. For instance, we excluded anonymity affordance because it is more suitable for discussing whistleblowing [57, 76]. In the reviewed papers, the major research emphasis is on the visibility affordance of social media. Other affordances

Table 3. Distribution of the sources used to examine social media use in different phases of disaster management

Source	Type of social media	Disaster management phase	Purposes	Affordances identified
Alexander [54]	Facebook, Twitter	Response	Monitor and listen (gather) disaster information. Initiate crowdsource	Visibility, Persistence, Interactivity, Collaboration
Al-Saggaf, Simmons [55]	Facebook, YouTube	Response, Recovery	Disaster information dissemination and damage assessment after disaster	Visibility
Brandt et al. [56]	Twitter	Preparedness, Response, Recovery	Provide description of disaster; Provide information on resource distribution; Provide information on volunteerism; Provide information on action to reduce threats	Persistence, Visibility, Association, Collaboration
Chumani, Leon [57]	Twitter	Response	Provide disaster response information, Public engagement	Visibility, Persistence, Metavoicing, Interactivity
Dong et al. [58]	Twitter	Response	Sentiment analysis	Persistence
Ehnis, Bunker [59]	Facebook	Response	Information dissemination, warning broadcast, appeal for information	Visibility, persistence, interactivity
Kongthon et al. [60]	Twitter	Response, Recovery	Provide disaster information and news coverage analysis	Visibility
Landwehr et al. [61]	Twitter	Preparedness, Response	Data collection and early warning	Visibility, Persistence, Interactivity
Loni et al. [62]	Twitter	Response	Disaster information dissemination	Visibility
Loni et al. [63]	Twitter	Response	Provide vital information to affected communities	Visibility
Lovari, Bowen [64]	Facebook, Twitter	Response	Provide disaster information	Visibility
Liu, Zhang, Zhang [65]	Sina Weibo	Response	Disaster information dissemination and communication of emotions	Visibility
Mauroner, Heudorfer [66]	Facebook, Twitter	Preparedness, Response	Provide early warning and information on relief activities	Visibility

(*continued*)

Table 3. (*continued*)

Source	Type of social media	Disaster management phase	Purposes	Affordances identified
Martin, Li, Cutter [67]	Twitter	Response	Provide and gather disaster response information	Visibility, Persistence, Metavoicing, Communicability, Interactivity
Mitcham, Taylor, Harris [68]	Facebook, Twitter	Response	Disseminate essential information	Visibility
Niles et al. [69]	Twitter	Preparedness, Recovery	Information sharing and dissemination	Visibility
Oo-Tha, Pan, Sandeep [70]	Facebook	Response	Provide crisis awareness, relief coordination, volunteer mobilization and initiate crowdsourcing	Visibility, Communicability, Interactivity
Page-Tan [71]	Nextdoor	Preparedness, Response, Recovery	Provide disaster information and warning and resource mobilozation	Visibility
Panagiotopoulos et al. [72]	Twitter	Response	Dissemination of critical disaster information	Visibility
Tagliacozzo, Magni [73]	Facebook, Twitter, YouTube	Recovery	Provide information about housing, infrastructure and financing	Visibility
Sutton et al. [74]	Twitter	Response	Provide disaster information and warning about the event, its impact, and guidelines for self-protection	Visibility
Xiao, Huang, Wu [75]	Twitter	Recovery	Correlate disaster information to disaster impact	Visibility, Persistence
Yates, Paquette [10]	SharePoint, MediaWikis	Response	Knowledge sharing	Visibility

such as persistence, association, metavoicing, generability, communicability, interactivity, and collaboration have been given limited consideration. Thus, the awareness of DMOs is limited to only the visibility affordance of social media, therefore, disaster managers are not able to utilize the full potential of social media. Social media is a powerful tool that has the potential to be implemented in all phases of disaster management and not only in the response and recovery phases as largely reported in the literature. Therefore, in the next section we propose a social media framework for understanding the different affordances that exist in the disaster management phases which can assist DMOs to realize the full potential of social media in managing disaster events.

4 Social Media Affordances for Disaster Management

One of the objectives of this paper is to develop a framework for understanding the use of social media functionalities (affordances) for disaster management. Our framework offers an all-inclusive view of various affordances of social media which can be used to understand social media functionalities in all the four phases of disaster management. Largely, existing frameworks (e.g., Liu et al. [65], Stewart, Wilson, [77]) were developed to be used only at the response and recovery phases of disaster management. The proposed framework fills the gap by showing how social media can be used in the other phases. In Fig. 3, we present the proposed social media affordances framework for disaster management.

Description of the framework is provided in the section below, explaining social media affordances in the various phases of disaster management.

5 Affordances in the Mitigation Phase

The mitigation phase of disaster management includes the shaping of public plans to modify the causes of disasters and lessen their effects on lives, property, and infrastructure. Thus, mitigation measures try to prevent hazards from developing into disasters altogether or ease their effects when they occur. From the data analysis, we identified three affordances of social media in the context of disaster mitigation and when these are implemented will lead to affordances in *visibility, communicability*, and *interactivity*. Opportunities that can be manipulated in these affordances include public education, facilitating information exchange with at risk communities while developing a bond with them, and gauging public sentiments to gather data when the disaster strikes. Thus, social media has the affordance of informing, educating and empowering communities with relevant knowledge to guide the general public towards disaster preparedness and risk reduction.

Affordances in the Preparedness Phase

The goal of disaster preparedness activities is to attain an adequate level of readiness to respond to any disaster event through programs that strengthen the capacities of disaster responders and communities. We identified five different affordances of social media at the preparedness phase which are *visibility, persistence, association, collaboration*, and *interactivity* and when actualized may provide listening capacity (follow public discussions on social media) in order to gather key disaster information and issue. The affordances at this phase offers the platform to develop public emergency preparation pages that people can follow to stay up to date on the latest disaster news and updates. Through *persistence, collaboration*, and *interactivity* affordances, social media offers the technology to detect disaster signals and issue early warning information and provide situation awareness to at risk communities.

Affordances in the Response Phase

The response phase aims to provide immediate assistance to save lives and support the morale of affected communities during disasters. The main task in this phase is provision of emergency assistance in the form of provision of first aid, food, water, clothing,

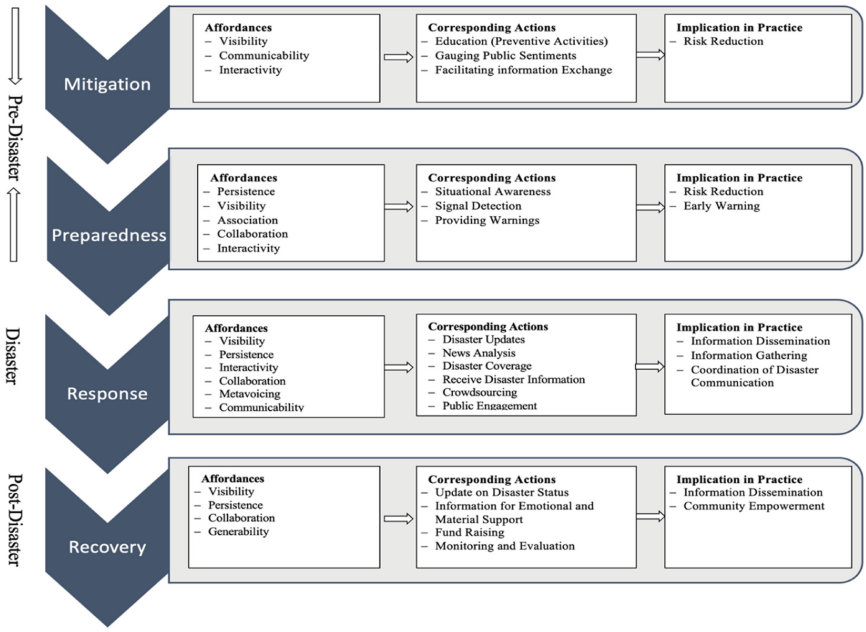


Fig. 3. Social media affordances for disaster management

search, rescue, and evacuation of affected people. In this phase, we identified six affordances which are *visibility, persistence, interactivity, collaboration, metavoicing,* and *communicability* which can be implemented for information dissemination, information gathering, and coordination of disaster communication. Thus, social media have the affordances of providing information on the basic facts about the disaster and the measures being taken by government agencies and non-governmental organizations to address the disaster. Affordances of social media can also be implemented to provide information about the extent of the disaster, losses incurred and the current situation of the disaster. Social media offers affordance of crowdsourcing where disaster respondents can utilize to gather information about people who might be affected in certain areas in order to provide assistance and rescue. One of the key affordances of social media in this phase is using it as a tool for providing information and instructions with real-time alerts and warning and posting of various emergency information such as locations of hospitals in need of blood donations, emergency mobile numbers and security information. Risk communication is another major action that can be implemented using social media in this phase of disaster management. Social media provides a platform for efficient disaster risk communication which can aid in preventing losses.

Affordances in the Recovery Phase

Recovery phase entails measures taken after a disaster has occurred in order to restore post-disaster living conditions of affected people and their communities. This phase focuses on reconstruction and rehabilitation of damaged infrastructure, environment, livelihoods, and social and economic systems. From the data analysis, we identified four

affordances in the recovery phase which are *visibility, persistence, collaboration*, and *generability* and when implemented may lead to providing monitoring and evaluation capacity. Hence, social media affordances can be used to monitor the activities and postings of affected communities to establish various situational awareness and make damage estimates. Social media can also be used after a disaster has occurred to facilitate the lessons learned and process useful risk and disaster research. It can also be used to inform the public about the efforts, relief funds and donations available since the affected communities will need basic necessities like food, clothing and shelter.

6 Conclusion, Implications, and Future Research

The affordances of social media make it an effective tool in the coordination and management of communication throughout the phases of disasters. This study shows social media affordances exist in all the phases of disaster management and not just the response phase only as being practice largely by disaster responders. Though the extensive use of social media during the response phase is promising, its minimal use in the other phases of disaster management is a cause for concern. To address the knowledge gap, we suggested a holistic framework to aid in the understanding of the various social media affordances in all phases of disaster management which has implications for both theory and practice. Theoretically, the study contributes to literature on use of technology in disaster management. Studies on the use of social media for disaster management is largely an under-researched area [78, 79] and this study produces a better understanding of utilization of social media services for managing disaster events.

Practically, it provides DMOs and disaster responders with an idea of current and holistic usage of social media in disaster management. The findings of this study show that the full potential of social media has not been used in the domain of disaster management. The study provides policy-makers in disaster management with an understanding of the different disaster management phases and the social media affordances (functionalities) that are best suited for each phase while helping disaster responders to focus on each of the phases. Additionally, the proposed framework in this study offers the developers of social media tools with a picture of the functionalities in each of the disaster management phases that can help them to optimize social media services for each of the phases.

Essentially, the holistic use of social media functionalities in all disaster management phases can minimize the impact of disasters and help in managing them better towards achieving sustainable development goals (SDGs) especially SDG 9 (Industry, Innovation and Infrastructure) and SDG 11(Sustainable Cities and Communities). The increasing number of disaster events in the past decade (e.g., COVID-19, floods, bush fires, etc.) has underscored the need for sustainable disaster management. In sustainable disaster management approaches, it is necessary to support coordination among the diverse actors including technology, industry, policy and society. This implies that it is critical to align disaster management with the needs of society, anticipate environmental, ethical, and societal concerns and be responsive to them, and enhance these efforts by engaging with different stakeholders. Information and Communication Technology

(ICT) in general and social media in particular are essential tools for drawing sustainable disaster management framework [80] and the proposed framework in this article contributes to such efforts.

Directions for future research include testing the proposed framework to examine how various social media platforms play roles in the various disaster management phases and their associated usage benefits in each of the phases.

Acknowledgements. This work was financially supported by Hasso Plattner Institute for Digital Engineering, through the HPI Research School at UCT.

References

1. Centre for Research on the Epidemiology of Disasters: 2021 Disasters in numbers. Brussels. Brussels: CRED. https://cred.be/downloadFile.php?file=sites/default/files/2021_EMDAT_report.pdf (2022)
2. Baas, S., Ramasamy, S., DePryck, J.D., Battista, F.: Disaster Risk Management Systems Analysis. Food and Agriculture Organization, Rome, Italy (2008)
3. Shahbazi, M., Ehnis, C., Shahbazi, M., Bunker, D.: Tweeting from the shadows: Social media convergence behaviour during the 2017 Iran-Iraq earthquake. ISCRAM Asia Pacific. Wellington. http://onimap.citizenlab.org/filtering-pol.html (2018). Last accessed 23 Apr 2022
4. Mirbabaie, M., Youn, S.: Exploring sense-making activities in crisis situations. In: 10th Multikonferenz Wirtschaftsinformatik.,pp. 1656–1667. Luneburg, Germany. http://mkwi2018.leuphana.de/wp-content/uploads/MKWI_269.pdf (2018). Last accessed 14 Apr 2022
5. Ehnis, C.: Social media within emergency management organisations – a case study exploring social media utilisation for emergency and disaster management. University of Sydney. https://ses.library.usyd.edu.au/handle/2123/17938 (2017). Last accessed 13 Mar 2022
6. Mirbabaie, M., Ehnis, C., Stieglitz, S., Bunker, D., Rose, T.: Digital nudging in social media disaster communication. Inf. Syst. Front. **23**(5), 1097–1113 (2021). https://doi.org/10.1007/s10796-020-10062-z
7. Abdul Mueez, A.S., Mardiana, O., Rosliza, A.M.: Role of social media in disaster mangement. Int. J. Public Health Clin. Sci. **6**(2), 77 (2019). https://doi.org/10.32827/ijphcs.6.2.77
8. Ngamassi, L., Ramakrishnan, T., Rahman, S.: Use of social media for disaster management: a prescriptive framework. J. Organ. End User Comput. **28**(3), 122–140 (2016). https://doi.org/10.4018/JOEUC.2016070108
9. Comunello, F., Mulargia, S.: Social Media in Earthquake-Related Communication: Shake Networks. Emerald Publishing Limited (2018). https://doi.org/10.1108/9781787147911
10. Mitcham, D., Taylor, M., Harris, C.: Utilizing social media for information dispersal during local disasters: the communication hub framework for local emergency management. Int. J. Environ. Res. Public Health **18**, 10784 (2021). https://doi.org/10.3390/ijerph182010784
11. Effing, R., Spil, T.A.: The social strategy cone: towards a framework for evaluating social media strategies. Int. J. Inform. Manage. **36**(1), 1–8 (2016)
12. United Nations International Strategy for Disaster Reduction (UNISDR): UNISDR Terminology on Disaster Risk Reduction (2009). https://www.undrr.org/publication/2009-unisdr-terminology-disaster-risk-reduction
13. FEMA (Federal Emergecny Managenet Agency): FEMA Pub 1 (2010). http://www.fema.gov/pdf/about/pub1.pdf. Last accessed 15 Apr 2022
14. Chikoto, G., Sadiq, A., Fordyce, E.: Disaster mitigation and preparedness: comparison of nonprofit, public and private organizations. Nonprofit Voluntary Sector Q. **42**(2), 391–410 (2012)

15. Islam, S., Chik, Z.: Disaster in Bangladesh and management with advanced information system. Disaster Prev. Manag. **20**(5), 521–530 (2011). https://doi.org/10.1108/09653561111178952
16. Srinivas, H.: Environmental Management and Disaster Risk Reduction. GDRC Reseaarch Output- Continuing Research Series E-018, Global Development Research Center, Kobe, Japan (2015)
17. Kunguma, O., Skinner, J.: Mainstreaming media into disaster risk reduction and management. South Africa. Disaster Advances **10**, 1–11 (2017)
18. Artman, H., Brynielsson, J., Johansson, B. J., Trnka, J.: Dialogical emergency management and strategic awareness in emergency communication. In: 8th International Conference on Information Systems for Crisis Response and Management, ISCRAM 2011. Lisbon (2011)
19. Bruns, A., Burgess, J., Crawford, K., Shaw, F.: Crisis Communication on Twitter in the 2011 South East Queensland Floods. QUT, Brisbane (2012)
20. Ehnis, C., Bunker, D.: Social Media in Disaster Response: Queensland Police Service Public Engagement During the 2011 Floods. ACIS 2012 Proceedings (2012). https://aisel.aisnet.org/acis2012/107. last accessed 2022/03/05.
21. Procter, R., Crump, J., Karstedt, S., Voss, A., Cantijoch, M.: Reading the riots: what were the police doing on Twitter? Policing Soc. **23**(4), 413–436 (2013). https://doi.org/10.1080/10439463.2013.780223
22. Teresa, A.S.: Emergency Managers' Perspectives on Social Media Use for Situational Awareness During Disasters. Walden University (2012)
23. San, Y., Clarence, W., Thorkildsen, Z.: Social media in the emergency management field 12 survey results. http://cdrmaguire.com/emkey/Resources/Social/SocialMedia_EmergencyManagement.pdf (2013). Last accessed 16 Mar 2022
24. Potter, E.: Balancing Conflicting Operational and Communications Priorities: Social Media Use in an Emergency Management Organization. In: ISCRAM 2016 Conference. Rio de Janeiro, Brazil (2016)
25. Dufy, N.: The use of social media in countrywide disaster risk reduction public awareness strategies. Aust. J. Emerg. Manage. **30**(1), 12–16 (2015)
26. Lendrum, D.A.: How People Affected by Disaster use Social Media: A Study of Facebook usage during the 2017 Garden Route Fires. Stellenbosch University, Stellenbosch (2019)
27. Subba, R., Bui, T.: Online convergence behavior, social media communications and crisis response: an empirical study of the 2015 Nepal earthquake police twitter project. In: Proceedings of the 50th Hawaii International Conference on System Sciences, Hawaii, USA (2017)
28. Kankanamge, N., Yigitcanlar, T., Goonetilleke, A.: How engaging are disaster management related social media channels the case of Australian state emergency organisations. Int. J. Disaster Risk Reduction **48**, 101571 (2020)
29. Power, R., Kibell, J.: The social media intelligence analyst for emergency management. In: Proceedings of the 50th Hawaii International Conference on System Sciences, Hawaii, USA (2017)
30. Gibson, J.J.: The Ecological Approach to Visual Perception. Houghton Mifflin, Boston (1979)
31. Norman, D.A.: The Psychology of Everyday Things. Basic Books, New York (1988)
32. Norman, D.A.: The Design of Everyday Things: Revised and, Expanded Doubleday, New York (2013)
33. Markus, M.L., Silver, M.S.: A foundation for the study of IT effects: a new look at DeSanctis and Poole's concepts of structural features and spirit. J. Assoc. Inform. Syst. **9**(10), 609–632 (2008)
34. Zammuto, R.F., Griffith, T.L., Majchrzak, A., Dougherty, D.J., Faraj, S.: Information technology and the changing fabric of organization. Organ. Sci. **18**(5), 749–762 (2007)

35. Majchrzak, A., Markus, M.L.: Technology Affordances and Constraints in Management Information Systems (MIS). Encyclopedia of Management Theory. Sage Publications (2012)
36. Gaver, W. W.: Technology affordances. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM, pp. 79–84 (1991)
37. Hafezieh, N., Eshraghian, F.: Affordance theory in social media research: systematic review and synthesis of the literature. In: Twenty-Fifth European Conference on Information Systems (ECIS), Guimaraes, Portugal (2017)
38. Fayard, A.L., Weeks, J.: Affordances for practice. Information and Organization **24**(4), 236–249 (2014)
39. Strong, D.M., et al.: A theory of organization-EHR affordance actualization. J. Assoc. Inform. Syst. **15**(2), 53–85 (2014). https://doi.org/10.17705/1jais.00353
40. Ellison, N.B., Steinfield, C., Lampe, C.: Connection strategies: social capital implications of Facebook-enabled communication practices. New Media Soc. **13**(6), 873–892 (2011)
41. Mohajerani, A., Baptista, J., Nandhakumar, J.: Exploring the role of social media in importing logics across social contexts: the case of it SMEs in Iran Technological importing logics across social contexts: the case of IT SMEs in Iran. Technol. Forecast. Soc. Change **95**, 16–31 (2015)
42. Majchrzak, A., et al.: The contradictory influence of social media affordances on online communal knowledge sharing. J. Comput.-Med. Commun. **19**(1), 38–55 (2013). http://www.scopus.com/inward/record.url?eid=2-s2.0-84885897490&part-nerID=tZOtx3y1
43. Zheng, Y., Yu, A.: Affordances of social media in collective action: the case of free lunch for children in China. Inform. Syst. J. **26**(3), 289–313 (2016)
44. Chen, Q., Xu, X., Cao, B., Zhang, W.: Social media policies as responses for social media affordances: the case of China. Gov. Inform. Quart. **33**(2), 313–324 (2016). https://doi.org/10.1016/j.giq.2016.04.008
45. Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G.: The PRISMA Group: Preferred reporting items for systematic reviews and meta-analyses: the PRISMA Statement. PLoS Med **6**(6), e1000097 (2009). https://doi.org/10.1371/journal.pmed1000097
46. Kitchenham, B., Pearl, B.O., Budgen, D., Turner, M., Bailey, J., Linkman, S.: Systematic literature reviews in software engineering – A systematic literature review. Inf. Softw. Technol. **51**, 7–15 (2009)
47. Nightingale, A.: A guide to systematic literature reviews. Surg. - Oxf. Int. Ed. **27**, 381–384 (2009)
48. Denyer, D., Tranfield, D.: Producing a systematic review. In: Buchanan, D.A., Bryman, A. (eds.) The Sage Handbook of Organizational Research Methods, pp. 671–689. Sage, Los Angeles, CA (2009)
49. Dumay, J., Cai, L.: Using content analysis as a research methodology for investigating intellectual capital disclosure. J. Intell. Capital **16**(1), 121–155 (2015)
50. Lombardi, R., Secundo, G.: The digital transformation of corporate reporting – a systematic literature review and avenues for future research. Meditari Accountancy Res. **29**(5), 1179–1208 (2020). https://doi.org/10.1108/MEDAR-04-2020-0870
51. Tranfield, D., Denyer, D., Smart, P.: Towards a methodology for developing evidence informed management knowledge by means of systematic review. Brit. J. Manage. **14**(3), 207–222 (2003)
52. Petticrew, M., Roberts, H.: Systematic Reviews in the Social Sciences: A Practical Guide. Blackwell Publishing, Malden, MA (2008)
53. Mergel, I.: Social media adoption and resulting tactics in the U.S. federal government. Gov. Inform. Quart. **30**(2), 123–130 (2013). https://doi.org/10.1016/j.giq.2012.12.004
54. Alexander, D.E.: Social media in disaster risk reduction and crisis management. Sci. Eng. Ethics **20**, 717–733 (2014). https://doi.org/10.1007/s11948-013-9502-z

55. Al-Saggaf, Y., Simmons, P.: Social media in Saudi Arabia: exploring its use during two natural disasters. Technol. Forecast. Soc. Chang **95**, 3–15 (2015). https://doi.org/10.1016/j.techfore.2014.08
56. Brandt, H.M., et al.: Examining the role of Twitter in response and recovery during and after historic flooding in South Carolina. J. Publ. Health Manage. Pract. **25**(5), E6–E12 (2019). https://doi.org/10.1097/PHH.0000000000000841
57. Chumani, W., Leon, T.G.: Social media use for public engagement during the water crisis in Cape Town. Inform. Polity **26**, 441–458 (2021). https://doi.org/10.3233/IP-200273
58. Dong, Z.S., Meng, L., Christenson, L., Fulton, L.: Social media information sharing for natural disaster response. Nat. Hazards **107**(3), 2077–2104 (2021). https://doi.org/10.1007/s11069-021-04528-9
59. Ehnis, C., Bunker, D.: Social media in disaster response: queensland police service – public engagement during the 2011 floods. In: ACIS 2012 Proceedings, vol. 107 . https://aisel.aisnet.org/acis2012/107(2012). Last accessed 21 May 2022
60. Kongthon, A., Haruechaiyasak, C., Pailai, J., Kongyoung, S.: The role of social media during a natural disaster: a case study of the 2011 Thai Flood. Int. J. Innovation Technol. Manage. **1**(03), 14400124 (2014). https://doi.org/10.1142/S0219877014400124
61. Landwehr, P.M., Wei, W., Kowalchuck, M., Carley, K.M.: Using tweets to support disaster planning, warning and response. Saf. Sci. **90**, 33–47 (2016). https://doi.org/10.1016/j.ssci.2016.04.012
62. Loni, H., Thomas, K., Stephen, N., Nic, D., Claudia, R.: Crisis communications in the age of social media: a network analysis of zika-related tweets. Soc. Sci. Comput. Rev. **36**, 523–541 (2017)
63. Loni, H., Stephen, N., Ryan, S., Thomas, E.K.: Social Media Use for Crisis and Emergency Risk Communications During the Zika Health Crisis. Digit. Gov. **1**, 13 (2020). https://doi.org/10.1145/3372021
64. Lovari, A., Bowen, S.A.: Social media in disaster communication: a case study of strategies, barriers, and ethical implications. J. Public Aff. **20**, e1967 (2019). https://doi.org/10.1002/pa.1967
65. Liu, T., Zhang, H., Zhang, H.: The impact of social media on risk communication of disasters-a comparative study based on Sina Weibo blogs related to Tianjin explosion and typhoon pigeon. Int. J. Env. Res. Public Health **17**(3), 883 (2020). https://doi.org/10.3390/ijerph17030883
66. Mauroner, O., Heudorfer, A.: Social media in disaster management: how social media impact the work of volunteer groups and aid organisations in disaster preparation and response. Int. J. Emerg. Manage. **12**(2), 196–217 (2016)
67. Martin, Y., Li, Z., Cutter, S.L.: Leveraging twitter to gauge evacuation compliance: spatiotemporal analysis of hurricane matthew. PLoS ONE **12**(7), 1–22 (2017). https://doi.org/10.1371/journal.pone.0181701
68. Niles, M.T., Emery, B.F., Reagan, A.J., Dodds, P.S., Danforth, C.M.: Social media usage patterns during natural hazards. PLOS ONE **14**(2), e0210484 (2019). https://doi.org/10.1371/journal.pone.0210484
69. Oo-Tha, K. K., Pan, S. L, Sandeep, M.S.: Exploring social media affordances in natural disaster: case study of 2015 myanmar flood. In: ACIS 2017 Proceedings, vol. 109. https://aisel.aisnet.org/acis2017/109 (2017). Last accessed 12 May 2022
70. Page-Tan, C.: The role of social media in disaster recovery following hurricane harvey. J. Homeland Secur. Emerg. Manage. **18**(1), 93–123 (2021). https://doi.org/10.1515/jhsem-2018-0054
71. Panagiotopoulos, P., Barnett, J., Bigdeli, A.Z., Sams, S.: Social media in emergency management: Twitter as a tool for communicating risks to the public. Technol. Forecast. Soc. Change **111**, 86–96 (2016). https://doi.org/10.1016/j.techfore.2016.06

72. Tagliacozzo, S., Magni, M.: Government to Citizens (G2C) communication and use of social media in the post-disaster reconstruction phase. Environ. Hazards **17**(1), 1–20 (2017). https://doi.org/10.1080/17477891.2017.13390
73. Sutton, J., League, C., Sellnow, T.L., Sellnow, D.D.: Terse messaging and public health in the midst of natural disasters: the case of the boulder floods. Health Commun. **30**(2), 135–143 (2015). https://doi.org/10.1080/10410236.2014.974124
74. Xiao, Y., Huang, Q., Wu, K.: Understanding social media data for disaster management. Nat. Hazards **79**, 1663–1679 (2015). https://doi.org/10.1007/s11069-015-1918-0
75. Yates, D., Paquette, S.: Emergency knowledge management and social media technologies: a case study of the 2010 Haitian earthquake. Int. J. Inform. Manage. **31**(1), 6–13 (2011). https://doi.org/10.1016/j.ijinfomgt.2010.1
76. Lam, H., Harcourt, M.: Whistle-blowing in the digital era: motives, issues and recommendations. New Technol., Work Employ. **34**(2), 174–190 (2019)
77. Stewart, M.C., Wilson, B.G.: The dynamic role of social media during hurricane# sandy: an introduction of the STREMII model to weather the storm of the crisis lifecycle. Comput. Human Behav. **54**, 639–646 (2016)
78. Graham, M.W., Avery, E.J., Park, S.: The role of social media in local government crisis communications. Public Relat. Rev. **41**(3), 386–394 (2015)
79. Potts, L.: Social media in disaster response: how experience architects can build for participation. Routledge, Taylor, Francis Group, New York (2014)
80. D'Andrea, A., Grifoni, P., Ferri, F.: Discussing the role of ICT in sustainable disaster management. Sustainability **14**(12), 7182 (2022). https://doi.org/10.3390/su14127182

Adopting Smart Technologies of Industry 4.0 to Formulate Data for Enhanced Business Intelligence

Sean Kruger(✉) and Adriana A. Steyn

University of Pretoria, Pretoria Gauteng 0083, South Africa
sean.kruger@tuks.co.za

Abstract. Emerging technologies of Industry 4.0 (I4.0), often referred to as "smart", have been shown to offer advantages across several business sectors when effectively adopted. Despite the adoption thereof being a well-established field, gaps remain on how these smart technologies have been combined to formulate insightful data to enhance business intelligence (BI), particularly in non-western contexts. To address this gap, a case study showing how smart technologies were adopted for the agriculture sector in South Africa to generate data towards improved BI is presented, including the processes that enabled this. The findings show how technologies were integrated to develop an artificially intelligent system that could identify diseased animals more effectively. The study does not argue that the system itself is revolutionary. Instead, it demonstrates that technology adoption remains relevant, even in the context of the study to enable decision making. Using these insights, and data from semi-structured focus group sessions, a conceptual model is presented to replicate such solutions in developing regions towards enhanced data driven decisions.

Keywords: Artificial intelligence · Big Data · Business intelligence · Data driven decisions · Fourth industrial revolution · Industry 4.0 · Information systems · SDG8

1 Introduction

With the expansion of disruptive technologies integrating into physical, digital and biological facets of our lives, several challenges are facing business leaders, government and academia [1–3]. Nonetheless, there are several opportunities that can be grasped if these technologies can be effectively combined or integrated into existing systems and develop needed data [4]. This advancement can be attributed to the paradigm known as the Fourth Industrial Revolution (4IR), also referred to as Industry 4.0 (I4.0) [5, 6]. Notwithstanding, for organisations, as well as individuals to innovate within this paradigm, the decision to adopt I4.0 technologies, such as the Internet of Things (IoT), cloud computing, big data and analytics, artificial intelligence (AI) and robotics is critical [1, 7].

P. Ndayizigamiye et al. (Eds.): IDIA 2022, CCIS 1774, pp. 154–171, 2023.
https://doi.org/10.1007/978-3-031-28472-4_10

One area of opportunity to drive productivity, efficiency or even create new forms of value using these emerging "smart" technologies is business intelligence (BI) [8]. BI has assisted leaders in making more accurate, timely and smart decisions to increase profitability, productivity and effectiveness of business operations [9]. This further increases the potential for organisations to reduce certain workforce areas and eliminate inadequacies across systems [10, 11]. A core pillar of BI is new, reliable and insightful forms of data which I4.0 technologies have shown to enable. However, "traditional models" to improve BI with smart technologies has varied between technologies, regions, applications and systems, making the decision on which is the most applicable difficult to navigate. Moreover, the models have focused on developed regions, leaving limited indigenous examples. Perhaps not an oversight, as several models do exist confirming tacit understanding that local systems and mechanisms do support data development for improved decisions [12]. This study acknowledges this and assumes that the model's used can be applicable in this study's context due to the extensive application to date. One model in particular has seen usage in this paradigm and in some developing regional contexts, the Technology Acceptance Model (TAM) [13–16]. For this reason, it is used within this study.

Notwithstanding, practical insights into I4.0 specific usages remains lacking, especially on the interconnected applications across the physical and digital within developing regions to create BI for improved decision making. As a result, despite the possibilities, adoption cases that demonstrate how to leverage smart technologies of this paradigm remain limited, particularly in the academic field [17]. As a result, the following question is considered: "*How are smart technologies of I4.0 being used by entrepreneurs to develop BI in developing regions such as South Africa (SA)?*".

To address this question, a case study of an entrepreneurial firm that developed a solution for the agriculture sector was conducted to develop insights to begin addressing this gap. Both entities, the entrepreneurial company that developed the solution and farm it was tested on were registered in SA at the time of writing. The case study data was split into two sections. The first section demonstrates how smart technologies were integrated to create an artificially intelligent system that could detect diseased animals. The systems accuracy, processes and applications are then shown. The second section uses qualitative data from semi-structured focus group sessions to capture the processes used to integrate the smart technologies towards BI. The purpose; provide insights for industry practitioners, business leaders and entrepreneurs on how I4.0 technologies can be used, but also the processes to formulate data for BI. The main contribution is the proposed conceptual model that incorporates the TAM, so it can be tested in other contexts and industries towards data driven decisions.

To achieve this, the paper is structured as follows. Firstly, there is a review of existing literature that briefly covers smart technologies stemming from I4.0 which are applicable to this study. Innovation and BI is then reviewed noting how entrepreneurs are key stakeholders of innovation. The research methodology is then presented, followed by a review of findings where the conceptual model is presented. Finally, conclusions are drawn, and directions for further research presented.

2 Theoretical Background

The Global Innovation Index recognizes business development, intellectual property and patent cooperation treaty applications as key areas towards improving a countries innovation performance [18]. For a country like SA, initiatives that can support such outcomes has been noted as critical, where innovation is seen as a strategic mechanism to support economic growth [19]. For example, there has been extensive focus and investment in I4.0 technology development within the country. Over 75% of technology start-ups have been funded by government at higher education institutions and science councils between 2008 and 2014 [20]. Innovation enablement through technology adoption can be argued to be at a critical point [21], including that of SA as a developing country, to stimulate decent work and economic growth [22], showing relation to Sustainable Development Goal (SDG) 8 [23]. The key stakeholders to drive this are entrepreneurs as they are strongly associated with innovation, including the capacity to innovate using smart technologies of I4.0 [1]. One enablement pillar to do so is data, where improved BI is a key outcome [17]. This case study, although limited in scope, provides insights into processes to generate data using smart technologies for BI to further engage in this paradigm.

2.1 Enabling Technologies of I4.0 in Agriculture

One market sector where innovation has been driven by various entrepreneurial endeavours to increase profits, and is critical for a sustainable environment, is agriculture. However, because of rapid technology development, existing business models are becoming redundant in this sector [24]. To address these redundancies, monitoring equipment, predictive analytics, forecasting, smart logistics and warehousing are being optimized by embedding smart technologies across systems to increase efficiency and yield [25]. This requires a combination of smart devices and systems towards intelligence, where they are purposely designed to automate either physical, technical or intellectual functions [17].

One of which is the Internet of Things (IoT) as it links physical infrastructure to computational capabilities of an organisation, enabling data collection, interaction and connectivity on a physical level [25, 26]. This allows devices to monitor and sense the physical environment permitting an interaction between the digital and physical realm [27]. By putting intelligence into everyday objects, they can collect information from the environment and interact, or even control the physical world [28]. Considering the agriculture sector, it is said the world will need to produce 70% more food in 2050 than it did in 2006, to feed the growing population [29]. With the need to overcome water and land shortages whilst managing costs in a sustainable way, “smart agriculture” has been crafted through advanced technologies to deliver solutions to these problems [30]. Smart agriculture is replacing older, more time-consuming farming techniques with new ones to ensure maximum yield with less wastage [31]. This is supported by IoTs, as sensors provide real-time feedback on minerals, humidity levels and soil moisture to predict the best time to harvest, aligned with meteorological changes and resource availability for better decision making [32]. To date this has occurred and been well adopted in

developing regions. Although indigenous examples do exist, they are more limited and also note several barriers on a macroeconomic level that affects uptake [33, 34].

Cloud systems are seeing an increase to manage and store data from IoT devices, but also facilitate digital technologies. They are data-driven applications that allow real-time data processing, access to resources and infrastructure without requiring hardware [21]. These systems enable the storage, evaluation and access to data [35] which can facilitate elastic resource scaling with minimum costing provision. This can include processing of Big Data, which is data that would typically be too expensive to store, manage, and analyse using traditional database systems and software. [36]. In the case study used, unstructured data allowed intelligence to be generated in areas needed for the entrepreneurs to create a more effective agriculture environment whilst factoring in constrained resources [37]. Despite the setbacks and needs for effective management protocols, the data offered real value in producing meaningful information which allowed the improvement of BI [8], at the core of which was AI.

AI is a novel and smart technology, as it can learn from previous situations, provide input, automate complex decision processes and produce predictive feedback for smarter decisions based on data and past experiences [38]. Despite these advantages, concerns around AI have been raised [39]. This includes considering who is accountable to build and govern its functions, associated job losses where humans can be replaced and its activities [6]. Despite this, modern AI has changed the paradigm of people instructing machines through rules to solve problems, to one where the machines can solve problems themselves. This is based on algorithms, where an algorithm is a self-contained sequence of instructions performed by a computational machine [38]. At the basis of AI is data, and with Big Data, new methods to analyse data was identified as a need. To address this, ML was developed. ML is aimed towards the generation of knowledge by increasing the accuracy of data and learning from it [24]. It differs from AI in that it focusses on the accuracy of data [40]. ML has expanded to allow computers to process languages and images, which was not before possible [41]. With this enablement, AI can operate with increasing autonomy and capability as it can base decisions on more accurate data [7]. By using or combining technologies such as ML and cloud systems to process data, AI could be incorporated on varying organizational levels for effective decisions to be taken [38], as was the instance in this case study.

While these smart technologies have been shown as beneficial, with several areas that have enhanced innovation, it has proved challenging to navigate without the correct skills and resources [1, 34]. There are several studies that are driving the movement and usage of these very technologies to an inclusive Society 5.0 [42]. Notwithstanding, there is evidence arguing that digital exclusion is a growing concern resulting from the 4IR paradigm, especially in the Sub-Saharan region. The authors advocate the usage and advancements brought on by these technologies. However, note that the lower levels of understanding to integrate them alongside existing ecosystems and cultural dynamics poses a major risk for the region. For example, lower skills have seen lower adoption rates, as the usefulness alongside level of effort required is often not understood, stifling innovation and engagement within the paradigm. The result of this is a country falling behind within a globally competitive arena.

3 Methodology

The purpose of this study was to demonstrate an instance of successful adoption using smart technologies in a developing region and capture the processes that enabled this. The specific outcome was an artificially intelligent system that could identify diseased animals more effectively. The study does not argue that the system created is revolutionary. Instead, it demonstrates that technology adoption is relevant, even in the context of the study which is in the developing region of Sub-Sahara Africa. If done well, sectors such as agriculture can be improved with data insights generated. To capture a real-life example of such application, and identify how data was generated to enhance BI, a case study approach was used [4, 17, 43]. The case study, although analysed in depth, does have several limitations such as the reliance and scalability of the system developed. Moreover, due to ethical limitations, the naming, specific location, participant demographics and exact revenue of the farm could not be disclosed.

The case study focused on a bounded entity in SA, an animal farm within the Gauteng province. The farm trades primarily with dairy cattle and domestic pigs. It also produces maize for feed. The farm has over one thousand hectares utilised and can be considered a Small-to-Medium (SMME) based on its revenue. The geographic profile is mixed, with approximately 40% of the farm having difficult terrain, with limited access by road. The farm had limited resources in terms of staffing during the time of assessment, with lower skilled employees who were not always able to take note of diseased livestock. The requirement was data to support the identification of diseased animals, as they were considered the assets of the business, and needed to be healthy to have higher value. The farm was selected as the owner requested assistance in developing an innovative solution and allowed the tracking of the processes captured. For the purposes of the study, the cross-sectional time horizon was applied [38].

Data collection for the case study was two-fold [4]. The first section reviews the artificially intelligent system that could detect diseased animals' level of accuracy, processes and applications. The data to build the system was based on quantitative data from IoT devices and images collected. The control units (IoT devices) selected were M5Stacks with LTE Modules (M8321). They were configured to communicate and transmit live data within a frequency band of LTE to GSM (Mhz), depending on the area. This allowed for monitoring and feedback on power usage, temperature and time. Geographical location using a Global Satellite Positioning System (GPS) NEO-M8N module was also added to know where the information was gathered, and a photoresistor light-controlled variable resistor to measure the level of light. The light sensor provided ambiance and calibration data. The units were configured using the Arduino Integrated Development Environment (IDE), which is based on C++ text to code. To deploy the M5 units efficiently and collect the data across geographical locations, they were attached to a drone. The drone allowed for access over difficult terrain, especially where foliage was apparent and deployment at a higher altitude was needed. The drone used was a DJI Phantom 4 Pro v2. This offered efficient propulsion as well as a noise reduction over the previous model. The unit has Taply™ and ActiveTrack™ to control flying and return to an original position when the battery is low. This equipment ensured a flight time of around 30 min to cover the necessary area of around 500 hectares at the time of testing. In South Africa, drones are aircraft which means users need to adhere

to the South African Civil Aviation Authority (SACAA) of Remotely Piloted Aircraft System (RPAS) Part 101 regulation when they are used for commercial purposes [44]. Despite the drone unit having a camera, infrared thermography (IRT) imaging camera was added. This allowed for a non-destructive testing technology, which has been used to determine the superficial temperature of objects [45]. The visualization of temperature distribution has been shown to detect peripheral blood flow, which is a tool used to detect stressed or diseased animals [45]. There are also various other studies that attribute IRT technology application towards the identification of diseased animals across several species [45]. For the ML tool to be built, the parameters of body temperatures and thermal exchanges were incorporated into the system, where the material emissivity coefficient was assumed to be constant. Various versions and organizations provide IRT equipment, but for the case study, a Flir Ultra Force 275 was selected due to its optimum combination between capability, weight, power and footprint-constrained platforms. The system delivered required high-definition (HD) images in lowlight conditions. For the data to be accessible and usable, it was decided to be stored on a cloud platform. The platform selected allowed Software as a Service (SaaS), which includes AI and ML, as well as options to pre-process data in the pipeline through relevant transform functions. The transmission of data from the IoT devices could be stored directly onto the cloud from the M5Stack using the LTE module. This also included metadata from the IRT device. The full images, however, were first stored on a local device, and then transmitted onto the cloud. This conserved data fees.

The second set of data was qualitative, to capture the processes of creating, testing and formulating insights from the data for BI. This was attained from semi-structured focus group sessions. The group sessions were structured to note how the solution was developed, the technical aspects of the solution, the data insights generated and how I4.0 technologies were integrated to create BI. The key processes that enabled this were noted and categorised into activity and function. There were 24 participants, all of whom were directly part of the solutions development team and included the farms owner. 23 participants were from the entrepreneurial firm who collaborated with an academic makerspace to create the solution presented. This was how the researchers were able to engage within the process and collect the data. The reason for this number was that the participants all formed part of the team involved in development, usage and understanding of the solution. Farm workers who were not involved or did not understand the technology were excluded. Although the sample is limited, it did provide a range of insights due to the participants active involvement in the case, where barriers, successes and nuances of the value created could be noted.

The analysis was based on the TAM model from Davis (1985) [46]. TAM has been used to explain several underlying factors of user acceptance of technologies, especially in the I4.0 [47]. Furthermore, it has been proven useful in empirical investigations as well as theory development for new technologies to explain the adoption process. Although it is considered a strong non-native assessment tool, it has proven useful and seen application in various situations, hence its selection for this indigenous case study [13, 48, 49]. Furthermore, because the case study involved the integration of I4.0 technologies, TAM was considered a useful and applicable theory [50]. From TAM, two constructs were used to identify behavioural intention of users. The first was perceived usefulness (PU),

which explains the belief of an individual user of how such a technology would be useful to them to improve their job or abilities to do their job. Secondly, perceived ease of use (PEoU), as it explains the level of effort needed to use the system or technology [51]. Both these have been shown to be affected by skills of the users, making it relevant for the case study's context. Figure 1 demonstrates the overview of the methodology to extract needed data from the case based on the TAM model. From the research methodology, a holistic overview of the functions of BI using I4.0 technologies could be determined towards improved data driven decisions.

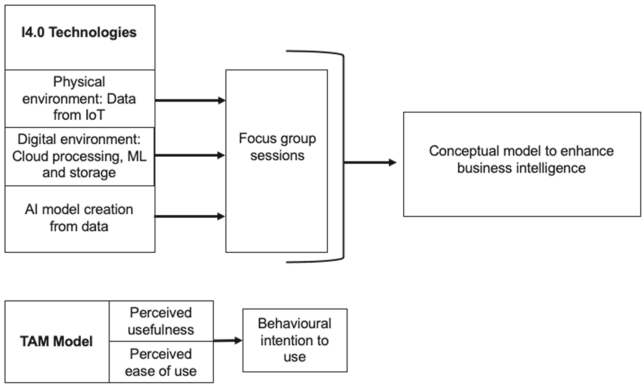


Fig. 1. Overview of methodology to create a conceptual model

4 Findings

It was found from the processed images and associated metadata gathered from the IoT devices, that useful patterns in the data could be identified using ML to enhance accuracy for the identification of diseased animals. The patterns were based in a Cloud Vision API to detect multiple objects from an image. This allowed the translation of thermal images to decide which animals are infected. Boundary rates were used to note the instances with similar features to include or exclude them, allowing classes to be set. Based on the cloud object localization, ML was able to analyse images and assign objects between the animal and its surroundings. One example of the above analysis and level of accuracy is demonstrated in Table 1 below. This shows the identified object, association through mid-code, where the score associated refers to accuracy. Despite the accuracy scores achieved (which were relatively high), the identification and parameters of diseased animals remain under investigation.

It should be noted that before the image data could begin to be optimized, data cleaning was required. This used prediction, detection and classification parameters to ensure that only relevant images were assigned in the cloud for ML purposes. The parameters themselves for image processing involved manipulating digital images to improve quality, correct lighting problems and assign required metadata. This started with low-level processing, which processed grayscale adjustment, focus correction, contrast and

Table 1. Outputs of ML analysis

Identifier	Mid code	Classification accuracy score	Classification boundary rates
Animal	/m/01k0	0.845448587	(0.32468076266, 0.78459418387), (0.47838912272, 0.7894741387), (0.43812272, 0.9733661065)
Bovine	/m/0489g	0.8442345	(0.312, 0.661647671), (0.634583853, 0.661647871), (0.638380053, 0.9705800882)
Landscape	/m/01thn0	0.6353456275	(0.5125398, 0.760708), (0.6256646, 0.760708), (0.6256646, 0.94601655)

noise reduction. The middle-level processing involved segmentation, description and classification of objects present in the image. Finally, high-level processes, or image classification, involved the recognition and classification of regions of interest that are usually performed by statistical classifiers or neural networks. The process is shown in Fig. 2.

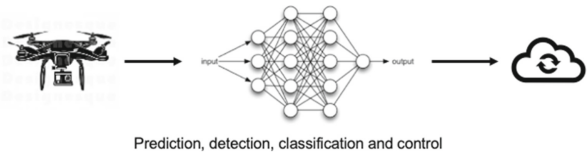


Fig. 2. Process overview of images from the collection to processing, classification and storage on a cloud system

The scores were based on the ability of the ML algorithm to classify assigned categories (such as “domestic pig” or “bovine” and “landscape). Over 12 000 processed images from a range of 10 dairy cattle were used. Of the 10 bovines, one was diagnosed by a veterinarian with oestrus. The system was later deployed to test the distinguishment between bovine and wildlife. This was a limited feature due to budgetary constraints and need to focus on diseased animals. For the diseased animal to be identified, classification was assigned based on the thermoregulation of animals, where there are energy balances between them maintaining homeostasis and their surroundings. Where thermal changes occur, psychological mechanisms are changed, increasing energy expenditure rates, for example, temperature. Bovine research has seen IRT used for diagnosis purposes for oestrus, mastitis and lameness. The focus has been on both bodies and eye temperatures, but for this study it was for the overall temperature relative to the environment. The classification accuracy ratio used was the number of correct predictions to the total number of input samples [52]. The reason the investigation will continue is the cost of misclassification of the minor class samples could be very high, as the classification considers diseased animals and that associated risk.

From the ML outputs, an AI could then be developed to flag risk areas. This model for selection is conceptually not limited to animals such as sheep, pigs or cattle, but can be extended to wildlife. The IRT parameters are based on research and are routinely used in veterinary practice to identify problem areas [45]. The basis of this is that mammals maintain a constant body temperature regardless of their external environment, as they are endothermic. The AI success does depend heavily on the ML sensitivity ratings, as environmental conditions need to be improved. This can be improved over time with metadata collected from the IoT devices. Figure 3 demonstrates the overall process used to collect, process, store and analyse data for BI in this case.

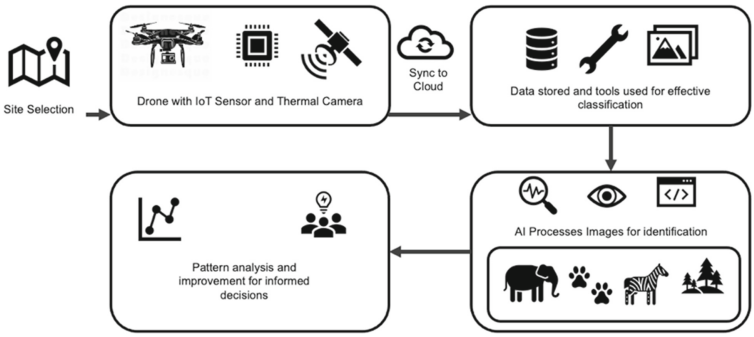


Fig. 3. Overview of technical process

The system has several limitations and should not be taken as a replacement of traditional diagnosis methods. The traditional hands-on methods were not suspended, instead integrated to identify the diseased animal. The input sample has several areas for expansion, where weather conditions, lightning and various other breeds were controlled as far as possible but could impact the classification accuracy ratio. The investigation and system development should continue but was limited to this deliverable due to costing and the project ending.

The second set of data was from semi-structured focus group sessions with 24 participants. It was found that there are certain processes needed to create the system that could enhance needed BI. The data findings from the focus groups are shown in Table 2, which indicates the activities alongside associated functions that enabled the adoption of these technologies. Key to which was the support the of the academic makerspace that developed the solutions alongside the entrepreneurial firm. This is because it forms part of the region's innovation ecosystem.

Table 2. Process flow from focus group sessions to enhance BI with smart technologies

Business activity	Function
Business needs analysis	Identify what the actual need is. Define the information needed to deliver BI that aligns with strategy
Current system analysis	Analyse current systems to avoid infrastructure misalignment. This can also assist with reducing duplication of existing information
Data needs analysis	Define what data is required and the exact methods to collect it
Assign technologies	From the data collection needs, determine the most appropriate technology. Considerations to include are budget, accuracy, flexibility, design needs and resource availability in terms of skills. This was noted as the most vital section in terms of which novel technology to apply
Develop solutions for the collection: Assess collaborative partnerships	This includes the development of a solution through rapid prototyping. Various technology levels need to be considered and defined, as well as drive innovative solutions relevant to the business context. Assessment and collaborative partnerships can be leveraged within ecosystems to aid in development
Optimize solutions and test	Once a technology has been tested, the actual collection needs to occur. Data sets with main variables need to be captured and stored to analyse and design the model. This includes Big Data with associated metadata (GPS location, temperature, weather condition, time, date) needed for ML
Obtain data	Once data has been obtained, it can be stored in data warehousing for analysis to create BI
Train data sets	Data is usually prepared by being collected in a semi-automated way, where there are various software tools applicable. This is the primary area where ML is used to ensure accurate data is available to train the AI model for decision making
Test data sets	Testing is required to determine the model quality

(*continued*)

Table 2. (*continued*)

Business activity	Function
Validation test suites and scenarios for analysis	This requires building suites that verify the models and that the model algorithm can analyse the data provided to identify specific patterns and results. This allows for needed information of BI to ensure knowledge is created
Communicate results	Stakeholders need to be informed of the information and ensure it delivers on business needs, and that decision-makers have the needed BI. The AI model can be trained to distribute the relevant and actionable information
Enhance and further develop	This is to ensure the services or modifications can deliver effective changes for business enhancement and performance management

5 Discussion

The case study provided an I4.0 technology adoption case within a South African context, where such adoption remains a relevant pillar in supporting innovation. The TAM model was selected due to its extensive application abroad, and it appears applicable within non-western regions. The purpose of the system developed was not to reduce workforce areas, rather it enhanced the stakeholder's ability to identify diseased animals and be advised on how to act accordingly. This generated not only positive outcomes for the animals, but from a business perspective, assets (livestock) could be managed in a more efficient manner through BI. This was done despite the resource limitations of the case study.

The case study showed how the technologies of I4.0 including AI, IoT and cloud systems can be integrated to generate data insights for improved BI. The application of the technologies within the study was dependent not only on the usefulness or stage of the technologies maturity, but also how they would integrate into existing system. The context of the study was a key point in addressing a gap in literature, which was a developing region that supported smart agriculture towards sustainable economic and environmental benefits. They were combined to identify livestock health in difficult areas. The images processed demonstrated acceptable accuracy in terms of animal identification and temperatures relative the surrounding environments. The temperature being an important indicator for their phycological status. In practice, the animals were identified to a level of acceptable accuracy for the project deliverable, but also accurate temperature measured for skin levels. The ranges of these parameters are seeing expansion in the field

of veterinary practices, where the forecasting can be integrated into such systems. A positive variable in this case was that of the reducing prices of thermographic equipment as well as their ability to connect to platforms for evaluating surface temperatures. The derivatives and other fields of application could prove worthwhile not only for the region, but in several other contexts. This is vital for the developing region as smart agriculture can look to address ever growing food security and environmental concerns.

Practically, the case study captured the process used regarding how entrepreneurs can drive data decisions through key technology adoption. It also demonstrated the capability in such regions does exist, where adoption is occurring. It also noted that collaboration is a key point, where an innovation mechanism was used in the adoption process, as it offered access to needed skills to adopt technology as noted in Table 2. In this instance, the AI system that could improve detection of diseased animals. At the basis of this was IoT, where computational abilities were leveraged to connect devices and collect new forms of data that were required to develop the system. This in turn shows a practical example of how these objects were integrated to create new forms of value. This is a likely reason to why there is an exponential interest in smart technologies. Moreover, although not novel from a global perspective, within a developing country perspective, by creating a system where these technologies integrated, a needed product could be formed. The purpose was to support decisions in an industry that is vital for the country of SA, smart agriculture, by showing how resource monitoring and tracking could improve decisions based on BI in real-time.

Theoretically, within the study context, several barriers are apparent, such as individual levels of bias. When dealing with smart technologies, skills levels saw a direct influence on the behavioural intention to adopt. Notwithstanding, there are a few contributions. Firstly, skills to develop such solutions is critical. In this sense, research and development (R&D) areas and innovation environments such as academic makerspaces are still relevant where mutually beneficial collaborative efforts align and they themselves have highly skilled individuals to enhance PU and PEoU. This aligns with developed regions which use such mechanisms to develop innovation capability [53] but also developed regions who use Digital Innovation Hubs [54]. Secondly, smart technologies have a key role in improving innovation with I4.0 technologies, even in a developing region. These technologies have furthered our ability to innovate by being able to integrate not only on a physical level, but a digital one as well. In so doing, certain dimensions of smart technology adoption from an information systems perspective are noted in Table 3.

Although the sample pertains to one case study, and results cannot be generalized for a larger population, a formal procedure of the activities discussed has been formulated to potentially replicate these processes. The results show that there is a definite cycle followed with the aim to innovate and create needed BI. This is shown in the conceptual model in Fig. 4. The purpose is to illustrate how data can be generated and tested in other contexts based on the findings. Going forward, as the evolution of the system can interact proactively, integration of new smart modules (modularity and scalability software and hardware) can occur as well as the introduction of non-proprietary hardware

Table 3. Dimensions of I4.0 integration

Dimensions	Practical integration	Theoretical TAM Construct
Needs analysis within socioeconomic environment (Context)	Sensing and collecting needed data or usage of existing data to characterize relevance to other entities and ecosystems	PEoU of current system: External factors of the TAM model
Current system analysis for diagnosing errors and solutions development	Actuation and communication to monitor the internal state and diagnose internal issues through novel analysis	Assessing the Perceived usefulness (PU) of smart technology
Identify technologies and assign needed controls	Used to modify internal parameters and while adjusting automatically for deviations to meet goals. Methods and rules need to be assigned for actions between non-human and human entities for specific functions and roles	Assessing the perceived ease of usage (PEoU) that impacts technologies adoption
Develop solutions	Various technologies are available, based on the previous dimensions, use the analysis outputs to generate viable solutions with pre-defined rules	Actual adoption

and software to reduce dependencies. The model incorporates adoption points during product development that is applicable for South Africa. Part of which was a lack of skills, monetary resources and larger socioeconomic issues. The skills related strongly to the TAM model, as this has been shown not only in global areas, but locally as well, where skills effect the PU and PEoU that in turn affect adoption levels. There was also resistance in terms of the loss of jobs or trust levels of what the system generated. This was not the intended case but could be a concern for future projects or expansion. Finally, the "smart" accept is noted, but does add a level of complexity to the TAM model that requires further study.

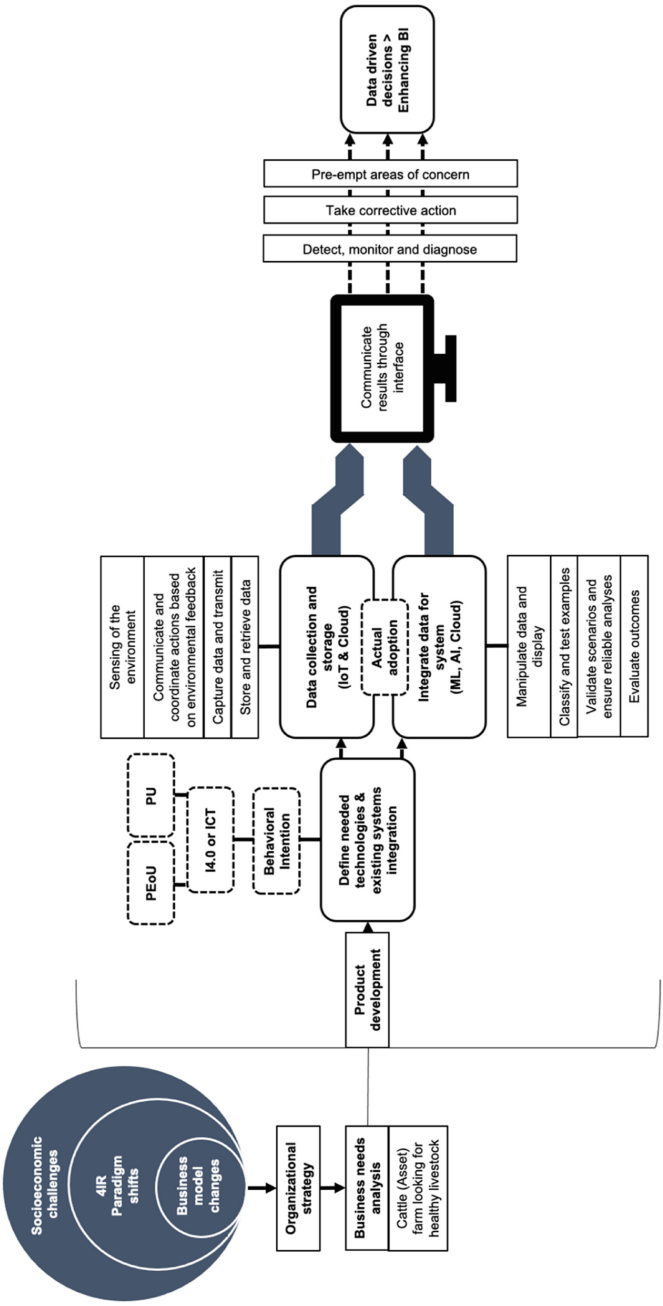


Fig. 4. Proposed conceptual model

6 Conclusion

This study investigated how smart technologies of I4.0 have been used to collect information, process it internally and develop data insights for enhanced BI. To achieve this, data was collected using I4.0 technologies and shown how it can be used to train AI models and enhance BI. From focus group sessions, activities used to embed these kinds of technologies in a practical way was noted to enable complex systems in varying industries. To add to the body of knowledge on information systems on how this can be achieved and enhance BI, a conceptual model as a tool to apply such technologies effectively is presented. The purpose of which is to provide insights for entrepreneurs, industry practitioners, academics and leaders to innovate through technology adoption within the I4.0 paradigm to support decision making in an ever-digitised world. Importantly to the study, was its context and the sector this was developed in.

6.1 Research Limitations

Despite striving for rigor within this methodology, there are several limitations to this study. Firstly, previous literature reviewed only certain technologies of I4.0. These are continually developed and there could be more viable options for similar case studies at a later stage. When collecting data and using the technologies, awareness of limitations, as well as the varying levels of applications was considered. However, biases in sessions can be present despite all efforts made to reduce this but was required to gain the needed insights. The case study was limited to one organization within South Africa, specifically the agriculture sector. Finally, it was assumed that the TAM model was applicable to this context and region, but as noted, the smart integration can have inter-playing effects, requiring further investigation.

6.2 Future Research

From the above study, there are several areas of future research. First and foremost, the application of the conceptual model and enhancement thereof for other sectors. Secondly, a skills assessment framework within the I4.0 paradigm to enhance innovation. Finally, a more generally acceptable TAM which specially encompasses the rapid integration of smart technologies which stem from I4.0.

Acknowledgements and Declarations. This research received no funding. There are no conflicting interests present to the knowledge of the authors. Appreciation is attributed to the privately registered company in SA, which allowed the access to information and processes followed; and the innovation environment, a makerspace, which facilitated the rapid prototyping used for this study.

References

1. Kruger, S., Steyn, A.A.: A conceptual model of entrepreneurial competencies needed to utilise technologies of Industry 4.0. The Int. J. Entrepreneurship Innovation **22**(1), 56–67 (2020). https://doi.org/10.1177/1465750320927359

2. Liu, G., Gao, P., Chen, F., Yu, J., Zhang, Y.: Technological innovation systems and IT industry sustainability in China: a case study of mobile system innovation. Telematics Inform. **35**, 1144–1165 (2018). https://doi.org/10.1016/j.tele.2018.01.012
3. Steenhuis, H.J., Fang, X., Ulusemre, T.: Global diffusion of innovation during the fourth industrial revolution: the case of additive manufacturing or 3D printing. Int. J. Innov. Technol. Manag. **17**, 1–34 (2020). https://doi.org/10.1142/S0219877020500054
4. Luthra, S., Kumar, A., Zavadskas, E.K., Mangla, S.K., Garza-Reyes, J.A.: Industry 4.0 as an enabler of sustainability diffusion in supply chain: an analysis of influential strength of drivers in an emerging economy. Int. J. Prod. Res. **58**(5), 1505–1521 (2020). https://doi.org/10.1080/00207543.2019.1660828
5. Park, H.S.: Technology convergence, open innovation, and dynamic economy. J. Open Innov.: Technol., Mark., Complexity **3**(1), 1–13 (2017). https://doi.org/10.1186/s40852-017-0074-z
6. Nankervis, A., Connell, J., Montague, A., Burgess, J. (eds.): The fourth industrial revolution. Springer, Singapore (2021). https://doi.org/10.1007/978-981-16-1614-3
7. Park, S.-C.: The Fourth Industrial Revolution and implications for innovative cluster policies. AI Soc. **33**(3), 433–445 (2017). https://doi.org/10.1007/s00146-017-0777-5
8. Saucedo-Martínez, J.A., Pérez-Lara, M., Marmolejo-Saucedo, J.A., Salais-Fierro, T.E., Vasant, P.: Industry 4.0 framework for management and operations: a review. J. Ambient. Intell. Humaniz. Comput. **9**, 789–801 (2017). https://doi.org/10.1007/s12652-017-0533-1
9. Ishaya, T., Folarin, M.: A service oriented approach to business intelligence in telecoms industry. Telematics Inform. **29**, 273–285 (2012). https://doi.org/10.1016/j.tele.2012.01.004
10. Autor, D.H.: Why are there still so many jobs? the history and future of workplace automation. J. Econ. Perspect. **29**, 3–30 (2015). https://doi.org/10.1257/jep.29.3.3
11. Jain, A., Shah, D., Churi, P.: A review on business intelligence systems using artificial intelligence. In: Smys, S., Tavares, J.M.R.S., Balas, V.E., Iliyasu, A.M. (eds.) ICCVBIC 2019. AISC, vol. 1108, pp. 1023–1030. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-37218-7_107
12. Matthews, J.R.: Understanding indigenous innovation in rural West Africa: challenges to diffusion of innovations theory and current social innovation practice. J. Hum. Dev. Capabilities **18**, 223–238 (2017). https://doi.org/10.1080/19452829.2016.1270917
13. Moeuf, A., Lamouri, S., Pellerin, R., Tamayo-Giraldo, S., Tobon-Valencia, E., Eburdy, R.: Identification of critical success factors, risks and opportunities of Industry 4.0 in SMEs. Int. J. Produc. Res. **58**(5), 1384–1400 (2020). https://doi.org/10.1080/00207543.2019.1636323
14. Chatterjee, S., Rana, N.P., Dwivedi, Y.K., Baabdullah, A.M.: Understanding AI adoption in manufacturing and production firms using an integrated TAM-TOE model. Technol. Forecast. Soc. Change **170**, 120880 (2021). https://doi.org/10.1016/j.techfore.2021.120880
15. Rahman, M.M., Lesch, M.F., Horrey, W.J., Strawderman, L.: Assessing the utility of TAM, TPB, and UTAUT for advanced driver assistance systems. Accid Anal Prev. **108**, 361–373 (2017). https://doi.org/10.1016/j.aap.2017.09.011
16. Salimon, M.G., Kareem, O., Mokhtar, S.S.M., Aliyu, O.A., Bamgbade, J.A., Adeleke, A.Q.: Malaysian SMEs m-commerce adoption: TAM 3, UTAUT 2 and TOE approach. J. Sci. Technol. Policy Manage. **14**, 98–126 (2021). https://doi.org/10.1108/JSTPM-06-2019-0060
17. Dalmarco, G., Ramalho, F.R., Barros, A.C., Soares, A.L.: Providing industry 4.0 technologies: the case of a production technology cluster. The J. High Technol. Manage. Res. **30**(2), 100355 (2019). https://doi.org/10.1016/j.hitech.2019.100355
18. Boudreaux, C.J.: Institutional quality and innovation: some cross-country evidence. J. Entrepreneurship Public Policy **6**, 26–40 (2017). https://doi.org/10.1108/JEPP-04-2016-0015
19. World Economic Forum: The Future of Jobs Report 2020 (2020)
20. Department of Science and Technology South Africa: White Paper on Science, Technology and Innovation: Science, technology and innovation enabling inclusive and sustainable South African development in a changing world (2019)

21. Andrzej, M.: Uncertainty in the sphere of the Industry 4.0 – potential areas to research. Bus., Manage. Educ. **14**(2), 275–291 (2016). https://doi.org/10.3846/bme.2016.332
22. Fleacă, E., Fleacă, B., Maiduc, S.: Aligning strategy with sustainable development goals (SDGs): process scoping diagram for entrepreneurial higher education institutions (HEIs). Sustainability (Switzerland) **10**, 1–18 (2018). https://doi.org/10.3390/su10041032
23. Zivdar, M., Imanipour, N., Talebi, K., Hosseini, S.R.: An explorative study of inputs for entrepreneurs' decision-making to create new venture in a high-tech context. Int. J. Entrepreneurship Innov. **18**, 243–255 (2017). https://doi.org/10.1177/1465750317723218
24. Oztemel, E., Gursev, S.: Literature review of Industry 4.0 and related technologies. J. Intell. Manuf. **31**(1), 127–182 (2018). https://doi.org/10.1007/s10845-018-1433-8
25. Serpanos, D.: The cyber-physical systems revolution. Computer (Long Beach Calif). **51**, 70–73 (2018). https://doi.org/10.1109/MC.2018.1731058
26. Alter, S.: Making sense of smartness in the context of smart devices and smart systems. Inf. Syst. Front. **22**(2), 381–393 (2019). https://doi.org/10.1007/s10796-019-09919-9
27. Pham, T.N., Tsai, M.-F., Nguyen, D.B., Dow, C.-R., Deng, D.-J.: A cloud-based smart-parking system based on Internet-of-Things technologies. IEEE Access **3**, 1581–1591 (2015). https://doi.org/10.1109/ACCESS.2015.2477299
28. Borgia, E.: The internet of things vision: key features, applications and open issues. Comput Commun. **54**, 1–31 (2014). https://doi.org/10.1016/j.comcom.2014.09.008
29. United Nations: The Sustainable Development Goals Report, pp. 1–56. United Nations Publications (2017). https://doi.org/10.18356/3405d09f-en
30. Saha, A.K., Saha, J., Ray, R., Sircar, S., Dutta, S., Chattopadhyay, S.P., Saha, H.N.: IOT-based drone for improvement of crop quality in agricultural field. In: 2018 IEEE 8th Annual Computing and Communication Workshop and Conference, CCWC 2018, pp. 612–615 (2018). https://doi.org/10.1109/CCWC.2018.8301662
31. Witkowski, K.: Internet of things, big data, industry 4.0 – innovative solutions in logistics and supply chains management. Procedia Eng. **182**, 763–769 (2017). https://doi.org/10.1016/j.proeng.2017.03.197
32. Ferrández-Pastor, F.J., García-Chamizo, J.M., Nieto-Hidalgo, M., Mora-Martínez, J.: Precision agriculture design method using a distributed computing architecture on internet of things context. Sensors **18**(6), 1731 (2018). https://doi.org/10.3390/s18061731
33. Sutherland, E.: The fourth industrial revolution-the case of South Africa. Politikon **47**, 233–252 (2020). https://doi.org/10.1080/02589346.2019.1696003
34. Prinsloo, J., Vosloo, J.C., Mathews, E.H.: Towards industry 4.0: a roadmap for the South African heavy industry sector. South Afr. J. Ind. Eng. **30**, 174–186 (2019). https://doi.org/10.7166/30-3-2237
35. Paek, B., Lee, H.: Strategic entrepreneurship and competitive advantage of established firms: evidence from the digital TV industry. Int. Entrepreneurship Manage. J. **14**(4), 883–925 (2017). https://doi.org/10.1007/s11365-017-0476-1
36. Qin, J., Liu, Y., Grosvenor, R.: A categorical framework of manufacturing for Industry 4.0 and beyond. Procedia CIRP **52**, 173–178 (2016). https://doi.org/10.1016/j.procir.2016.08.005
37. Ustundag, A., Cevikcan, E.: Industry 4.0: Managing The Digital Transformation. Springer International Publishing, Cham (2018)
38. Deng, L.: Artificial intelligence in the rising wave of deep learning: the historical path and future outlook. IEEE Signal Process. Mag. **35**, 177–180 (2018). https://doi.org/10.1109/MSP.2017.2762725
39. Makridakis, S.: The forthcoming Artificial Intelligence (AI) revolution: its impact on society and firms. Futures **90**, 46–60 (2017). https://doi.org/10.1016/j.futures.2017.03.006
40. Monostori, L., et al.: Cyber-physical systems in manufacturing. CIRP Ann. Manuf. Technol. **65**, 621–641 (2016). https://doi.org/10.1016/j.cirp.2016.06.005

41. Sandino, J., Pegg, G., Gonzalez, F., Smith, G.: Aerial mapping of forests affected by pathogens using UAVs, hyperspectral sensors, and artificial intelligence. Sensors (Switzerland) **18**, 1–17 (2018). https://doi.org/10.3390/s18040944
42. Bartoloni, S., et al.: Towards designing society 5.0 solutions: the new Quintuple Helix – Design Thinking approach to technology. Technovation **113**, 102413 (2022). https://doi.org/10.1016/j.technovation.2021.102413
43. Liu, J., Timsina, P., El-Gayar, O.: A comparative analysis of semi-supervised learning: the case of article selection for medical systematic reviews. Inf. Syst. Front. **20**(2), 195–207 (2016). https://doi.org/10.1007/s10796-016-9724-0
44. South African Civil Aviation Authority: National Legislation, South Africa. http://caa.mylexisnexis.co.za/. Last accessed 20 Dec 2019
45. Nääs, I.A., Garcia, R.G., Caldara, F.R.: Infrared thermal image for assessing animal health and welfare. J. Anim. Behav. Biometeorol. **2**(3), 66–72 (2014). https://doi.org/10.14269/2318-1265/jabb.v2n3p66-72
46. Davis, F.: A technology acceptance model for empirically testing new end-user information systems: Theory and results (1985). https://doi.org/10.1126/science.146.3652.1648
47. Sohn, K., Kwon, O.: Technology acceptance theories and factors influencing artificial Intelligence-based intelligent products. Telematics Inform. **47**, 101324 (2020). https://doi.org/10.1016/j.tele.2019.101324
48. Lee, Y., Kozar, K.A., Larsen, K.R.T.: The technology acceptance model past, present, and future. Commun. Assoc. Inform. Syst. **12**, 752–780 (2003). https://doi.org/10.17705/1CAIS.01250
49. Gangwar, H., Date, H., Ramaswamy, R.: Understanding determinants of cloud computing adoption using an integrated TAM-TOE model. J. Enterp. Inf. Manag. **28**, 107–130 (2015). https://doi.org/10.1108/JEIM-08-2013-0065
50. Masood, T., Sonntag, P.: Industry 4.0: adoption challenges and benefits for SMEs. Comput. Ind. **121**, 103261 (2020). https://doi.org/10.1016/j.compind.2020.103261
51. Venkatesh, V., Morris, M.G., Davis, G.B., Davis, F.D.: User acceptance of information technology: toward a unified view. MIS Q. **27**, 425–478 (2003)
52. Sethi, P., Sarangi, S.R.: Internet of things: architectures, protocols, and applications. J. Electr. Comput. Eng. **2017**, 1–25 (2017). https://doi.org/10.1155/2017/9324035
53. Kruger, S., Steyn, A.A.: Enhancing technology transfer through entrepreneurial development: practices from innovation spaces. J. Technol. Transf. **45**(6), 1655–1689 (2019). https://doi.org/10.1007/s10961-019-09769-2
54. Crupi, A., et al.: The digital transformation of SMEs – a new knowledge broker called the digital innovation hub. J. Knowl. Manag. **24**, 1263–1288 (2020). https://doi.org/10.1108/JKM-11-2019-0623

Exploring Healthcare Workers' Perceptions of Digital Healthcare Support Platforms: The Case of NurseConnect

Ronaldo Nombakuse, Pitso Tsibolane(✉), and Sumarie Roodt

University of Cape Town, Cape Town, South Africa
{pitso.tsibolane,sumarie.roodt}@uct.ac.za

Abstract. The provision of professional and psycho-social support for healthcare workers (HCWs) is integral to the delivery of high-quality care worldwide. New digital health platforms (DHPs) are being developed to support healthcare workers in delivering patient care during stressful times. However, as these new support technologies become more widely available, healthcare workers may need clarification about how to integrate them effectively into daily practice. This qualitative study explores the perceptions of nurses and midwives towards a nationally rolled out DHP called NurseConnect, which is aimed at improving maternal and newborn clinical knowledge of HCWs to understand better how digital healthcare support platform tools might be integrated into daily patient care. In-depth semi-structured interviews of purposively sampled three accoucheurs (male midwives), three midwives, and three public hospital nurses were conducted. An inductive thematic analysis approach highlighted overall pessimism, manual workarounds, and abandonment of the digital platform due to negative perceptions of lack of awareness, low usefulness, and poor usability. The study proposes the need for integrating the NurseConnect platform into the employee work practice policies, targeted promotion of platform benefits, and the digitization of guidelines and textbooks that workers use to improve the current healthcare process. A longitudinal nationwide mixed-methods study should be undertaken to probe further the value and impact of digital healthcare support platforms.

Keywords: Digital health platforms · Non-adoption · Healthcare workers · Nurseconnect

1 Introduction

There has been a rise of digital health platforms (DHPs) in the past few years due to their usefulness to healthcare workers and patients worldwide [1]. The broad digital health platforms category encompasses mobile health (mHealth), telehealth, telemedicine, and wearable devices [24]. The goal of these platforms varies widely. In South Africa, two national digital healthcare platforms are intended to support patients; the MomConnect messaging service and the NurseConnect support platform to help nurses and midwives [3, 21]. Generally, digital health platforms are an example of healthcare innovation, as

P. Ndayizigamiye et al. (Eds.): IDIA 2022, CCIS 1774, pp. 172–184, 2023.
https://doi.org/10.1007/978-3-031-28472-4_11

they provide a platform in which digital technologies facilitate patients' and healthcare workers' participation in the healthcare delivery process [14]. However, there is a small number of DHPs in general in low and middle-income countries (LMICs) [9]. Among the reasons is the lack of funding, digital skills, and technological capabilities amongst healthcare workers (HCWs), despite DHPs having the potential to improve the quality of their work [3]. The lack of healthcare workers' involvement on the development phase of the platform ensures that the platform is fit-for-purpose to increase their digital capability [11]. The lack of digital health policy and strategy plays a significant role in non-adoption, resistance, and negative attitude towards DHPs [6, 30].

This paper aims to qualitatively explore the perceptions of HCWs, specifically nurses and midwives, towards the nationally rolled-out digital healthcare support platform, NurseConnect, which aims to improve maternal and newborn clinical knowledge through the provision of professional and psycho-social support for healthcare workers in South Africa. The rest of the paper is structured as follows: Sect. 2 is the literature review, Sect. 3 will cover the research design, Sect. 4 will present findings for this study, and the last section will be the conclusion and recommendations for future work.

2 Literature Review

2.1 Digital Health Platforms

The implementation of DHPs needs to be integrated into the daily routines of healthcare workers to eradicate the manual work they are expected to perform to succeed [26]. DHPs intended for patients should be educational, enable independence, and ease patients' frustrations in public and private healthcare facilities during processes such as appointment booking [11, 28]. DHPs for patients and healthcare workers should be based on the actual access criteria evaluation, which must be integrated into the patient or healthcare worker's day-to-day activities [8]. The criteria for access evaluation include, but are not limited to, a) physical access to the app to evaluate whether the initiative is accessible to the users and highlights barriers that may hinder their accessibility, b) appropriateness of the app to evaluate whether the app is suitable for the needs of the intended users [28], c) integration to daily activities to evaluate whether DHPs have any disturbances on the daily activities of the users and d) relevant content to evaluate the content of the app [9]. Further research, education, and evidence-based tools are required to support nurses in helping patients select, initiate, and sustain the use of digital health platforms to improve health outcomes [11].

Even though DHPs have an advantage in the overall healthcare practice, healthcare workers in LMICs need to receive the necessary support [5]. DHPs are helpful as they add value to healthcare facilities, healthcare workers and patients with educational content that empowers them to overcome challenges that they encounter on a day-to-day basis [29]. The benefits encompasses; improvement of healthcare process since majority of the processes are manual, seamless sharing of electronic healthcare records to eradicate paper-based, sharing of healthcare information through the utilization of telehealth, easy access to medical records for patients, patient-to-patient online portals for communication, support for healthcare workers who practice in areas with inadequate resources with important health information to raise awareness to patients and citizens, as well as

support for patients who need remote treatment to alleviate healthcare facility visitation to patients with limited transportation.

2.2 mHealth Adoption in South Africa

Mobile health technology is significant to healthcare workers who practice in areas where patients have inadequate access to quality healthcare and substantial health information to raise awareness [1]. The purpose of mHealth applications is to assist healthcare workers with educational content, create awareness about various health topics, support diagnostic treatment and provide telemedicine to patients who reside in areas with inadequate access to healthcare services [11]. mHealth technologies must emphasize the type of communication and the level of interaction, as shown in Table 1.

Table 1. Types of mHealth Applications [3].

Type of communication	Level of Interaction (with the user)
Text Messaging	• Appointment reminders • Health Education & Awareness notifications • Emergency button connected to a health facility
Mobile-site services	• Read the full educational article on the site • Complete patient surveys • View patient medical records
Voice messages	• Staffed information lines • Read messages to visually impaired patients
A mix of voice & video messages	• Instruction videos, for example to expecting mothers or patients with chronic illness • Mobile telemedicine
Text messaging	• Appointment reminders • Health Education & Awareness notifications • Emergency button connected to a health facility
Mobile-site services	• Read the full educational article on the site • Complete patient surveys • View patient medical records
Voice messages	• Staffed information lines • Read messages to visually impaired patients

mHealth initiatives lack value and innovation for the targeted population, which hinders the advantages of mHealth utilization to individuals who reside in low and middle-income countries (LIMICs) [7]. The possibility of the patient data breach to malicious attacks [20] indicated that health facilities do not have technologies that protect patient data due to the non-existence of response from the security, risk, and compliance. Scholars also discovered that "more experienced healthcare workers seem to be less digitally capable compared with their younger counterparts" thus non-adoption of these

[1, 28]. The non-adoption of digital health platforms is problematic since healthcare workers are expected to perform manual work for capturing, accessing, and storing patient data. This replicated effort triggers frustration to healthcare workers, which has an impact on quality health delivery to patients [12]. Finally, DHPs have the potential to improve the lives of patients who reside in areas with inadequate access to healthcare [10]. Thus, it is imperative for LMICs to adopt DHP strategies that are currently implemented in HICs, to improve the digital capability of healthcare workers and patients [11].

2.3 Digital Applications in Health

Real Access to technology is put into effective use to analyze all issues for lack of access to technology [17, 18, 25]. This centers on several elements that ICT developers must incorporate for users, to eradicate non-adoption of DHPs. The first elements are physical access to technology which aims to understand whether the initiative is available to the users, by highlighting the barriers that may hinder accessibility e.g., network coverage and access to smartphones. Appropriateness of technology aims to understand whether the mHealth initiative is suitable for the needs of the intended users. Suitability can be assessed through capacity, user friendly and language use. Language use aims to understand whether the messages sent are clearly understood by users. Affordability of technology aims to understand whether the users can afford DHPs in the context of mobile data since South Africa is experiencing a challenge of high data prices. Integration into daily routines aims to evaluate whether the new DHP will not have any disruption to the daily life of the user.

High income countries (HICs) such as Canada, the United States of America, the United Kingdom, and Australia have shown success in DHP adoption, through the application of the above-mentioned elements [27]. In 2015 Canada launched the Omama Project which assists expecting and existing mothers with maternal and child health information [1]. In 2017 United Kingdom introduced the peanut app to help new and existing mothers with peer-to-peer education where they converse about antenatal and postnatal care [13]. The center of perinatal excellence (COPE) supports mothers and fathers with emotional support of being a parent and was launched in Australia. The app supports parents until the child is one year old, which has similar features with the MomConnect messaging service [6].

There are more than a few DHPs that have been implemented in LMICs with an intent of assisting healthcare workers and patients [13]. While the majority of these DHPs were in the pilot phase, a minuscule number of them have not been enhanced ever since [9]. MomConnect messaging service was launched in South Africa in 2014 to assist pregnant women and first-time mothers with antenatal and postnatal care information [28]. The messages are sent via SMS and WhatsApp three times a week, the initiative is implemented in line with the principles of mobile alliance for maternal action (MAMA) [2]. Integrated healthcare information service through mobile telephony (IHISM) is a mobile phone-based patient record and assists citizens with general health queries via text messages [3]. Free AIDS Test was launched in Ethiopia to educate, raise awareness, and encourage citizens using text messages to test for HIV/AIDS [15]. The author further elaborated about CellPhones4HIV, which was launched in South Africa to send

HIV/AIDS text messages to patients and raise awareness about mother-to-child HIV transmission.

2.4 The Nurseconnect Support Platform

The South African National Department of Health (NDoH) implemented a national health digital strategy to support patients, citizens, and healthcare workers for quality healthcare delivery [19]. The Nurseconnect support platform is an association of the MomConnect messaging service that was officially presented to the public by NDoH in 2016. The platform aims to support nurses and midwives in their everyday routines with clinical knowledge and professional and social support [21]. Users subscribe by following unstructured supplementary service data (USSD) or Mobile-sites guidelines. On successful registration, a short messaging service (SMS) is sent to the user to confirm the registration. A clinical code is mandatory for all midwives and nurses who anticipate registering to the platform. As shown in the below illustration derived from the mobile platform site (Fig. 1).

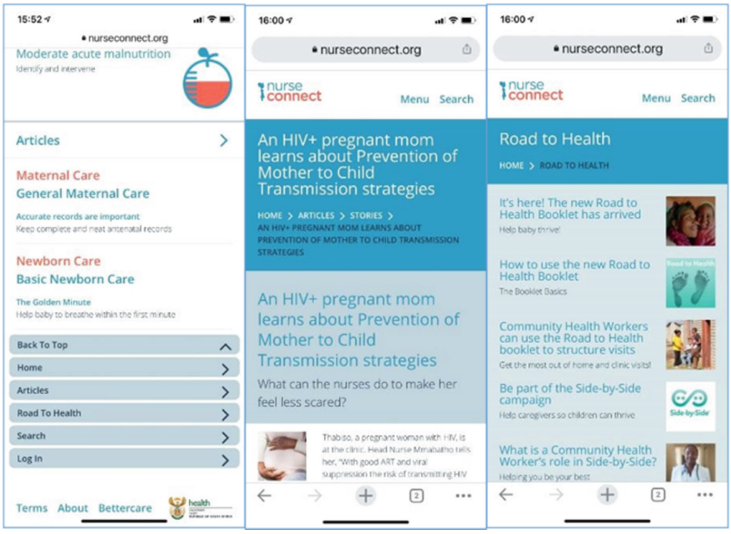


Fig. 1. The nurseconnect support platform – mobile site [21].

The Nurseconnect support platform messages are sent two times a week via SMS. Each educational or motivational message entails a link to redirect users to read the complete article on the Nurseconnect mobile site. The mobile site has the following features on the homepage; daily topic, search, road to health, back to the top, login, terms, about and better-care.

3 Research Design

3.1 Methodology

This study is exploratory and follows a qualitative research methodology [22]. The qualitative, inductive, interpretive research methodology was deemed suitable for this study as it enhances a deeper understanding of the perceptions and experiences of healthcare workers regarding digital health platforms [16].

3.2 Participants

This study was conducted with nurses, and midwives in maternal and child healthcare or family planning in South Africa. A semi-structured interview instrument was employed, and a mobile phone was utilized to collect data from the participants. A purposive sampling technique was used to purposefully choose nurses and midwives within the healthcare fraternity since Nurseconnect is intended for them. The audio data gathered from the interviews was transcribed from audio to text, and no automated voice transcribing tools were used. A total of 9 participants were interviewed (see Table 2). All personal information of the participants was removed from the data using audio editing software. The illustration below depicts the summary of the participants.

3.3 Data Collection

Face-to-face interviews were conducted over three months in 2019. The duration for each interview was between 30–60 min on average; each interview transcript was three to seven pages long. Participants were allowed to use their home languages thus four South African languages and English was used by participants. Figure 2 illustrates an example of an interview transcript from NVivo for Midwife four and five with a combination of English & isiZulu.

The data gathered from the interviews was safely stored in the UCT OneDrive Cloud Storage. The data gathered was first analyzed manually using Microsoft excel packages and the second round of analysis was done in NVivo. The reason for analyzing data in two steps was to have an overall understanding of the data collected by going through each line of the transcript iteratively, so that no code words or phrases are omitted if the researcher used a computer-assisted qualitative data analysis software (CAQDAS) like NVivo, ATLAS.ti, HyperRESEARCH, MAXqda2 and QDA Miner [23].

3.4 Data Analysis

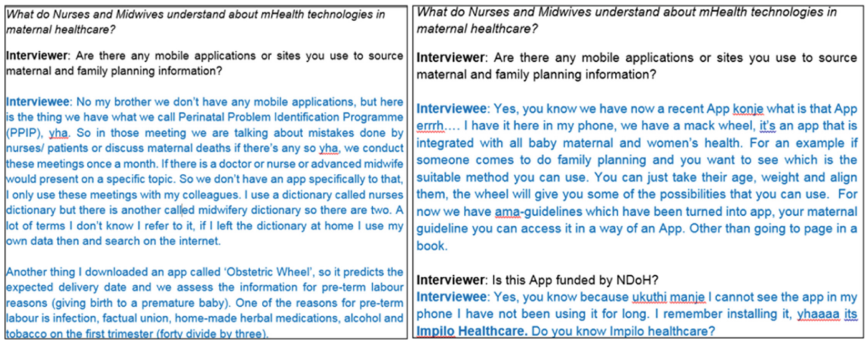
Thematic analysis is an approach used in qualitative research to analyze data to categorize themes after data has been gathered from the sample [16]. Hence thematic analysis was used to analyze the collected data from the nurses and midwives. This approach follows six phases to identify themes which emerged during data analysis for this study, namely, familiarize yourself with the gathered data, generate initial codes, search for themes, review themes, define and name themes, and produce a report [4].

Table 2. Semi-structured interviews participants

Unique name	Occupation	Language	Municipality	Analysis tools
NUR001	Nurse	English & isiXhosa	City of Cape Town Metropolitan Municipality	Microsoft Excel & NVivo
NUR002	Nurse	English & Tswana	Ngaka Modiri Molema District Municipality	Microsoft Excel & NVivo
MID001	Accoucheur *(Male Midwife)*	English	City of Cape Town Metropolitan Municipality	Microsoft Excel & NVivo
MID002	Midwife	English & isiXhosa	City of Tshwane Metropolitan Municipality	Microsoft Excel & NVivo
MID003	*Midwife*	English & isiXhosa	Inxuba Yethemba Local Municipality	Microsoft Excel & NVivo
NUR003	*Nurse*	English	Sedibeng District Municipality	Microsoft Excel & NVivo
MID004	Accoucheur *(Male Midwife)*	English & isiZulu English & isiXhosa	Cape Winelands District Municipality	Microsoft Excel & NVivo
NUR004	*Nurse*	English & isiXhosa	Inxuba Yethemba Local Municipality	Microsoft Excel & NVivo
MID005	Accoucheur *(Male Midwife)*	English & isiXhosa	City of Cape Town Metropolitan Municipality	Microsoft Excel & NVivo

3.5 Ethical Considerations

The researcher had a duty to safeguard consent and confidentiality for all participants, therefore a letter of consent was circulated amongst all participants to voluntarily participate in this study [23]. The researcher utilized five components with all nurses and



What do Nurses and Midwives understand about mHealth technologies in maternal healthcare?

Interviewer: Are there any mobile applications or sites you use to source maternal and family planning information?

Interviewee: No my brother we don't have any mobile applications, but here is the thing we have what we call Perinatal Problem Identification Programme (PPIP), yha. So in those meeting we are talking about mistakes done by nurses/ patients or discuss maternal deaths if there's any so yha, we conduct these meetings once a month. If there is a doctor or nurse or advanced midwife would present on a specific topic. So we don't have an app specifically to that, I only use these meetings with my colleagues. I use a dictionary called nurses dictionary but there is another called midwifery dictionary so there are two. A lot of terms I don't know I refer to it, if I left the dictionary at home I use my own data then and search on the internet.

Another thing I downloaded an app called 'Obstetric Wheel', so it predicts the expected delivery date and we assess the information for pre-term labour reasons (giving birth to a premature baby). One of the reasons for pre-term labour is infection, factual union, home-made herbal medications, alcohol and tobacco on the first trimester (forty divide by three).

What do Nurses and Midwives understand about mHealth technologies in maternal healthcare?

Interviewer: Are there any mobile applications or sites you use to source maternal and family planning information?

Interviewee: Yes, you know we have now a recent App konje what is that App errrh.... I have it here in my phone, we have a mack wheel, it's an app that is integrated with all baby maternal and women's health. For an example if someone comes to do family planning and you want to see which is the suitable method you can use. You can just take their age, weight and align them, the wheel will give you some of the possibilities that you can use. For now we have ama-guidelines which have been turned into app, your maternal guideline you can access it in a way of an App. Other than going to page in a book.

Interviewer: Is this App funded by NDoH?
Interviewee: Yes, you know because ukuthi manje I cannot see the app in my phone I have not been using it for long. I remember installing it, yhaaaa its **Impilo Healthcare.** Do you know Impilo healthcare?

Fig. 2. Interview transcript for MID004 & MID005

midwives who participated on this study; (I) any personal information like name, a surname of the participants was be kept confidential therefore each participant was given a unique name MID/NUR with a number (all indicators that could suggest the identity of the participants were deleted from the data, using an audio editing software), (II) the questions for the interviews were not sensitive or instigated any harm to participants (III) participants were informed prior to the commencement of the interview that they will be recorded and a letter of consent was issued to obtain their signature (IV) participants had a right to withdraw after the interview and their answers were removed by researcher from this study and, for example NUR005, was removed as per her request.

4 Findings and Analysis

The selected participants were interviewed through a semi-structured interview research instrument via WhatsApp, Microsoft teams, Google Hangout and Zoom. The interviews were conducted on participant convenience to ensure that the researcher acquired more in-depth information until no further significant insight was yielded by the analysis. Five themes emerged from the data that was collected, the following sections present the themes and findings.

4.1 Perceived Usefulness

Participants stated that the Nurseconnect support platform is unresponsive, particularly when nurses and midwives require clarity on the information that is sent to them, *"…. They take very long to respond sometimes, and it made me wonder – because sometimes I would need an answer now-now(immediately)"* **(MID005).** NUR003 stated that it is essential for National Department of Health (NDoH) to recruit more employees on the helpdesk to assist with all related queries *"……delegate personnel who will act as point of contact to assist nurses / midwives with emotional and psychological support"* **(NUR003).** The current communication process is one-way, and this created a barrier for users who want more details on the messages they receive, *"…I send this you send this, if I don't understand anything I send you an SMS you reply to me, so it was a one-way thing, they would send us SMS's, but we couldn't have communication wa bona"*

(NUR002). Thus, most users opted to use their guidelines and textbooks to access maternal and family planning information. *"…. I don't have an App; I use my textbooks and the guidelines"* **(NUR002; MID001; MID002; MID003).** Healthcare workers cannot distinguish between the Nurseconnect support platform and the MomConnect messaging service, thus, they cannot realize the usefulness of the platform.

4.2 Willingness to Make Use of the Platform

Restricted accessibility to hospital or clinical resources such as Wi-Fi to acquire information is another factor that discourages users from willingly utilize the Nurseconnect, *"…we do not have free Wi-Fi that is one of the things that need to be improved, if you put free Wi-Fi for us"* **(NUR001; NUR002; NUR003; MID002; MID005; NUR004).** The healthcare workers who are most impacted by the challenge are those who practice in rural areas with inadequate access to maternal healthcare resources *"…. we are living in these poor areas there's no internet access, the landlines are not working, you use your own phone to get an ambulance or 'ntho eo' other personnel's there's nothing, literally nothing"* **(NUR002; NUR004).** As a result, if the platform was zero-rated like the COVID-19 portal, a small number of participants exhibited high-level of willingness to put into effective use the mobi-site, *"…. Having to walk from this block to another to ask my colleagues about something I am not sure about, if NDoH can provide us with Wi-Fi I think a lot can improve"* **(MID004).** Majority of participants exhibited willingness to utilize the platform if there was provision of free Wi-Fi in public healthcare facilities, the platform was zero-rated, and the management was fully involved in encouraging healthcare workers to make use of the platform.

4.3 Nurseconnect Usability Challenges

Users stated that the platform has one way communication *"…it was a one-way thing, they would send us SMS's, but we couldn't have communication 'wa bona'"* **(NUR002).** Communication is supposed to be two-way until there is an alignment between the sender and the receiver. Thus, the continuation of a way one communication from the platform with the users might have a negative impact on the NDoH vision and strategy, particularly when users necessitate emergency assistance, *"…they too very long sometimes to answer"* **(NUR004; MID005).** However, MID005 stated that the content on the mobile site needs to be expanded and include other key topics that are discussed in the guidelines and textbooks, *"…I stopped using it, I am now using my guideline for information purposes actually"*. Due to the complexities of this platform, one participant opted to use a similar platform to the Nurseconnect to source maternal information, *"…. It was kind of helping, but I am now using Mpilo Healthcare App, it has all the guidelines you know, you can actually see how many people are reporting cases, it has complaints and requests"* **(MID004).** While other participants specified that the NurseConnect is time-consuming and they work under immense pressure, *"…if you introduce this new process this will be time-consuming because this might need more hands we need to deal with patients"* **(NUR004).**

4.4 Status of mHealth Adoption in Maternal Healthcare and Family Planning

Healthcare workers will still be dependent on their guidelines and textbooks, whereas there is a gap to digitize these materials to improve the current ways of working. Therefore, NDoH needs to implement initiatives and in-service training with an aim of educating healthcare workers about DHPS as an enabler to their day-to-day activities. The majority of the participants have little or no knowledge about mHealth applications, *"…I don't have an App at all, and I just use the internet. There is no specific App that speaks to the information that I might need as an Accoucheur"* **(MID005).** Resistance to change is another factor for lack of mHealth adoption amongst nurses and midwives. Users are at ease in following the current ways of working than adopting new, *"…. I don't have an App; I use my textbooks and the guidelines at work"* **(MID003).** On the other hand, one participant had a perception regarding the use of mHealth applications, the user highlighted that patients might perceive nurses and midwives as bad-mannered if they are constantly on their mobile phones, *"…. Because now if I am always on my phone is regarded as rude, patient might not know I am trying to access information they might think I am on WhatsApp or Facebook"* **(MID001).** Furthermore, MID005 cannot distinguish between a mobile app and a maternal healthcare initiative, *"…. No my brother we don't have any mobile applications, but here is the thing we have what we call Perinatal Problem Identification Programme (PPIP)".*

4.5 Improvement of Manual Processes in Maternal Healthcare and Family Planning

There is a lot of manual work that is expected to be performed by nurses and midwives, this is time consuming and might lead to incorrect reporting if the clinical data is captured incorrectly on the NDoH system SINJANI. Therefore, there is a need to move away from the use of paper-based records. *"…We do everything manually and only when you are submitting a statistics then you handover to our clerk to capture on our system"* **(NUR003; NUR004; MID001).** Manual record capturing is seen as a dual work which is another factor that contributes to behavioral change amongst nurses and midwives, *"…We basically do 'double job' in Excel when you do calculations to capture the data"* **(MID004).** In addition, MID002 further explains her pain points on the current manual process at the healthcare facility she practices in, *"…the issue of records also comes to the issue of us keeping records for a certain period of time and throwing others away".* The participants further emphasized that the current process is also taking a lot of their time, *"…we must document everything that is time-consuming. Let's say you do a reproductive treatment, and they have STI symptoms. If they want family planning, then you must do the documentation for all that. Then you make sure you put your stats on the system"* (NUR002) As a result, participants expressed the benefits of digitizing maternal and family planning records, *"…when they are in a software method you can't easily disport, you can save, don't use a lot of space we can minimize"* **(MID001; MID002; MID003; NUR002).** The continuance of manual capturing, storing and processing of maternal and family planning data is a serious problem that needs to be eradicated by NDoH. Thus, there is a need to develop a Nurseconnect App to integrate some of the activities that healthcare workers perform, for instance patient record and appointment bookings.

5 Recommendations and Future Work

5.1 Recommendations

It is recommended that the National Department of Health (NDoH) must; delegate trained representatives from the department of health who will perform as a point of contact at different provincial and national healthcare facilities to support healthcare workers to subscribe to the Nurseconnect platform. The NDoH must organize an international day for midwives and nurses, to raise awareness about the Nurseconnect support platform. Improve the current communication process by allowing users to send enquiries about the messages they receive from the platform every morning or afternoon. The platform needs to be consistent in terms of the time it sends messages to the users. NDoH must introduce other mediums of communication like WhatsApp, where users can forward messages on various WhatsApp groups like 'Nurses Who Care'. Nurseconnect content needs to be updated to cater for nursing and midwifery information needs. The current content was last updated in 2016/17, the content can include eclampsia, anaemia during pregnancy and ophthalmia neonatorum information. NDoH must implement a zero-rated Nurseconnect App for nurses and midwives. In order to integrate the guidelines, textbooks and dictionaries that the nurses and midwives utilize when they refresh their memories on key terms in practice.

5.2 Future Work

The Nurseconnect support platform is a decent initiative especially in a developing country like South Africa, where there are little or no mHealth initiatives for the maternal, child healthcare and family planning practice. Therefore, there is possibility for future research to investigate various topics in practice, the research area that can be conducted by scholars is as follows; development of mHealth application to support healthcare workers in maternal, child healthcare and family planning practice: a case of the Nurseconnect support platform. The researcher can visit several public healthcare facilities in different provinces of the country to collect data from all workers of these practices, to produce better results.

Acknowledgement. This work is based on the research supported in part by the National Research Foundation of South Africa (MND210607608912).

References

1. W. H. Organization: Digital health platform handbook: building a digital information infrastructure (infostructure) for health (2020)
2. Aderounmu, G.A., Akintola, A., Asare, S.D., Ogwu, F.J.: Information and communication technology (ICT) and globalization of engineering education: the Nigeria experience. In: International Conference on Intelligent and Innovative Computing Applications (ICONIC) (2018)
3. Botha, A., Booi, V.: mHealth implementation in South Africa. In: 2016 IST-Africa Week Conference. IEEE, pp. 1–13 (2016)

4. Braun, V., Clarke, V.: Using thematic analysis in psychology. Qual. Res. Psychol. **3**(2), 77–101 (2006)
5. Brown, J., Pope, N., Bosco, A.M., Mason, J., Morgan, A.: Issues affecting nurses' capability to use digital technology at work: an integrative review. J. Clin. Nurs. **29**(15–16), 2801–2819 (2020)
6. Chen, J., Lieffers, J., Bauman, A., Hanning, R., Allman-Farinelli, M.: The use of smartphone health apps and other mobile health (mHealth) technologies in dietetic practice: a three country study. J. Hum. Nutr. Diet. **30**(4), 439–452 (2017)
7. N.P. Commission. National development plan vision 2030 (2013)
8. DuBoff, M., Futrell, E.: The mHealth Planning Guide: Key Considerations for Integrating Mobile Technology into Health Programs. The Johns Hopkins University Center for Communication Programs, Baltimore, MD (2013)
9. Ferguson, C., Jackson, D.: Selecting, appraising, recommending and using mobile applications (apps) in nursing. J. Clin. Nurs. **26**, 3253–3255 (2017)
10. Fischer, A.E., Sebidi, J., Barron, P., Lalla-Edward, S.T.: The MomConnect nurses and midwives support platform (NurseConnect): a qualitative process evaluation. JMIR Mhealth Uhealth **7**(2), e11644 (2019)
11. Gordon, W.J., Landman, A., Zhang, H., Bates, D.W.: Beyond validation: getting health apps into clinical practice. NPJ Dig. Med. **3**(1), 1–6 (2020)
12. Henson, P., David, G., Albright, K., Torous, J.: Deriving a practical framework for the evaluation of health apps. The Lancet Dig. Health **1**(2), e52–e54 (2019)
13. Hussain, T., Smith, P., Yee, L.M.: Mobile phone–based behavioral interventions in pregnancy to promote maternal and fetal health in high-income countries: systematic review. JMIR Mhealth Uhealth **8**(5), e15111 (2020)
14. Iyawa, G.E., Herselman, M., Botha, A.: Digital health innovation ecosystems: from systematic literature review to conceptual framework. Procedia Comput. Sci. **100**, 244–252 (2016)
15. Kebede, D.: Use Of Mobile Network-Based Apps In HIV/AIDS Treatment and Counseling In Ethiopia. St. Mary's University (2018)
16. Kumar, R.: Research Methodology: A Step-by-Step Guide for Beginners. Sage (2018)
17. Maphumulo, W.T., Bhengu, B.R.: Challenges of quality improvement in the healthcare of South Africa post-apartheid: a critical review. Curationis **42**(1), 1–9 (2019)
18. Marshall, C., Lewis, D., Whittaker, M.: mHealth Technologies in Developing Countries: A Feasibility Assessment and a Proposed Framework, pp. 1–47. University of Queensland, Herston, Australia (2013)
19. Mustafa, A.: Information Practices in Midwifery: A Case Study of an Antenatal and Intrapartum Care Environment in the Western Cape, South Africa. Cape Peninsula University of Technology (2015)
20. Naidoo, S., Bottomley, D., Naidoo, M., Donnelly, D., Thaldar, D.: Artificial intelligence in healthcare: proposals for policy development in South Africa. South Afr. J. Bioethics Law **15**(1), 11 (2022)
21. Phaneuf, A.: How mHealth apps are providing solutions to the healthcare market's problems. Business Insider 4 (2019). https://www.businessinsider.com/mhealth-apps-definition-examples
22. Sargeant, J.: Qualitative research part II: Participants, analysis, and quality assurance. J. Grad. Med. Educ. **4**(1), 1–3 (2012)
23. Saunders, M., Lewis, P., Thornhill, A.: Research Methods for Business Students Eight Edition. QualitativeMarket Res.: An Int. J. (2019)
24. Seebregts, C., et al.: Designing for scale: optimising the health information system architecture for mobile maternal health messaging in South Africa (MomConnect). BMJ Glob. Health **3**(Suppl 2), e000563 (2018)

25. Simblett, S., et al.: Barriers to and facilitators of engagement with mHealth technology for remote measurement and management of depression: qualitative analysis. JMIR Mhealth Uhealth **7**(1), e11325 (2019)
26. Swerts, M., Westhof, E., Bogaerts, A., Lemiengre, J.: Supporting breast-feeding women from the perspective of the midwife: a systematic review of the literature. Midwifery **37**, 32–40 (2016)
27. Taherdoost, H.: Sampling methods in research methodology; how to choose a sampling technique for research. Int. J. Acad. Res. Manag. **5**, 18–27 (2016). https://doi.org/10.2139/ssrn.3205035
28. van Graan, A.C., Williams, M.J.S., Koen, M.P.: Professional nurses' understanding of clinical judgement: a contextual inquiry. Health SA Gesondheid **21**, 280–293 (2016)
29. Watson, A., Wilkinson, T.M.: Digital healthcare in COPD management: a narrative review on the advantages, pitfalls, and need for further research. Ther. Adv. Respir. Dis. **16**, 17534666221075492 (2022)
30. Kapepo, M.I., Van Belle, J.-P., Weimann, E.: Towards a theoretical understanding of workarounds emerging from use of a referral mobile application: a developing country context. Procedia Comput. Sci. **196**, 533–541 (2022). https://doi.org/10.1016/j.procs.2021.12.046

An IoT Based Helopeltis Sp Pest Control System

Kannole E. Veronica[1(✉)], Rushingabigwi Gerard[1], and Diwani Abubakar[2]

[1] University of Rwanda, CST Campus Nyarugenge, Kigali, Rwanda
verokan1@yahoo.com
[2] State University of Zanzibar, Zanzibar, Tanzania
abubakar.bakar@suza.ac.tz

Abstract. Agriculture is the main activity of residents in the southern part of Tanzania. Cashew nuts being the most cash crop brought by Portuguese, has never been yielding the optimal yields due to the existence of pests especially Helopeltis Sp. The Internet of Things (IoT) based Pest Control system aims to implement a system which will be able to capture, identify and store the Helopeltis pest using Google Colab and Proteus simulation tools. Pest recognition process has been done in the Google Colab Pro platform using tensor-flow library, python 3.7 programming language and Faster-RCNN InceptionV2 model. The accuracy of 97.87% was obtained while the mechanical part of wiping pests into container, wiping them from the top of the container lid to the environment and farmer notification SMS (Short Message Service) has been facilitated using Proteus. The control of the cashew nut pests will now be in a green way by discouraging the use of pesticides which also destroys pollinators and degrades the quality of the soil, the crop and the environment. It has been found that image training requires adequate resources including high performance graphic card, memory and processing power, the use of Google Colab has catered for all of them. Also it has been noted that the more the images are used for the training the more accuracy the detection will be.

Keywords: IoT · Faster-RCNN · Pest recognition

1 Introduction

Growing crops has been the main activity for the most of the East Africans, be it for business or for family consumption. The produce has never been enough to sustain families due to the poor practice of small scale farming [1]. The agriculture sector is still struggling with various problems even to those practicing large scale farming, these include pests and diseases [2, 3]. Agriculture has been a good contributor in the economy in Tanzania especially through cash crops unfortunately it has been producing poor yields and causing farers to be poor. Speaking of Manyoni and Liwale districts in Tanzania, regardless of the being blessed with vast and rich soil suitable for Cashew nut and other crops, they have never obtained best cashew yields at the maximum potential due to the problem of the existence of Helopeltis Sp, Pseudotheraptus wayi and Aphis Sp pests [4]. The existence of Helopeltis Sp pest refer to Fig. 1, in cashew nut farms attacks the

P. Ndayizigamiye et al. (Eds.): IDIA 2022, CCIS 1774, pp. 185–199, 2023.
https://doi.org/10.1007/978-3-031-28472-4_12

flowers, leaves refer to Fig. 2, shoots, nuts and cashew nut apples ultimately resulting in low yields. There are natural enemies (parasitoids) that feed on the eggs and the pests as well for example; Telenomus cuspis, Erythmeles helopeltidis, Chaetostricha sp., Ufens sp. And Gonatocerus sp they feed on the Helopeltis eggs while the spiders, reduviid bugs, ants, praying mantids, mantispid flies, robber flies, and pentatomid bugs feed on the adult pests [4]. Parasitoids and predators are not sufficient enough to eliminate the problem and so the black lesions, leaf damage and crop dieback are inevitable. This study designed and implemented a system that will control Helopeltis Sp pest using Internet of Things and machine learning. System will be sending notification message to the farmer upon the detection of the container status being full. The intention of the study is to use deep learning techniques to recognize Helopeltis Sp pest by employing the pre-trained Faster-RCNN InceptionV2 models on the collected data from Manyoni and Liwale districts in Tanzania. This system will play a vital role on the control of the pests by discouraging the use of chemicals, and will ensure better crops, soil and air quality preservation and thus will make it a green solution.

Fig. 1. Helopeltis sp pest

The paper has been organized into six sections; Related work, Methodology, Results and Discussion, Conclusion and Recommendation, Future Works and References.

2 Related Work

The use of chemicals (pesticides) to control pest has posed negative impacts to the soil and air. More efforts have been done by researcher. A Sneha et al., to reduce the use of chemicals by employing the WSN (Wireless Sensors Network) system that will inform the farmer on the only pest affected region to be treated [5]. Comprehensive work has also been done by Y. Shi et al., involving various technologies including GPRS (General Packet Radio Service), 3S (Global positioning and remote Sensing technology), WSN and intelligence data processing and decision making, to provide farmers and stakeholders with the early warning alerts for crop diseases and pests existence. The same system also provides online consultation services on crop diseases and pests, this

Fig. 2. Damages caused by Helopeltis sp pest.

allows that the agricultural experts to advise farmers on the proper usage of the pesticides [6]. Other researchers have been writing on the ways and the reasons for the diseases and pests to grow, without discussing pest control methods [7]. More agricultural cashew nut best practices have also been discussed along with the ecological requirements [12]. The electromagnetic exposure methods have been implemented by some researchers as the means of controlling pests by exposing the pests to electromagnetic and warm them up to fifteen degrees before the wheat flourishing time so that they feel free reproduce themselves and also fly out to look for food where there is none, and this process will consume their energy and die before the flourishing or harvesting time. Unfortunately this method is no longer implementable due to cancer suspicions to human beings [8]. Cutworm pests control mechanism using the robot to apply barriers in the grape vines by applying wrapper rings to the trunks has been discussed. This method encourages the use of chemicals that are placed in the packing cloth of the wrapper [10]. Integrated pest management has been suggested by several researchers including A. Chougule et al., where all data about the pests and diseases are stored in the database, and the real time information from the sensors are being used by the agricultural experts to give out necessary procedures to be taken to control the situation. The procedures can include the information on the type of pesticides to be used, the location of farm to be applied and the time to apply [11]. K Saranya et al., used IoT to control pests and diseases by using the PIR (Passive Infrared motion sensor) and ultrasonic generator, after the pests being observed sound will be generated 38–44 kHz for beetles and mosquitos while 60 kHz was used to repel the rodents [9]. The main drawback of this method is that even the non-targeted insects and pollinators like bees, ant, yellow banded special wasp and moths will also repel thus discouraging pollination process. B. Vijayalakshmi has based on the threat detection side by early detecting the pests attack, for any discovered disease the alert is sent to the farmer along with the advisory pesticides to be used [26].

Temperature, humidity sensors and camera were used by S. Zhang et al., to capture the present leaves conditions and being compared with the healthy leaves images, once deference is noted then message is sent to the farmer to inform on the changes, possible causes and guidance on what is to be done to overcome the threat. The motor for the water or pesticides spray will be actuated depending on the value obtained from the moisture sensor and the computed results from the Arduino Pi 3B microcontroller respectively. Big data platforms have been used to handle real time data collected from the sensors placed in farms whereby the analysis is done online and the agricultural experts answer questions from the farmers through their smart mobile phones or computers on a real time basis [13]. This method also encourages the use of the pesticides.

M. Wadhai used video processing techniques to detect pests and alert farmer on the location of the observed pests as early as possible without suggesting any means to overcome them [14].

More technologies have been applied to solve the pest's threat, F. A. Faria et al., used machine learning only to identify the type of pest by considering the wing and the aculei parts without suggesting the way to control them [15] just like N.R. Gavai et al., used machine learning for flower classification [16], while X. Liu et al., used the same technology for pest prediction on the stored grains [17].

The comparison study was conducted by Siswanto et al., to determine the best Helopeltis sp control method between the three various chemical formulas, the results obtained showed that all the three formulas improved the yields at different percentages but non targeted insects like natural enemies were also killed [18].

Online system was developed by C. Liu and K. Cai with the capabilities to read images posted by the farmer, analyzing them and giving out the insights and advising the farmer what should be done in a real time way basing on China's agricultural guidelines [19]. Likewise K. A. Upendra et al., developed a system that accept image data to identify pests along with the location coordinates to predict on spreading areas possible for the disease outbreak [20]. All the two papers did not discuss any chemical free mechanisms to control pest.

S.F. Qi et al., and F. Qiao et al., at different times discussed on pests specifically at the storage areas. S.F. Qi et al. designed a wireless sensor trap [21], while F. Qiao et al., used STM32 microcontroller integrated with capacitance detection circuit both aimed to detect pest species and counting them out in the grain storage so that pest species are identified along with their quantity [22], none of the two aimed to eliminate pests from the grains.

P. Seung et al., observed that the use of heating as the pest control mechanism affects the taste of the crops although it can kill all the pests at a hundred percent efficiency [23]. This is the major drawback as taste and nutrients are vital aspects to be maintained.

E. Suganya et al., used image processing techniques for plant's disease discovery, to allow farmers to control it before it spreads out to more crops [24], likewise Q. Dai et al., have put their efforts on the best quality images generated from the camera for accurate pest classification through deep learning methods [25]. Both of them did not provide the mechanism to control plant disease and pest respectively.

In this paper a green pest control solution has been provided using IoT and machine learning technologies without using chemicals.

3 Methodology

Various mechanisms were collectively used to accomplish the main objective of the research, the activities done followed the Rapid Application Development (RAD) model please refer to Fig. 3.

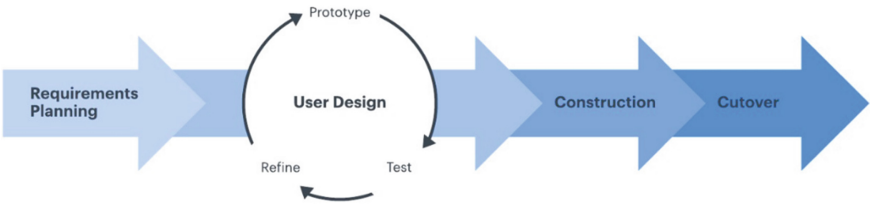


Fig. 3. Rapid application development model [27].

This model allows simultaneously execution of the project tasks hence saves time.

The system is comprised of two main sections namely; pest identification section, where Faster-RCNN InceptionV2 was used to train the model and identify pest along with Google Colab Pro, Tensor-Flow library and python 3.7 programming language as shown in Table 1. The second section is the system actuation, which was implemented using Proteus simulation tool and other tools as shown in Table 2.

Table 1. Object recognition tools

Tool	Resource
Operating system	Windows 10
Programming language	Python 3.7
Labeling software	Labelimg
GPU card	Google Colab Pro
Processor	Google Colab Pro
Memory	Google High Memory
Machine learning	Tensorflow library
Plotting tool	Tensor board

Google Colab GPU card was used to replace the weak graphic card of the computer used to run the image recognition process.

3.1 Data Collection

The total of 8 farms from the two districts were visited. In each field 10 cashew nut trees were randomly selected to collect Helopeltis pests from, and a total of 185 images

Table 2. System actuation tools.

Tool	Resource
Simulation	Proteus software
Sensing	IR obstacle sensor library
Texting	SIM900D library
Wiping	Servo motor library
Obstacle	Potentiometer
Display	LCD, Virtual terminals
Microcontroller	Arduino UNO

obtained from the collected pest by placing them on various backgrounds so as to create a powerful model. The CSV (Comma-Separated Value) file had a total of 119 labeled images used for training and 43 labelled images for testing. The validation process used 23 images of the pest and 30 non-pest images.

The total of 180 pests were observed in one time observation from the visited farms in Liwale district refer to Fig. 4, and at least 124 pests were observed from Manyoni district fields refer to Fig. 5.

The blue color represents Black Lesions (BL) while Orange is for Leaf Damage (LD).

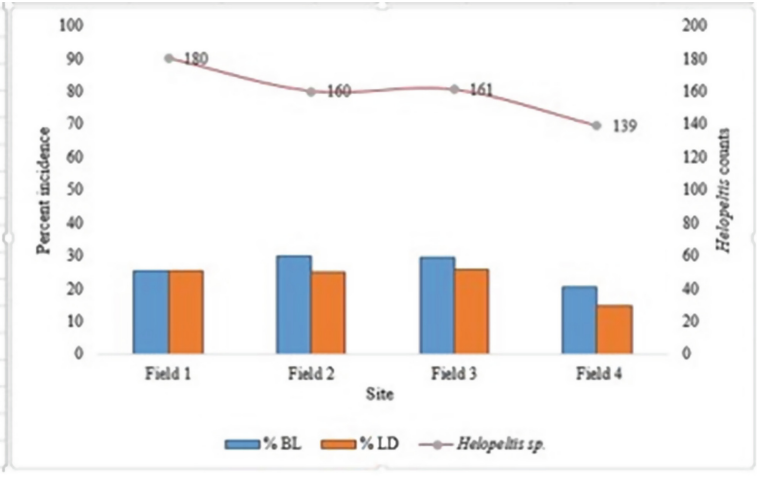


Fig. 4. Pest counts in Liwale district (Color figure online).

Liwale is at the southern part of Tanzania where cashew nuts are highly grown.

The blue color represents Black Lesions (BL) while Orange is for Leaf Damage (LD).

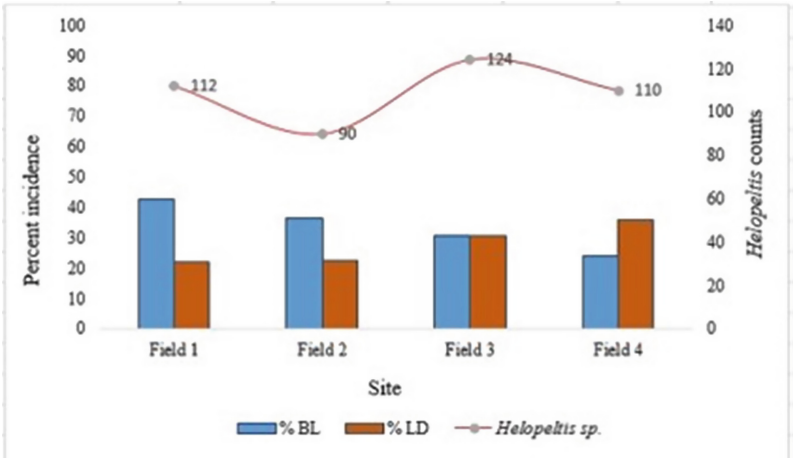


Fig. 5. Manyoni pest counts (Color figure online).

Manyoni is available at the central part of Tanzania despite the fact that cashew nuts are grown at a low scale compared to the southern and coastal areas pest counts are also high.

3.2 System Architecture

The system architecture has been described using the block diagram which is divided into two sections namely;

The Pest Identification Section.
This part has been accomplished using the Google Colab Pro platform, referring to Fig. 6 the block diagram is comprised of the following blocks;

- Image gathering – Camera was used to take Helopeltis pictures from the eighty cashew trees visited.
- Image labeling – In a laptop the software called Labelimg used to label all the images
- Generating training data – The obtained data are not suitable for training until they are converted into to Tensorflow (TF) records for the training model to handle them. The python scripts were used to complete this conversion task.
- Pre-Training stage – These are the procedures after generating the TF records but before data training taking place. They include creating a label map and employing the Faster-RCNN InceptionV2 model to serve the research needs.
- Data training – The python command is run to initiate the training process. Data training was done for at least three hours.

Exporting Inference graph – The frozen graph is generated by running the python script, this graph contains the classifier that will be classifying every detected object.

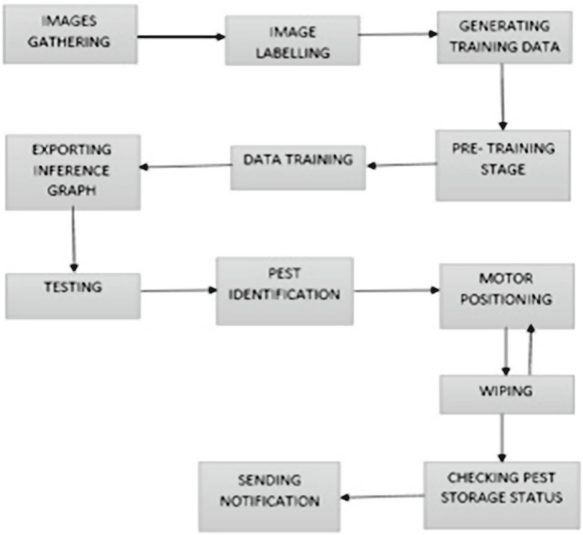


Fig. 6. System block diagram.

- Testing – The exact image name is written in the script to be tested for the identification.
- Pest Identification – The matched object will be covered with the rectangle along with the percentage of accuracy.

The system Actuation Section

- Motor Positioning – The servo motor will return to its initial position after moving the wiper to the specified direction.
- Wiping – The wiper will move in clockwise direction to mimic the wiping into the container action and anticlockwise for wiping out to the environment.
- Checking Pest storage status – The container will have the sensor to sense once it is full, the PIR(Passive Infrared) sensor in Proteus will be controlled by potentiometer that acts as the obstacle to represent the fullness or emptiness of the storage.
- Sending Notification – This process will be triggered by the storage container status especially when it is full so that the farmer does the emptying.

All two sections have been also elaborated using the flow chart shown in Fig. 7

3.3 Evaluation of the Model

$$\begin{aligned} \text{Precision} &= \text{TP}/(\text{TP} + \text{FP}) * 100 \\ &= 23/(23 + 1) * 100 = 95.83\% \end{aligned} \tag{1}$$

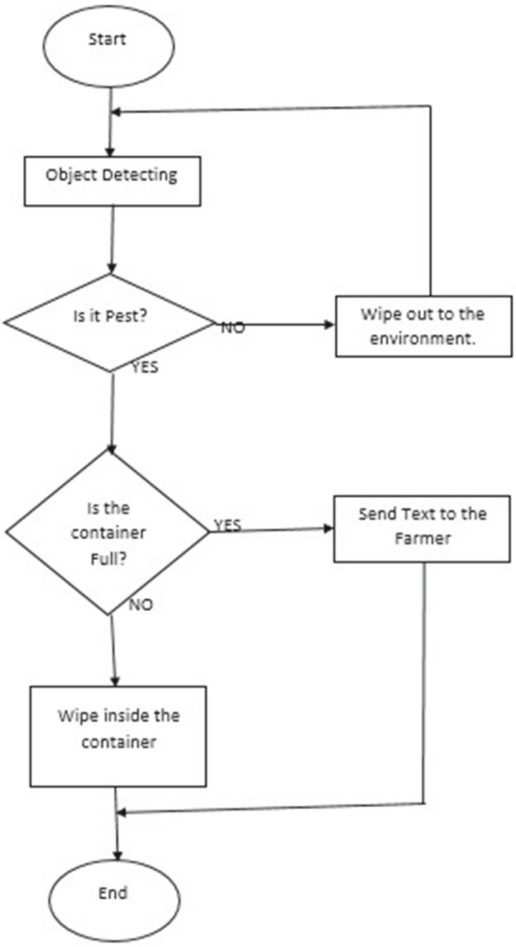


Fig. 7. Pest identification and System Actuation data flow diagram

where TP stands for true positives and they are actual pests. TN stands for true negatives and they are not pests. FP stands for False Positives These were detected as positives but they are not pests. FN stands for False Negative. These are pests but the system didn't recognize them. The value from (1), (2) and (3) is used to calculate F1 factor

$$\begin{aligned} &\text{Accuracy} = (\text{TP} + \text{TN})/(\text{TP} + \text{TN} + \text{FP} + \text{FN}) * 100 \\ &= (23 + 30)/(23 + 30 + 1 + 0) * 100 = 98.15\% \end{aligned} \tag{2}$$

$$\begin{aligned} &\text{Recall} = \text{TP}/(\text{TP} + \text{FN}) * 100 \\ &= 23/(23 + 0) * 100 = 100\%. \end{aligned} \tag{3}$$

$$\begin{aligned} &\text{F1Score} = 2 * (\text{Precision} * \text{Recall})/(\text{Precision} + \text{Recall}) \\ &= 2 * (1) + (2) + (3) = 2 * (95.83 * 100)/(95.83 + 100) = 97.87\% \end{aligned} \tag{4}$$

The value from Eq. (4) is used to determine how well the model performs. Therefore the mode performed with the 97.87% of accuracy.

4 Results and Discussion

Simple Random Sampling technique was used to select 80 cashew nut trees from the 8 farms to collect the pests from. This study used only 1 class for model training where images used were captured by Apple IPhone 6, Tecno Cammon 12 and Infinix Hot 10 cameras. Training was done for more than 3 h to obtain minimum loss of 0.004 at 54,868 epochs.

4.1 Sensor Results

Sensor module will be sensing the level of the pest collector so that the emptying process could be done once full. This part is simulated using IR (Infrared) obstacle sensor in Proteus whereby potentiometer is used to control it by acting as the obstacle by varying voltages. For visibility purposes the yellow LED will be on when the container is full refer to Fig. 8 and will be off for the empty status of the container.

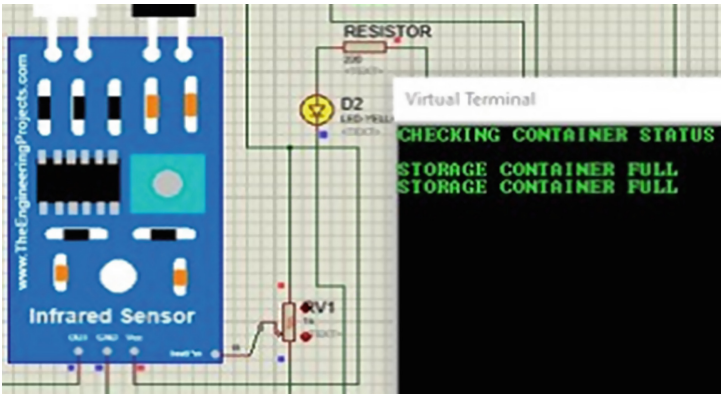


Fig. 8. Infrared obstacle sensor (Color figure online).

4.2 Pest Identification Results

This was done in the Google Colab platform by running a series of python scripts in the Google Colab platform to detect Helopeltis Sp pest, as shown in Fig. 9. The Faster-RCNN InceptionV2 model is among of the models that are more accurate and fast.

Fig. 9. Pest detection.

4.3 System Actuation Results

The physical observation of the pest movements proves that 1 s response time is adequate to capture the pest, to increase efficiency the Servo motor with the ability to actuate 180 degrees with 480 ms is good enough. This unit is comprised of three processes;

- Wiping into the container – The lid which is controlled by the servo motor will be actuated to turn 180 degrees then the wiper will wipe them in the container. Servo motor in Proteus will rotate 180 degrees clockwise to mimic turning the lid and wiping into the container upon the detection of the Helopeltis pest.
- Wiping out to the environment – This system design is intending to leave free other non-targeted insects for they could be of most important like cashew nut pollinators, the wiper will be activated to wipe them out back to the environment without turning the lid, the anticlockwise movement of the motor is for the empty container status and the wiping out process refer to Fig. 10.

- Sending notification to the farmer – Farmer will be notified in case the container is full or there is an obstacle on the lid by which system failed to wipe. The "Container is Full" text messages will be sent to farmer once the potentiometer has very low voltage, for the higher values the "Container is Empty" text will be shown on the virtual terminal but no message will be sent to the farmer. These messages will be visible through the two virtual terminals connected to the micro controller, sensor and the GSM (Global System for Mobile communications) module as shown in Fig. 11.

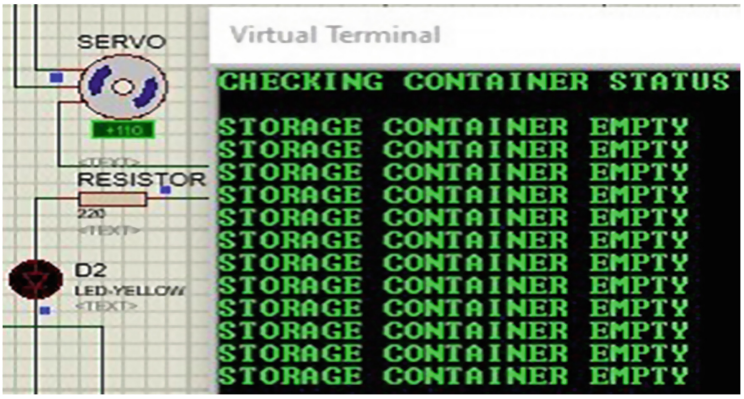


Fig. 10. Empty status of the container.

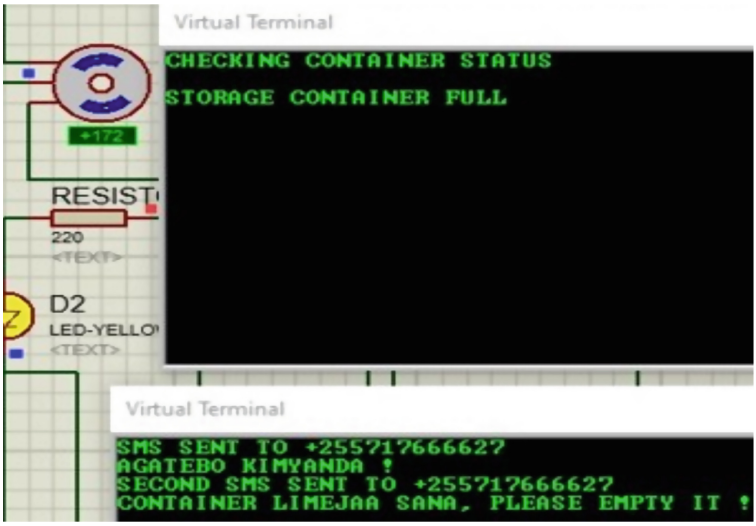


Fig. 11. Notification message.

5 Conclusion and Recommendation

Agriculture sector plays a pivotal role in human beings and therefore it is worthy of conducting research on it, apart from this sector providing materials for the industries but also feeding people giving them various benefits including blood pressure and blood sugar regulation through cashew nuts. This sector faces pest and disease problems causing the production of poor yields that directly affects the availability of the nutrients. This study designed the Helopeltis Sp pest control mechanism using Internet of Things and machine learning, also will be sending notification messages to the cashew nut farmer when a sensor indicates the container status being full. Pest detection was successfully implemented with the accuracy of 97.87 % using the Google Colab Pro platform and

other events were successfully done in Proteus. This system design is open for future implementations, and can also be used to solve the pest problem with the similar characteristics as Helopeltis Sp. It has been noted that there is no limit for the number of images to be used for training because the more images the more accuracy will be achieved but this is at the expense of time and internet connection charges.

6 Future Works

The system is open for future works by;

- Carrying out the implementation in real physical devices and software tools
- Using Arduino Nano BLE Sense micro controller to perform the intended job but also collecting environmental data including light intensity, humidity, temperature, and pressure for further agricultural insights
- Adding the module that will send data to the cloud for further analysis.
- Choosing the container material that allow air to pass through for the survival of the pollinators mistakenly wiped with the pests, also adding another sub system for emptying the container automatically.

References

1. Pack, M., Mehta, K.: Design of affordable greenhouses for east Africa. In: 2012 IEEE Global Humanitarian Technology Conference, pp. 104–110. Seattle, WA, USA (2012)
2. Waweru, B.W., Rukundo, P., Kilalo, D.C., Miano, D.W., Kimenju, J.W.: Effect of border crops and intercropping on aphid infestation and the associated viral diseases in hot pepper (Capsicum sp.). Crop Protect. **145**, 105623 (2021). https://doi.org/10.1016/j.cropro.2021.105623
3. Freire, F.C.O., Cardoso, J.E., dos Santos, A.A., Viana, F.M.P.: Diseases of cashew nut plants (Anacardium occidentale L.) in Brazil. Crop Prot. **21**(6), 489–494 (2002). https://doi.org/10.1016/S0261-2194(01)00138-7
4. Boniface, B.: Status of sucking insect pests in cashew growing locations of South and Central Zones, Tanzania (2020). https://doi.org/10.12692/ijb/16.4.34-45
5. AdelineSneha, J., Chakravarthi, R., Glenn, J.A.: A review on energy efficient image feature transmission in WSN for micro region pest control. In: 2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT), pp. 4859–4862. Chennai, India (2016). https://doi.org/10.1109/ICEEOT.2016.7755643
6. Shi, Y., Wang, Z., Wang, X., Zhang, S.: Internet of things application to monitoring plant disease and insect pests. In: Proceedings of the 2015 International Conference Application Science Engineering Innovation, vol. 12, no. Asei, pp. 31–34 (2015). https://doi.org/10.2991/asei-15.2015.7
7. Gupta, D., Belwal, M.: Pest identification and control of diseases in crop fields through image processing and tracking of atmospheric parameters. In: 2018 2nd International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), 2018 2nd International Conference on, Palladam, India, pp. 289–294 (2018). https://doi.org/10.1109/I-SMAC.2018.8653695

8. Anis, M., et al.: Agricultural applications. Nanovate 235–242 (2017). https://doi.org/10.1007/978-3-319-44863-3-13
9. Saranya, K., Uva Dharini, P., Uva Darshni, P., Monisha, S.: IoT based pest controlling system for smart agriculture. In: Proceedings 4th International Conference Communication Electronics System ICCES 2019, no. ICCES, pp. 1548–1552 (2019). https://doi.org/10.1109/ICCES45898.2019.9002046
10. Lee, T., Hudson, S., Chang, J.-Y.: Design and development of an automated band wrapper robot for grapevine pest control. In: 2009 4th International Conference on Autonomous Robots and Agents, Wellington, pp. 10–14 (2009). https://doi.org/10.1109/ICARA.2000.4803971
11. Chougule, A., Jha, V.K., Mukhopadhyay, D.: Using IoT for integrated pest management. In: 2016 International Conference on Internet of Things and Applications (IOTA), pp. 17–22. Pune, India (2016). https://doi.org/10.1109/IOTA.2016.7562688
12. Kapinga, F.A., Kasuga, L.J.F., Kafiriti, E.M.: Growth and production of cashew nut. Soils, Plant Growth Crop Prod. Prod. Cashew Nut **1**(1), 1–10 (2010)
13. Zhang, S., Chen, X., Wang, S.: Research on the monitoring system of wheat diseases, pests and weeds based on IOT. In: 2014 9th International Conference on Computer Science and Education, pp. 981–985. Vancouver, BC, Canada (2014) https://doi.org/10.1109/ICCSE.2014.6926609
14. Wadhai, M.: Technique, pp. 544–547 (2015)
15. Faria, F.A., et al.: Automatic identification of fruit flies (Diptera: Tephritidae). J. Vis. Commun. Image Represent. **25**, 1516–1527 (2014). https://doi.org/10.1016/j.jvcir.2014.06.014
16. Gavai, N.R., Jakhade, Y.A., Tribhuvan, S.A., Bhattad, R.: MobileNets for flower classification using TensorFlow. In: 2017 International Conference on Big Data, IoT and Data Science (BID), pp. 154–158. Pune, India (2017). https://doi.org/10.1109/BID.2017.8336590
17. Liu, X., Wang, C., Su, Y., Hu, T.: cIntegrating multimedia and artifical intelligence for pest prediction and aeration control of stored grain bins. In: ICEMI 2009 – Proceedings of the 9th International Conference Electronics Measurements Instruments, pp. 880–883, (2009). https://doi.org/10.1109/ICEMI.2009.5274158
18. Siswanto, Trisawa, I.M., Karmawati, E., Suhesti, S.: Control of Conopomorpha cramerella, Helopeltis sp., and Phytophthora palmivora using botanical and biological pesticides. IOP Conf. Ser. Earth Environ. Sci. **418**(1), 012086 (2020)
19. Liu, C., Cai, K.: Design and implement of web-based intelligent decision support system for prevention and control of fruit tree diseases and pests. In: Proceedings of the 2009 4th International Conference Computer Science Education ICCSE 2009, pp. 1269–1271 (2009)
20. Upendra, K.A.N., Lakmali, W.L.A.T.A., Wickramaarchchi, H.W., Wikramanayake, G.N., Goonatillake, J.: Social life network for disease and pest control. In: 2014 14th International Conference on Advances in ICT for Emerging Regions (ICTer), pp. 266. Colombo, Sri Lanka (2014)
21. Qi, S.F., Li, Y.H.: A new wireless sensor used in grain pests detection. In: Proceedings of the 2012 International Conference Control Engineering Communication Technology ICCECT 2012, pp. 755–758 (2012). https://doi.org/10.1109/ICCECT.2012.97
22. Qiao, F., Ji, C., Zeng, X., Zhang, J.: A capacitive pest detection approach based on STM32 microcontroller. In: 2019 6th International Conference System Informatics, ICSAI 2019, no. Icsai, pp. 1035–1039 (2019). https://doi.org/10.1109/ICSAI48974.2019.9010239
23. Saeung, P., Santalunai, S., Thosdeekoraphat, T., Thongsopa, C.: Improved efficiency of insect pest control system by SSPA. In: 2018 5th International Conference Industrial Engineering Applications ICIEA 2018, pp. 179–183 (2018). https://doi.org/10.1109/IEA.2018.8387092
24. Suganya, E., Sountharrajan, S., Shandilya, S.K., Karthiga, M.: IoT in agriculture investigation on plant diseases and nutrient level using image analysis techniques. In: Balas, V.E., Hoang Son, L., Jha, S., Khari, M., Kumar, R. (eds.) Internet of Things in Biomedical Engineering, pp. 117–130. Academic Press (2019)

25. Dai, Q., Cheng, X., Qiao, Y., Zhang, Y.: Agricultural pest superresolution and identification with attention enhanced residual and dense fusion generative and adversarial network. IEEE Access **8**, 81943–81959 (2020). https://doi.org/10.1109/ACCESS.2020.2991552
26. Vijayalakshmi, B., Ramkumar, C., Niveda, S., Pandian, S.C.: Smart pest control system in agriculture. In: 2019 IEEE International Conference on Intelligent Techniques in Control, Optimization and Signal Processing (INCOS), pp. 1–4. Tamilnadu, India (2019). https://doi.org/10.1109/INCOS45849.2019.8951351
27. 4 Phases of Rapid Application Development Methodology. https://www.lucidchart.com/blog/rapid-application-development-methodology. Accessed 10 Oct 2022

Privacy and Security in Digital-for-Development Ecosystems

Phishing Attack Victims and the Effect on Work Engagement

Matthew James Werner and Kennedy Njenga(✉)

Department of Applied Information Systems, University of Johannesburg, Johannesburg, South Africa
knjenga@uj.ac.za

Abstract. Most of the research carried out in cybersecurity considers the technical aspects of the security of an organisation's systems. This work highlights the importance of considering the 'softer' social side of cybersecurity that looks at the lived experiences of phishing attack victims and the effect of such attacks on work engagement. In order to understand these effects, the study adopted the grounded theory (GT) approach to collecting and analysing data elicited from participants. The participants were theoretically sampled from the metropole area of Johannesburg and presented lived experiences of phishing attacks in their unique contexts. The data were transcribed and coded using GT techniques. From the codes, categories derived, and substantive theory that explains the effects of phishing attack victims on work engagement was generated. The implications of this theory to previous theories and scholars and practitioners are discussed.

Keywords: Cybersecurity · Phishing · Work engagement · Grounded theory

1 Introduction

1.1 Background on the Rise in Cyber Attacks

The 21st century saw a significant rise in mobile computing, with most mobile smartphones and tablets having advanced processing power compared to older mainframe systems. These advanced mobile computational devices brought unique sets of cybersecurity issues and concerns as these were connected to the Internet. In the discipline of Information Systems, the concern with how these mobile and other computational devices are to be protected is known as cybersecurity. Cybersecurity looks at the overall measures implemented within organisations and at an individual level to ensure electronic devices are secure from external threats such as hackers, viruses, malware, and, most notably, phishing attacks [1]. Cybersecurity also ensures that all of the security properties of the organisation are intact and up-to-date from these attacks [2].

Cybersecurity research in the social sciences is related more to human than technical aspects and is evolving [3]. Understanding Cybersecurity is essential to this study because cybersecurity issues are now considered multilateral, consisting of social, physical, and technical aspects. Cyber-attacks continue to evolve over the years, from traditional attacks which emphasised 'hacking' or gaining unauthorised access to data stored

P. Ndayizigamiye et al. (Eds.): IDIA 2022, CCIS 1774, pp. 203–217, 2023.
https://doi.org/10.1007/978-3-031-28472-4_13

on a computer system. Newer innovative attacks now involve phishing and smishing, which take the softer qualitative aspect of attacks [3]. Social approaches play a critical role in successful social engineering attacks.

1.2 Need for Research into Phishing Attacks

There is a justifiable need to look at the lived experience of a phishing attack victim and how the victim approaches work engagement after the attack. While many studies have focused on phishing attack ramifications [4, 5] and preventative measures [6, 7] few focus on the human dimension. This research considers the human dimension of work engagement and sees an employee's willingness and desire to passionately perform more than just the fundamental activities associated with their job role. Phishing attack victims will not necessarily display these qualities, and it is necessary to understand how these qualities are affected. It is important to note that phishing attack victims will, after an attack, not address tasks with pride and certainty and to the best of their ability after an attack.

This work, therefore, problematises this human dimension, emphasising the effects of phishing attacks on employee work engagement. These insights are essential because if organisations pay minimal attention to how phishing attacks affect their employees, they may not point to why productivity is declining. Importantly, this work not only points out the ramifications of the employee work engagement effects but importantly proposes measures that organisations may undertake to mitigate and manage the employee work engagement effects by coming up with a theory that explains this. The work, therefore, outlines the research objective in the following section.

1.3 Research Objective

In order to gain insights into the ramifications of employee work engagement from phishing attack victims, this research work explored the social, contextual and lived experiences elicited from cyber-attack victims. A qualitative grounded theory approach was undertaken for this purpose. To gain this understanding, the research addresses the following question.

1. What are the lived experiences of phishing attack victims, and to what extent is work engagement affected?

It is felt that South African organisations would benefit from research that addresses this question. South African organisations need to be aware of how phishing attacks affect work engagement so that they can device required support to victims. Importantly, understanding work engagement addresses the appropriate response measures to enhance cybersecurity.

To meet this stated research objective, the work is presented as follows: the introduction has laid context for understanding that the work engagement of victims of phishing attacks can be affected. The extent of this is unknown. A review of literature on cybersecurity then follows after the introduction section and points to the ongoing discourse on phishing as an innovative form of a cybersecurity attack. The methodology section

that follows after the literature review outlines the adaption of the grounded theory approach to this research work. The justification for using GT to elicit, analyse and develop a substantive theory is motivated. The data analysis section that follows the methodology points to how GT techniques of coding, interpreting and categorising data were employed. The discussion and implication of the results follow in the penultimate section, and the work is concluded.

2 Literature Review

2.1 Literature Search

In order to obtain a clear understanding of the literature on phishing attacks in cybersecurity, a systematic literature search was carried out. Using Boolean operators, AND, OR, NOT [8] and keywords on select databases such as Association of Computing Machinery Digital Library (ACM DL), Association for Information Systems (AIS eLibrary), Compendex Plus (Access via Engineering Village), Emerald, Engineering Village, Henery Steward Talks, IEEE Xplore, IG Global and Inspec, as depicted in Table 1, a literature search was carried out on relevant work in order to establish a preliminary understanding of how cyber-attacks influence behaviour.

Table 1. Example of word search in databases by thematic area

Thematic area	Keyword search	Database
Cybersecurity	"Cybersecurity"	ACM DL, AIS eLibrary Science direct, IEEE Xplore, Science direct, IEEExplore, Compendex Plus, IG Global and Inspec
Attacks	"Stealth Botnet", "Computer worms", "Spear phishing", "Malware", "Trojan"	
Behaviour	"susceptibility & phishing", "fear & phishing", "motivation", "work", "Employee engagement", "Employee commitment", "Employee involvement", "benefits", "Factors reducing employee involvement"	

Results from the keyword search produced numerous articles, which the researcher evaluated to determine the relevance and appropriateness of addressing the research objective. Once relevant articles were obtained, an understanding of phishing attacks in cybersecurity was gained, and this is explained in the next section.

2.2 Background on the Rise of Phishing Cyber Attacks

Cybersecurity describes the efforts made by individuals and organisations to protect their hardware and software from continually evolving threats, such as viruses, worms, phishing, and clickjacking, as well as understanding the nature of these diverse threats to

computerised systems [9]. Phishing attacks are cybersecurity threats where the attacker utilises electronic communication channels to convey socially engineered messages to entice the individual or individuals to perform actions beneficial for the attacker. The more sophisticated phishing is spear phishing, where attackers send specially crafted emails targeting a specific individual or organisation. The attacker will carefully construct an email asking individuals to disclose their personally identifiable information (PII), such as credit card details. These emails are created in such a way as to obfuscate the actual source [10] and are tailored to an individual or the organisation the attacker wants to exploit [11].

According to a report by Verizon, phishing was the top threat in breaches in 2020. Phishing was also ranked 2nd as the top threat in incidences. Furthermore, email links were the principal malware vector used in breaches [12]. Phishing attacks severely affect organisations as they often result in the loss of confidential data, specifically PII and passwords. The attackers can sell this data on the black market to the highest bidder. Furthermore, because of a phishing attack, an organisation's reputation could be damaged, and its brand value could be negatively affected. Lastly, an organisation could also experience a drop in its share price because investors and suppliers do not want to be associated with data leaking and negative publicity [4].

2.3 Phishing Attacks on Work Engagement

Work engagement considers organisational staff members involved in their job roles physically, emotionally, and cognitively [13]. Furthermore, work engagement indicators include employees' emotional, physical, and cognitive behaviour [14]. Work engagement and motivation are closely tied to greater job satisfaction, leading to greater organisational productivity [14]. A demotivated employee is less job satisfaction, and this can have a devastating impact on any organisation. These employees could be a liability to the organisation as they are unwilling to carry out routine tasks or, if they do, will execute them poorly [15].

Studies have shown that phishing attacks affect employees of organisations negatively and that employees' trust in the Internet could be destroyed [4]. Phishing victims will shy away from interacting with Internet-based applications such as banking, procurement systems, and electronic commerce (e-commerce) [4].

Furthermore, the consequences of experiencing a phishing attack extend far beyond financial losses [16]. For example, social ramifications include loss of trust, doubt, shamefulness, and hesitation to seize opportunities. In addition, participants/employees noted that they were less likely to respond to businesses or job opportunity ads after a phishing attack as the participants were unsure whether it was a legitimate or a job opportunity.

In order to understand these ramifications, localised to the South African context, research was undertaken. The methodology used for this research is described in the next section.

3 Methodology

3.1 Grounded Theory Approach to Understanding the Impact of Phishing

This research applied the Grounded theory (GT) approach considered the most suitable to understand phishing attacks and resultant behaviour in a work engagement. GT is a well-known methodology that has been used in numerous behaviour research studies in information systems security. One of the benefits of GT is that it is compatible with both qualitative and quantitative data generation techniques. GT is used to discover or construct a theory from data. The data is obtained systematically and is analysed using comparative analysis. Despite grounded theory being a complex methodology, it is intrinsically flexible [17].

The primary principle of GT is that a theory will be constructed from actual data. Unstructured questions are asked to elicit rich and unique lived experiences [18]. These questions tend to be more flexible as the researcher wishes to understand specific aspects of the phenomenon. It avoids situations where responses to questions are convoluted and complex to analyse [18]. The approach used is illustrated by Fig. 1.

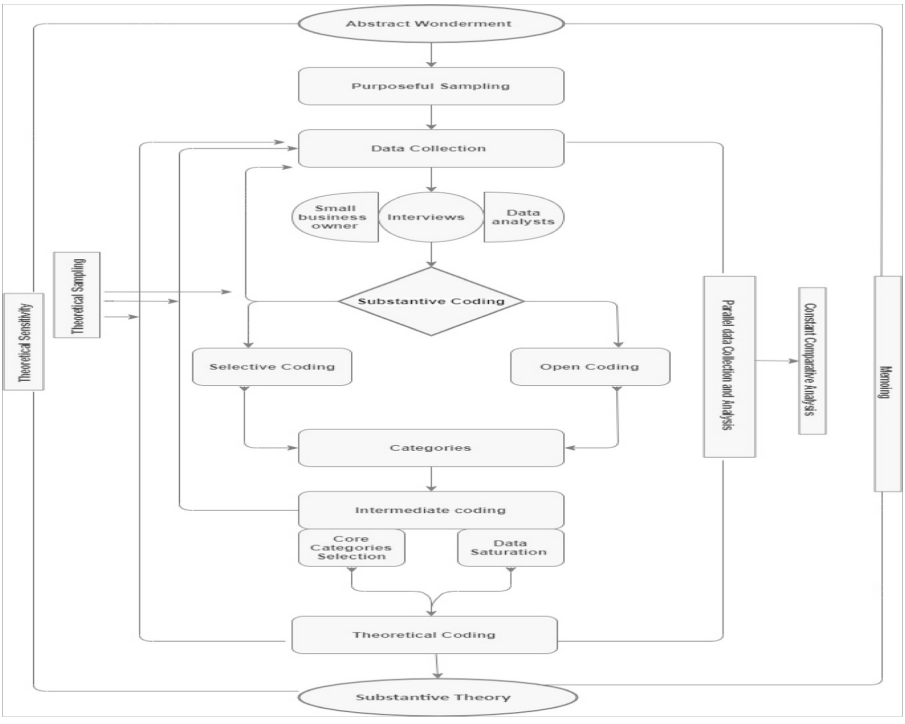


Fig. 1. The grounded theory approach. Source: Chun Tie, Y., Birks, M., & Francis, K. (2019). Grounded theory research: A design framework for novice researchers. SAGE open medicine, 7 (Page 3).

A researcher will conduct preliminary unstructured interviews to obtain general information surrounding a phenomenon. This process of gathering general information is referred to as Abstract Wonderment. This wonderment is considered the first surprise that researchers face in conducting research despite the researcher identifying their limitations in knowledge surrounding the phenomenon under study [19]. What follows is the application of the purposeful sampling technique that asks questions such as whether the study participants have ever been victims of phishing attacks.

The purposeful sampling technique took the form of asking questions such as

- If the study participants have ever been a victim of a cyberattack? Yes/No
- If (yes) - Did the cyberattack occur at a place of work? Yes/No
- If (yes), Was the cyberattack a "phishing attack"? i.e., a type of cyberattack involving social engineering usually aims to gain confidential information, i.e., login credentials or passwords.

If the individual indicated that they had been a victim of a cyberattack, it occurred at their place of work, and it was a phishing attack, i.e., yes, then the researcher scheduled an interview with the individual to obtain the worker engagement effects to work that they experienced because of phishing attacks.

3.2 Interviews, Coding and Interpretation

Ten participants were pre-selected for an interview. Before any interviews or data were collected, ethical clearance was obtained. Interviews were scheduled at a mutually agreed time. Participants were told that should they feel uncomfortable answering questions. They were not obligated to continue. All interviews were transcribed. The raw data were coded on the transcripts and consisted of two levels, the first level being open coding, followed by selective coding [20]. The study adopted the Glaserian grounded theory method. The coding process was split into two stages; In vivo coding, where participants' exact words are used, and coding based on researcher interpretation (descriptive coding). Following this process, the lived experiences of victims of phishing attacks were documented. Progressively the codes started forming a higher level of abstraction and understanding. Following this, codes were numbered numerically, i.e., CD^1, CD^2….CD^n, representing either an in vivo code or a descriptive code. The coding process is for the first participant (P1), shown in Table 2.

The coding process continued until theoretical saturation had been reached. Constant comparison of codes was made to ensure that the codes reasonably described the experiences, which would lead to the development of a substantive theory. Coding bias was avoided by allowing for systematic data analysis and the emergence of new ideas and concepts based on the interpretation that was grounded on data. The description of the categories that emerged from the coding process is detailed in the following analysis.

Table 2. Example of the coding process employed in the study

Extracts of transcripts	Open coding	
	In vivo coding	Descriptive coding
"I had to follow up something for him, so he forwarded a link…I was also crazy busy and of course because he send it on I thought it was all legit. It looked a bit weird CD[4] It was from the post office and he said I must track the parcel because he doesn't have time, and of course you know I obviously helped him with stuff…"	*It looked a bit weird*	*Suspicion*

4 Data Analysis

4.1 Codes and Categories Generated

From the transcripts of the ten participants, sixty-one codes were generated, and from these codes, ten categories were developed, and finally, five core categories emerged. This is illustrated by Fig. 2. The process of generating categories from codes and core categories from categories is the essence of the GT approach.

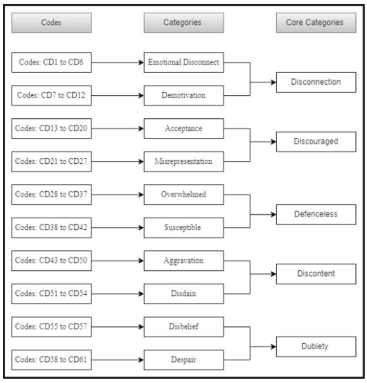


Fig. 2. Generating codes, categories and core-categories

The discussions that follow highlight how the above categories were derives from codes interpreted from participants' qualitative data. Many codes (underlined) were derived for each category, but for illustrative purposes, few selected codes are presented in the following narration.

4.2 Emotional Disconnect

It was noted and observed that victims of phishing attacks could not consequently manage and handle the stress that followed, hindering their emotional well-being. The category "Emotional Disconnect" was therefore derived from open codes from the following participants: Participant 1 (P1) mentioned that "*I don't know*[CD1] *what do you do" as a way of divorcing from commitment*. P2 explained that without support from management after the incident, commitment would remain low, thus advocating their intervention. "*Your emotional commitment*[CD2] *[will only] remain [high] by them being there for you and providing support*" P7 shared a similar sentiment and resounded this support by advocating raising answers through education, "*Basic education for non-technical people [is needed]*[CD3]". However, P4 shared a different opinion suggesting that emotional commitment need not be affected when one is confident in handling an attack suggesting, *"… give somebody that [has] the experience or … more confidence in handling these situations*[CD4]". P3 provided an additional measure proposing that employees be "*…allowed or provided with content*[CD5] *regarding the [nature] of attack…*" to raise awareness. As mentioned earlier, P4 motivated for "*… educating their staff on an emotional level*[CD6]" I think what the company…needs to do that. Six codes were derived from these interviews and were aggregated to form the category Emotional Disconnect.

4.3 Demotivation

The category 'demotivation' was derived from open codes as follows: P3 suggested that management support was crucial in encouraging a victim of a phishing attack to bounce back to normal after a phishing incident. P3 states, *"…a manager encouraging you to [continue working]*[CD7]". Moreover, P8 contended that motivating a phishing attack victim was necessary since "*threats continue evolving*[CD8]*, understanding different threats and also understanding their business…is required*". P8 added an interesting perspective alluding to a demotivated and pessimistic outlook after a phishing attack "*I don't think there is much more that the organisation can*[CD9] *do… Right?*" Additionally, P1 emphasised the importance of receiving support from management, given that a phishing attack victim could feel as they are going in the wrong direction or that is nothing is going right, "*Yeah, we did speak about it, but definitely it made a huge difference to get back on track and be as motivated as before…Because like I said, it felt like you know you're going 10 steps backwards*[CD10]" P7 shared a contentious view holding that while phishing attacks were inevitable, there were no outright solutions: "*I think it is a matter of understanding that firstly when we talk about phishing attacks….There is no way*[CD11] *that you can have 100% bulletproof solution to that*". This view could perhaps motivate employees indirectly in order not to allow this to hamper work engagement. While it may not be possible to have a 100% bulletproof solution for email phishing attacks, as alluded to by one of our participants, another participant, P1, believed that the basics work best, while not a solution to email phishing attacks, it could potentially provide a bit of reprieve from them, and stipulated that "*I think in hindsight we all just said…that we have to have a better/proper backup*[CD12] *On all the systems*". Six codes were derived from these interviews and were aggregated to form the category Demotivation.

4.4 Acceptance

The category 'acceptance' was derived from open codes as follows: P4 contended that it would be better to find continuity and accept that while the attack has happened, this should not hold back on work engagement. "*I'm talking about living with it once it's happened*[CD13]". P1 also contended that it would be better to move on. "*I don't think there's anything that [one] could have done to be honest*[CD14]" and that "*there's nothing that they could*[CD15] *actually do to help*". Similarly, P7 echoed similar views suggesting that "*I don't think there's anything that they can do*[CD16]". P3 was more insight and suggested a possible way forward after an attack by introspecting on existing weaknesses. "*Allow us to interact with our business structure*[CD17] *to understand the business structure*". In addition, P5 shared the same perspective mentioned earlier as they said, "*I wouldn't say my work performance was affected…Because I just took it as one of these…Hard lessons to learn from*[CD18]". P10 echoed similar sentiments as they stipulated that "*I expect this to happen [receipt of phishing emails is considered "normal"]*[CD19]" Lastly, P1 provided additional clarification and agreed with the other perspectives put forward as they said, "*Well, I think we try to motivate each other, and you know the nature of the beast we all like miff and upset and so you just have to get on*[CD20]*, you know…And if you don't want to, you can pack your stuff and leave*". Eight codes were derived from these interviews and were aggregated to form the category Acceptance.

4.5 Misrepresentation

The category 'misrepresentation' was derived from open codes as follows: P4 gave an example of what could happen when facts were misrepresented: "*[They sent] fictitious purchase orders*[CD21] *to everybody that was [currently] in the mailbox and anyone who had been sent an email in the past*" Additionally, P4 highlighted the reason for misrepresentation was for personal gain, cautioning against these kinds of attacks:"[It is one way of] "*sourcing new business…you [now] think twice when you get an email*[CD22] *from any-body*" P3 shared the concern that the email misrepresentation and phishing attacks was widespread. "everybody's [computer] is compromisedCD23 …" Additionally, P4 described a probable cause as to why email phishing attacks could be increasing, by stipulating that "*People are desperate for business*[CD24]*, so they just see a purchase order and they Click to open and it causes absolute chaos for us*" Lastly, P7 said misrepresentation could adversely affect planning and decision making, "*also affects your planning because there were certain things [tasks] that you planned… but now you have to change*[CD25]". A potential solution to this increasingly frequent phenomenon to curb email phishing attacks and their associated misrepresentation was stipulated by, P7 who said "*basic education for non-technical people*[CD26]" This potential solution was expanded upon by P8, who said "*I think Preventative, as in you know, pre getting it…There's lots of E learnings and courses and even attendance….Classroom based training that's forced upon you as mandatory training to take right…every organisation I've worked for, they very diligent….In terms of….Ensuring all employees go through sufficient training*[CD27]*…to be aware…Of these type of things and how…To deal with it, so you know regarding their processes and so forth so*". Seven codes were derived from these interviews and were aggregated to form the category Misrepresentation.

4.6 Overwhelmed

The category 'overwhelmed' was derived from open codes as follows: P4 suggested that as a victim of a phishing attack, this whole experience was overwhelming, impacting on work engagement (performance). "*So yeah, I'd say for a good week and a half [this event] impacted our performance*[CD28]". P1 pointed out similar concerns on the impact on performance. "*So, you're almost like nervous*[CD29] *in a way to…do any [work] because of…, you know these guys [Perpetrators] are obviously very clever*[CD30]". Furthermore, P4 shared similar sentiments by indicating that their actions were rather frantic as opposed to calm and collected "*Obviously you are desperately trying to put out fires*[CD31]" P3 shared a potentially devastating impact on work engagement "*But after the…. phishing attack… my action plan file, which I store on my desktop…that file was also compromised…and those files are also no [longer] …readable*[CD32]". P3 also noted that while work engagement was unaffected before the phishing attack, once the attack was initiated, the whole experience was overwhelming to the point of abandoning work. "*So, before everything is like very smooth and fast, I knew where I stored all my files, all the files and folders are very well organised on my system",* but after the attack, *"I felt like running away*[CD33]". P1 contended that "*you actually feel …demotivated completely*[CD34]". P5 pointed out that "*it took a bit longer to complete certain tasks*[CD35]" P6 shared a different view by saying, "*I don't think it's really affect it*[CD36]". Lastly, P4 shared the harsh reality by saying, "*Our output because of all the extra efforts we've got to put in to scrutinise emails and all the…Rest of it…Quite possibly 10% [performance loss] on a daily basis*[CD37]". Ten codes were derived from these interviews and were aggregated to form the category Overwhelmed.

4.7 Susceptible

The category 'susceptible' was derived from open codes as follows: P3 perceived himself as always being the cautious type, only to fall victim to a phishing attack leaving him feeling susceptible and vulnerable. "*Before…I only heard about this…. but I never realized*[CD38] *I could [be a victim] and once this happened [to] me, then my interest on cyber security increased as I felt vulnerable*[CD39]. *I don't want to sleep*". P1 provided more insight regarding how to overcome susceptibility and suggested the importance of being more alert and cautious in the future. "*…yeah, we did speak about it…it made a huge difference to get back on track*[CD40] *and be motivated as before*". Lastly, P3 shared a potential solution to reducing susceptibility to email phishing attacks by ensuring that from both an organisation and individual perspective. Once a victim has received training on how to identify and not engage with phishing emails, there needs to be some form of additional training or assessment that revaluates the individual's understanding regularly. As they said, "*Accepted the fact that this happened…And now he [the manager] wants you [the victim/participant] to learn from it [phishing email, which the participant interacted with]… and he says, you know, if this were to happen again [receiving another phishing email], what would we [the participant] do differently*[CD41]". P10 confirmed the need for training of non-technical individuals and overall awareness about phishing emails and suggested it could be a potential solution for reducing susceptibility, by saying that "*Arrange training and make sure that this stuff [phishing attacks] is aware*

of[CD42]". Five codes were derived from these interviews and were aggregated to form the category Susceptible.

4.8 Aggravation

The category 'Aggravation' was derived from open codes as follows: P7 was left feeling frustrated or annoyed after they had been a victim of an email phishing attack. As they indicated that their planning had to be reworked to address vulnerabilities that had been uncovered by stipulating that "*but in so doing it also affects your planning because there were certain things [tasks] that you planned on doing, but now you have to change back and start saying, OK, let's park this [postpone this task] for now*[CD43]." *Let's expedite this [implement this task as soon as possible]*[CD44]". P1 shared similar sentiments as they indicated that they had to complete mundane tasks, such as restoring data, due to an email phishing attack, as they stipulated "*so everybody had to reload [data/information] everything*[CD45]". Next, P2 provided another mundane task by saying that "*Phoned AMD… But then still she felt that it would be best if we update my password and everything*[CD46]". As if having to deal with mundane tasks was not undesirable enough, P4 further added that they had to resolve the direct damage of the email phishing attack and the indirect damage by stating that "*Obviously you are desperately trying to put out fires*[CD47] *[various issues that are arising due to the email phishing attack]*". P4 also provided a harsh reminder of how devasting phishing attacks can be, adding that, "*So I don't know How much actual business…I'm deleting because I don't want to risk opening up that email…you cautious and that, but at the same time you kind of thinking okay but I could potentially be losing a new customer*[CD48]". Although P9 did not necessarily have to complete mundane tasks, they still expressed their overall frustration with the situation, by stating that "*and I think it's like makes me feel irritated and frustrated*[CD49]". Conversely, one positive outcome that could be salvaged from an individual experiencing a phishing attack, as indicated by P8, as they said "*you kind of use that historical….Experience to now start to be a little bit more proactive*[CD50]*, so even though things have happened*". Seven codes were derived from these interviews and were aggregated to form the category Aggravation.

4.9 Disdain

The category 'disdain' was derived from open codes as follows: P1 was left feeling disdain after experiencing an email phishing attack. As they stipulated that they were unable to use their digital infrastructure or company computer network, as the email phishing attack had caused the system to become unusable, unless a ransom was paid, by stipulating that "*They can't use the actual system*[CD51]*, which makes it easier because you have to wait until all the work is captured again*". P1 went on to provide additional information, indicating that if the company had implement the necessary protections and trained their staff about phishing attacks, the consequences of the outcome could have been minimised, which would not have resulted in their employees expressing these emotions, as they stipulated that "*Yeah MJ, I don't think there's anything that they could have done to be honest, it was just a matter of because like I said if the proper stuff was in place before this happened*[CD52]*, I think less damage would have taken… Place and if we*

were trained properly[CD53]*…To lookout for these things…But when it actually happened and after …I mean there's nothing that they could actually do to help*". Furthermore, P3 shared similar sentiments as they also had the unpleasant experience of experiencing resentment towards their organisation as they had to reset all their credentials, by stating that "*Every password is compromised*[CD54] *and that time just because of my computer already have like firewalls and all that so they don't access my personal account*". Four codes were derived from these interviews and were aggregated to form the category Disdain.

4.10 Disbelief

The category 'disbelief' was derived from open codes as follows: P1 was left in disbelief after experiencing an email phishing attack, as they could not believe the extent of the damage caused, by stating that "*It was terrible. It was the worst feeling ever*[CD55]*…the damage was done…it started already from the morning*". P3 echoed similar sentiments as they could not access any content on their Personal Company Computer (PCC), as they stipulated that "*I will tell you like this is a very horrible experience*[CD56]". P6 also expressed their concern after an email phishing attack, stating, "*I was confused, stressed*[CD57]". Three codes were derived from these interviews and were aggregated to form the category Disbelief.

4.11 Despair

The category 'Despair' was derived from open codes as follows: P1 was left in despair after they had experienced an email phishing attack. As they were unaware and did not have the knowledge, skills or training to deal with a phishing attack, they stated that "*I sent him [an expert in Information Technology] a screenshot from my phone*[CD58] *[to show what was happening with the computer system]*". Similarly, P3 shared similar sentiments as they stated that "*Immediately inform [victim informs] our [their company's] incident response team*[CD59]". P4 echoed a familiar perspective, as they stated, "*the only thing [no other viable option] that I could do was close down [revoke the license] that email address*[CD60]". However, P8 indicated they were not left in despair as they managed to recognise and halt the file being downloaded before any severe damage could be done, as they "*cancelled it [file being downloaded] quick enough*[CD61]". Four codes were derived from these interviews and were aggregated to form the category Despair.

5 Discussion

5.1 Deriving Core-categories from Categories to Formulating a Substantive Theory

From the 6 categories generated from open coding, the next phase according to GT approaches was to compare the codes with codes and codes with these categories. This process in known as constant comparative analysis which involved sorting and organising data into groups in a structured way to formulate a new grounded theory. Glaser, Strauss, & Strutzel [21], suggest that it is important to revise core-categories called "normative expectation" as data moves to a higher level of abstraction.

5.2 Emergence of Substantive Theory

The normative expectation as data analysis moved to higher levels of abstraction was for a theory to emerge. The theory that emerged then is shown by Fig. 3.

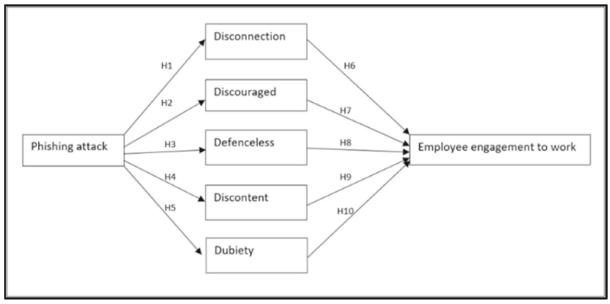


Fig. 3. Emergence of Substantive Theory (of Phishing effects on work engagement)

Figure 3 presents the following propositions:

- A phishing attack incident is likely to create a disconnection to work.
- A phishing attack incident is likely to leave the employee felling discouraged.
- A phishing attack incident is likely to leave the employee feeling defenseless.
- A phishing attack incident is likely to leave the employee feeling discontent.
- A phishing attack incident is likely to leave the employee having dubiety (doubtfulness).
- Feeling disconnected to work will negatively influence work engagement.
- Feeling discouraged will negatively influence work engagement.
- Feeling defenseless will negatively influence work engagement.
- Feeling discontent will negatively influence work engagement.
- Having dubiety (doubtfulness) will negatively influence work engagement.

Since the aim of a GT approach is to develop a theory, the outcome of this research as shown by Fig. 3, has elucidated the use of GT in context of employees working around Johannesburg area and fulfills this criteria. The substantive model shows five important core categories namely, *disconnection*, *discouraged, defenseless, Discontent, and Dubiety* as outcomes of a phishing attack. These five emotive feelings will likely determine work engagement. Organisations are therefore encouraged to adapt measures that can help employees manage these emotions. What is important is that the employees should receive the support that they require. Organisations can also adapt additional measures such as educating employees on risks imminent in the cyber space, remaining current with existing attacks, and developing policies and procedures to handle phishing attacks. These measures could alleviate the feelings of *disconnection*, *discouraged, defenseless, discontent, and dubiety.*

5.3 Implications on Information Systems Theory

The theoretical strength of this research lies in showing through original work, how GT can be used to formulate a theory that can predict work engagement from phishing attacks. The research work has developed some theoretical propositions from the substantive theory which can be generalised outside the Johannesburg business contexts to the broader national and international contexts. The research work was done is such a way as to minimise research bias for this purpose. It is from the deep interaction with the study participants, that the theoretical basis to predict work engagement was determined. The research work, therefore offers new insights and perspectives that can shape and improve the existing body of knowledge in cybersecurity and specifically phishing attacks.

5.4 Implications for Organisations and Practice

Reflecting of the substantive theory developed in this research, potential use of the substantive theory is identified as follows: Professionals 'workspace is increasingly becoming Internet driven and computing technology centric. This increases the likelihood of these professionals experiencing email phishing attacks. The overarching benefit of this work is that professionals as well as organisations could be able to aid their employees in returning to their optimal level of performance if the issues around *disconnection*, *discouraged, defenseless, discontent, and dubiety* are addressed. Practitioners can therefore use this theory as a guide to phishing attacks management.

6 Conclusion

This paper has demonstrated that beyond the technical aspects of cybersecurity that a lot of research has been undertaken, the social contextual aspects has often been overlooked. This paper shows the importance of this social contextual aspect namely, the emotions phishing attacks elicits on employees and the effects these emotions have on work engagement. From work carried out using the grounded theory approach, five emotions namely, *disconnection*, *discouraged, defenseless, discontent, and dubiety* were developed that formed the basis of the substantive theory for this work. The development of this substantive theory was underpinned by a rigorous process of coding and categorising of data supports these findings. It is envisaged that researchers and practitioners will find value and benefit from this research.

References

1. Craigen, D., Diakun-Thibault, N., Purse, R.: Defining cybersecurity. Technol. Innov. Manag. Rev. **4**(10) (2014)
2. Von Solms, R., Van Niekerk, J.: From information security to cybersecurity. Comput. Secur. **38**, 97–102 (2013)
3. Benson, V., McAlaney, J., Frumkin, L.A.: Emerging threats for the human element and countermeasures in current cybersecurity landscape. In: Cyber Law, Privacy, and Security: Concepts, Methodologies, Tools, and Applications, pp. 1264–1269. IGI Global (2019)

4. Jain, A.K., Gupta, B.B.: A survey of phishing attack techniques, defence mechanisms and open research challenges. Enterp. Inf. Syst. **16**(4), 527–565 (2022)
5. Mihai, I.C.: Overview on phishing attacks. Int. J. Inf. Sec. Cybercrime **1**, 61 (2012)
6. Singh, A.P., Kumar, V., Sengar, S.S., Wairiya, M.: Detection and prevention of phishing attack using dynamic watermarking. In: Das, V.V., Thomas, G., Lumban Gaol, F. (eds.) AIM 2011. CCIS, vol. 147, pp. 132–137. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-20573-6_21
7. Chaudhry, J.A., Chaudhry, S.A., Rittenhouse, R.G.: Phishing attacks and defenses. Int. J. Secur. Appl. **10**(1), 247–256 (2016)
8. Yushtina, A.: Ontology-based data extraction in the scholarship-related content. MS thesis (2013)
9. Gupta, C.P., Goyal, K.K.: Cybersecurity: a self-teaching introduction. Mercury Learning and Information (2020)
10. Cui, Q., et al.: Tracking phishing attacks over time. In: Proceedings of the 26th International Conference on World Wide Web (2017)
11. Shakela, V., Jazri, H.: Assessment of spear phishing user experience and awareness: an evaluation framework model of spear phishing exposure level (spel) in the namibian financial industry. In: 2019 International Conference on Advances in Big Data, Computing and Data Communication Systems (icABCD). IEEE (2019)
12. Bhardwaj, A., et al.: Why is phishing still successful? Comput. Fraud Secur. **2020**(9), 15–19 (2020)
13. Abu-Shamaa, R., Al-Rabayah, W.A., Khasawneh, R.T.: The effect of job satisfaction and work engagement on organisational commitment. IUP J. Organ. Behav. **14**(4) (2015)
14. Kuok, A.C.H., Taormina, R.J.: Work engagement: evolution of the concept and a new inventory. Psychol. Thought **10**(2) (2017)
15. Winkler, K., Saur, C.: Employee retention management: long-term retention of employees–a comparison of generations. J. Appl. Leadersh. Manag. **7**, 96–111 (2019)
16. Kelley, C.M., et al.: Something smells phishy: exploring definitions, consequences, and reactions to phishing. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 56, no. 1. SAGE Publications, Los Angeles (2012)
17. Strauss, A., Corbin, J.M.: Grounded Theory in Practice. Sage (1997)
18. Engward, H., Davis, G.: Being reflexive in qualitative grounded theory: discussion and application of a model of reflexivity. J. Adv. Nurs. **71**(7), 1530–1538 (2015)
19. Stoupe, D.: Understanding abstract wonderment: the reflections of a novice researcher. Grounded Theory Rev. **15**(2) (2016)
20. Glaser, B.G., Strauss, A.L.: The Discovery of Grounded Theory: Strategies for Qualitative Research. Routledge (2017)
21. Glaser, B.G., Strauss, A.L., Strutzel, E.: The discovery of grounded theory; strategies for qualitative research. Nurs. Res. **17**(4), 364 (1968)

Information Security Framework Adoption for South African Small and Medium Enterprise

Michael N. Moeti(✉), Makhulu R. Langa, and Khuliso Sigama

Tshwane University of Technology, Polokwane 0699, South Africa
{moetimn,langarm,sigamak}@tut.ac.za

Abstract. The small and medium enterprises (SMEs) sector is the backbone of the economy, a big employment creator, poverty reducer and the main engine for economic growth. SMEs rely heavily on technology to conduct their business but are constraint by lack of information security knowledge to protect themselves against cyber-attacks that may render their business an easy prey. This paper investigates factors that may influence SMEs in South Africa in the adoption of a security framework. These factors were applied in the development of an information security adoption framework for SMEs. The study collected qualitative data by means of interviews. ISO/IEC 27002 information security framework formed the theoretical basis for this study. Thematic analysis was employed for analysis of the collected data. The results indicate that close to 90% of SMEs do not have security policies or a risk management strategy. The study identified new, important themes critical in the development of an information security adoption framework for SMEs: ISM Best Practices and technical security architecture. The study contributes a conceptual framework that illustrates how the identified themes relate to concepts essential to information security adoption among SMEs.

Keywords: SMEs · Security policy · Thematic analysis

1 Introduction

The consensus among policy makers, practitioners, and economists and business experts is that small and medium enterprises (SMEs) are drivers of economies globally (Shin and Hwang 2015; Sungkawati et al. 2021). The SME sector is akin to the backbone of the economy, a big employment creator, poverty reducer and constitutes the main 'engine' for economic growth (Lekhanya 2015; Lekhanya et al. 2017). These sways are true for both developed and developing countries. However, SMEs are also encountering competitive pressure fueled by globalization, compliance with legislation, resource scarcity, the explosion of knowledge pertaining to information management and an increase in market expansion due to emerging technologies and innovation (Lekhanya et al. 2017). Information technology (IT) is a critical organizational resource supporting the competitive strategy of the organization. ICT has become the driving force behind businesses and is changing the way businesses operate. In today's high-speed global business world, organizations need to be always-on, always-connected computing for

P. Ndayizigamiye et al. (Eds.): IDIA 2022, CCIS 1774, pp. 218–233, 2023.
https://doi.org/10.1007/978-3-031-28472-4_14

employees and for their business partners and customers (Weinman 2017). Most business activities are conducted by computer networks, which facilitate communications among organizations.

The rapid improvements in technology have led to a rise in information security concerns. Aman et al. (2021) argues that technology has also brought unnecessary complexities and boundless opportunities for slip-ups. Technology development out-paces the development of control practices and adequate in-house expertise for small businesses. Moreover, the increased usage of the Internet and telecommunication devices has brought risks, threats, and vulnerabilities to small organizations (Saffady 2020). Most of the small organizations have insufficient resources, skills and knowledge to safeguard their information management systems (Valli et al. 2021). Although security is a well-recognized issue, it has been noted that many organizations, particularly SMEs, do not apprehend all the steps needed to shield themselves, Valli et al. (2021) further alluded, which makes them an easy prey to information technology bullies. With cyber-crime on the increase, security threats are becoming increasingly sophisticated and harder to detect. This may in turn lead to data leakage, down-time and reputation loss, which in turn lead to loss of existing customers resulting in a negative impact on the organization's bottom line and ultimately profit margins (King 2021).

Numerous studies (Jørgensen 2015; Nagahawatta et al. 2021; Shojaifar and Järvinen 2021) ranging from the academia to the practitioners have contributed to this research area, providing several methods, frameworks and models to define security policy adoption in SMEs. The terms SME or SMME are used interchangeably. SMEs needs to have a secure IT environment but very often this is upended by other priorities. This shortcoming makes small business entities insecure and prone to hackers and various forms of malpractices that threaten effective information management in small business entities. SMEs regularly find the task of keeping their business functions aligned with their security process highly challenging.

Unlike large organizations, SMEs frequently lack security measures and policies to secure their computer resources. Their network infrastructure is often maintained by end-users with limited computer technical skills and knowledge (Nagahawatta et al. 2021). Smith and Ali (2019) stated some of the common worst mistakes committed by computer users as follows:

- Opening email attachments from unverified sources.
- Failing to install software security patches for commonly used applications such as Microsoft Office.
- Writing their confidential information such passwords on paper.
- Downloading and installing games and screen savers from untrusted sources.
- Failing to run regular backups and or verifying the integrity of backups.

SMEs are not always aware of the security threats that may negatively affect their companies (Shojaifar and Järvinen 2021), as they don't fathom that security is more than just pre-venting viruses and blocking spam. They need a proper awareness on information security and the right measures of preventing security breaches. The aim of this study was to conceptualize a security framework for SMEs.

Key Concepts

SME in South African Context

Rehman and Anwar (2019) pointed to the fact that the geographical placement of the SME as well as country specific legislation influence the numerous small or medium enterprise definition. A small to mid-size enterprise (SME) organization upholds revenues, assets and employs less than 250 employees. The standards of measuring whether an organization is an SME, also vary among countries and industries. In this paper South African standards of determining an SME are used. Defining an SME can be a challenging task (Lekhanya et al. 2017). SMEs comprise over 95% of the economy globally. They are the driving force behind many innovations, which contributes to employment creation, investments and exports (Lekhanya 2015). SMEs are naturally born survivors; they are content to survive if they make a decent existence (Alonso-Almeida et al. 2018). Most of their functions are often 'patched up together' lacking any degree of integration and sophistication. Policies and frameworks for information planning and disaster recovery are usually non-existent which makes them more vulnerable to cyber security issues (Srivastava et al. 2019).

As is the case in most countries, the definition of SMEs is a challenge. South African authors and researchers have exploited multiple ways to define an SME by looking at different perspectives such as the size, environment of operation, owner managed level and whether it is semi-formal or formal in relation to the economy development. The South African National Small Business Act 102 of 1996 officially provided definitions for small businesses which were revised by the National Small Business Amendment Acts of 2003 and 2004 (DTI 2008). Table 1 below, shows the enterprise definitions given in the National Small Business Act, for the classification of micro, small or medium enterprises.

According to Table 1 it is very clear that South African definitions are like that of other developing countries such as Ghana, Kenya and Nigeria (Mambula 2002; Quartey et al. 2017), whereby SMEs are classified as start-up enterprises or survivalist enterprises. Furthermore, most of the definitions are categorized according to the size of the enterprises, the total number of employees paid, the total annual turnover and total assets fixed value as well as the impact of the SMEs to the economy.

Generally, small and medium enterprises (SMEs) are regarded as the main engines for economic development since they play a pivotal role towards the growth of the country (Amoah et al. 2022). Furthermore, they are labor intensive, capital saving and capable of helping to create most of the employment the world needs by the end of the century (). In developing countries small organizations produce a considerable share of their gross domestic product (GDP) and are considered as a key source of employment generation, innovative business ideas as well as viewed as a breeding ground for entrepreneurship (Chipunza and Naong 2020).

In Nigeria SMEs, make up to 90% of all the available businesses and at least account for 70% of the country's employment. In Kenya, SMEs employ almost 87% of the workforce (Murithi 2021). In South Africa, the Department of Trade Industry (2020) indicated that more than half of all the employment comes from organizations having less than 200 workers and it contributes about 50% to 60% to the gross domestic product (GPD) with a possibility to increase regularly (Abor and Quartey 2010; Taiwo and

Table 1. Classification of micro, small or medium enterprises

Sector or sub-sectors in accordance with the standard industrial classification	Size or class	Total full-time equivalent of paid employees less than	Total annual turnover less than: total annual turnover less than	Total gross asset value (fixed property excluded) less than
Agriculture	Medium Small Very small Micro	100 50 10 5	R 4.00 m R 2.00 m R 0.40 m R 0.15 m	R 4.00 m R 2.00 m R 0.40 m R 0.10 m
Mining and quarrying	Medium Small Very small Micro	200 50 20 5	R 30.00 m R 7.50 m R 3.00 m R 0.15 m	R 18.00 m R 4.50 m R 1.80 m R 0.10 m
Manufacturing	Medium Small Very small Micro	200 50 20 5	R 40.00 m R 10.00 m R 4.00 m R 0.15 m	R 15.00 m R 3.75 m R 1.50 m R 0.10 m
Electricity, gas and water	Medium Small Very small Micro	200 50 20 5	R 40.00 m R 10.00 m R 4.00 m R 0.15 m	R 15.00 m R 3.75 m R 1.50 m R 0.10 m
Construction	Medium Small Very small Micro	200 50 20 5	R 20.00 m R 5.00 m R 2.00 m R 0.15 m	R 4.00 m R 1.00 m R 0.40 m R 0.10 m
Retail and motor trade and repair services	Medium Small Very small Micro	100 50 10 5	R 30.00 m R 15.00 m R 3.00 m R 0.15 m	R 5.00 m R 2.50 m R 0.50 m R 0.10 m
Wholesale trade, commercial agents and allied services	Medium Small Very small Micro	100 50 10 5	R 50.00 m R 25.00 m R 5.00 m R 0.15 m	R 8.00 m R 4.00 m R 0.50 m R 0.10 m
Transport, storage and communications	Medium Small Very small Micro	100 50 10 5	R 20.00 m R 10.00 m R 2.00 m R 0.15 m	R 5.00 m R 2.50 m R 0.50 m R 0.10 m

Falohun 2016). Furthermore, the 2019 StatsSA report estimated that more than 95% of South African businesses were SMEs and that these SMEs contributed over 60% of total employment in the country.

a. ***Information Security practice within SMEs***

Most of the information security frameworks and strategies that were originally developed for large organizations or government institutions may not be beneficial or practical in SMEs (Kaila and Nyman 2018). Small organizations generally lack resources such as finances, technology, skilled human resource, and proper infrastructure. With the lack of funds SMEs cannot have the same resources large organizations possess to secure their networks appropriately. Additionally, many owner-managed SMEs deem information security to be a low priority.

However, SMEs and large organizations have a lot in common despite their small footprint in the global marketplace. They use the same technology to conduct business, offer services to clients of which some are relatively located in remote areas, and they provide access to their database systems (Wang 2019). Large organization have the necessary platform for handling security issues within their human resources capabilities and funds galore to mitigate risks better than SMEs. Within SMEs it is believed that security can be provided with the help of configuring a firewall and deploying an anti-virus program (Lenhard 2022). On top of this, their ICT infrastructure is maintained by end-user, usually someone who has limited computer technical skills or knowledge of IT security.

The emerged technologies and information system have enabled work activities to be accomplished more efficient and effectively than ever before. This technology has also brought unnecessary complexities and a great room for error and developed much quicker than the development of security controls and adequate in-house expertise for both SMEs and large organizations (Saffady 2020).

b. ***Information Security policies within SMEs***

Information security policy is often referred to as a living document which is used to specify a set of rules, guidelines and behaviors on organizational assets, but it is continuously modified as technology and employee requirements change (Bryson 2018). This definition is directly in line with the international standard on information security management as it plans to secure the organization's physical and technological resources as well as the information it handles. Furthermore, an organization's security policy may include allowed or disallowed behavior, broad guiding protocols to be used to achieve goals, backup procedures, compliance and enforcement with the policy to ensure that necessary procedures are followed.

The fact that an organization has a security policy is an indication that it is committed to protect the confidentiality, integrity and availability of all the systems they use as well as the information it handles. However, if employees of the organization are not keen or are unwilling to follow security polices, the efforts to develop one are wasted (Spurling et al. 2018). Lenhard (2022) emphasized that in organizations without policies, particularly SMEs, security practices will be developed without clear ambitions and responsibilities. For example, most large organizations establish computer security technologies and practices to secure information resources, although security cannot be attained through technical tools only and their impact and effectiveness are far from over.

The recent concerns on security policies continues to grow in academic literature (Feng et al. 2019; Wu et al. 2021) and point out the need of empirical investigation to the security incidents. While in the last decade, IT security policies were never talked about and unheard to the outside world, currently they are regarded by organization of all sizes as one of the most important cornerstones of IT security. Organizations are now taking action to protect their own information and information entrusted to them by customers, suppliers and partners. They are forming structures and programs to address and evaluate both internal and external risks and threats to their electronic information (Menard et al. 2017). However, most academics and practitioners' communities, as Balozian et al. (2019) assert, have generally focused on common issues related to information security policies which include the awareness and behavior of employees with information systems.

2 Methodology

This paper employed an interpretivist approach to support the belief that reality is constructed by subjective perceptions and predictions (Littlejohn and Foss 2008). Researchers who select this paradigm are only interested in the social construction of meaning. People have a free will, purposes, goals, and intentions. Therefore, people should be studied as active agents. As a result, the researcher can draw themes from within a particular context (Farrugia 2019). It can then be said that human beings create their own reality associated with their own belief systems and values (Bakker and Lelkes 2022).

The reason behind applying an interpretivist approach in this paper was to understand the opinions of the key stakeholders from different SMEs in SA through interviews with an effort to understand reasons that may lead their organizations to adopt an information security framework. This approach allowed the researcher to understand the complexity of social and human variables within the information security adoption phenomenon and where it is located. Since these variables are difficult to measure, the researcher applied processes with an intense enquiry to understand and analyze the data. A focus group was selected for sampling to get the in-depth views of the participants (Subiantoro 2018). In total, twenty-four in-depth interviews were conducted with key stakeholders from different SMES (see Table 2). Participants experience and social factors are instrumental to the concepts that is being investigated, as well as the views that are generated during the interviews.

As suggested by McIntosh (2017), participants' identities should always remain anonymous to afford participants the freedom to share information regarding the phenomenon under investigation. A total number of ten participants were selected for this study. All the participants held an influential position in their organizations. There was a good mix of CEOs, senior managers, middle-level managers, and employees amongst the interviewees. To preserve the confidentiality of participants, each interviewee was assigned a special code. All the interviews were recorded and their duration varied between 45 to 60 min each. The interview guide was piloted to ensure validity and integrity.

Table 2. Participants details per SME

Background information of the participants (Source: Author) Code	Unit/Division	Position	Years of experiences within the current position
A	CE-I	Manager	10
B	CE-I	CEO	2
C	CE-I	Employee	9
D	CE-I	Manager	1
E	CE-I	Manager	2
F	CE-I	Manager	10
G	CE-I	CEO	5
H	CE-I	Employee	1
I	CE-I	Project Manager	5
J	CE-I	Project Manager	5

2.1 Thematic Analysis Process

The goal of this study was to conceptualize a security framework for SMEs that could be implemented to improve information systems security. This study used content analysis to identify themes affecting the adoption of information security within SA SMEs. To validate the identified themes, interviews were conducted, and thematic analysis was used to confirm the themes. The content analysis was used to review high-quality literature (Webster and Watson 2002). The approach considered only the Senior Scholars' Basket of Journals, and the top 50 journals, such as ACM/IEEE transactions of the AIS journal ranking (Vom Brocke et al. 2015; Fink 2019). Themes that were identified and new ones that emerged were analysed using thematic analysis. This was for verification of the themes, and determining their influence on the adoption of the security framework. During further analysis of the interviews, several related concepts emerged which guided the researcher to construct a final information security adoption framework.

This became even more evident after transcribing and reading through the first few interviews. t For this reason the researcher then employed a more data driven approach during the initial phases of analysis. The data driven approach allowed the themes to emerge primarily from the data, as opposed to using a theoretical framework upfront to seek out predetermined themes from data. After the identification of any relevant themes from the raw data, the researcher was able to relate this back to the theory obtained in the related literature. The hermeneutic cycle of Myers (2019) was employed numerous times across all six phases of analysis, with a strong focus on the identification of latent as opposed to semantic themes. The mechanisms that constituted the actual process of analysis consisted of the following phases:

Phase One: The transcription of all the interviews was conducted with great care. Phase one is recognized as an important part of the analysis process, due to the fact that it is the

foundation of all the analysis work and has given the researcher the opportunity to get acquainted with the data. Transcribing all the interviews further aided this familiarization process. During this phase the researcher made an initial list of relevant concepts that may be part of all possible themes. This in turn assisted in the execution of phase two, since at least some initial analysis had been performed.

Phase Two: Phase two execution required coding the entire data corpus, which resulted in a coding framework containing information beyond the core concepts of this research paper. This process involved analyzing each interview transcript bearing in mind the list that was created in the initial phase. Phase one coding was more data than theory driven, so as not to miss any information that might have been of interest later. This included the creation of codes for specific data extracts. According to phase two its mechanisms and data extracts could be coded multiple times. This is illustrated with examples in Table 3, where column two contains multiple codes for the data extract it is associated with (in column one). The third column in Table 3 allows for easy navigation of each participants' transcript.

Some of the codes could also be associated with more than one data extract. This type of coding is illustrated in Table 4, where the first row contains an example of three data extracts associated with one code (in column two). During the execution of phase two, the researcher was cautious not to interpret the data extracts, but to rather create a coding framework based on that which was said.

Phase Three: In phase three, the researcher identified candidate themes and associated sub-themes from the coded data extracts. An extract of one such candidate theme (and sub-themes) is given in Table 4. The alphabetic column three, is used to identify the participant from whom the code originated and subjected to further analysis. Using this form of data organization became useful during phase four where the identified themes had to be refined and their associated extracts collated. The researcher has taken great care not to eliminate any themes at this stage, but rather to form as many candidate themes as possible.

Phase Four: Phase four consisted of a dual process whereby the candidate themes were refined on two levels. Firstly, the collated data extracts had to undergo scrutiny as to whether they tied into the candidate themes with which they were associated. Secondly, once complete, evaluating the themes across the entire data set occurred. This ensured that the identified themes were valid in relation to the data set as a whole and that it captured the meanings as they were portrayed by the participants (Table 5).

This two-step process resulted in some themes being eliminated, renamed, or merged with other candidate themes. An extract of one such theme, together with the data extracts collated under its name, is illustrated in Table 6.

Phase Five: After reading the entire interview data corpus the researcher used the output of phase four (refined themes) to construct the final thematic map (see Fig. 1). As suggested by Braun and Clarke (2005), these themes were organized to avoid overlapping, which is illustrated by the fact that there is no association between the two main themes

Table 3. Data extracts that have been coded multiple times.

Data extract	Code	Line in transcript
"The major thing as far as our lot is concerned is protection and that's the thing, we get worried about from the most. Being exploited. And that is kind of foremost in our minds about if were not secured that might destroy the organization reputation and integrity. Assurance of protection increases confidents and leads to better decision making"	Concerned about protection from intruders	1
"A lot of people see security policy as something they don't use. I would not even just talk about the security threats etc. but take it from the most basic. Let's understand information security then work through the security issues etc. I think there is a lot of hype. Unnecessary hype in terms of the security. I think we are so many other things we are making a mountain out of a mole hill"	Understanding policy first then security Information security surrounded in hype	2
"Yes, and the reason being I want to know, because it firstly would be a test in terms of how they react to things. The fact that it didn't affect me would be a good sign. So, it's part of understanding how they react in terms of when they're at risk. Secondly if there's consistent breach, I would possibly want to improve the security in organization"	Training and awareness on security issues should be considered Insight into their incident response practices Constant breach prompts improved	3

("Trust in information security" and "Views as Subscribers"). As phase five primary contained the main output, it was the final thematic map which enabled the researcher to interpret the data extracts associated with these themes.

Phase Six: Phase six concluded the process of analysis resulting in the creation of a narrative based on the researcher's interpretations of -

- the identified themes and the data extracts associated with them;
- the context within which these data extracts were embedded; and
- each SME's operational and security context.

Table 4. Data extracts classified under the same code

Data extract	Code	Line in transcript
Ineffectiveness of top management is a major concern The only concern that has ever been raised is lack of resources and support from management. What happens if management's involvement is absent?	Top Management's influence	1
Combine your knowledge combine your skills, combine your understanding, and then come with recommendation[s]. Whereas [an] individual SME you might feel isolated you might be scared even financially is it the right way to go Why do you have recreated, at each SME re-establish, redevelop you know why you have to have your skill… You can't have one SME have the complete skill set to serve all the needs of the organization We know that's the truth. So, what do we do? Rather combine those strengths	Advantages ISM best practices	2

Table 5. Data extracts classified under the same code

Candidate theme	Sub-themes	Code [associated participants]
Knowledge of information security	Technical security architecture Information Security Incident Response Information security training and awareness	1

Table 6. Redefined theme with collated extracts

Refined theme	Data extracts [associated participants]
Security policy	

Throughout this narrative the researcher highlighted the contextual differences between the participating SMEs. This involved not only providing the reader with participatory statements to support these arguments, but also interpretations of the relationship between these statements and operational context of these participants and SMEs. Where

applicable, references to the literature were made and thus also formed part of the interpretations made within that context. In the following section the reader is presented with the resultant narrative, wherein all the main and sub-themes are discussed.

2.2 Theoretical Framework

The ISO 27001 standard was published in 2005 under the title "Information technology—Security techniques—Information security management systems—Requirements". It describes the requirements that an ISMS must fulfill to achieve certification. As a theoretical framework, the standard is aimed at organizations from all sectors and of all sizes. However, there is some doubt over the suitability for SMEs (Fanta 2016). Concrete measures for the fulfillment of requirements are not stipulated by the standard but must rather be developed and implemented on a company-specific basis. Certification requirements of ISO 27001 are elucidated through the elaboration of terms and concepts and supplemented with an implementation guideline within ISO 27002. The focal point of ISO 27001 is the requirement for planning, implementation, operation, and continuous monitoring and improving of a process-oriented ISMS. The approach should be aligned with the PDCA cycle (Fig. 4). The coverage and scope of an ISMS should be defined for planning and implementation. Risks should be identified and assessed (Torten 2018) and control objectives should be defined for the information and information systems. Suitable measures for protecting operations should be derived from these.

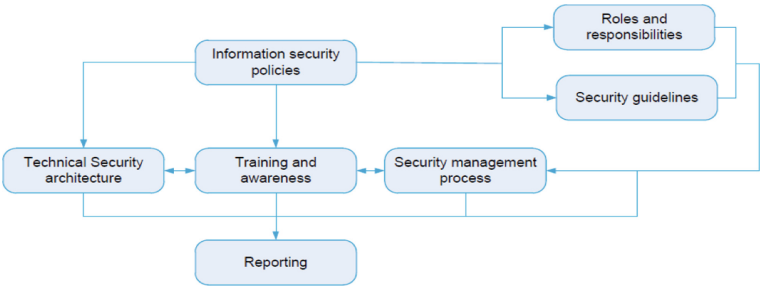


Fig. 1. ISO/IEC 27002 information security framework (Disterer 2013).

ISO 27000, "Information Technology, Security Techniques, Information Security Management Systems, Overview and Vocabulary," International Organization for Standardization ISO, (Disterer 2013). Figure 1 depicts the importance of information security policies for the roles and responsibilities, and security guidelines. They play a huge role in the development of technical architecture, training and awareness, security management, SMEs' size, and top manager's intentions. However, a lack of data security, lack of data privacy and size of IT resources hinder CC adoption.

3 Research Findings and Discussion

During further analysis of the illustrated themes (see Fig. 3), several related concepts emerged. From this the researcher decided to create a new conceptual framework. Used

in combination, these new conceptual frameworks guided the interpretation process. To better illustrate these new frameworks the conceptual framework is provided before interpreting the main theme with which it is associated (Fig. 1). It is anticipated that this will not only assist the reader in understanding how the concepts are related, but more importantly why they are related (Fig. 2).

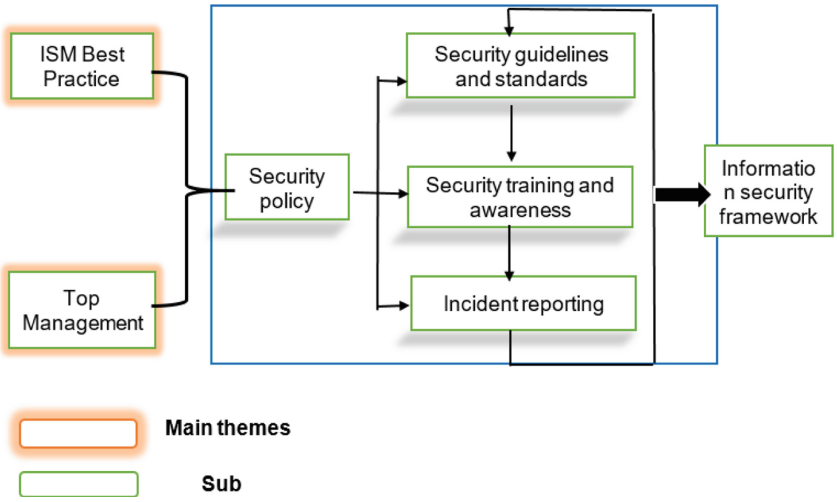


Fig. 2. Conceptual security framework for SMEs.

The proposed technical security framework and all its components developed by this study, could be utilized to mitigate the risks that could destruct the reputation and integrity of an organization. The benefits of this framework are that it is built or customized according to the SME's needs to solve specific information security problems. The modern economy has required organizations of all sizes to have information security practices, policies, frameworks and procedures in place. Without this they are vulnerable and could be exploited by threats. Furthermore, Gbadeyan, Butakov and Aghili (2017) emphasized that information security frameworks at least provide the skeletal structure an organization can look to for its security initiatives.

The developed framework is categorized into three main sections. The highest (third) level consists of living documents (policies) which are aligned with the SME's vision and mission and must be continually updated on a regular basis. This is to ensure compliance and provides rules for securing resources within an SME. The intermediate (second) level includes standards which describe statements of what must be done to comply with policy. Finally, the lowest (first) level contains practices, procedures and guidelines, including details of how employees should comply with information security policy and standards of the organization (Whitman 2016).

The findings of this study indicate that management influence has a positive impact on the overall security effectiveness of an organization. SMEs therefore, need to leverage the empirical evidence obtained from this study to improve their IT usage and security.

4 Conclusion

Most SMEs in South Africa lack strong and adequate security measures to protect their information against internal and external security threats. SMEs face security threats that cause potential loss of organization information (Jahankhani et al. 2022). These losses might be caused by the absence of security policies or ineffectual security policies within SMEs. SMEs have experienced vulnerabilities on their networks, such as data losses and theft of data. This might have been caused by a lack of security awareness among employees and managers (Saffady 2020). Information security practices, strategies and techniques for any organization, whether big or small are a necessity to protect and provide a clear road map of information security. This entails that SMEs which do not have information security in place, will have difficulty in dealing with security issues as well as managing people and protocols within the organization. This will lead to inconsistent decision making and jeopardizing of security by both internal and external members of the society. Sharing information security policies with all stakeholders is a crucial step. A training session would openly engage employees in a positive attitude to information security, which will ensure that they obtain an understanding of the procedures and instruments in place to protect the information, for example confidentiality and information sensitivity issues. Such training awareness should investigate vital topics, namely how to use/collect/disseminate/delete data, maintain data integrity, confidentiality, appropriate usage of IT technologies and correct usage of social networking.

5 Recommendations

The researcher would like to highlight several recommendations for South African SMEs who have either started or are considering a move towards employing information security practices. As with most projects a successful information security strategy adoption process should be founded on frequent communication. This should not only be done internally amongst key stakeholders, but most definitely also by the employees of the intended security solution. Regarding communication, the researcher would like to highlight the following recommendations:

Engage with Users: Key stakeholders should do this as soon as possible. After all, these are the people who will be making use of the system and in turn either deem it a useful service or not. Such engagements should include awareness campaigns with a specific focus on information security.

Engage with Other Key Stakeholders: This can take the form of regular meetings or the establishment of a forum where matters of urgency can be discussed. It is important to note that these discussions should be widely attended. If possible, engage all stakeholders. If the SME is in a pre-adoption phase, these types of discussions are vital, since they allow for the formation of information policies, guidelines, requirements and any further strategic decision making. Once adopted, such meetings may not be needed as frequently and should take place to monitor what has been implemented and make changes as required.

Employ Specialised Staff: The presence of an information security officer enhances the adoption and operation of information security practices. In fact, from the information gathered during this study the researcher deems the presence of such a staff member as not just a recommendation, but rather a requirement. Such members of staff should be tasked with the creation of security incident response procedures as well as communicating with origination members at all levels. This would entail the sharing of information that might directly affect the users, such as disclosing breaches in security. If no awareness campaigns have been conducted and the user base is uneducated, such forms of disclosure could be counterproductive, hence the need to educate and make users aware of information security.

Perform a Threat Assessment: Many key stakeholders were not aware of the specific threats. This highlights the need to not only become familiar with these threats, but to also identify the likelihood of them occurring within the operational context of their SME. Measures need to put in place to address internal threats, since most of the participants' regard this as a concern.

Although these recommendations are aimed at both start-ups and well-develop SMEs, it would be wise for organizations to investigate them further before starting an information security adoption campaign. The researcher anticipates that such investigations will allow future researchers to break these recommendations up into even more detailed components.

References

Abor, J., Quartey, P.: Issues in SME development in Ghana and South Africa. Int. Res. J. Financ. Econ. **39**(6), 215–228 (2010)

Alonso-Almeida, M.D.M., Bagur-Femenias, L., Llach, J., Perramon, J.: Sustainability in small tourist businesses: the link between initiatives and performance. Curr. Issue Tour. **21**(1), 1–20 (2018)

Aman, A.H.M., Shaari, N., Ibrahim, R.: Internet of things energy system: Smart applications, technology advancement, and open issues. Int. J. Energy Res. **45**(6), 8389–8419 (2021)

Amoah, J., Belas, J., Bélas, J., Dziwornu, R., Khan, K.A.: Enhancing SME contribution to economic development: a perspective from an emerging economy. J. Int. Stud. (2022)

Balozian, P., Leidner, D., Warkentin, M.: Managers' and employees' differing responses to security approaches. J. Comput. Inf. Syst. **59**(3), 197–210 (2019)

Bakker, B.N., Lelkes, Y.: The structure, prevalence, and nature of mass belief systems. Cambridge Handb. Political Psychol. **89** (2022)

Bryson, J.M.: Strategic Planning for Public and Nonprofit Organizations: A Guide to Strengthening and Sustaining Organizational Achievement. Wiley, Hoboken (2018)

Chipunza, L.T., Naong, M.N.: Demographic variables as drivers of innovation in small accommodation businesses: a case of South Africa and Zimbabwe. Afr. J. Sci. Technol. Innov. Dev. 1–9 (2020)

Cooper, D.R., Schindler, P.S., Sun, J.: Business Research Methods, vol. 9, pp. 1–744. McgrawHill, New York (2006)

Da Veiga, A., Martins, N.: Improving the information security culture through monitoring and implementation actions illustrated through a case study. Comput. Secur. **49**, 162–176 (2015)

Disterer, G.: ISO/IEC 27000, 27001 and 27002 for information security management (2013)
Fanta, A.B.: Complementarity between relationship lending and collateral in SME access to bank credit: evidence from Ethiopia. J. Afr. Bus. **17**(3), 308–318 (2016)
Farrugia, L.: WASP (write a scientific paper): the ongoing process of ethical decision-making in qualitative research: Ethical principles and their application to the research process. Early Hum. Dev. **133**, 48–51 (2019)
Felderer, M., Katt, B.: A process for mastering security evolution in the development lifecycle (2015)
Feng, N., Wang, M., Li, M., Li, D.: Effect of security investment strategy on the business value of managed security service providers. Electron. Commer. Res. Appl. **35**, 100843 (2019)
Fink, A.: Conducting Research Literature Reviews: From the Internet to Paper. Sage Publications (2019)
Gbadeyan, A., Butakov, S., Aghili, S.: IT governance and risk mitigation approach for private cloud adoption: case study of provincial healthcare provider. Ann. Telecommun. **72**(5–6), 347–357 (2017). https://doi.org/10.1007/s12243-017-0568-5
Hartmann, S.B., Nygaard, L.Q.V., Pedersen, S., Khalid, M.S.: The potentials of using cloud computing in schools: a systematic literature review. Turkish Online J. Educ. Technol. **16**(1), 190–202 (2017)
Jahankhani, H., Meda, L.N.K., Samadi, M.: Cybersecurity challenges in small and medium enterprise (SMEs). In: Jahankhani, H., Kilpin, D.V., Kendzierskyj, S. (eds.) Blockchain and Other Emerging Technologies for Digital Business Strategies. ASTSA, pp. 1–19. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-98225-6_1
Jørgensen, K.: Integration of safety in management tasks in onshore transport SME's. In: 8th International Conference on Working on Safety: Smart Prevention for Sustainable Safety, WOS 2015, pp. 50–62. Scientific Committee (2015)
Kaila, U., Nyman, L.: Information security best practices. Technol. Innov. Manag. Rev. (2018)
King, E.E.: Bring your own device security awareness and security behavior: a quantitative explanatory study. Doctoral dissertation, Capella University (2021)
Kotler, P., Burton, S., Deans, K., Brown, L., Armstrong, G.: Marketing. Pearson Higher Education (2015)
Lekhanya, L.M.: Public outlook on small and medium enterprises as a strategic tool for economic growth and job creation in South Africa. J. Gov. Regul. (2015)
Lekhanya, L.M., Olajumoke, N.G., Nirmala, D.: Exploring fast moving consumer goods (FMCG) Small, Medium, and Micro enterprises manufacturers' need for innovation to achieve growth. Economics **8**(2), 8–16 (2017)
Lenhard, T.H.: Configuration of security systems. In: Lenhard, T.H. (ed.) Data Security: Technical and Organizational Protection Measures against Data Loss and Computer Crime, pp. 87–92. Springer Fachmedien Wiesbaden, Wiesbaden (2022). https://doi.org/10.1007/978-3-658-35494-7_18
Littlejohn, S.W., Foss, K.A.: Theories of Human Communication, 9th edn. Thomson Higher Education, Belmont (2008)
Mambula, C.: Perceptions of SME growth constraints in Nigeria. J. Small Bus. Manag. **40**(1), 58–65 (2002)
McIntosh, M. (ed.): Globalization and Corporate Citizenship: The Alternative Gaze: A Collection of Seminal Essays. Routledge (2017)
Menard, P., Bott, G.J., Crossler, R.E.: User motivations in protecting information security: protection motivation theory versus self-determination theory. J. Manag. Inf. Syst. **34**(4), 1203–1230 (2017)
Myers, M.D.: Qualitative Research in Business and Management. Sage (2019)
Oppong, S.: Between Bandura and Giddens: structuration theory in social psychological research? (2014)

Nagahawatta, R., Warren, M., Lokuge, S., Salzman, S.: Security Concerns Influencing the Adoption of Cloud Computing by SMEs: A Literature (2021)
Quartey, P., Turkson, E., Abor, J.Y., Iddrisu, A.M.: Financing the growth of SMEs in Africa: what are the constraints to SME financing within ECOWAS? Rev. Dev. Financ. **7**(1), 18–28 (2017)
Rehman, A.U., Anwar, M.: Mediating role of enterprise risk management practices between business strategy and SME performance. Small Enterp. Res. **26**(2), 207–227 (2019)
Rodriguez, A.R., de Sevilla Müller, L.P., Brecha, N.C.: The RNA binding protein RBPMS is a selective marker of ganglion cells in the mammalian retina. J. Comp. Neurol. **522**(6), 1411–1443 (2014)
Saffady, W.: Managing Information Risks: Threats, Vulnerabilities, and Responses. Rowman & Littlefield Publishers (2020)
Shin, D.W., Hwang, E.: A Lagrangian multiplier test for market microstructure noise with applications to sampling interval determination for realized volatilities. Econ. Lett. **129**, 95–99 (2015)
Shojaifar, A., Järvinen, H.: Classifying SMEs for approaching cybersecurity competence and awareness. In: The 16th International Conference on Availability, Reliability and Security, pp. 1–7 (2021)
Smith, D.T., Ali, A.I.: You've been hacked: a technique for raising cyber security awareness. Issues Inf. Syst. **20**(1) (2019)
Spurling, G., Felton-Busch, C., Larkins, S.: Aboriginal and Torres Strait Islander health. Aust. J. Prim. Health **24**(5), i–ii (2018)
Srivastava, S.R., Dube, S., Shrivastaya, G., Sharma, K.: Smartphone triggered security challenges—issues, case studies and prevention. In: Cyber Security in Parallel and Distributed Computing: Concepts, Techniques, Applications and Case Studies, pp. 187–206 (2019)
StatsSA website. https://www.statssa.gov.za/?p=13900. Accessed 17 Dec 2020
Subiantoro, I.H.: Pertunjukan Ritual Seren Taun Di Cigugur Kabupaten Kuningan Jawa Barat. Doctoral dissertation, PPS ISI Yogyakarta (2018)
Sungkawati, E., Suarniati, N.W., Hernanik, N.D., Anugerah, R.: SMEs creative economy in the Covid-19. Arch. Bus. Rev. **9**(1) (2021)
Taiwo, J.N., Falohun, T.O.: SMEs financing and its effects on Nigerian economic growth. Eur. J. Bus. Econ. Accountancy **4**(4) (2016)
Torten, R.J.: A quantitative regression study of the impact of security aware-ness on information technology professionals' desktop security behavior. Doctoral dissertation, Capella University (2018)
Vagle, M.D.: Crafting Phenomenological Research. Routledge (2018)
Valli, C., Martinus, I., Stanley, J., Kirby, M.: CyberCheck.me: a review of a small to medium enterprise cyber security awareness program. In: Daimi, K., Arabnia, H.R., Deligiannidis, L., Hwang, M.-S., Tinetti, F.G. (eds.) Advances in Security, Networks, and Internet of Things. TCSCI, pp. 233–242. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-71017-0_17
Vom Brocke, J., Simons, A., Riemer, K., Niehaves, B., Plattfaut, R., Cleven, A.: Standing on the shoulders of giants: challenges and recommendations of literature search in information systems research. Commun. Assoc. Inf. Syst. **37**(1), 9 (2015)
Wang, E.S.T.: Role of privacy legislations and online business brand image in consumer perceptions of online privacy risk. J. Theor. Appl. Electron. Commer. Res. **14**(2), 59–69 (2019)
Webster, J., Watson, R.T.: Analyzing the past to prepare for the future: writing a literature review. MIS Q. xiii–xxiii (2002)
Weinman, J.: The evolving cloud. IEEE Cloud Comput. **4**(3), 4–6 (2017)
Wu, Y., Tayi, G.K., Feng, G., Fung, R.Y.: Managing information security outsourcing in a dynamic cooperation environment. J. Assoc. Inf. Syst. **22**(3), 2 (2021)
Yeomans, L.: Qualitative methods in business research (2017)

The Effectiveness of School Anti-cyberbullying Policies and Their Compliance with South African Laws: A Conceptual Framework

Simo-Sihle Ganca and Michael Kyobe(✉)

University of Cape Town, Cape Town, South Africa
gncsim001@myuct.ac.za, michael.kyobe@uct.ac.za

Abstract. With the rise in cyberbullying incidents that sometimes result in high school student fatalities, schools are under pressure to curb the aggression. Students have been identified as ill-equipped to manage cyberbullying. At the same time, school anti-cyberbullying policies that are key to eradicating the social issue have been noted as lacking, weak and inadequate – often due to teachers' need for more experience in developing such policies and the complexity of legislation for this group. This article proposes a conceptual framework that can be used to develop effective school anti-cyberbullying policies that are compliant with laws. A systematic literature review methodology was adopted to synthesise the literature. The literature review revealed that the policy content should include attributes such as cyberbullying definition and reporting. Such policies should be developed in a consultative approach with stakeholders in the school environment and external to the school, such as policy experts, participating. This article further identifies communication of the anti-cyberbullying policy as a key element to ensuring effectiveness. Lastly, we identified the regular review of the policy as critical in keeping it effective and compliant with laws. When changes are made, these should be communicated further.

Keywords: Cyberbullying · Anti-cyberbullying · School anti-cyberbullying policy · Students

1 Introduction

Technological advances have bought incredible comforts to modern society. However, these new technologies have also bought novel challenges. Chief amongst these is cyberbullying. Although consensus on a common definition of it has not been reached, cyberbullying is characterised by its repetitive nature, the intention to cause harm and the use of technology to achieve anonymity [1, 2]. Cyberbullying is a global concern. Unsurprisingly, studies have established a correlation between increased access to mobile technology among high school students and high incidents of cyberbullying in the South African context. Notably, a study among 149 high school students in Tshwane (Gauteng Province) revealed that although traditional bullying is still slightly more prevalent (45.1%), cyberbullying has increased (44.9%) [3]. The increase is compared to studies

P. Ndayizigamiye et al. (Eds.): IDIA 2022, CCIS 1774, pp. 234–248, 2023.
https://doi.org/10.1007/978-3-031-28472-4_15

between 2012 and 2014 where cyberbullying was noted (2012: 37%, 2013: 20.9%, 2014: 15.2%) [4–6]. More recently, the concerns over rising cyberbullying incidents among South African high school students were expressed by Nyoni and Lidzhegu [7].

Due to the wide-ranging effects of cyberbullying, pressure and responsibility have been placed on schools to deal with the scourge of cyberbullying effectively. These effects often extend to the school environment, even when the incident occurred off-campus. Schools have thus been identified as the ideal location to address the growing problem [7]. Schools must implement effective measures to combat the escalating cyberbullying incidents. The Department of Basic Education (DBE) has mandated schools to implement measures such as adopting or developing anti-cyberbullying policies to mitigate the aggression [8]. These policies must comply with South African laws governing the use of electronic devices. In addition, the policies must be in line with the national Code of Conduct framework that specifically addresses cyberbullying and the security of students on and off school premises [8].

2 Literature

Although technology plays an important role in academic and social lives, it can be used maliciously. The latter is often the case for adolescents. The study by [9] posits that South African students are ill-equipped to use technology safely, leading to a high risk of cyberbullying or being cyberbullied. The uptick in technology usage among high school students has further exposed these adolescents to this threat [10].

[11] suggests that to eradicate any social issues such as cyberbullying, policies must be developed, particularly to promote a safe environment that combats cyberbullying in schools [12]. [9] surveys 169 schools across South Africa (SA) and finds that 50% of the surveyed schools had anti-cyberbullying policies in place. However, these policies were merely prohibiting the usage of cell phones on the school premises. Only a few schools, mostly developed (based on the Living Standard Measure score), call out cyberbullying and have provisions to deal with and curtail the behaviour [9]. Despite the advancement of anti-cyberbullying policies in recent years, cyberbullying is still growing [7]. In the few schools in SA with anti-cyberbullying policies, little is known regarding the effectiveness of these policies, with researchers stressing the need to evaluate the policies for effectiveness [7, 9].

2.1 Literature Review Methodology

In selecting literature to include in this study, the researchers applied the guidelines for conducting a review as proposed by [44, 46]. The guidelines propose (1) identifying the research questions, (2) identifying relevant studies, (3) evaluating and selecting literature and (4) organising and executing a literature review [52].

First, the researchers developed keywords and phrases to identify relevant literature based on the research questions.

Research Question 1 (RQ1): Are the policies adopted in South African high schools to regulate cyberbullying effective?

Research Question 2 (RQ2): Do the policies adopted in South African high schools to regulate cyberbullying comply with South African laws?

Research Question 3 (RQ3): Are the policies adopted in South African high schools to regulate cyberbullying reviewed regularly to ensure adequacy?

The keywords included 'Cyberbullying', 'Cyberbullying in High Schools', 'Cyberbullying Management', 'Anti-cyberbullying' and 'Policies'. Further, Boolean search phrases and filters were applied to limit the results, including operators 'AND' and 'OR' coupled with the keywords used.

Second, the search results were critically evaluated for appropriateness and relevancy to the study. This was achieved by reviewing the identified literature to determine whether the articles aided in answering the following research questions.

The selection criteria included an applicable date; literature from the last ten years was selected. A filter was applied based on the articles' publication date to ensure that recent work was reviewed. This was done to prove that the phenomenon investigated is persistent [45] (i.e., it fulfils Basic Claim One: 'The question that the researcher proposes to investigate is relevant and persisting in the field of study'). The researchers argue that in addition to explicitly stating the research questions and embedding the study in the body of work in the field, citing recently published literature further supports the need for the current study. Table 1 below summarises the inclusion criteria used. Although some literature before 2012 is included, these are mainly articles about traditional bullying to depict the evolution of bullying and how some bullying management issues identified with traditional bullying are still relevant to cyberbullying.

Table 1. Selection criteria

Selection criteria	Criteria overview	Inclusion rationale
Study content and methodology	Cyberbullying and anti-cyberbullying policies measures. The study context was high school settings or adolescents. Both qualitative and quantitative studies were included	Prior research indicated issues with adequately responding to cyberbullying. Further, the researchers recognised the need to record outlier and contradictory cases and competing viewpoints; hence, a wider study context was used, and there were no restrictions on the methodology used in the studies
In scope period	Studies from 2012 to 2022	To evidence the maturity of the phenomenon and the relevance of the research questions
Publication sources	Journal articles, conference papers, reputable electronic sources	These sources include in-depth, carefully analysed research and literature reviews

Third, the selected literature was synthesised, paying close attention to themes within each article. This aided the researchers in organising the literature into focus areas and subtopics. Conflicting discourse amongst scholars was also noted to ensure that the researchers highlighted where scholars disagreed. Lastly, in synthesising the literature, the researchers noted the gaps and executed the literature review below.

2.2 Theoretical Frameworks Underpinning This Study

Various theories are employed to explain and explore the cyberbullying and anti-cyberbullying phenomenon, and their factors and characteristics are explained in the different sections. These include the general strain theory (GST), the socio-ecological system theory (SEST) and the agency theory (AT). These theories are explained and operationalised in Sects. 2.3, 2.4 and 2.5, respectively.

2.3 The Extent of Cyberbullying and the General Strain Theory

Developed by Robert Agnew in 1992, the GST posits that when people are placed under undue strain, they react to this by using unhealthy and sometimes criminal coping mechanisms [13]. The GST explains how cyberbullying strains high school students, causing poor academic performance, self-esteem problems, criminal behaviour and, in extreme cases, suicide [14]. The impact is often permanent and significant [15]. The extent of cyberbullying is not limited to the bullied individual but extends to the school climate, where students feel its effect collectively [15]. Teachers are also affected directly by the disruption of school activities, whereby [14] points out that schools are the leading environment where bullying occurs. School leadership and teachers are further indirectly affected by establishing authority over cyberbullying for matters on and off campus [16].

2.4 Anti-cyberbullying Policy and the Socio-ecological System Theory

A practical, inclusive policy that sets clear procedures and guidelines enables schools to curb cyberbullying, which is fundamental to any plan to combat cyberbullying [17]. Bronfenbrenner's SEST describes how people's surroundings interact with them over their lives. That is, a person's growth is greatly influenced by their dynamic interactions with their family, friends and the community where they reside. Over time, this impacts the social environment [50]. Developed in 1977, the SEST theorises that factors across various systems influence child cyberbullying and anti-cyberbullying policies. With the child at the core, the SEST divides the levels of interaction into five distinct but nested systems/structures: the microsystem, mesosystem, exosystem, mesosystem and chronosystem. Effective anti-cyberbullying policies as an intervention cannot be achieved unless interaction with family, school and community is established [17]. As such, for this study, the SEST was used to understand anti-cyberbullying policies through the different socio-ecological systems that impact the students that such policies are intended to protect.

Anti-cyberbullying Policy Content. Anti-cyberbullying policies should enforce discipline in schools and provide a positive school environment [18]. Using the macrosystem construct, the development and implementation of anti-cyberbullying policies, a measure taken by schools to curtail the cyberbullying problem, are most effective when certain attributes are considered [19]. Ontario legislation requires school anti-bullying policies to include 1) an agreed definition of cyberbullying, 2) clear rules regarding the use of technology, 3) measures which will be followed in preventing and intervening if bullying occurs or is reported, and 4) a clear statement that cyberbullying is prohibited [20]. Researchers support this globally, including in SA [14, 16, 19]. [21] further emphasise the importance of having a policy document that is clear and concise, attributing the contents to the effectiveness of the policy.

Additionally, studies by [20, 22] and [23] show that for anti-bullying policies to be effective and compliant, the school boards should adhere to legislature enacted by provincial and national governments (macrosystem). However, teachers and students find that embedding legal jargon in anti-cyberbullying policies makes them difficult to comprehend and often results in underreporting incidents when they occur [19], perhaps partly because the laws themselves do not explicitly address cyberbullying (see Table 2). Despite the lack of cyberbullying laws and a limited understanding of how the laws can be applied, schools must still have compliant policies [9, 24]. This clearly shows the macrosystem's influence on school anti-cyberbullying policies and, consequently, on the students.

Another complexity in ensuring compliance can be seen as a double-edged sword. On the one hand, schools need to ensure that their policies and procedures uphold student safety and human rights. On the other hand, they must ensure that these policies are not infringing on laws such as the freedom of expression and privacy rights of both the victims and the perpetrators.

Table 2. Adapted from [25] and [51].

Law Identified	Overview	Factors to be considered in Anti-cyberbullying Policy Development
Bill of Rights	The constitution declares acts that involve maltreatment, abuse, violence and degradation in all forms, including cyberbullying, against children a violation of their basic rights. Refer to Section 28(1)d – every child has the right to be protected from maltreatment, abuse or degradation. At the same time, Section 28(2) requires that children's best interests be the core principle in any responses or policies. Section 10 details human dignity as one of the children's rights, which cyberbullying, by its nature, contravenes	Schools are obligated to ensure students are protected from cyberbullying There is a need to ensure policy does not infringe on human rights such as the freedom of expression right of the perpetrator (as protected by Section 16[2] of the constitution) or the privacy right of the victim (as protected by Section 14[d] which regulates communication, including the use of technology) The responses applied to address cyberbullies need to protect the bully, i.e. be appropriate for their age. Thus, the reformation and rehabilitation of the bully rather than punishment should take precedence (refer to Sections 28[2] and 33). Further, Section 33 stipulates specific attributes to be communicated to the perpetrator when disciplinary decisions are made; therefore, schools should ensure that their policies include these in their provisions Schools are further required to promote the right to education by firstly teaching students about cyberbullying and further ensuring cyberbullying is prohibited, as it impacts the education of the victim
The African Union Convention on Cyber Security and Personal Data Protection (AUC-CS)	The South African Constitution recognises that SA is to obey international laws and conventions such as the AUC-CS. The AUC-CS identifies sexting as a form of child pornography and cybercrime; therefore, the distribution of such violates the legislation	Article 8 requires the development of policy to address cyberbullying, which is cybercrime, which should detail the procedures for dealing with perpetrators (Article 25) Procedures to audit/inspect technological devices and safe-keep the evidence collected to allow for the evidence to be used in the disciplinary hearings of the perpetrator should be documented in the policy
Protection from Harassment Act 17 of 2011	This act provides ways to redress cyberbullying for students who are victims of this aggression. Under this act, cyberbullying, including cyberstalking, is seen as harassment. This act further makes provisions that allow for the victim's identity to be protected when requested to ensure retaliation is not encountered	School reporting channels should allow for anonymous reporting, which is provided for by Section 8

(*continued*)

Table 2. (*continued*)

Law Identified	Overview	Factors to be considered in Anti-cyberbullying Policy Development
Films and Publications Act 65 of 1996	This act provides legal consequences for sexting, the distribution of images or footage of child pornography to third parties or information regarding any sexual conduct with a child (see Section 27[1]). This act requires bystanders or individuals with knowledge of these acts to report these to authorities	Policies should have a provision that stipulates that bystanders should report knowledge of cyberbullying While the act prohibits the further distribution of cyberbullying (e.g. videos depicting violence or sexual assault), it is unclear whether school policies should regulate redistribution, how this will be monitored and what measures should be put in place
Cybercrimes and Cybersecurity Bill	This act criminalises the creation, distribution and participation in discriminatory messages or those purporting violence	While the act prohibits the further distribution of cyberbullying (e.g. videos depicting violence or sexual assault), it is unclear whether school policies should regulate redistribution, how this will be monitored and what measures should be put in place
Electronic Communications and Transaction Act 25 of 2002	Cyberbullying can be seen as cybercrime because it involves the humiliation and degradation of individuals using electronic devices. This act regulates cyberbullying by criminalising the act. Therefore contravening this act can lead to penalties or imprisonment	
South African Schools Act 84 of 1996	All forms of bullying are unacceptable in South African schools. Although there is no specific mention of cyberbullying, the act states that schools should be a safe environment for students by allowing them to remain abuse free. Further, the adoption of a code of conduct is mandatory, which should be completed in consultation	Developing anti-cyberbullying policies is mandatory to establish a disciplined, safe school environment that allows for education. At a minimum, the school Code of Conduct must have a regulation for cyberbullying The challenge will be regulating off-campus bullying incidents as a contract between students and the school The search and seizure of student cell phones should be done in a manner that does not infringe on students' privacy. Thus the policies should be clear on what steps are to be taken when doing this (including possibly having parents present)
The Guidelines for Prevention and Management of Sexual Violence in Public Schools	This makes provisions to address violence in schools and ensure a safe learning environment by defining which actions constitute this form of harassment	
Bill of Responsibilities for the Youth of SA	It explicitly prohibits cyberbullying and makes it the responsibility of every student, including bystanders, to report cyberbullying	Policies should have a provision that stipulates that bystanders should report knowledge of cyberbullying

Based on the above and in line with the SEST, this study recognises that the various stakeholders and factors could influence the effectiveness of anti-cyberbullying policies, and these are examined below.

Anti-cyberbullying Policy Development. According to [19], the DBE, which can be seen as the school community, the parents' community and the experts (collectively referred to as the advice community), such as the police, lawyers and policy consultants, fall under the exosystem. Teachers regard the involvement of the DBE as important in ensuring the standardisation of anti-cyberbullying policies across schools [9]. This is because most schools do not have the capabilities (i.e. expertise and experience) to develop such policies independently. The DBE has disseminated standards and frameworks to prevent cyberbullying and intervene in schools. This clearly shows the influence the DBE (an exosystem) has on schools' development of anti-cyberbullying policies. However, merely stipulating requirements is insufficient, and more is required of the DBE in developing such policies to be effective.

[16] further suggests that developing a school anti-cyberbullying policy should explore consulting with policy development and cyberbullying experts. These policies must be broad enough to cover all cyberbullying possibilities [26]. Consultation during development will ensure that the anti-cyberbullying policy adheres to relevant legislation clearly and understandably [27]. Consulting is permissible under Section 30(1) of the South African Schools Act. The Guidelines (Paragraph 1.5) require that in developing an anti-cyberbullying policy, consensus must be reached before the policy can be adopted.

Developing an anti-cyberbullying policy should also involve the broad school community, including parents, teachers and students [17]. This can be achieved by making the draft policy available to these stakeholders for analysing [28]. [29] highlighted that students' commitment to the school environment and, thereby, to schools' rules has a direct influence on the level of cyberbullying. Therefore, it could be inferred that robust anti-cyberbullying policies may not reduce cyber aggression without students' engagement.

Policy Communication. Proper communication of anti-cyberbullying policy is essential because only after parents and students are aware of such a policy can responsive procedures be taken [16]. Furthermore, it is highlighted that legislation mandates that anti-bullying policies be reviewed and that the policies are to be communicated to the school community, which includes staff, students and parents [30]. The policy should also be communicated to stakeholders when changes are made, as required by the DBE's guidelines.

Communicating this policy to parents will also increase their awareness of the seriousness and danger of cyberbullying. A knowledgeable parent will, in turn, teach their children about online threats, thus reducing the rate of children's exposure to cyberbullying [7, 31, 32].

2.5 Need to Review the Policy

[20] stresses that monitoring and revising anti-cyberbullying policies is necessary to ensure they are relevant and effective in addressing bullying. This is because having such policies does not guarantee their effectiveness [17]. Further, scholars have identified the review of policies as one of the features of the anti-cyberbullying programme that positively impacts the success of such programmes [33]. [16] recommends that strategies such as a regular review of anti-bullying policies can effectively address challenges in SA. Schools that do not review their anti-cyberbullying policies to ensure they are still relevant and reflect emerging facets of cyberbullying and technology ignore the risks posed by cyberbullying [34].

[21] note a lack of auditing and reviewing implemented anti-cyberbullying policies in schools; thus, the effectiveness and compliance of these policies still need to be discovered, despite a notable increase in interest in understanding whether policies are effective. Similarly, scholars have noted a lack of empirical evidence for evaluating anti-cyberbullying policies [35]. Literature and best practice recommend conducting an annual policy audit to ensure relevance and sufficient coverage [14].

The AT theory is used to highlight further the significance of periodically evaluating policies and the requirement for independent reviews to ensure the review's objectivity. This theory is further used to explain how good school governance practices create a better relationship between schools and the stakeholders such as the DBE, the school governing body and parents. This, in turn, results in the well-being of students through reduced cyberbullying.

Agency Theory. According to the AT, a working relationship between a principal (such as the DBE) and an agent (such as teachers) is characterised by opportunistic or self-interested behaviour. [48] suggest that this relationship creates an 'agency problem' where schoolteachers may report that they have implemented and adopted anti-cyberbullying policies that comply with DBE requirements. However, the extent to which the policies are fit-for-purpose may be overestimated. Hence, to reduce this agency problem, scholars suggest a need to review the policies to determine their adequacy and revise policy content to meet its objective of reducing cyberbullying in schools [49].

[36] and [14] suggest that the evaluation of anti-cyberbullying policies should be conducted with an investigational approach. This is evidence-based, which is a characteristic of an audit. Adopting this approach would improve on studies such as [37], which is a review and not evidence-based. Similarly, other studies include interviewing teachers and students to assess the efficacy of policies and content analysis of cyberbullying policies. While these are beneficial, they are prone to bias (the agency problem) and neglect consideration of how the policies were developed. [38] postulates that teachers and school management report more success in implementing anti-cyberbullying strategies than the students; this is attributed to the lack of teacher awareness [18]. In 2020, an evaluation of the maturity of South African cybersafety programmes was conducted and revealed the need for a post-anti-cyberbullying evaluation, where the results are used to update the policy [39].

The need for an evidence-based approach is also highlighted by the requirement of policymakers and stakeholders outside the school for quality information on the anti-cyberbullying policy to make effective decisions and improve their judgement of the policy's usefulness. This is supported by the lending credibility theory (LCT), which explains that audits enhance the confidence of stakeholders in audited policies [40]. This study adopts an audit approach to the review of the anti-bullying policies in schools.

2.6 Conceptual Model

Using the literature review as a foundation, the following conceptual model is suggested to further empirically examine the effectiveness of anti-cyberbullying policies adopted in schools. The proposed conceptual model presented in Fig. 1 depicts that (1) policy content, (2) policy development, (3) policy communication and (4) policy compliance and review influence the effectiveness of anti-cyberbullying policies. In turn, the effectiveness of anti-cyberbullying policies positively influences the extent of cyberbullying in schools and schools' compliance with regulations and legislation.

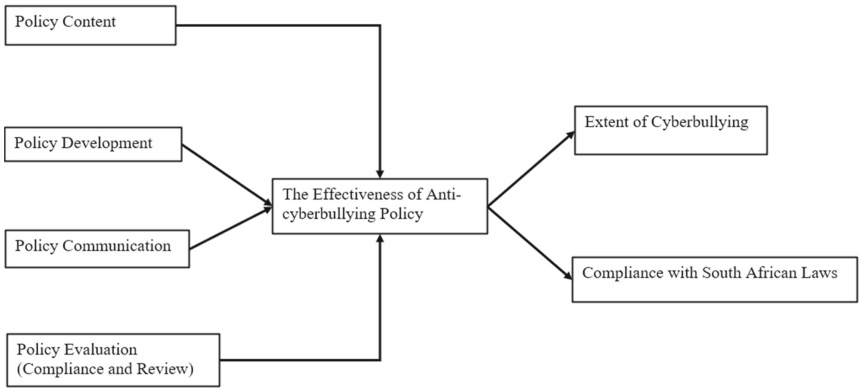


Fig. 1. The conceptual framework developed by the authors.

Table 3 below summarises the identified constructs and briefly describes the items relevant to each, as depicted in Fig. 1.

Table 3. Conceptual framework constructs definitions

Theory	Construct	Construct definition	References
SEST	Policy Content	In this paper, 'policy content' refers to the components of school bullying policies. These components guide the effectiveness of the policy by ensuring that all important aspects are covered	[9, 14, 16, 18–23]
SEST	Policy Development	In this paper, 'policy development' is defined as a guidance that has been given in establishing the document and the involvement of school board associations, the government and experts in the field to ensure that the documents are effective in addressing cyberbullying	[9, 16, 17, 19, 26–29]
SEST	Legislation and guidelines	This refers to how cyberbullying law requirements have been incorporated into the anti-cyberbullying policy	[22, 27, 30, 41–43]
SEST	Communication	In this paper, 'communication' refers to making the anti-cyberbullying policy known to parents and children	[9, 17, 19, 26, 28, 29]
AT, LCT	Compliance and Review	'Compliance' in this paper refers to conforming to legal requirements outlined in an attempt to eradicate cyberbullying in schools	[7, 16, 30–32]

3 Application of the Proposed Conceptual Framework

First, the framework identifies the relevant attributes to be included in the policy to ensure its comprehensiveness. This breakdown is especially useful for teachers finding it difficult to develop these policies.

Second, this framework highlights that teachers or schools cannot on their own develop comprehensive policies, and further be used to leverage the DBE for contribution and support in developing these. As set out by the SEST, the successful integration of the microsystem, macrosystem, exosystem and mesosystem enable collaboration between schools developing policies and stakeholders with varying interest in reducing

cyberbullying. This allows for the effective management of cyberbullying through comprehensive policies. The framework further aids in identifying stakeholders to include in policy development and ensures that policies developed by schools are not contravening South African law and applicable international law. The factors to be considered in this matter are presented in Table 2.

Last, the framework identifies the importance of regularly reviewing the policies to ensure continued relevance and extensiveness with changes communicated. Thus, reducing the agency problem that agency theorists posit. The framework is not limited to cyberbullying but can also be used to enhance traditional bullying policies.

The researchers will validate the framework by testing the developed hypotheses. The hypotheses will be tested through two questionnaires, one evaluating the policies (content, developers, communication and review) and the other determining management's view of the effectiveness of their policies and understanding of incidents that have occurred in their schools.

4 Conclusion

Although traditional bullying is still slightly more prevalent, the rise of cyberbullying incidents among vulnerable high school students is concerning. The rise in accessibility to mobile technology, especially among students, exacerbates the issue. Schools have been identified as the ideal location to address the phenomenon, with various laws, guidelines and principles placing responsibility squarely on schools. This article provided a view into the existing literature, identifying anti-cyberbullying policies' attributes and theories supporting these. This was used to ground the proposed framework in theoretical work. The applicable laws and factors that should be included in anti-cyberbullying policies were identified and presented. The article further identified that teachers are struggling with developing ant-cyberbullying policies and need assistance. Thus, anti-cyberbullying policies should be developed in a consultative manner, with stakeholders across the school environment participating. These stakeholders include teachers, students, parents, experts and the DBE. When the development of policy is left to schools, inadequate policies result.

Limitations. Two main methodological limitations were identified in this study. First, the researchers only considered literature published in English. [47] found that in conducting systematic literature reviews, non-restrictive searches yielded higher-quality scores than those that restricted their search to English. Further, the evaluation of the suitability of the literature for inclusion or exclusion in the study was performed by one researcher.

References

1. Smith, P.K., Kupferberg, A., Mora-Merchan, J.A., Samara, M., Bosley, S., Osborn, R.: A content analysis of school anti-bullying policies: a follow-up after six years. Educ. Psychol. Pract. **28**(1), 47–70 (2012). https://doi.org/10.1080/02667363.2011.639344
2. Tokunaga, R.S.: Following you home from school: a critical review and synthesis of research on cyberbullying victimization. Comput. Hum. Behav. **26**(3), 277–287 (2010). https://doi.org/10.1016/j.chb.2009.11.014

3. Du Preez, C., Prinsloo, J.: An exploration of bullying in school context with specific focus on cyberbullying as a form of cybercrime in Tshwane, Gauteng. CTA Criminol. South. Afr. J. Criminol. **30**(3), 103–120 (2017)
4. Burton, P., Leoschut, L.: School Violence in South Africa Results of the 2012 National School Violence Study CENTRE for JUSTICE and CRIME PREVENTION MONOGRAPH SERIES, No. 12 CAPE TOWN (2013)
5. Popovac, M., Leoschut, L.: Cyber bullying in South Africa: impact and responses. Centre Justice Crime Prev. **13**(6), 1–16 (2012)
6. Tustin, D., Zulu, G., Basson, A.: Bullying among secondary school learners in South Africa with specific emphasis on cyber bullying. Child Abuse Res. S. Afr. **15**(2), 13–25 (2014)
7. Nyoni, J., Lidzhegu, E.M.: Description search results page ARTICLE the whole-school approach to manage cyberbullying: lessons from South African schools. J. Educ. Stud. **20**(1), 55–75 (2021)
8. Western Cape Government. Bullying at School. https://www.westerncape.gov.za/general-publication/bullying-school. Accessed 16 July 2022
9. Kritzinger, E.: Short-term initiatives for enhancing cyber-safety within South African schools. S. Afr. Comput. J. **28**(1) (2016). https://doi.org/10.18489/sacj.v28i1.369
10. Mkhize, S., Gopal, N.: Cyberbullying perpetration: children and youth at risk of victimization during Covid-19 lockdown. Int. J. Criminol. Sociol. **10**, 525–537 (2021). https://doi.org/10.6000/1929-4409.2021.10.61
11. Dasgupta, K.: Youth response to state cyberbullying laws. N. Z. Econ. Pap. **53**(2), 184–202 (2018). https://doi.org/10.1080/00779954.2018.1467959
12. Dorio, N.B., Clark, K.N., Demaray, M.K., Doll, E.M.: School climate counts: a longitudinal analysis of school climate and middle school bullying behaviors. Int. J. Bullying Prev. **2**(4), 292–308 (2019). https://doi.org/10.1007/s42380-019-00038-2
13. Moon, B., Morash, M.: General strain theory as a basis for the design of school interventions. Crime Delinq. **59**(6), 886–909 (2012). https://doi.org/10.1177/0011128712466949
14. Gillespie, G.L., Willis, D.G., Amar, A.F.: Review and application of the national academies of sciences, engineering, and medicine bullying or cyberbullying recommendations for screening and lesbian, gay, bisexual, and transgender youth. Nurs. Outlook **66**(4), 372–378 (2018). https://doi.org/10.1016/j.outlook.2018.03.003
15. Cilliers, L., Chinyamurindi, W.: Perceptions of cyber bullying in primary and secondary schools among student teachers in the eastern cape Province of South Africa. Electron. J. Inf. Syst. Dev. Ctries. **86**(4) (2020). https://doi.org/10.1002/isd2.12131
16. Smit, D.: Cyberbullying in South African and American schools: a legal comparative study. S. Afr. J. Educ. **35**(2), 1–11 (2015). https://doi.org/10.15700/saje.v35n2a1076
17. Chalmers, C., et al.: School policies on bullying and cyberbullying: perspectives across three Australian States. Educ. Res. **58**(1), 91–109 (2016). https://doi.org/10.1080/00131881.2015.1129114
18. Mucherah, W., Finch, H., White, T., Thomas, K.: The relationship of school climate, teacher defending and friends on students' perceptions of bullying in high school. J. Adolesc. **62**, 128–139 (2018). https://doi.org/10.1016/j.adolescence.2017.11.012
19. Pennell, D., Campbell, M., Tangen, D.: What influences Australian secondary schools in their efforts to prevent and intervene in cyberbullying? Educ. Res. **62**(3), 284–303 (2020). https://doi.org/10.1080/00131881.2020.1795701
20. Roberge, G.D.: Countering school bullying: an examination of policy implementation in Ontario and Saskatchewan. Int. J. Educ. Policy Leadersh. **6**(5) (2011). https://doi.org/10.22230/ijepl.2011v6n5a305
21. Siyam, N., Hussain, M.: Cyber-safety policy elements in the era of online learning: a content analysis of policies in the UAE. TechTrends **65**(4), 535–547 (2021). https://doi.org/10.1007/s11528-021-00595-8

22. Smith, P.K., Smith, C., Osborn, R., Samara, M.: A Content analysis of school anti-bullying policies: progress and limitations. Educ. Psychol. Pract. **24**(1), 1–12 (2008). https://doi.org/10.1080/02667360701661165
23. Swearer, S., Limber, S., Alley, R.: Developing and implementing an effective antibullying policy. Bullying Prev. Interv. Realistic Strateg. Sch. 39–52 (2009)
24. Barry, B.: Schools and the Law: A Participant's Guide. Juta, Cape Town (2006)
25. Hills, C.A.: Developing a Law and Policy Framework to Regulate Cyber Bullying in South African Schools. Doctoral dissertation, University of South Africa, Pretoria (2017)
26. Caruso, N.: What would you do if… your board were considering an antibullying policy? Am. Sch. Board J. (2009)
27. Lerman, B.: Addressing bullying: policy and practice. Princ. Leadersh. **11**(1), 34–37 (2010)
28. Sullivan, K.: The Anti-Bullying Handbook, 2nd ed. Sage Publications Ltd., Thousand Oaks (2010)
29. Kowalski, R.M., Limber, S.P., McCord, A.: A developmental approach to cyberbullying: prevalence and protective factors. Aggress. Violent. Beh. **45**, 20–32 (2019). https://doi.org/10.1016/j.avb.2018.02.009
30. Batterbee, W.: Cyber Bullying: Bare Schools Filling the Legal Gap. Doctoral Dissertation, North-West University (2014)
31. De Lange, M., Von Solms, R.: An E-safety educational framework in South Africa. In: Proceedings of the Southern Africa Telecommunication Networks and Applications Conference (SATNAC) (2012)
32. Kortjan, N., Von Solms, R.: A conceptual framework for cyber security awareness and education in SA. S. Afr. Comput. J. **52** (2014). https://doi.org/10.18489/sacj.v52i0.201
33. Ansary, N.S., Elias, M.J., Greene, M.B., Green, S.: Best practices to address (or Reduce) bullying in schools. Phi Delta Kappan **97**(2), 30–35 (2015). https://doi.org/10.1177/0031721715610088
34. Hinduja, S., Patchin, J.W.: Cyberbullying: a review of the legal issues facing educators. Prev. Sch. Fail. Altern. Educ. Child. Youth **55**(2), 71–78 (2011). https://doi.org/10.1080/1045988x.2011.539433
35. Hatzenbuehler, M.L., Schwab-Reese, L., Ranapurwala, S.I., Hertz, M.F., Ramirez, M.R.: Associations between antibullying policies and bullying in 25 States. JAMA Pediatr. **169**(10), e152411 (2015). https://doi.org/10.1001/jamapediatrics.2015.2411
36. Ertesvåg, S.K.: Improving anti-bullying initiatives: the role of an expanded research agenda. J. Educ. Change **16**(3), 349–370 (2014). https://doi.org/10.1007/s10833-014-9232-2
37. Ttofi, M.M., Farrington, D.P.: Effectiveness of school-based programs to reduce bullying: a systematic and meta-analytic review. J. Exp. Criminol. **7**(1), 27–56 (2010). https://doi.org/10.1007/s11292-010-9109-1
38. Rigby, K.: How teachers address cases of bullying in schools: a comparison of five reactive approaches. Educ. Psychol. Pract. **30**(4), 409–419 (2014). https://doi.org/10.1080/02667363.2014.949629
39. Kritzinger, E.: Improving cybersafety maturity of South African schools. Information **11**(10), 471 (2020). https://doi.org/10.3390/info11100471
40. Ittonen, K.: A Theoretical Examination of the Role of Auditing and the Relevance of Audit Reports (2010)
41. Patchin, J.W., Hinduja, S.: Measuring cyberbullying: implications for research. Aggress. Violent. Beh. **23**, 69–74 (2015). https://doi.org/10.1016/j.avb.2015.05.013
42. Bridie, H.: Minister for Education and Skills Ruairí Quinn T.D., publishes new anti-bullying procedures for all primary and post primary schools. https://merrionstreet.ie/en/category-index/education/schools/minister-for-education-and-skills-ruairi-quinn-t-d-publishes-new-anti-bullying-procedures-for-all-primary-and-post-primary-schools.html#:~:text=The%

20Minister%20for%20Education%20and%20Skills%2C%20Ruair%C3%AD%20Quinn. Accessed 27 Aug 2022
43. Stuart-Cassel, V., Bell, A., Springer, J.F.: Analysis of State Bullying Laws and Policies; Office of Planning, Evaluation and Policy Development, US Department of Education (2011)
44. Creswell, J.W.: Educational Research: Planning, Conducting and Evaluating Quantitative and Qualitative Research. Pearson, Boston, Massachusetts, USA (2012)
45. Ngwenyama, O.: The ten basic claims of information systems research: an approach to interrogating validity claims in scientific argumentation. SSRN Electron. J. (2019). https://doi.org/10.2139/ssrn.3446798
46. Kitchenham, B.: Procedures for Undertaking Systematic Reviews; Joint Technical report, Computer Science Department, Keele University (TR/SE-0401) and National ICT Australia Ltd (0400011T.1) (2004)
47. Wang, Z., Brito, J., Tsapas, A., Griebeler, M., Alahdab, F., Murad, M.H.: Systematic reviews with language restrictions and no author contact have lower overall credibility: a methodology study. Clin. Epidemiol. **7**, 243–247 (2015). https://doi.org/10.2147/clep.s78879
48. Jensen, M.C., Meckling, W.H.: Theory of the firm: managerial behavior, agency costs and ownership structure. J. Financ. Econ. **3**(4), 305–360 (1976)
49. Franco-Santos, M., Nalick, M., Rivera-Torres, P., Gomez-Mejia, L.: Governance and well-being in academia: negative consequences of applying an agency theory logic in higher education. Br. J. Manag. **28**(4), 711–730 (2017). https://doi.org/10.1111/1467-8551.12249
50. Adu, J., Oudshoorn, A.: The deinstitutionalization of psychiatric hospitals in Ghana: an application of Bronfenbrenner's social-ecological model. Issues Ment. Health Nurs. **41**(4), 1–9 (2020). https://doi.org/10.1080/01612840.2019.1666327
51. De Wet, A.: Father to sue education department over rape'-civil remedies for sexual violence and harassment in public schools. Child Abuse Res. S. Afr. **14**(1), 22–35 (2013)
52. Veldheer, S., et al.: A systematic scoping review of how healthcare organizations are facilitating access to fruits and vegetables in their patient populations. J. Nutr. **150**(11), 2859–2873 (2020). https://doi.org/10.1093/jn/nxaa209

Human-Computer Interaction (HCI) for Digital Inclusion

Perceived Usability in Learning Management Systems' Adoption: A Study of Higher Education Institutions in Ghana

Adolph Sedem Yaw Adu[1,2(✉)] and Judy Van Biljon[2]

[1] Department of Computer Science, Ho Technical University, Ho, Ghana
aadu@htu.edu.gh
[2] School of Computing, University of South Africa, Pretoria, FL, South Africa
vbiljja@unisa.ac.za

Abstract. The digitalization of education has made learning management systems (LMSs) an indispensable tool for facilitating teaching and learning in higher education institutions (HEIs). Research studies on the factors that influence LMS usage in Sub-Saharan Africa is lacking specifically in considering both technology adoption and usability as theoretical lenses. This study investigates the factors that influence the use of LMSs in HEIs in Ghana. Specifically, what factors influence perceived usefulness (PU) and perceived usability (PUsab) and if PU and PUsab can be used together in modelling the usage of the LMS. This quantitative study used a validated questionnaire to conduct a survey with 300 respondents from three Ghanaian HEIs. The data was analyzed using the Partial Least Squares Structural Equation Modeling regression technique. The findings confirm that human factors (self-efficacy and learning style), social influence and access to technical support influence PUsab and the same factors also influence perceived PU. Both PUsab and PU influences LMS use among students in Ghana. The constructs PU and PUsab were found to be highly related when evaluating the factors influencing LMS usage and could thus not be adequately distinguished to be modelled together. Therefore, the construct PUsab could not replace perceived ease of use (PEU), as previous studies have suggested. While the constructs of PUsab and PEU seem semantically similar, PUsab includes the effectiveness construct, which could explain the relatedness between PUsab and PU. This finding has important implications for understanding how the technology adoption literature intersects with the usability literature.

Keywords: Learning Management System Perceived usability · Perceived ease of use · Perceived usefulness · LMS · TAM · Ghana

1 Introduction

Education is one of the main drivers of socio-economic development towards addressing the challenges of unemployment and inequality [1]. The impact of education on socio-economic development is enhanced by the integration of ICT into educational systems

P. Ndayizigamiye et al. (Eds.): IDIA 2022, CCIS 1774, pp. 251–266, 2023.
https://doi.org/10.1007/978-3-031-28472-4_16

[2]. Any form of education that requires the integration of Information and Communication Technologies (ICTs) to support the teaching and learning process is referred to as E-learning [3]. The benefits of adopting e-learning include providing a convenient and flexible learning environment and support for distance education [4] and increased access to information, improvement in the delivery of lessons, access to standardized content, learner-centeredness, on-demand availability and improved convenience for learners [5]. Learning Management Systems (LMSs) are tools for implementing ICT in formal educational institutions, specifically the deployment of LMSs for implementing e-learning in HEIs [6]. LMSs are currently the main platform for the delivery of knowledge in e-learning environments [7, 8]. LMS is an interactive web-based software that provides a platform for communication between lecturers and students and among students [9]. An LMS enables course creation and administration, uploading of lecture materials, submission of assignments and assessments and monitoring of students' progress. Popular LMSs include Moodle, Blackboard, Caroline, EdX and Sakai [10]. Moodle, Sakai and Blackboard are the most commonly deployed LMSs in Africa [11]. The benefits of LMS adoption in HEIs include improved student satisfaction, improved performance [12], lower student dropout rates [13] and reduction in the cost of providing study materials and the sharing of resources to learners [14].

Deployment of LMSs in HEIs is on the rise in countries across Africa [11, 15–17]. Unfortunately, this does not apply to the adoption and use of LMSs among students in African HEIs [18]. However, students believe LMSs can support teaching and learning and are willing to adopt them if the usability issues around these LMSs are addressed [15, 18]. Usability is defined as the "extent to which a system can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use" [19]. The constructs of usability include efficiency, effectiveness, satisfaction, memorability, errors, learnability and cognitive load [20]. Usability has been identified as one of the key determinants of LMS adoption among students [21, 22]. When LMSs have usability issues, students tend to spend time learning to use the LMS instead of focusing on learning the educational content [8, 23]; this impacts negatively on the adoption of the LMS and academic performance. Although technology adoption literature is mature, technology adoption of LMSs remains an essential issue that needs further investigation. The novelty of this study at the intersection of technology adoption and usability is the integration of technology adoption and PUsab by investigating the replacement of PEU with PUsab and the investigation of the relationship that exists between both concepts and PU in the context of LMS adoption in Ghana.

2 Problem Statement

A key measure of the success of any e-learning system deployment is the usage of that system and the lack thereof prevents the realization of the intended benefits of deployment [23]. LMSs are deployed in HEIs to achieve the benefits as stated in the previous section and when LMSs are unused by students these benefits are unrealized. Therefore it is essential to identify the factors influencing the use of LMSs and further devise strategies to overcome the barriers to LMS usage in HEIs [24]. Even though lecturers are the main implementers of LMSs in teaching and learning, success of LMSs

in education depends on their acceptance and use by students [25]. The Technology Acceptance Model (TAM) is the most frequently used framework in extant literature for identifying the factors influencing the usage of LMSs in HEIs [26]. In addition, usability has also been identified as one of the factors that influence the adoption of LMSs in extant literature [21, 22]. The current study, therefore, investigates the factors influencing the usage of LMSs in HEIs in Ghana through the integration of technology adoption and usability, specifically by replacing the PEU construct of TAM with PUsab and investigating whether PUsab can be modelled with PU in explaining LMS adoption.

3 Literature Review

The following differences between usability and technology adoption need to be considered:

a) Different origins: Usability encompasses the concepts of effectiveness, efficiency and user satisfaction [27], which is core to the Human-Computer Interaction field, the latter has been described as an interdisciplinary field involving Computer Science, Psychology and Education among other fields. The concept of technology acceptance originated in the field of Information Systems literature, while technology adoption originated in Sociology. Therefore, the theorizations (models, frameworks and theories) involving usability, technology acceptance and technology adoption developed independently in their respective contexts to a large extent.
b) Different level of detail: Technology Adoption encompasses perceived usefulness and PEU [29, 30] constructs that are measured by considering users' opinions. Usability measures are more objective, i.e., effectiveness can be measured by considering how many users completed the task (with or without) assistance. Efficiency can be measured by considering the time to complete the task and other measures such as mouse clicks or eye gaze patterns.
c) Different purposes: Technology Adoption predicts acceptance, while usability has the interim goal of improving the system toward acceptance.

In this study, we will focus on usability and technology adoption, thus excluding technology acceptance. The following definitions are presented as a point of departure:

The following studies in explored the relationship between TAM and usability.

Holden & Rada [30] extended the PEU construct of TAM to include usability, which proved more significant to technology adoption and acceptance than the initial construct of PEU. Lin [31] explored the relationship between perceived usefulness of TAM and effectiveness and efficiency attributes of usability as well as PEU of TAM and learnability and memorability of attributes of usability empirically. The causality between perceived usefulness and usability (effectiveness and efficiency) was inconclusive, while PEU was established to be linked to the usability attributes of learnability and memorability. In a similar study, Pal and Vanijja [32] confirmed that the PEU construct of TAM is more linked to PUsab than the PU through the integration of the System Usability Scale (SUS) with a revised TAM. The constructs of PUsab were limited to efficiency, effectiveness and satisfaction, not considering the impact of human factors, organizational and social

Table 1. Differences between technology adoption and PUsab

Technology adoption constructs	Perceived usability constructs
Perceived usefulness: the degree to which a person believes that using a particular system would enhance his or her job performance	**Effectiveness:** Accuracy and completeness with which users achieve specified goals
Perceived ease of use: The degree to which a person believes that using a system would be free of effort	**Efficiency:** Resources expended in relation to the accuracy and completeness with which users achieve goals
	Satisfaction: Freedom from discomfort, and positive attitudes towards the use of the product
	Learnability: Ease with which a new users can use the system to accomplish a task on the first attempt
	Memorability: Ease with which users can use a system after a period of not using it
	Errors: Frequency at which users make errors while using the system to accomplish a task
	Cognitive Load: Amount of cognitive processing required by a user to use an application

factors in investigating the relationship between the revised TAM and PUsab. The study proposed that future studies could consider replacing the PEU construct of TAM with PUsab. In a study conducted by Lah, Lewis, and Šumak [33], PUsab was linked to the PEU construct of TAM.

In summary, in extant literature the PEU construct in TAM is often linked to PUsab. Departing from the definitions as stated in Table 1, perceived usefulness seems related to effectiveness with effectiveness being on a finer grained (more detailed) level. For example, a system may be perceived as effective in accomplishing a task but considering all the other factors that influence usefulness it may not be considered useful due to personal, social, institutional factors or some other mediating factors. Ease of use would be related to efficiency but again efficiency is on a finer grained level considering resource usages such as time and effort. Another option is to consider replacing ease of use with PUsab and then measuring that in terms of effectiveness, efficiency, satisfaction, learnability, memorability, errors and cognitive load.

3.1 Theoretical Framework

The adoption and use of LMSs by students is a complex endeavor and is influenced by several factors [26]. Further, it is important to understand from the perspective of students the significant factors that influence the usage of LMSs in HEIs [24]. TAM in

its original and extended form has been used in several studies to explore the various factors that influence the usage of LMSs in HEIs, as evident in the following studies:

- Kaewsaiha and Chanchalor [24] studied the factors that influence the use of LMSs among students and teachers in Thailand. They Extended TAM with subjective norm, job relevance and perceived resource and concluded that all three factors were determinants of LMS usage in HEIs.
- Dlalisa and Govender [34] studied the acceptance and use of LMSs among academics in South Africa and found the original TAM factors of perceived usefulness, PEU, attitude and Intention were predictors of LMS usage.
- Alomari, El-Kanj, Alshdaifat, & Topal [28] investigated the human factors influencing the effective use of LMSs among students and faculty members in Saudi Arabia. The study identified three main indicators, technological and psychological characteristics as well as student-instructor interaction that influence the use and effectiveness of LMSs. Technological factors include control and attitudes, psychological factors include learning style, teaching style, self-efficacy, enjoyment, usefulness, and training, while student-instructor interaction included promptness, feedback, support and mentoring, availability and encouraged interaction and expectation.
- Coleman & Mtshazi [29] developed a framework for identifying the factors influencing the use of LMS by academic staff in a South African university. Internal factors include the ease of use and the usefulness of the LMS, self-efficacy anxiety, teaching style, personal innovativeness and lack of time. External factors include organizational, technological, social and contextual former usage and practices.
- Abdallah, Ahlan, & Abdullah [35] developed a theoretical framework of factors influencing instructors' adoption of LMSs through literature review. The study identified instructor factors, organization, society, belief and technology as the key determinants of LMS adoption from the perspectives of instructors. Instructor factors include attitude towards the LMS, personal innovativeness, self-efficacy, instructor confidence, computer experience, computer anxiety and knowledge. Organizational factors include facilitating conditions, management support, user training, appropriate infrastructure, technical support and technology alignment. Social factors include social influence and subjective norm. Belief factors include perceived usefulness, motivation, and PEU.
- Fung & Yuen [25] developed a framework for the factors affecting students' and lecturers' use of LMSs through a literature review. The study identified students, lecturers, technology, resources, pedagogy and curriculum as the six (6) dimensions of factors affecting the use of LMSs among students and lecturers. Individual factors include demographics, learning styles, learning strategies, particular life characteristics, access to technical resources and the need for interaction and internet connectivity. Individual factors influencing the use of LMSs by lecturers include demographics and variables related to technology (attitude towards technology, control, self-efficacy, technology literacy and teaching style). External factors include technology (features, speed, user interface design, PEU and the intention to use), resources (facilitating environment, training and support), pedagogy and curriculum.
- Asiri, Mahmud, Abu Bakar, & Mohd Ayub [36] developed a theoretical framework for the factors influencing the adoption of LMSs by faculty members in Saudi Arabia.

The study identified attitudes towards LMS, belief in e-learning and competence in using LMSs as internal factors that affect the use of LMSs by faculty members. The external factors include demographics, organizational barriers, technological barriers and social barriers.

In summary, several factors have been explored in identifying the factors influencing the adoption of LMSs, most of these studies adopted TAM factors. Some of these factors were also closely related to the PUsab construct. It was concluded that some of the factors named differently, remain conceptually the same. It was also identified that the factors from literature could be organized into TAM, human, organizational and social factors. Different from many previous studies, this study involved only students, therefore, factors selected for this study were considered relevant to students. The factors selected are depicted in Table 2.

3.2 Conceptual Framework

Figure 1 represents the conceptual framework proposed in this study based on Table 2. The TAM based factors include perceived usefulness and PEU, while the others include self-efficacy, learning style and LMS literacy categorized under human factors. Institutional access to Technical Resources, Training and Access to Technical Support are classified under Organizational support, while Social Influence is categorized under the social factor. This study introduced the PUsab factor which have usability concepts such as effectiveness, efficiency, satisfaction, learnability, memorability, errors and cognitive load.

3.3 Hypothesis

Based on the conceptual framework in Fig. 1, the following hypothesis were be evaluated.

1. H1a. Human Factors: Self-efficacy, learning style and LMS literacy positively influences perceived usefulness (PU).
2. H1b. Human Factors: Self-efficacy, learning style and LMS literacy positively influences PUsab.
3. H2a. Organizational Support: Training, access to technical support and access to technical resources positively influences perceived usefulness (PU).
4. H2b. Organizational Support: Training, access to technical support and access to technical resources positively influences PUsab.
5. H3a. Social influence positively influences perceived usefulness.
6. H3b. Social influence positively influences PUsab.
7. H4a. Perceived usefulness positively influences LMS usage
8. H4b. PUsab positively influences LMS usage.

Table 2. Factors investigated in the current study.

Factor	Sub-Factor/Component	Evidence
Human	Self-efficacy	Abdallah et al., 2016; Al-Busaidi & Al-Shihi, 2009; Alomari et al., 2020; Coleman & Mtshazi, 2017; Fung & Yuen, 2012
	Learning style	Al-Busaidi & Al-Shihi, 2009; Alomari et al., 2020; Coleman & Mtshazi, 2017; Fung & Yuen, 2012
	LMS literacy	Abdallah et al., 2016; Al-Busaidi & Al-Shihi, 2009; Asiri et al., 2012; Fung & Yuen, 2012
Organizational support	Institutional access to technical resources	Abdallah, Ahlan, & Abdullah, 2016; Al-Busaidi & Al-Shihi, 2009)
	Training	Alomari et al., 2020; Abdallah et al., 2016; Al-Busaidi & Al-Shihi, 2009
	Access to institutional support	Abdallah, Ahlan, & Abdullah, 2016; Al-Busaidi & Al-Shihi, 2009
Social	Social influence	Abdallah, Ahlan, & Abdullah, 2016; Al-Busaidi & Al-Shihi, 2009
Technology adoption related	Perceived usefulness	Alomari, El-Kanj, Alshdaifat, & Topal, 2020; Coleman & Mtshazi, 2017
	Perceived ease of use	Alomari, El-Kanj, Alshdaifat, & Topal, 2020; Coleman & Mtshazi, 2017
Perceived usability	Efficiency	Nielsen,1994; ISO 9241-11, 2018; Harrison et al., 2013
	Effectiveness	Nielsen,1994; ISO 9241-11, 2018; Harrison et al., 2013
	Satisfaction	Nielsen,1994; ISO 9241-11, 2018; Harrison et al., 2013
	Learnability	Nielsen,1994; Harrison et al., 2013

(*continued*)

Table 2. (*continued*)

Factor	Sub-Factor/Component	Evidence
	Memorability	Nielsen,1994; Harrison et al., 2013
	Errors	Nielsen,1994; Harrison et al., 2013
	Cognitive load	Harrison et al., 2013

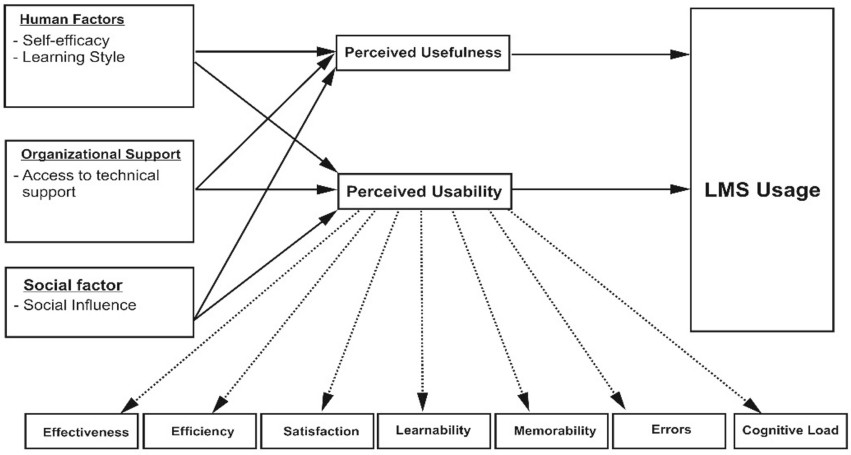


Fig. 1. Conceptual framework

4 Methodology

Towards exploring the factors influencing the use of LMSs in Ghanaian Higher Education Institutions from the perspective of both students, three universities were sampled from the southern, middle belt and Northern parts of the country. Participants in the study were hundred (100) students from each sampled university. The questionnaire was administered to only respondents that have used an LMS in the last three (3) months or the past semester. Unfortunately, the sampled university in the Northern part of the country had not used an LMS since 2013, thus this university was replaced with another university in that part of the country. Ethical clearance for the data collection was obtained from the UNISA College of Science, Engineering and Technology's Ethics Review Committee.

The developed online questionnaire was shared with participants via email and social media platforms. The social media platform used was mainly WhatsApp where the link to the online questionnaire was sent to the WhatsApp group platforms that students belong to for completion. To ensure a high completion rate, the online questionnaire was in some cases printed and physically administered among the respondents. The data obtained from the survey was analyzed using R version 4.2.0 with the lavaan library. To ensure the reliability of the developed instrument, a pilot test was done using twenty

(20) participants who did not form part of the final sample for the study. From the pilot, only factors with Cronbach alpha values greater the 0.7 were included in the final questionnaire.

The survey instrument (available from Google Drive) consisted of four (4) parts; the first part introduced respondents to the aim of the study and obtain their consent; the second part focused on participant demographics; the third part focused on the LMS adoption factors identified from literature, while the final part contained the usability factors. The factors identified from literature in Table 2 were re-worded to align with the context of this research. The responses to the questions were measured on the Likert scale of strongly Disagree, Disagree, Neutral, Agree and Strongly Agree. A total of forty-four (44) items were used to measure thirteen (13) factors influencing the use of LMSs in the Ghanaian HEI context. These factors include LMS usage (5), Self-Efficacy (3), learning style (3), Access to Technical Support (3), Social Influence associated with using LMSs (3), Perceived Usefulness (3), Effectiveness (3), LMS Efficiency (4), Satisfaction (4), LMS Learnability (4), Memorability (3), Errors (3) and Cognitive load (3).

5 Analysis

5.1 Measurement Model

The Structural Equation Model (SEM) was fitted using the lavaan library of the R language. Maximum likelihood parameter estimates (MLM) with standard errors and a mean-adjusted chi-square test statistic that are robust to non-normality was used. The MLM chi-square test statistic is also referred to as the Satorra-Bentler chi-square with robust standard errors. The latent factors were standardized, allowing free estimation of all factor loadings. Convergent validity checks how the items within the constructs are measuring what they are intended to measure, while discriminant validity checks whether items for measuring different constructs are statistically different [37]. Based on the recommendations of Gefen, Straub, and Boudreau [38], composite reliability (CR) and average variance extracted (AVE) was used to determine the reliability and convergent validity of the various constructs. For good reliability, a construct must have CR above 0.70; for convergent, validity AVE should be above 0.50 [38]. Constructs (LMS literacy, access to technical resources, and training) were dropped based on the recommendations of [38], as indicated in Table 3. The CR values for all the constructs were above 0.70; AVE values above 0.50, to indicate good reliability and convergent validity of the constructs. Discriminant validity was accessed by the Heterotrait-Menotrait (HTMT) method based on the recommendation of [38]. Most constructs had HTMT values less the 0.9 except for learnability and memorability as well as errors and cognitive load that had very high values for the HTMT. The constructs learnability and memorability and then errors and cognitive load were therefore combined in the final proposed model in Fig. 2. Furthermore, these concepts are semantically related.

5.2 Structural Model

The logical next step after establishing good reliability, convergent and discriminant validity for the model is to assess the structural part of the SEM model. The proposed

Table 3. Construct validity and reliability

Item	Cronbach alpha	AVE	CR reliability
LMS Use	0.89	0.64	0.90
Self-Efficacy	0.89	0.73	0.89
Learning style	0.88	0.71	0.89
Access to technical support	0.91	0.78	0.91
Social influence	0.86	0.69	0.87
Perceived usefulness	0.89	0.74	0.90
Effectiveness	0.88	0.78	0.88
Efficiency	0.72	0.50	0.75
Satisfaction	0.88	0.79	0.88
Learnability and memorability	0.83	0.50	0.83
Errors and Cognitive load	0.91	0.62	0.90

model could not include both PU and PUsab in the same model, because the two constructs were highly related (0.90) and therefore it was decided to fit two separate models. The fit indexes for the PUsab model indicate a good fit: the relative Chi-square = 1.46, Robust Comparative Fit Index [CFI] = 0.96, Robust Tucker-Lewis Index [TLI] = 0.95, Robust Root Mean Square Error of Approximation [RMSEA] = 0.044 and Standardized Root Mean Square Residual [SRMR] = 0.064, as indicated in Table 4. The recommended range of values used in the model fit index is based on the recommendation of [39].

Table 4. Model fit indexes

Fit Index	Recommended range	Measurement Model (PUsab)	Measurement Model (PU)	Remark
Chi-Square χ^2	Non-significant	Chi-square (542) = 793.13; p = 0.000	Chi-square (292) = 605.82; p = 0.000	*Not Good
Relative χ^2 (χ^2/df) CMIN in AMOS	<5	1.46	2.07	Good
Robust CFI	>0.90	0.96	0.958	Good
Robust TLI	>0.90	0.95	0.93	Good
Robust RMSEA	<0.08	0.044	0.044	Good
SRMR	<0.08	0.064	0.064	Good

* The relative chi-square is rather used in practice

The fit indexes for the perceived usefulness structural model also indicate a good fit: the relative Chi-square = 2.07, Robust Comparative Fit Index [CFI] = 0.95, Robust Tucker-Lewis Index [TLI] = 0.93, Robust Root Mean Square Error of Approximation [RMSEA] = 0.044 and Standardized Root Mean Square Residual [SRMR] = 0.064 as indicated in Table 4. The structural model consists of regression equations as indicated in Table 5 which shows that Human factors, access to technical support and social factor influences the PUsab of LMSs, while PUsab of LMSs influence the usage among students. Specifically, from the structural model, the following can be concluded:

1. Human factors (H1b) have a statistically significant influence on LMS's PUsab (standardized Beta = 0.6, p-value = 0.000).
2. Access to technical support is not statistically significant in influencing LMS's PUsab (standardized Beta = −0.1, p-value = 0.179).
3. Social influence (H3b) has a statistically significant influence on LMS's PUsab (standardized Beta = 0.5, p-value = 0.001).
4. PUsab (H4b) has a statistically significant influence on LMS use among students (standardized Beta = 0.8, p-value = 0.000).

Table 5. Structural modeling of PUsab

	Estimate	Std. Err	z-value	P(>|z|)	Std.lv	Std.all
Usability~						
Human factors	2.075	0.497	4.177	0.000	0.624	0.624
Access technical support	−0.385	0.287	−1.343	0.179	−0.116	−0.116
Social influence	1.649	0.491	3.356	0.001	0.496	0.496
LMS use~						
Usability	0.447	0.103	4.356	0.000	0.830	0.830

Figure 2 indicates the structural model for perceived usability. The CFA factor loadings of the various constructs (effectiveness [0.85], efficiency [0.85], satisfaction [0.65], learnability and memorability [0.38] and errors and cognitive load [0.37]) appropriately loaded onto PUsab.

The structural model for perceived usefulness indicates that Human factors, access to technical support and social factor influences the perceived usefulness of LMSs while perceived usefulness also influences LMSs usage among students. Specifically, from the structural model, the following can be concluded:

5. Human factors (H1a) have a statistically significant influence on LMS's perceived usefulness (standardized Beta = 0.3, p-value = 0.000).
6. Access to technical support has a statistically significant influence on LMS's perceived usefulness however a negative influence (standardized Beta = −0.2, p-value = 0.024).

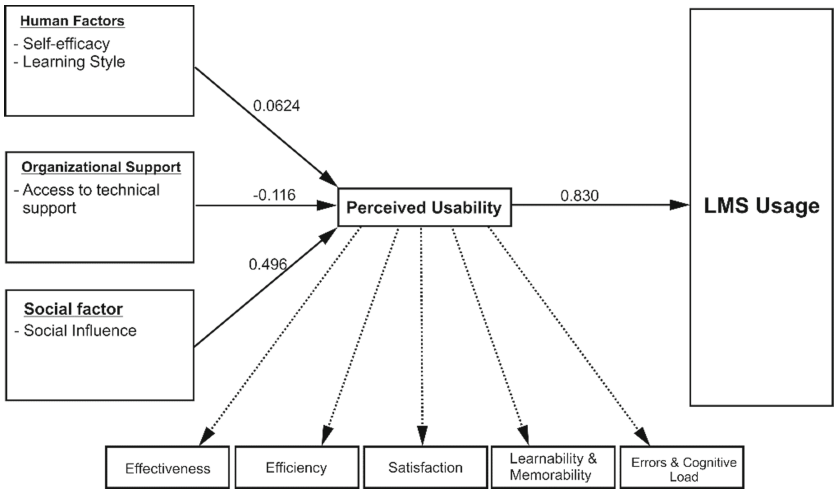


Fig. 2. Modeling PUsab

7. Social influence (H3a) has a statistically significant influence on LMS's perceived usefulness (standardized Beta = 0.8, p-value = 0.000).
8. Perceived usefulness (H4a) has a statistically significant influence on LMS use among students (standardized Beta = 0.7, p-value = 0.000) (Fig. 3 and Table 6).

Table 6. Structural modeling of perceived usefulness

	Estimate	Std. Err	z-value	P(>|z|)	Std.lv	Std.all
Usefulness~						
Human factors	0.757	0.181	4.191	0.000	0.346	0.346
Access technical support	−0.479	0.212	-2.264	0.024	−0.219	−0.219
Social influence	1.693	0.327	5.176	0.000	0.775	0.775
LMS use~						
Usefulness	0.495	0.059	8.411	0.000	0.734	0.734

5.3 Discussion

The findings confirm that human factors (self-efficacy and learning style), access to organizational support and social influence positively influence PUsab, while PUsab positively influences LMS use among students in Ghanaian HEIs. Furthermore, human factors (self-efficacy and learning style) and social influence positively influence PU, while PU positively influences LMS use among students in Ghanaian HEIs. These findings are consistent with previous studies that identified self-efficacy [28, 35], learning

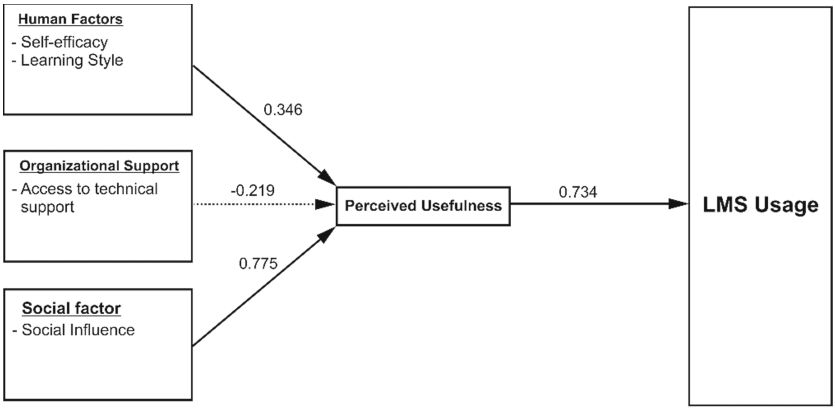


Fig. 3. Modeling perceived usefulness

style [26, 29], access to technical support [29, 35], social influence [36, 38] and PU [29, 30] as factors influencing LMS use in HEIs. Interestingly there was a negative correlation between access to technical support and PUsab and PU. We could not find any literature to support or explain this relationship and therefore suggest further investigation. The findings of this study also revealed that PU and PUsab are highly related, which is in contrast with the findings of [32–34]. PU and PUsab being related imply that these cannot be modelled together and therefore PUsab cannot replace PEU in modelling LMS adoption.

To explain this finding, we consider the components of the PUsab construct, i.e., effectiveness, efficiency and user satisfaction [40] whereby the effectiveness component could be per definition related to the PU construct of TAM, i.e., if technology is effective in allowing a user to complete a task it would be considered useful.

6 Conclusion and Future Work

The purpose of this study was to consider the possibility of integrating PUsab into the technology adoption literature. The definitions of technology adoption and usability have originated in different fields but more recently there has been studies suggesting that PUsab could replace PEU when considering the adoption of LMSs. While these constructs may seem similar, PUsab is a compound construct which can be decomposed into effectiveness, efficiency, satisfaction, learnability, memorability, errors and cognitive load. We found that PUsab and PU are highly related and thus PUsab cannot be sufficiently distinguished from PU to include those in the same model. Therefore, PUsab cannot replace PEU in measuring technology adoption of LMS's in HEI's in Ghana. We acknowledge the limitation on the generality of our findings as that being from only one country. More research is needed to consider these factors from the perspective of both students and lecturers and to do so in different contexts. Furthermore, usability measures like observation, usability testing and heuristic evaluation of LMSs should be considered. Despite the limitations, we propose this finding as important in theorizing the nexus between the usability and technology adoption literature.

Acknowledgement. This paper is based on the research supported by the South African Research Chairs Initiative of the Department of Science and Technology and National Research Foundation of South Africa (Grant No. 98564).

References

1. Beegle, K., Christiaensen, L., Dabalen, A., Gaddis, I.: Poverty in a rising Africa' overview (2016). https://doi.org/10.1596/978-1-4648-0723-7
2. Adamu, M.S.: Developing a mobile learning app: A user-centric approach. In: ACM International Conference Proceeding Series, vol. 21–25-Nove, pp. 139–143 (2016). https://doi.org/10.1145/2998581.2998602
3. Ssekakubo, G., Suleman, H., Marsden, G.: Issues of adoption: have e-learning management systems fulfilled their potential in developing countries? In: ACM International Conference Proceeding Series, pp. 231–238 (2011). https://doi.org/10.1145/2072221.2072248
4. Denan, Z., Munir, Z.A., Razak, R.A., Kamaruddin, K., Sundram, V.P.K.: Adoption of technology on e-learning effectiveness. Bull. Electr. Eng. Inform. **9**(3), 1121–1126 (2020). https://doi.org/10.11591/eei.v9i3.1717
5. Bhuasiri, W., Xaymoungkhoun, O., Zo, H., Rho, J.J., Ciganek, A.P.: Critical success factors for e-learning in developing countries: a comparative analysis between ICT experts and faculty. Comput. Educ. **58**(2), 843–855 (2012). https://doi.org/10.1016/j.compedu.2011.10.010
6. Lwoga, E.T., Komba, M.: Understanding University students' behavioural continued intentions to use e-learning in Tanzania. In: Proceedings and Report of the 7th UbuntuNet Alliance Annual Conference, pp. 167–188 (2014)
7. Alturki, U.T., Aldraiweesh, A., Kinshuck, D.: Evaluating the usability and accessibility of LMS blackboard at king Saud university. Contemp. Issues Educ. Res. **9**(1), 33–44 (2016). https://doi.org/10.19030/cier.v9i1.9548
8. Lehong, S., Van Biljon, J., Sanders, I.: Open-distance electronic learning environments: supervisors' views on usability. In: 2019 Conference on Information Communications Technology and Society ICTAS 2019 (2019). https://doi.org/10.1109/ICTAS.2019.8703605
9. Alfadly, A.A.: The efficiency of the learning management system (LMS) in AOU, Kuwait, as a communication tool in an E-learning system. Int. J. Educ. Manag. **27**(2), 157–169 (2013). https://doi.org/10.1108/09513541311297577
10. Harrati, N., Bouchrika, I., Mahfouf, Z.: E-Learning: on the uptake of modern technologies for online education. In: Proceedings of the 6th International Conference on Information Communication and Management ICICM 2016, pp. 162–166 (2016). https://doi.org/10.1109/INFOCOMAN.2016.7784235
11. Mtebe, J.S.: Learning management system success: increasing learning management system usage in higher education in sub-Saharan Africa. Int. J. Educ. Dev. Using Inf. Commun. Technol. **11**(2), 51–64 (2015). https://www.learntechlib.org/p/151846/
12. Jo, I.H., Kim, D., Yoon, M.: Analyzing the log patterns of adult learners in LMS using learning analytics. In: ACM International Conference Proceeding Series, pp. 183–187 (2014). https://doi.org/10.1145/2567574.2567616
13. Naveh, G., Tubin, D., Pliskin, N.: Student satisfaction with learning management systems: a lens of critical success factors. Technol. Pedagog. Educ. **21**(3), 337–350 (2012). https://doi.org/10.1080/1475939X.2012.720413
14. Kunene, K.N., Petrides, L.: Mind the LMS content producer: Blackboard usability for improved productivity and user satisfaction. In: 2017 Proceedings of the EDSIG Conference on Austin, pp. 1–14 (2017). http://proc.iscap.info/2017/pdf/4387.pdf

15. Mgeni, M.S., Ismail, M.J., Yunus, S.A., Haji, H.A.: The contribution of LMS to the learning environment: views from the state university of Zanzibar. In: Mendy, G., Ouya, S., Dioum, I., Thiaré, O. (eds.) AFRICOMM 2018. LNICSSITE, vol. 275, pp. 305–314. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-16042-5_27
16. Mtebe, J.S., Kissaka, M.M.: Heuristics for evaluating usability of learning management systems in Africa. In: 2015 IST-Africa Conf. IST-Africa 2015, pp. 1–13 (2015). https://doi.org/10.1109/ISTAFRICA.2015.7190521
17. Ssekakubo, G., Suleman, H., Marsden, G.: Issues of adoption: have e-learning management systems fulfilled their potential in developing countries? (2011). https://doi.org/10.1145/2072221.2072248
18. Asampana, I., Akanferi, A.A., Ami-Narh, J.: Reasons for poor acceptance of web-based learning using an LMS and VLE in Ghana. Interdiscip. J. Inf. Knowl. Manag. **12**(June), 189–208 (2017). https://doi.org/10.28945/3742
19. "ISO 9241-11:1998(en), Ergonomic requirements for office work with visual display terminals (VDTs)— Part 11: Guidance on usability. https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-1:v1:en. Accessed 30 Sep 2022
20. Harrison, R., Flood, D., Duce, D.: Usability of mobile applications: literature review and rationale for a new usability model. J. Interact. Sci. **1**(1), 1 (2013). https://doi.org/10.1186/2194-0827-1-1
21. Gunesekera, A.I., Bao, Y., Kibelloh, M.: The role of usability on e-learning user interactions and satisfaction: a literature review. J. Syst. Inf. Technol. **21**(3), 368–394 (2019). https://doi.org/10.1108/JSIT-02-2019-0024
22. Nakamura, W.T., De Oliveira, E.H.T., Conte, T.: Usability and user experience evaluation of learning management systems a systematic mapping study. In: ICEIS 2017–Proceedings of the 19th International Conference on Enterprise Information Systems, vol. 3, no. Iceis, pp. 97–108 (2017). https://doi.org/10.5220/0006363100970108
23. Altalbe, A.L.I.: Antecedents of actual usage of e-learning system in high education during COVID-19 pandemic: moderation effect of instructor support. IEEE Access **9**, 93119–93136 (2021). https://doi.org/10.1109/ACCESS.2021.3087344
24. Kaewsaiha, P., Chanchalor, S.: Factors affecting the usage of learning management systems in higher education. Educ. Inf. Technol. **26**(3), 2919–2939 (2020). https://doi.org/10.1007/s10639-020-10374-2
25. Fung, H., Yuen, A.: Factors affecting students' and teachers' use of LMS – towards a holistic framework. In: Cheung, S.K.S., Fong, J., Kwok, L.-F., Li, K., Kwan, R. (eds.) ICHL 2012. LNCS, vol. 7411, pp. 306–316. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-32018-7_29
26. Fathema, N., Shannon, D., Ross, M.: Expanding the technology acceptance model (TAM) to examine faculty use of learning management systems (LMSs) in higher education institutions. MERLOT J. Online Learn. Teach. **11**(2), 210–232 (2015). https://jolt.merlot.org/Vol11no2/Fathema_0615.pdf
27. Kumar, B.A., Goundar, M.S.: Usability heuristics for mobile learning applications. Educ. Inf. Technol. **24**(2), 1819–1833 (2019). https://doi.org/10.1007/s10639-019-09860-z
28. Alomari, M.M., El-Kanj, H., Alshdaifat, N.I., Topal, A.: A framework for the impact of human factors on the effectiveness of learning management systems. IEEE Access **8**, 23542–23558 (2020). https://doi.org/10.1109/ACCESS.2020.2970278
29. Coleman, E., Mtshazi, S.: Factors affecting the use and non-use of learning management systems (LMS) by academic staff. South Afr. Comput. J. **29**(3), 31–63 (2017). http://www.scielo.org.za/pdf/sacj/v29n3/04.pdf
30. Holden, H., Rada, R.: Understanding the influence of perceived usability and technology self-efficacy on teachers' technology acceptance. J. Res. Technol. Educ. **43**(4), 343–367 (2011). https://doi.org/10.1080/15391523.2011.10782576

31. Lin, C.-C.: Exploring the relationship between technology acceptance model and usability test. Inf. Technol. Manag. **14**(3), 243–255 (2013). https://doi.org/10.1007/s10799-013-0162-0
32. Pal, D., Vanijja, V.: Perceived usability evaluation of Microsoft Teams as an online learning platform during COVID-19 using system usability scale and technology acceptance model in India. Child. Youth Serv. Rev. **119**, 105535 (2020). https://doi.org/10.1016/j.childyouth.2020.105535
33. Lah, U., Lewis, J.R., Šumak, B.: Perceived usability and the modified technology acceptance model. Int. J. Hum. Comput. Interact. **36**(13), 1216–1230 (2020). https://link.springer.com/article/10.1007/s10799-013-0162-0
34. Dlalisa, S.F., Govender, D.W.: Challenges of acceptance and usage of a learning management system amongst academics. Int. J. Ebus. eGov. Stud. **12**(1), 1–16 (2020). https://doi.org/10.34111/ijebeg.202012105
35. Abdallah, N.A.O., Ahlan, A.R., Abdullah, O.A.: Factors affecting instructors' adoption of learning management systems: a theoretical framework. In: Proceedings–6th International Conference on Information and Communication Technology for the Muslim World, ICT4M 2016, pp. 13–18 (2016). https://doi.org/10.1109/ICT4M.2016.016
36. Asiri, M.J., Mahmud, R.B., Abu Bakar, K., Bin Mohd Ayub, A.F.: Factors influencing the use of learning management system in Saudi Arabian higher education: a theoretical framework. High. Educ. Stud. **2**(2), 125–137 (2012). https://doi.org/10.5539/hes.v2n2p125
37. Al-Busaidi, K.A., Al-Shihi, H.: Instructors' acceptance of learning management systems: a theoretical framework. In: Knowledge Management and Innovation in Advancing Economies: Analyses and Solutions–Proceedings of the 13th International Business Information Management Association Conference, IBIMA 2009, vol. 3, pp. 1199–1207 (2009). https://www.researchgate.net/profile/Kamla-Ali-Al-Busaidi/publication/42539726_Instructors'_Acceptance_of_Learning_Management_Systems_A_Theoretical_Framework/links/09e4151095b5c3be36000000/Instructors-Acceptance-of-Learning-Management-Systems-A-Theoretical-Framework.pdf
38. Hair, J., Black, W.C., Babin, B.J., Anderson, R.E.: Multivariate data analysis: Pearson new international edition. Br. Libr. Cat. Data 201–225 (2014). https://www.worldcat.org/title/multivariate-data-analysis/oclc/900353065
39. Hooper, D., Coughlan, J., Mullen, M.: Structural equation modelling: guidelines for determining model fit. Electron. J. Bus. Res. Methods **6**(1), 53–60 (2008). https://academicpublishing.org/index.php/ejbrm/article/view/1224
40. Nielsen, J.: Usability Engineering. Morgan Kaufmann, Burlington (1994). https://uweb.engr.arizona.edu/~ece596c/lysecky/uploads/Main/Lec9.pdf

A Framework to Capture the Factors that Influence the Adoption of Digital Platforms in E-Government

Setsi Mamabolo, Marita Turpin(✉), and Tendani Mawela

Department of Informatics, University of Pretoria, Pretoria 0002, South Africa
marita.turpin@up.ac.za

Abstract. This study investigates platformisation in the public sector and the impact of platformisation on e-government. The study considers how digital platforms are currently used in e-government in multiple countries. The advantages of using digital platforms in e-government are discussed. Following this, it is suggested that a framework is developed that captures the factors that influence the adoption of digital platforms in e-government. To develop the conceptual framework for the adoption of digital platforms, a systematic literature review (SLR) is performed to search the literature to identify current frameworks used to adopt technologies in e-government. The TOE framework is used as a lens in the thematic analysis to determine factors influencing digital platform adoption in e-government. The contribution of this study is the suggested conceptual framework that captures the factors that influence the adoption of digital platforms in e-government.

Keywords: E-government · Technology adoption · Digital platform · TOE framework · Systematic literature review · Platformisation

1 Introduction

The world is currently undergoing a fourth industrial revolution, in which information technology and digital devices are propelling the world toward a new digital transformation paradigm [1]. Any institution that wants to stay relevant and competitive must go digital in order to improve the quality and value of its service or product delivery [2]. The public sector is not exempted. It is also undergoing digital transformation.

Due to the success of digital platforms such as Uber, Amazon and Airbnb in the private sector, public sector organizations are turning to platformisation as a transformation strategy [3]. Digital platformisation is a result of the ongoing transformation of the public sector through the process of combining existing IT silos with new technologies to create digital platforms [3]. In sectors such as finance, health, education, and agriculture, embracing digital platform services can play an important role in the dissemination of information, improving service delivery, and reaching out to customers more effectively and efficiently [2]. In the public sector, digital transformation may also imply new ways

P. Ndayizigamiye et al. (Eds.): IDIA 2022, CCIS 1774, pp. 267–282, 2023.
https://doi.org/10.1007/978-3-031-28472-4_17

of collaborating with stakeholders, the development of new service delivery platforms, and the formation of new types of relationships [4].

Simultaneously, public services struggle to adapt to innovations such as digital platforms. To assist the struggling public sector to adapt and innovate, the current digital transformation paradigm could benefit from a framework that captures the factors that influence the adoption of digital platforms. A SLR is conducted to assess the factors that influence the adoption of digital platforms in e-government. The framework containing factors influencing digital platform adoption would alert governments to consider the factors in their transformation journeys. The literature has previously argued for the potential benefits of e-government adoption for development. There is also a need to understand the developmental implications of digital platforms [9].

This study adds to the discourse in Information and Communication Technologies for Development (ICT4D) research, about the contribution of digital platforms towards socio-economic development. The study aims to develop a framework to capture the factors that influence the adoption of digital platforms in e-government. The study assists government adopters of digital platforms by outlining factors that influence adoption and how they relate to e-government.

This paper is structured as follows. A background section provides literature context on digital transformation in the public sector. This is followed by a report on the SLR that was performed, including its planning and outcomes. The results of the SLR are analysed through the TOE framework. After this, a conceptual framework for the factors that influence the adoption of digital platforms in e-government is proposed. Lastly, a conclusion is presented.

2 Background

The governance, business intelligence, and profit generation aspects of digital platforms have been studied extensively in the private sector [5]. However, studies in the public sector are limited. There is a growing perception that digital platforms can contribute to socio-economic development, leading to digital platforms becoming a new focus for research related to ICT4D [5]. To get a clear picture of digital platforms' contribution to socio-economic development we reviewed the extant e-government literature.

Initial discussions on the digital platforms for e-government were led by O'Reilly [6] who defined "Government as a Platform" as "the use of collaborative technologies to solve collective problems at the local, state, national, and international levels". The O'Reilly [6] discussion stems from the emergence of Government 2.0 and Web 2.0, which advocated transparent, participatory, and collaborative government. Government 2.0 is defined as the adoption of digital technologies such as cloud computing, wikis, crowdsourcing, mobile applications, mashups, developer contests, or any other Web 2.0 phenomenon to the operations of government [6].

2.1 Digital Transformation in the Public Sector

The strategy of digital platformisation is propelling technology forward in the public sector [3]. Digital platforms are the primary drivers of the platform economy and are critical to digital transformation [7].

The New Public Management (NPM) approach to public sector transformation introduced modern business thinking, including competition, to the public sector, resulting in a shift from monopolistic control of service delivery to a quasi-market [8]. The era of the internet and websites followed NPM public sector transformation to digitalise operations using ICTs to improve internal and external processes as well as structures to promote *efficiency*, transparency, and accountability [3]. The integration of existing IT silos in public sector ICT operations can be enabled by digital platformisation.

2.2 Digital Platforms in E-Government

Simply put, digital platforms are online software platforms that enable businesses to build apps and provide a trading environment for market participants. From a technological perspective, digital platforms are layered modular ICT-enabled architectures with stable core components and flexible complementary modules [3]. Economically, digital platforms serve as hubs for interactions and value creation among multiple actors [3]. Digital platform research in information systems focuses on the socio-technical aspects of digital platforms, such as their impact on organizational structures or international standards [9]. Table 1 summarizes various definitions of digital platforms in e-government found in the literature as part of ongoing information systems research on digital platforms for development.

Table 1. Definitions for digital platforms in e-government

Definition	Source
Digital platforms are regarded as electronic hubs that enable interactions among actor groups from a socio-technical standpoint	[3]
"Government as a Platform", or "GaaP", is an attempt to conceptualize the extent to which the internet could change the government. It describes the use of collaborative technologies to enable more participatory government and better collective problem-solving at the city, region, national, and international levels	[6]
Government as a Platform is a business model in which the government manages and facilitates a civic marketplace where citizens can participate actively in the co-creation of services	[10]
A public Service Platform (PSP) is the technology that supports both the demand side of the marketplace (citizens searching among public offerings) and the supply side (the public and private sectors providing publicly funded services in quasi-markets)	[8]
Digital platforms are defined as the extensible codebase of a software-based system shared by applications that interoperate with it, as well as the interfaces through which they interoperate, and their ability to offer third-party developers, or compliments, the ability to build compliments on them	[5]
Digital platforms are a collection of digital resources, such as services and content, that enable providers and users to engage in value-added interactions	[11]

(*continued*)

Table 1. (*continued*)

Definition	Source
"Government as a Platform" is defined as an artifact that should emerge as a distinct organizational form, support market dynamics, and (technically) support a platform ecosystem	[12]
A digital platform within the state is a source of open data collected for the authorities' feedback on the population and the development of state economic policy	[13]
Digital platforms are a unique type of information technology (IT) artefact with socio-technical phenomena that necessitate careful consideration of how they function in a social context, providing unique development opportunities	[9]
Digital platforms are a value creation system that connects actors around a common activity or need, allowing them to collaborate, better allocate and use resources, and co-create value for one another	[7]

According to the definitions in Table 1, in the context of e-government, digital platforms could be transactional platforms or innovative platforms that promote transparency, participation, and collaboration in the public sector. The dominant market logic of platforms is more focused on transactional platforms, whereas a human-centered logic includes various types of platformisation such as transaction, innovation, and integration platforms [14]. Digital platforms can be used in e-government as part of the ICT initiatives that harness and transform relationships with citizens, businesses, and government branches.

2.3 Digital Platforms and Socio-economic Development

In terms of how digital platforms affect people and various socio-economic development efforts, there are key differences. Digital platforms, for example, are improving and transforming people's lives by providing a unique mechanism for connecting and networking, as well as creating new employment opportunities, innovation, and income [7]. Table 2 shows how digital platforms are used in e-government and their effects on people's socioeconomic well-being, based on key literature. Table 2 expresses the use of digital platforms in e-government and how it affects citizens' well-being by improving the efficiency of government services. The e-government platforms used in different countries for sharing of information, citizen reporting, government service delivery, citizen mapping, identification registration, providing agricultural services, and port logistics are listed in Table 2. Digital platforms have created and expanded market space through highly flexible manufacturing processes, and they are regarded as critical in recalibrating and mediating social trust and interactions by improving information flow and access between individuals [7].

Table 2. The use and effects of digital platforms on development in e-government

Research project	Design attributes/Platform type	Outcomes of research projects and their contributions	Source
Ethiopian WoredaNet government project. Regional government administrative district	Collaborative consumption of resources in the G2G domain is being explored using a government-owned network infrastructure	Operational improvement that leads to improved efficiency	[11]
Altinn, a Norwegian public-sector platform for the mandatory reporting of company financial statements	Core government-business ecosystem of innovation and collaboration to support key intrastate interactions between government, citizens, and service providers	Inter-organisational collaboration leads to improved efficiency	[12]
MyGov.in India, the municipal platform NYC.gov in New York and the portal for government services Gosuslugi.ru in Russia	Legislation is passed to support digital platforms and sustain economic growth	Regulatory process consolidation to improve efficiency and digital innovation	[13]
UK Government platform initiatives	Network-enabled business transformation enables companies to rethink how they interact with customers and suppliers by providing architecture of related open standards (e.g. Internet standards like TCP/IP)	Open standards provide the rules that ensure component compatibility, allowing platform ecosystems to evolve and improve innovation	[10]
India's Aadhaar, the world's largest digital identity platform (social protection system)	The digital platform ecosystem orchestrated to address multiple societal challenges stemming from identity and its management	Real-time identity verification infrastructure presents a notable innovation as national identification schemes. It enables access to myriad government services	[5, 15]

(*continued*)

Table 2. (*continued*)

Research project	Design attributes/Platform type	Outcomes of research projects and their contributions	Source
Agriculture sector, Ghana's Esoko. Activist mapping examples like Ushahidi or Map Kibera enable local users to report events	Digital Platforms enable seemingly distant people and organisations to exchange and share information. This fundamental characteristic has socio-economic development as well as commercial settings	Examples of digital platform value being co-created with users' generated content	[9]
Ghana's paperless port	To automate previously manual services and encourage knowledge sharing among stakeholders, digital platforms were used to transform the port	Improved efficiency and value creation as transformational outcomes	[3]

Table 2's first column highlights projects in which digital platforms were used to improve public institutions. Digital platforms were used to improve regulatory processes, open standards, efficiency, digital innovation, and value creation. Platform value is co-created with user-generated content in activist mapping examples like Ushahidi or Map Kibera in Kenya, which allow local users to report events related to human rights and election monitoring [9].

A definition for digital platforms in the context of this study is derived from Tables 1 and 2 based on the impact and attributes of digital platforms in e-government. For this research, a digital platform refers to digital resources that enable participatory government by combining existing IT silos with new technologies to enable a civic marketplace that engages in value-added interactions. Addo [15] outlines the benefits of combining existing IT silos in the Aadhaar platform. Although Aadhaar does not guarantee benefits, services, or citizenship in and of itself, its real-time identity verification infrastructure allows access to a plethora of government services such as utilities, the National Rural Employment Guarantee Scheme (NREGS), benefits under the National Social Assistance Program (NSAP), and subsidized rations of daily necessities such as kerosene and food from the Indian Public Distribution System (PDS) [15]. It also allows access to social-economic services such as banking and telecommunications that were previously unavailable due to a lack of identity proof.

Concerns have arisen over how digital platform architecture and governance design enable organizations to scale the platform ecosystem through network effects; and how boundary resources can be configured to generatively combine resources [5]. Government agencies may face a variety of unanticipated design challenges and tensions that must be managed to focus on the larger and more holistic task of service innovation in order to provide more productive and better services [16]. To address the concerns and challenges raised in the literature, a SLR is used to search the literature and develop a conceptual framework for the factors that influence adoption of digital platforms in e-government.

The interoperability of a collection of heterogeneous components to interact, exchange and process information, is essential for the effectiveness of digital platforms [17]. Interaction between actor groups on a digital platform requires the interoperability of various technologies. Adoption of digital platforms includes the implementation of additional ICT artifacts. As a result, the adoption of digital platforms in e-government is investigated using the general adoption of technologies in e-government. Platformisation in e-government is still emerging, hence we surveyed the broader e-government adoption literature to understand the factors that may inform the approach towards the adoption of platforms in government.

3 Methodology for SLR

3.1 Data Sources

The primary goal of a SLR is to produce a comprehensive review by reducing non-objectivity through extensive literature searches using systematic, credible, and replicable processes with an audit trail of the researcher's findings [18]. The first step in a SLR is to use a credible source to find relevant and eligible literature that meets the research objectives.

This section describes the techniques and methodologies used in identifying and acquiring our eligible articles, which are the basis for the analysis presented in this paper. Given that the literature on digital platforms is new but interdisciplinary, we systematically searched the Web of Science and IEEE Xplore databases for relevant published (peer-reviewed) articles across all disciplines between 2012 and 2022. We only included studies that were published in English, with no restrictions on the study's subject or location. Table 3 lists the keywords that were used in the search. The keywords were chosen to support the research goal of identifying a framework that capture the factors that influence adoption of digital platforms in e-government. Search keywords are combined with Boolean operators and results from the Web of Science and IEEE Xplore databases are shown in Table 3.

Table 3. Search keywords

Search keywords	Web of science	IEEE Xplore
(Frameworks OR models OR frame OR schema OR structure OR method* OR strateg*) AND (adopt* OR implement* OR utiliz* OR use OR employ* OR deploy*) AND e-government	41	558
(Frameworks OR models OR frame OR schema OR structure OR method* OR strateg* OR employ* OR deploy*) AND (adopt* OR implement* OR utiliz* OR use OR use OR employ* OR deploy*) AND "digital government"	10	319
TOTAL (928)	51	877

3.2 Inclusion and Exclusion Criteria

The results of the online search using the selected search keywords generated many articles that needed to be sorted to determine relevant articles for the study. The final chosen articles were published between 2012 and May 2022. In total, 928 articles were chosen. The initial requirements were widened so as not to prematurely exclude papers that might be relevant [19]. The first results were obtained through automated searches based on title, abstract, and keywords. However, strict inclusion and exclusion criteria are used to ensure the papers' validity. The inclusion and exclusion criteria applied in the study are shown in Table 4.

Reporting and publishing protocols are crucial for improving the transparency of the research process and the reliability of published papers. Because it applies to all research, the value of the SLR is determined by what was done, what was discovered, and the clarity of reporting. To clarify reporting and transparency, the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement directs the theoretical and practical developments of the SLR method used in this study [20].

Table 4. Inclusion and exclusion criteria

Inclusion	Exclusion
• Peer-reviewed articles on technology adoption in the public sector published between 2012 and 2022 • A full research paper that follows PRISMA guidelines • Articles listed under IS Library Science, Computer Science, IS, and Public Administration on the Web of Science • Articles published in IEEE Xplore under the publication topics Government data processing, public administration, IS, and IT	• Articles that are unrelated to the study objectives or the primary focus • Articles that are not written in English • Articles from before 2012 • Articles released after June 2022 • Duplicate articles

The SLR phases of identification, screening, eligibility, and inclusion were applied as follows. In the Identification phase, 928 articles were identified of which 828 were left after removing duplicates. After screening the titles and abstracts, 172 articles were left that were assessed for eligibility. The assessment involved a full reading of the articles to see whether they addressed factors related to the adoption of technologies in e-government. During the assessment, 32 articles were identified for a full review and analysis.

3.3 Analysis and Quality

The literature was read and reread to become familiar with the data, inform the generation of initial codes, search for themes, review the themes, and define and name the themes [21]. The validity of the study, as well as the degree to which the review method and analysis reduced errors and biases, were all evaluated for quality by means of the following quality assessment criteria: [18, 22].

- Description and suitability of the inclusion and exclusion criteria for the review;
- The statistical likelihood that the literature search will include all relevant studies;
- Ensure the articles are applicable by checking that they contain certain keywords and by reading the entire article to see how it relates to the research perspective; and
- The reviewer's assessment of the included studies' validity and quality.

4 Findings and Results

4.1 Technology Organization Environment (TOE) Framework

A deductive thematic analysis was informed by the TOE Framework which was used as a lens to understand the factors that influence the adoption of digital platforms in e-government. The TOE framework has been widely used in technology adoption studies and serves as a useful lens for e-government cases [23]. TOE framework can be used to study how technological innovations are adopted and understand the crucial factors influencing the adoption of new technology in organizations [24].

Existing literature reveals that the TOE framework offered a thorough taxonomy with three dimensions (Technology, Organization and Environment) for examining the adoption of innovations at the organizational level. In comparison to other frameworks, it also emphasizes the psychological and social aspects and supports a large amount of empirical and theoretical research in the IS field [24]. The factors that influence the adoption of digital platform in e-government were classified into three TOE framework categories as presented in Table 5. The themes used in the thematic analysis of the data obtained from the articles were informed by the TOE framework. The factors that influence the adoption of technologies in e-government are shown in Table 5. Table 5 also shows the frequency of occurrence of the factors in the 32 articles analyzed. The number in the bracket, for example, Strategy (8), denotes that, out of the 32 articles examined, 8 of them discussed strategy as an adoption factor.

In the public sector, digital transformation is seen as a well-planned and executed organisation-wide change that affects not only the current information systems and business processes but also how human resources are allocated, how organizations are structured, how their information technology infrastructure is set up, and how their business and information technology strategies are implemented [25]. The three adoption factors categories listed in Table 5 are generalizations of various factors that influence the adoption of new technologies into e-government. The adoption factors are discussed in more detail below.

Technology. Infrastructure, security, compatibility, and quality are all technological factors that influence the implementation and adoption of innovative developments in e-government [3, 24]. The technological context refers to both the technological infrastructure and the development processes that an organization that intends to adopt a technology faces, and it describes both new technologies that are available to organizations and those that are currently in use [23, 26].

To enable the adoption of digital platforms, platformisation requirements such as interface standardization, interoperability, and integration frameworks that enable the automation of services and knowledge sharing among actors must be included [3].

Organization. The organizational factors focus on internal elements such as leadership, change management, and organizational readiness, and how technology aids the organization in the process [24]. The organizational context refers to elements critical to technology adoption within the organization, including management support, organizational compatibility, degree of coordination, and other resources and attributes [23].

Environment. The environmental factors involve external factors that influence how an organization operates, such as laws and regulations, industry structure, policies, levels of competition, economic conditions, and other external pressures [23, 24]. The three categories of adoption discussed above influence an organization's decision to incorporate an innovation, which, in turn, influences the organization's performance [23]. According to the 32 articles analysed, the adoption factors were analysed to examine their influence on the adoption of technologies in e-government.

4.2 Analysis of Adoption Factors

The influence of adoption factors was analyzed using the TOE lens codes that make up the three themes in Table 5. The 32 remaining articles used the presence of the codes in various models, frameworks, and tables to explain the factors that influence the adoption of technologies into e-government. After averaging the frequency of occurrence of the codes as per Table 5 in various models, frameworks, and tables in the 32 articles, the data showed that adoption factors grouped under environmental factors influence the adoption of new technologies in e-government by 51%, followed by technological factors by 28%, and organizational factors by 21%, based on the 32 articles examined. This could capture the factors that influence adoption of digital platforms in e-government, with

Table 5. Adoption factors: codes and themes from the literature

Technology	Organisation	Environment
Security (32)	Employee skills (26)	Trust (32)
Compatibility (24)	Policies (13)	Law and regulations (29)
Infrastructure (23)	Funding (12)	Perceived Ease of use (27)
System Quality (22)	Training (12)	Awareness (21)
Availability (19)	Collaboration (10)	Social Influence (19)
Privacy (18)	Employees (9)	Perceived usefulness (18)
Reliability (13)	Organizational Structure (8)	Users Literacy (18)
Data Issues (13)	Change Management (8)	Politics (13)
Complexity (10)	Risk (8)	Partnering Stakeholder (9)
Integration (8)	Strategy (8)	Socio-Culture (9)
Satisfaction with service (8)	Organizational readiness (8)	Attitude (8)
Network (7)	Change management (8)	Gender (8)
Interoperability (6)	Organizational Culture (7)	Competition (8)
Internet (6)	Leadership (7)	Age (8)
Value of data (3)	Management of resources (4)	Economics (7)
Mobile Devices (3)	Mismatch (2)	Language (6)
System Requirements (2)	Corruption (2)	Intensity of use (6)
Network Penetration (1)	Lack of budget (2)	Previous Experience (6)
Cloud (1)	Business Process (2)	Cost of internet (6)
Applications (1)	Re-organization (1)	National Culture (5)
	Public Service (1)	Participation (4)
		Location (4)
		Business (3)
		Media (2)
		Citizens (1)

environmental, technological, and organizational adoption factors receiving the most attention in that order.

The end of one initiative is followed by a new adaptation and introduction of new technologies because it is highly likely that the environment has already changed quickly due to the "no end in sight" nature of organisations' digital transformation process [25]. Digital platforms, as opposed to analog technologies, are more flexible because they have editable, reprogrammable, distributed, self-referential, and data-homogeneous features that enable innovation in response to environmental dynamics [3]. This explains why the adoption of digital technologies, such as digital platforms in e-government, is so strongly influenced by environmental factors.

According to the 32 articles examined, the two sub-factors with the greatest influence on environmental factors are trust followed by laws and regulations (Table 5). Trust is the conviction that other people will act in a predictable way. Citizens in many nations around the world worry that their personal information could be misused and violated by government agencies because interactions with e-government involve citizens exchanging personal information in order to access the services [27].

Citizens' trust is necessary for the interactions because they anticipate a dependable, private, and secure transaction. Perceptions of the high risks involved with online transactions and perceived poor regulatory security and privacy may contribute to low trust in e-government services [28]. Building user trust, which is dependent on data security regulations, has been identified as a factor in the successful adoption and use of e-government systems. The government must design and implement relevant laws and regulations to ensure the appropriate use of digital technologies, regulate data exchange services, and instill public trust in security and privacy standards [29].

Security, infrastructure compatibility, and service quality are three sub-factors that influence the technological factors of e-government adoption the most, according to the 32 articles under review (Table 5). Ensuring the security of the data exchanged during transactions is necessary to win user trust. Information security is a critical factor in the successful adoption of new technologies, including e-government systems, because it determines the intended adopters' and implementers' trust and security assurance for the new technology [30].

The degree to which an innovation is perceived to be compatible with existing values, beliefs, experiences, and adoption needs is referred to as compatibility [27]. Studies on the adoption of m-government in Arab countries have revealed that trust, citizens' perceptions of the compatibility of m-government with their lifestyles, culture, awareness, and system quality are important factors that influence adoption [31].

Technology must be compatible with platformisation requirements in order to enable the adoption of digital platforms. Gebeyehu and Twinomurinzi [11] used IBM's Post Study System Usability Questionnaire (PSSUQ) framework to assess the designed compatibility to platformisation in their study to assess the usability of collaborative consumption of digital platforms. A PSSUQ framework is used to assess Government Organizations' Collaborative Consumption of Digital Platforms. PSSUQ is a usability testing framework for software systems, applications, and digital platforms to assess their system usefulness, information quality, and interface quality [32]. The overall result of the PSSUQ survey questionnaire is calculated by averaging 16 questions scored on a 7-point scale with three subscales: system usefulness (questions 1–6), information quality (questions 7–12), and interface quality (questions 13 to 16) [32]. PSSUQ usability testing might be used to evaluate the technologies' platformisation compatibility.

According to the 32 articles reviewed, organizational leadership and employee skills and expertise are two sub-factors that have the greatest influence on organizational adoption factors (Table 5). Digital transformation in e-government, which includes the adoption of new technologies, must be understood from the perspective of the entire organization. This includes the notion that IT is not the means to support change; rather, processes, people, policies, and, most importantly, leadership must be fundamentally transformed in order to achieve digital transformation in e-government [4]. The success rate of digital transformation projects in the private sector is higher than in the public sector, owing to inherent influencing factors such as the political nature of decisions, poor leadership, insufficient ICT skills, bureaucracy, and resistance to change associated with public projects [28]. Top management guides and leads the digital transformation, which must be supported by providing employees with the necessary skills to support the change.

Apart from previous research that has used widely accepted technology acceptance models and theories to evaluate e-Government technology adoption, which may or may not be applicable for e-Government acceptance analysis, this study proposes a framework for factors that may influence the adoption of digital platforms in e-Government. The proposed framework will cater for the factors that influence the adoption of digital technologies in e-government and cover other aspects that are usually not sufficiently covered.

4.3 A Proposed Framework to Capture the Factors that Influence the Adoption of Digital Platforms in E-Government

In order to enhance the delivery of public services in public institutions, the proposed conceptual framework aims to capture factors that influence the adoption of digital platforms in e-government. The proposed conceptual framework considers factors that affect the positive adoption of digital platforms in e-government.

Figure 1 illustrates the conceptual framework that has been proposed to capture the factors that influence the adoption of digital platforms in e-government, which essentially calls for TOE adoption factors to be compatible with digital platforms. The PSSUQ usability testing framework can be used to test technology compatibility and the ability to provide participatory functionality of digital platforms.

In addition to the technology adoption factors already mentioned, the design of digital platforms may draw inspiration from participatory architecture. The organizational and technical designs, which are based on the logic of participatory architecture, encourage productive activities from a variety of participants who are not under the direct control of the government or the resource base, allowing complementors to make a variety of contributions [15].

Platform ecosystems are enabled by the participatory shared infrastructure of digital platforms. A new form of shared consumption known as the "sharing economy" or "collaborative consumption" is being brought about by the growth of digital platforms that allow the sharing of resources, underutilized assets, and services [11]. The platform's boundary and process rely on the interoperability of the Application Programming Interface (API) that enables innovation on digital platforms, which must be managed over time [10].

Platform ecosystems are typically made up of a combination of core technology components developed by a platform owner and a diverse range of external participants, who supplement the platform with applications and services that enhance and extend those developed by the original platform owner [10]. Digital platforms, which are made possible by digital technologies, combine artifacts like the internet, mobile, and emerging tools to offer unique innovations in socio-economic sectors like finance, health, education, and agriculture [3]. As a result, the adoption of digital platforms is dependent on and entails the adoption of other emerging digital technologies as well as a diverse range of external participants who supplement the platform to provide unique innovation effects.

By considering TOE adoption factors that are platformisation compatible and PSSUQ tested, a favorable environment can be created for the adoption of digital platforms that collaborate with other technologies in e-government to improve service delivery. These

innovations help organizations reach their customers more effectively and efficiently and disseminate information more effectively [2].

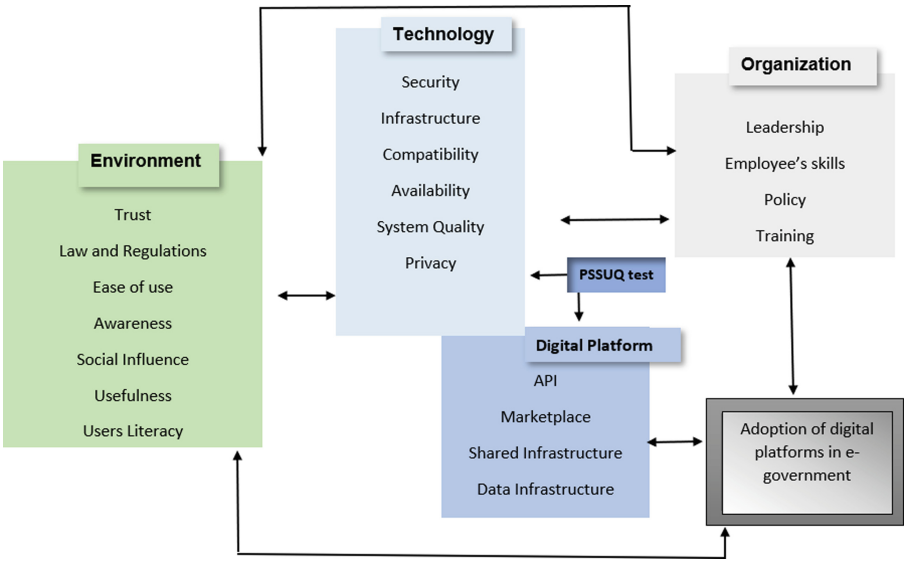


Fig. 1. A proposed framework to capture the factors that influence the adoption of digital platforms in e-government.

5 Conclusion, Limitations and Future Research

As public institutions increasingly will be using digital platforms to address socio-economic challenges, a better understanding of how this occurs is required. This paper's proposed conceptual framework contributes to a better understanding of how digital platforms can be better adopted in e-government. The study also sought to highlight the significance and effects of digital platforms for development for IS scholars to recognize that digital platforms must be supplemented by other digital technologies in order to be effective and efficient. In practice, this study helps government officials and policy-makers understand environmental factors, technological factors that are compatible with digital platforms, and organizational factors that influence platform implementation in government institutions.

The findings of the study are only limited to the assessed articles from the SLR. Given this limitation, further studies are recommended to assess the applicability of the proposed framework that captures the factors that influence the adoption of digital platform to e-government. We also need to assess, how the identified influencing factors affect the adoption and use of digital platforms in e-government.

References

1. Elmassah, S., Hassanein, E.A.: Digitalization and subjective wellbeing in Europe. Digit. Policy Regul. Govern. **24**, 52–73 (2022)
2. Okanga, P.A., Ho, T.T., Nguyen, T.B.T.: Assessment of efficiency in government services delivery through e-platform implementation? Evidence from Kenya. In: 14th Annual International Bata Conference for Ph. D. Students and Young Researchers (DOKBAT). Tomas Bata University in Zlín (2018)
3. Senyo, P.K., Effah, J., Osabutey, E.L.: Digital platformisation as public sector transformation strategy: a case of Ghana's paperless port. Technol. Forecast. Soc. Chang. **162**, 120387 (2021)
4. Bousdekis, A., Kardaras, D.: Digital transformation of local government: a case study from Greece. In: 2020 IEEE 22nd Conference on Business Informatics (CBI), pp. 131–140 (2020)
5. Masiero, S., Arvidsson, V.: Degenerative outcomes of digital identity platforms for development. Inf. Syst. J. **31**, 903–928 (2021)
6. O'Reilly, T.: Government as a platform. Innov. Technol. Govern. Global. **6**, 13–40 (2011)
7. Fu, X., Avenyo, E., Ghauri, P.: Digital platforms and development: a survey of the literature. Innov. Dev. **11**, 303–321 (2021)
8. Ranerup, A., Henriksen, H.Z., Hedman, J.: An analysis of business models in public service platforms. Gov. Inf. Q. **33**, 6–14 (2016)
9. Bonina, C., Koskinen, K., Eaton, B., Gawer, A.: Digital platforms for development: foundations and research agenda. Inf. Syst. J. **31**, 869–902 (2021)
10. Brown, A., Fishenden, J., Thompson, M., Venters, W.: Appraising the impact and role of platform models and government as a platform (GaaP) in UK government public service reform: towards a platform assessment framework (PAF). Gov. Inf. Q. **34**, 167–182 (2017)
11. Gebeyehu, S., Twinomurinzi, H.: A collaborative consumption digital platform for government organizations using design science. Digit. Govern. Res. Pract. **3**, 1–14 (2022)
12. D'Silva, F., Acando, A.: The sovereign digital platform-a strategic option for societal development (2018)
13. Stepanova, I., Vorotnikov, A., Doronin, N.: The potential of digital platforms for sustainable development using the example of the arctic digital platform 2035. In: IOP Conference Series: Earth and Environmental Science, p. 012004. IOP Publishing (2020)
14. Masiero, S., Nicholson, B.: Competing logics: towards a theory of digital platforms for socio-economic development. In: Bass, J.M., Wall, P.J. (eds.) ICT4D 2020. IFIP Advances in Information and Communication Technology, vol. 587, pp. 3–13. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-65828-1_1
15. Addo, A.: Orchestrating a digital platform ecosystem to address societal challenges: a robust action perspective. J. Inf. Technol. 02683962221088333 (2022)
16. Sun, P.-L., Ku, C.-Y., Shih, D.-H.: An implementation framework for E-government 2.0. Telemat. Inform. **32**, 504–520 (2015)
17. Ortiz-Bejar, J., Ortiz-Bejar, J.: Design and implementation of digital platform for e-government. In: 2021 IEEE URUCON, pp. 547–551. IEEE (2021)
18. Tranfield, D., Denyer, D., Smart, P.: Towards a methodology for developing evidence-informed management knowledge by means of systematic review. Br. J. Manag. **14**, 207–222 (2003)
19. Al-Zubidy, A., Carver, J.C.: Identification and prioritization of SLR search tool requirements: an SLR and a survey. Empir. Softw. Eng. **24**(1), 139–169 (2018). https://doi.org/10.1007/s10664-018-9626-5
20. Moher, D., et al.: Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst. Rev. **4**, 1–9 (2015)

21. Braun, V., Clarke, V.: Using thematic analysis in psychology. Qual. Res. Psychol. **3**, 77–101 (2006)
22. Kitchenham, B., et al.: Systematic literature reviews in software engineering–a tertiary study. Inf. Softw. Technol. **52**, 792–805 (2010)
23. Pudjianto, B., Zo, H., Ciganek, A.P., Rho, J.-J.: Determinants of e-government assimilation in Indonesia: an empirical investigation using a TOE framework. Asia Pac. J. Inf. Syst. **21**, 49–80 (2011)
24. Abied, O., Ibrahim, O., Kamal, S.N.I.M.: Proposing a conceptual model for cloud computing adoption in the Libyan E-government. In: 2021 7th International Conference on Research and Innovation in Information Systems (ICRIIS), pp. 1–6 (2021)
25. Jonathan, G.M., Hailemariam, K.S., Gebremeskel, B.K., Yalew, S.D.: Public sector digital transformation: challenges for information technology leaders. In: 2021 IEEE 12th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), pp. 1027–1033 (2021)
26. Abied, O., Ibrahim, O.: Cloud service adoption model in the Libyan e-government implementation. In: 2021 International Congress of Advanced Technology and Engineering (ICOTEN), pp. 1–7 (2021)
27. Freire, M., Fortes, N., Barbosa, J.: Decisive factors for the adoption of technology in E-government platforms. In: 2014 9th Iberian Conference on Information Systems and Technologies (CISTI), pp. 1–6 (2014)
28. Otieno, I., Omwenga, E.: Citizen-centric critical success factors for the implementation of e-government: a case study of Kenya Huduma Centres. In: 2015 IST-Africa Conference, pp. 1–9 (2015)
29. Thamjaroenporn, P., Achalakul, T.: Big data analytics framework for digital government. In: 2020 1st International Conference on Big Data Analytics and Practices (IBDAP), pp. 1–6 (2020)
30. Khanyako, E., Maiga, G.: An information security model for e-government services adoption in Uganda. In: 2013 IST-Africa Conference & Exhibition, pp. 1–11 (2013)
31. Alonazi, M., Beloff, N., White, M.: Developing a model and validating an instrument for measuring the adoption and utilisation of mobile government services adoption in Saudi Arabia. In: 2019 Federated Conference on Computer Science and Information Systems (FedCSIS), pp. 633–637 (2019)
32. Pinem, A.A., Yeskafauzan, A., Handayani, P.W., Azzahro, F., Hidayanto, A.N., Ayuningtyas, D.: Designing a health referral mobile application for high-mobility end users in Indonesia. Heliyon **6**, e03174 (2020)

Artificial Intelligence (AI) for Good

Towards an Artificial Intelligence Readiness Index for Africa

Rehema Baguma[1(✉)], Elizabeth Mkoba[2], Monica Nahabwe[3], Martin Gordon Mubangizi[4], Morine Amutorine[4], and Denis Wanyama[4]

[1] Makerere University, Makerere Hill Road, Kampala, Uganda
rbaguma@cit.ac.ug
[2] The Nelson Mandela African Institution of Science and Technology, Arusha, Tanzania
elizabeth.mkoba@nm-aist.ac.tz
[3] Ministry of Finance, Planning and Economic Development-Resource Enhancement and Accountability Program, Apollo Kagwa Road, Kampala, Uganda
monicanahabwe@gmail.com
[4] UN Global Pulse Lab Kampala, Plot 60, Malcolm X Drive, Kololo, Kampala, Uganda
{martin.mubangizi,morine.amutorine}@one.un.org, wanyamaden@gmail.com

Abstract. The applications and benefits of Artificial Intelligence (AI) for socio-economic development are immense. AI is projected to contribute approximately USD 15.7 trillion to the global Gross Domestic Product (GDP) by 2030. However, countries need to be prepared to harness such benefits. Hence, assessing the AI readiness of a country is paramount. Africa is currently the only continent without an AI readiness index tailored to its needs. It relies on the existing global indices, which may not accurately measure the progress attained by individual African countries because of the different levels of development and unique context. This paper proposes an AI readiness index for Africa. It starts by exploring what the AI readiness index needs of Africa are, examines the extent to which existing AI readiness indices meet the needs, and then looks at indicators that should constitute the AI readiness index for Africa.

The study employed a systematic literature review that aimed to explore the AI readiness needs for Africa and the extent existing indices meet these. The review focused on papers published on the AI readiness index between January 2018 to August 2022. The search strategy retrieved 301 papers, of which seven papers were selected for a detailed analysis. The study revealed that the existing indices partially meet AI readiness needs for Africa. The study also found that AI readiness index dimensions pertinent to Africa's requirements are: Vision, Governance and Ethics, Digital Capacity, Size of the Technology Sector, Research and Development, Education, Infrastructure, Data Availability, general level of employment, employment in Data Science and AI roles, and Gross Domestic Product-Per Capita Purchasing Power Parity. This study contributes to the knowledge of AI readiness for Africa and globally. The results of this study will benefit governments, researchers, and practitioners of AI and its applications.

Keywords: Artificial intelligence · Readiness index · Africa · Digital technologies · Digital capacity

P. Ndayizigamiye et al. (Eds.): IDIA 2022, CCIS 1774, pp. 285–303, 2023.
https://doi.org/10.1007/978-3-031-28472-4_18

1 Background

1.1 Introduction

There is no universally agreed definition of Artificial intelligence (AI). According to the Organization for Economic Co-operation and Development (OECD) and United Nations Conference on Trade and Development (UNCTAD), AI is the ability of machines and systems to acquire and apply knowledge and to carry out intelligent behaviour. This behaviour includes a variety of cognitive tasks such as sensing, processing oral language, reasoning, learning, making decisions, and demonstrating an ability to move and manipulate objects [1]. For [2], AI is a collection of Information and Communication Technologies (ICTs) that mimic human intelligence to enable machines facilitate jobs better, create greater efficiencies and drive economic growth. On the other hand, [3] defines AI as a scientific discipline aimed at building machines and systems that can perform many tasks that require human intelligence.

AI has been applied in various fields such as the automobile industry with the emergence of driverless cars, media, and communication through news feeds, business operations by influencing consumers' buying decisions and behaviours, and telecommunication through digital device functionality. Examples of sectors incorporating AI concepts in their operations include healthcare, law, transportation, education, retail, and financial firms, all of which have established and developed their in-house AI capabilities [4]. In the public service sector, a myriad of new AI technologies is being implemented including advancing the availability of education, detecting fraud, triaging health care needs, making payments to welfare recipients, speeding immigration decisions, and planning and implementing large urban and industrial infrastructure projects among others [5]. Examples of other applications include: sifting through massive troves of data and video captured by surveillance for alerts of abnormal or suspicious activity by the American military; medical imaging to detect lymph nodes in the human body in Computer Tomography (CT) images in German; detection and treatment of congestive heart failure, an illness that afflicts 10 percent of senior citizens and costs the United States $35 billion each year; analysis of people arrested for their risk of becoming future perpetrators by the city of Chicago; matching video images, social media activity, online purchases, travel records, and personal identity into a police cloud database in China to keep track of criminals [32], among others.

In Africa, AI has been applied in sexual and reproductive health monitoring chatbots in Kenya, smart farming in Nigeria, and tracking of illegal fishing in West Africa by AI-powered drones among others [6].

In 2016, the projected global economic impacts associated with using, developing, and adopting AI over ten years were found to be between $1.49 trillion and $2.95 trillion [9]. Price Water House Coopers (PWC), in their report "The Macroeconomic impact of Artificial Intelligence", projects AI to contribute approximately USD 15.7 trillion to global GDP by 2030, up from USD 2 trillion in 2019 with $6.6 trillion from increased productivity and $9.1 trillion from consumption effects. AI is expected to accelerate all sectors of the economy by a gain of at least 10% by 2030. The same report projects the services industry to gain the most (21%), with retail and wholesale trade as well as accommodation and food services also expected to see a considerable boost of about

15% [7]. Developing countries are to experience a modest increase due to the much lower rates of adoption of AI technologies. Thus Africa, Oceania, and Other Asian markets are estimated to have a total impact of 5.6% of GDP ($1.2 trillion) [7]. The AI contribution in Egypt alone is expected to be US$42.7 bn (7.7% of GDP) by 2030 [8].

AI offers new opportunities to reimagine how governments and the public sector can better serve sustainable development needs. Fast-evolving technologies can transform the traditional way of doing things across all government functions and domains [12]. The developments in AI can potentially disrupt and transform socio-economic activities across industries [2]. The World Economic Forum estimates that 85 million jobs will be replaced by machines with AI by 2025 globally. However, this will be offset by the creation of 97 million new jobs due to AI by the same year. According to PwC, 7 million existing jobs will be replaced by AI in the UK from 2017–2037, but 7.2 million jobs could be created [33].

AI promises to be a catalyst in accelerating development and allowing developing countries to leapfrog over some traditional obstacles [5]. In Africa, there is an increasing awareness of the positive impact that AI will have on developing countries, including sub–Saharan Africa, in sectors such as agriculture, health care, and public and financial services [10]. AI can drive economic growth, development, and democratization, thereby reducing poverty, increasing education, supporting healthcare delivery, increasing food production, expanding the capacity of the existing road infrastructure by increasing traffic flows, improving public services, and bettering lives [11].

Given AI's benefits and the fact that it is now a global phenomenon, it is important to assess whether Africa has the foundations necessary to reap what other continents are already harnessing hence the need for an AI Readiness Index specifically for Africa.

1.2 Motivation

An AI Readiness Index examines how ready a given government is to implement AI in the delivery of public services to its citizens [12]. It evaluates the different dimensions of a country's preparedness to implement AI and gives an idea of the progress so far, as well as insights into what needs to be done to reach the next level. Although there are several AI readiness indices at the global level and others at the continental level, existing indices may not accurately measure the progress attained by individual African countries because of the different levels of development and unique context.

The Oxford Insights and the International Development Research Centre (IDRC) started publishing the Government AI Readiness Index in 2017. The latest edition of the Government AI Readiness Index (2021 ed.) [13] provides an in-depth look at the state of AI readiness globally, ranking 160 countries using 42 indicators across ten dimensions [13]. However, it features 45 out of the 55 African countries and does not give reasons for leaving out the ten countries.

The second biggest global index is the Stanford University AI Readiness Report [14] published yearly since 2017. However, it has primarily focused on countries in the Global North and very few in Africa and Asia. In Africa, it usually features South Africa, Kenya, Egypt, Tunisia, Morocco, and Mauritius.

Although some of the indicators in the existing indices may be core and more widely applicable to measure AI readiness in different continents such as the existence of an AI

strategy or Data privacy and protection regulation, some indicators including the number of publications or attendance of highly competitive conferences like the Association for Advanced Artificial Intelligence (AAAI) and the number of patents, may not be well suited to the African context. Salesforce has since initiated an Asia Pacific AI readiness index to capture the unique context of the continent. In addition, [19] conducted a meta-analysis of the Oxford Government AI Readiness Index for its suitability for Latin America and the Caribbean.

To our knowledge, Africa is the only continent with no custom-tailored AI Readiness Index or publication presenting AI Readiness indicators that are appropriate for Africa - thus the motivation of this work. The authors of the Oxford Insights and IDRC Index acknowledge the dearth of relevant data repositories covering Africa which makes it difficult to get a full picture of developments on the continent or to explain national trends. The 2021 edition also recognizes infrastructural and regulatory challenges that are unique to the African continent namely: the perennial infrastructure limitations, and inconsistent policy and regulatory failure that slows progress towards AI readiness. The overall objective of this paper is to propose an AI Readiness Index for Africa. This objective is achieved through investigating the AI Readiness Index needs of Africa, the extent existing AI Readiness Indices meet these needs, and indicators that should constitute the AI Readiness Index for Africa.

The rest of the paper is organised as follows: Sect. 2 is the research methodology, Sect. 3 presents the proposed AI Readiness Index for Africa, including how it was arrived at, and Sect. 4 presents the conclusion, recommendations, and future work.

2 Methodology

The study reviewed literature, including grey literature published in peer-reviewed journals, technical reports, conference papers, books, and book chapters between January 2018 to August 2022. In this period, the applications and research in AI have been increasing in government, public and private sector operations, higher learning institutions, and various industries globally. Therefore, data sources included research databases like IEEE Xplore, ScienceDirect, The Association for Computing Machinery (ACM) Digital Library, SpringerLink, Google Scholar, and institutional websites, mainly for technical reports. These data sources were chosen based on the objectives of the research. In addition, the Preferred Reporting Items for Systematic Reviews and Meta-Analysis Protocol (PRISMA-P) was adopted as a review protocol [17].

A search string following Boolean operators derived from the research questions was used as the scope of searching papers: ((“AI” OR “Artificial Intelligence” or “artificial intelligence”) AND (“Readiness Index” OR “readiness index”)). The inclusion and exclusion criteria were used to reduce the number of studies by selecting the ones focusing on the study’s objective. This criterion covered: the relevance of the title, abstract and full test to the objective of the study, year of publication (2018–2022), Language (English), and type of publications: peer-reviewed journals, book chapters, technical reports, conference and books, and research method (qualitative, quantitative, and mixed method). This process resulted in selecting 301 papers (293 from the databases and eight reports from other sources). The papers were then screened based on titles and abstracts

and remained with 44 papers. Finally, the full texts of 44 papers were examined to identify content relevant to the objective of the study.

The quality assessment criteria by [18] was used to assess the quality of each paper using a scale of 0 for "No", 1 for "Partially", and 2 for "Yes". The minimum total score was 10. The results of the assessment are in Table 1 in the appendix.

3 Towards an AI Readiness Index for Africa

3.1 What Does Government AI Readiness in Africa Mean?

As developing countries continue to shift to more advanced digital platforms, they have adopted practices and policies that have a direct impact on the future of AI-based technology [19]. But for every country, it is only as prepared to take advantage of AI technology as its government and citizens [19]. Oxford Insights and DRC cite three key challenges to harnessing the use of AI for the common good i.e. policies, capacity, and adequate resources [20].

In simple terms, Government AI Readiness means having in place capabilities and enabling factors required for AI implementation. Annually, Oxford Insights and IDRC, produce the Government AI Readiness Index, a measure of capabilities and enabling factors required for a government to be ready for AI implementation. The Index is a tool to understand gaps and strengths in the capacity to implement AI [20].

In the 2021 Index [21], Sub-Saharan Africa scored lowest followed by Latin America and the Caribbean, and South and Central Asia. This reflects a persistent inequality in government AI readiness across continents with Africa in the lowest position. Few countries in Africa have national AI strategies providing a vision for the implementation of AI. By the time of the 2021 Oxford and IDRC index, only Mauritius had an AI strategy and Kenya was just in the process of developing one in the whole of Sub-Saharan Africa [20].

3.1.1 Existing Government AI Readiness Indices

The Oxford insights and the International Development Research Centre (IDRC) started publishing the Government AI Readiness index in 2018. The Oxford Insights 2021 Index provides an in-depth look at the state of AI readiness globally, ranking 160 countries using 42 indicators across 10 dimensions. It measures and ranks how ready a given government is to implement AI in the delivery of public services to its citizens. The metrics used have been evolving but the 2021 assessment has 33 indicators across 10 dimensions and three pillars. The expansion of the Index is aimed at giving a broader and deeper picture of government AI readiness. For example, the 2021 index includes indicators about data representativeness to capture how ready governments are to use AI in a way that avoids bias and a responsible use sub-index that ranks 34 countries according to how ready they are to use AI responsibly [20].

It uses indexed metrics on a scale of 0–10 based on different sources namely: UN World Economic Forum, Global Open Data Index, World Bank, Gartner, Nesta, and CrunchBase. The metrics are categorized into 3 pillars i.e. building blocks of Government AI Readiness namely: Government (Government needs to be willing to adopt AI, and

able to adapt and innovate), Technology Sector (Government needs a good supply of AI tools), and Data and Infrastructure (the tools need to be built and trained on high quality and representative data, and need appropriate infrastructure to be delivered to and used by citizens). Each of these pillars has dimensions that further specify how each of these conditions is fulfilled. A summary of the pillars, their description, and their indicators is given in Table 1 below.

Table 1. A summary of pillars, their dimensions, descriptions, and indicators for the Government AI Readiness Index from the Government AI Readiness Index 2020[1]

Dimension	Description	Indicator
Government Pillar		
Vision	Does the government have a vision for supporting the development and implementation of AI?	National AI Strategy
Governance and ethics	Are there the right regulations and ethical frameworks including data protection and privacy legislation to support implementation of AI in a way that builds trust and legitimacy?	• Data protection and privacy legislation • Cybersecurity legislation • National Ethics Framework • Legal framework's adaptability to digital business models
Digital capacity	What is the existing digital capacity within the government?	• Government procurement of advanced technology • ICT use and Government efficiency • Online services
Adaptability	Can the government change, adapt and innovate effectively?	• Effectiveness of Government • Government's responsiveness to change
Technology sector pillar		
• Size	• How large is the technology sector that will supply governments with AI technologies?	• Number of technology unicorns • Market value of public technology companies • Value of trade in ICT services (per capita) • Value of trade in ICT goods (per capita) • Computer software spending

(*continued*)

[1] https://static1.squarespace.com/static/58b2e92c1e5b6c828058484e/t/5f7747f29ca3c20ecb598f7c/1601653137399/AI+Readiness+Report.pdf.

Table 1. (*continued*)

Dimension	Description	Indicator
Innovation capacity	Does the technology sector have the right conditions to support innovation?	• Entrepreneurial culture • Ease of doing business • Research & Development spending • Company investment in emerging technologies
Human capital	Are there the right skills in the population to support the technology sector?	• Graduates in STEM • Quality of engineering and technology higher education • Digital skills • Knowledge-intensive employment
Data and infrastructure pillar		
Infrastructure	• Does the country have a good technological infrastructure to support AI technologies? • Availability of latest technologies • Internet bandwidth	• Telecommunications infrastructure • 5G infrastructure
Data availability	Is there good availability of data that could be used to train AI models?	• Open government data • Statistical capacity • Mobile-cellular telephone subscriptions • Internet users (% of adult population)
Data representativeness	Is the data available likely to be representative of the population as a whole	• Gender gap in Internet usage • Socioeconomic gap in Internet usage

However, Oxford insights and the IDRC index covers countries that they have enough data about to be confident in the score given [20].

For the 2021 edition, the highest-ranked African country (Mauritius) is in the 45th position followed by South Africa (59th), Seychelles (68th), Kenya (71st), and Rwanda (87th). The results are not a surprise given that African countries historically lag behind the rest of the world in technological development. The authors themselves acknowledge the dearth of relevant data repositories covering Africa which makes it difficult to get a full picture of developments on the continent or to explain national trends [20]. The authors also recognize the perennial infrastructure limitations in Africa, inconsistent policy, and regulatory failure that slows progress toward AI readiness. By January 2019, only Tunisia, Egypt, and Kenya showed signs of developing National AI strategies.

But on a good note, more African countries are progressively putting Readiness for AI Implementation on the national agenda. For example, Mauritius released its AI

strategy at the end of 2018, which included setting up a Mauritius AI council. Both the AI Strategy and the Mauritius 2030 Strategic Plan prioritize developing local talent, such as making programming a required university course. Other initiatives include: South Africa's Presidential Commission on the Fourth Industrial Revolution (4IR) tasked with developing a strategic plan for South Africa's 4IR vision, Uganda's presidential 4IR Committee and Data Privacy and Protection Act, 2019, and Rwanda's 2020 Data Protection Policy.

Furthermore, there are ongoing efforts to nurture the local AI talent and communities in Africa. This includes companies/groups of practitioners organizing AI conferences in Africa like Machine Learning Africa, and Data Science Africa to strengthen the AI knowledge community in Africa, and grassroots developments such as Deep Learning Indaba led by young innovators to support the development of locally grown AI expertise. Other initiatives towards making African countries ready for implementation of AI include:

- The African Digital Transformation Strategy (2020–2030) which mentions AI a few times, as an example of technologies to adopt or keep abreast of.
- The African Continental Free Trade Agreement could facilitate the growth of AI infrastructure and talent.
- AI4D Africa program, a 4-year CAD$ 20m partnership between IDRC and Swedish International Development Agency (SIDA) to support the creation of an ecosystem that supports responsible, home-grown development and deployment of AI through innovation, policy research, and skills.
- UNESCO's AI capacity building needs, an assessment survey to determine the current state of AI policy development and gaps in human and institutional capacity.
- The Moonshot for Africa program of the World Bank aimed at funding a range of ambitious digital connectivity programs in Africa.
- The global digital tax debate around OECD's proposal for a multilateral global digital tax regime where tech companies pay tax in the countries where they generate revenue, irrespective of physical presence.

3.1.2 The AI Index Report: Measuring Trends in Artificial Intelligence

The AI Index Report is an independent initiative at the Stanford Institute for Human-Centered Artificial Intelligence (HAI), produced by an interdisciplinary team of experts from academia and industry. The annual report tracks, collates, distils, and visualizes data on AI, enabling decision-makers to take meaningful action to advance AI responsibly and ethically [15].

The 2022 edition includes data from a broad set of academic, private, non-profit organizations and more self-collected data and original analysis than previous editions, including an expanded technical performance chapter, a new survey of robotics researchers, data on global AI legislation in 25 countries mostly from the Global North, and a new chapter on an in-depth analysis of technical AI ethics.

However, it is a general annual tracker of global developments in AI. In addition, the methodology used is not explicit in the report nor are the countries covered. But from the results, it is clear, the scope is based on the Global North and the more developed

Asian countries like China. Hence, regional balance in geographical scope is not one of the primary interests.

The latest edition [16] has 5 chapters on Research and Development, Technical Performance, The economy, AI Education, and Ethical Challenges of AI Applications. Like its title says, it just tracks AI trends globally based on information available in the public domain, particularly research output, where Africa is perennially weak. Hence, it was not considered of much relevance to Africa in terms of benchmarking to what extent it addresses Africa's AI Readiness needs.

The next section analyses dimensions that are pertinent to Africa's AI Readiness and compares this with what is covered by the Government AI Readiness Index by Oxford Insights & IDRC, so far, the only comprehensive Government AI Readiness Index in terms of subject and geographical scope.

3.2 Dimensions that are Pertinent for Africa's AI Readiness Index and to What Extent They are Covered by the Government AI Readiness Index by Oxford Insights and IDRC

ICT Infrastructure: The 2021 Government AI Readiness Index [21] recognized the infrastructural challenges of Africa as one of the reasons for low scores on the Government AI Readiness Index.

Internet Connectivity and Usage: Only 18% of the population in developing countries was online in 2019 compared to 87% in developed countries [22]. The majority of the offline populations are in Sub-Saharan Africa and South Asia with Sub-Saharan Africa being the only region where fixed broadband subscriptions are below 1 per 1000 inhabitants [23]. Additionally, more than 20% of the African population is not covered by a 3G or higher network, twice as much as in other regions of the world [24]. This impedes the benefits of participatory sensing, a process of massive sensor data collection via data-sensing devices that would address or reduce algorithmic colonization. This data is particularly useful in Big Data analytics and evidence-based decision-making due to its volume, granularity, and coverage [23].

End User Devices: A big number of the populations in Africa do not have Internet-enabled devices namely: smartphones, tablets, and computers. In 2019, Public authorities in sub-Saharan Africa still had challenges providing adequate ICT infrastructure with less than half of schools having access to the Internet and computers [25].

Regulation: Only 67% of African countries had data protection laws in 2019 [24]. Even in these countries, the laws are not robust due to financial and institutional capacity limitations.

Availability of Data: Infrastructure weaknesses (unreliable/limited connectivity, lack of/poor quality end-user devices) translate into a shortage of information as a great amount of data generated worldwide is a by-product of implementing technology [23]. In 2016, as many as 46 African countries did not have complete birth registration systems in place [26]. Africa still has several challenges hampering realizing the full potential

of Open Data namely: ICT infrastructure inequalities, unequal access to scientific education, and limited accessibility of private as well as public sector data [23]. Even in countries with the most advanced ICT infrastructure such as Mauritius, Cote d'Ivoire, and Rwanda, data is rarely or never released for reuse [23]. Yet in some Eastern European countries such as Estonia, Poland, and Georgia, more than 90% of incoming requests are fulfilled [23]. In developed countries, the major challenge is balancing data protection with freedom of information laws.

Skilled Workforce: AI readiness in Africa is also constrained by a lack of an appropriately skilled local workforce as AI solutions emerge from communities of researchers and entrepreneurs. The talent that exists in Africa faces numerous obstacles, including low visibility within the global community of AI researchers and entrepreneurs. For instance, in Rwanda, one of the highly ranked, the Permanent Secretary of the Ministry of ICT, reportedly estimated that the country had only about 10 AI engineers in November 2019 [20].

General Unemployment: The unemployment rate is the number of unemployed persons (from IMF) as a percentage of the total labour force as specified by the World Economic Forum. The Oxford Insights Ranking does not account for each country's unemployment rate. In developing countries, unemployment is often an indication of a country's economy but sometimes it is also an indicator of facts outside the Government's control such as conflict. Although unemployment rates and AI readiness are not directly correlated, unemployment must be considered before implementing AI technologies or automation to ensure there will be sufficient demand for AI-based/enabled products and services, and their associations in the ecosystem [19].

Employment in Data Science and AI Roles: Appreciation of the need for Data and AI roles in African government institutions is still low mainly due to a limited understanding of AI and its applications, as well as low data literacies among decision-makers. Hence, the number of AI and Data Science jobs are still few despite the need for these positions to utilize the power of AI and data science for more efficiency and effectiveness in public service delivery.

Overall Education Level: Although the Government AI Readiness Index by Oxford Insights and IDRC includes technological skills, it does not look at the overall education level of the country. This can be assessed by using the UNDP Education Index, an average mean of years of schooling, and expected years of schooling. Unequal access to education leads to interregional disparities in research productivity. Although Africa accounts for nearly 17% of the world's population, it generates only 3% of the global GDP and less than 1% of the world's research [27]. A good foundation for scientific education at the tertiary level requires enhancing digital literacy as early as kindergarten [28]. There is a need to strengthen the digital component of education by introducing programming skills, statistics, analytical and critical thinking skills as well as basics of data protection and intellectual protection laws [29]. These skills provide technological competency and creativity which are needed in the development of capabilities for Big Data analytics [30, 31]. Public authorities need to ensure that digital skills are embedded in education at all stages, including lifelong education for adults (SDG 4.4). Building a

wide range of ICT competencies (including programming, critical thinking, and basic statistical skills) is essential in the context of the robotification of the economy, as it prepares individuals for higher-end work in analysis, verification, or optimization of algorithm-based processes [24].

Research and Development: Other than education, Governments need to invest enough in Research and Development. Lack of/limited investment in Research and Development, together with unattractive pay, prestige, and working conditions, leads to brain drain. This makes it hard for local research institutions to reach their research potential and limits quality scientific research output to inform the development of home-grown technologies and solutions. This can be easily assessed by looking at how many locals undertake graduate studies in local institutions or abroad but are locally financed and working on local problems. Scientists often strike in most African countries and other developing countries, protesting research and development budget cuts [19]. If a country's researchers are underfunded and left to seek grants from outside sources or relocate to other countries, their research and the future of technological innovations will be influenced by outside sources. The government's acceptance and funding of AI technologies for Research Institutions and enterprises need to be prioritised and well-supported to boost AI implementation.

Gross Domestic Product Per Capita Purchasing Power Parity (GDP-PPP):
GDP-PPP is an individual's ability to buy the same quantity of an item in different countries. The GDP-PPP data published by the Central Intelligence Agency World Fact Book compares each Country's GDP on a purchasing power parity basis divided by the population as of July 1st for the same year. However, some countries with a high GDP-PPP may not score highly on the Government AI Readiness Index due to having small populations or specialised economies and lacking investment opportunities for high-impact technological innovation. This is common with countries that rely on tourism.

There are also cases of countries with low GDP-PPP ranking but high Government AI Readiness Index due to a growing or diversified economy combined with technological skills and data protection policies. The big Mac index, which compares the prices of a big Mac in different countries, illustrates the value of currencies that may be under or overvalued [19]. Table 2 summarises the dimensions that are key for Africa's AI readiness index and the extent to which they are covered by the Government AI Readiness Index by Oxford Insights and IDRC.

Table 2. Summary of dimensions that are key for Africa's AI readiness index and to what extent these are covered by the Government AI Readiness Index.

Dimensions for the Africa AI Index	Description	Indicators	Government AI Readiness Index by Oxford Insights & IDRC	
			What is covered	What is missing
Internet connectivity	Connectivity to the internet particularly broadband (broad band subscriptions) obtainable from ITU, 3G or higher network coverage	Percentage of the population connected to 3G or higher network 3G or higher network coverage in the country	Infrastructure dimension-telecom and 5G infrastructure	Internet use This is covered under data availability but it's also a key indicator of infrastructure
End user devices	Ownership of Internet enabled devices	Percentage of the population owning internet enabled devices	None	Add ownership of internet enabled end user devices to the infrastructure dimension
Regulation	Data protection and privacy and general AI implementation regulations	Availability of data protection and privacy regulations Availability of general AI implementation regulations	Governance & Ethics dimension- Ethics, rights, data protection and privacy laws	None. Adapt it as is
Availability of data	Availability of open data (government and private)	Availability of open data (government and private)	Data availability dimension-open government data, statistical capacity, mobile telephone subscriptions & internet users	None. Adapt it as is
Skilled workforce	Local talent and communities of researchers and entrepreneurs, to develop home grown AI solutions	Local AI Experts Local communities of researchers and entrepreneurs in AI technologies	Human Capital dimension: Graduates in STEM, quality of engineering and technology higher education, digital skills & knowledge-intensive employment	Emphasize nurturing of local skilled workforce as Engineers, Researchers and AI Tech entrepreneurs
General level of employment	Number of unemployed persons as a percentage of the total labour force as specified by the World Economic Forum	Level of unemployment	None	Add Level of employment as a new dimension to a new pillar called economy sector
Employment in data science and AI roles	Level of employment in Data Science and AI roles	Percentage of jobs advertised for data Science & AI related roles	New dimension	New dimension under economy sector

(*continued*)

Table 2. (*continued*)

Dimensions for the Africa AI Index	Description	Indicators	Government AI Readiness Index by Oxford Insights & IDRC	
			What is covered	What is missing
Overall education level	Overall education level of the country defined by the UNDP education index, level of digital literacy, digital component of the education system covering among others introduction to programming, statistics, and analytical and critical thinking, and basics of data protection and intellectual property	Overall education level Level of digital literacy Digital component of the education system	Human capital dimension-Graduates in STEM, quality of engineering & technology in HE, digital skills, knowledge intensive employment	Add general population literacy Digital component of non-STEM education system
Research & development	Government investment in Research & Development for AI, quantity and quality of research in AI output	Percentage of Government investment in Research & Investment in AI, quantity and quality of research in AI output, Number of AI start-ups, Attendance of data science & AI research and innovation events	Innovation capacity dimension-Entrepreneurship culture, ease of doing business, quality of engineering & technology in Higher Education (HE), digital skills, knowledge intensive employment, company investment in emerging technologies	Add to the innovation capacity dimension, quantity and quality of AI research
Gross Domestic Product-Per Capita Purchasing Power Parity (GDP-PPP)	An individual's ability to buy the same quantity of an item in different countries	Ability of the local population to buy the same quantity of an item in different countries	None	New dimension

The analysis summarized in Table 2 above shows several dimensions relevant to Africa's AI Readiness Index, their corresponding indicators, and the extent to which the Government AI Readiness Index by Oxford Insights and IDRC covers them. It also shows gaps in some dimensions of the Oxford Insights Index as well as new dimensions needed.

Based on the analysis in Tables 1 and 2, seven dimensions in the Government AI Readiness Index by Oxford Insights and IDRC need to be adapted as they are. These include vision, adaptability, data representatives, size, digital capacity, regulation, and availability.

On the other hand, the dimensions with gaps are 4 namely: infrastructure/internet connectivity, human capital, education, and innovation capacity.

The infrastructure dimension does not cover the quality and quantity of end-user devices and affordability of the Internet, which are both necessary for meaningful consumption of AI technologies and products/services, as well as generation of Big Data. Furthermore, Internet users should indicate data availability and infrastructure dimensions. On the other hand, the human capital dimension does not cover the general literacy of the population, which is necessary for effective demand and utilization of AI technologies, products, and services. In addition, the human capital dimension needs to emphasize nurturing the local skilled workforce, such as Engineers, Researchers, and AI Tech entrepreneurs that can develop and sustain homegrown AI solutions. The education dimension needs to include the general literacy level of the population and the digital component of non-STEM education programmes to mainstream nationwide development of digital literacy skills. Furthermore, the Research and Development dimension needs to include the quantity and quality of AI research and innovations in order to increase local knowledge and capacity for homegrown AI technologies and solutions.

The new proposed dimensions include level of general employment, level of employment in Data Science and AI roles, and Gross Domestic Product-Per Capita Purchasing Power Parity (GDP-PPP) which we recommend being categorized under a new pillar called the economy sector.

3.3 Towards an AI Readiness Index for Africa

Based on the analysis in Sects. 3.2 and 3.3, we propose an AI Readiness Index for Africa covering the adopted as is dimensions from the Government AI Readiness Index by Oxford Insights & IDRC, the dimensions that need modifications along with their recommended modifications, and the new dimensions recommended for the African context. Each dimension has a pillar/building block under which it falls, a description, indicators, and comment about whether it is adopted as is, modified or new.

It comprises of 4 pillars, three from the Oxford Insights and IDRC Index (Government, Technology sector, and Data and Infrastructure), and one new one i.e. Economy Sector.

Adapted as is from the Oxford Insights and IDRC Index: All 4 dimensions of the Government pilar for the Oxford Insights and IDRC Index are adapted as is namely: Vision, Governance and Ethics, Digital capacity, and Adaptability. From the Technology sector pillar, only 1 out of the 3 dimensions is adopted as is i.e. size while from the Infrastructure and Data pillar, also only 1 out of the 3 dimensions is adopted as is i.e. data representativeness.

Dimensions of the Oxford Insights and IDRC Index Recommended for Modification: Under the Technology sector pillar, we recommend renaming innovation capacity to Research and Development to make it broader, renaming human capital to education to also make it broader to cover general education/literacy level of the country's population, the quality and numbers of the STEM education system, and the digital component of the non-STEM education system.

Under the Data and Infrastructure pillar, we recommend the addition of 3G or higher network coverage, meaningful connectivity to the internet, and ownership of Internet-enabled devices to the infrastructure dimension.

New Pillar with New Dimensions: A new pillar called the economy sector with 3 new dimensions (level of general employment, level of employment in data science and AI roles, and GDP-Per Capita Purchasing Power Parity) is recommended. The 3 dimensions are determinants of the vibrancy of the AI-based economy in terms of the ability to consume commercial AI-based products and services, the quantity of local AI applications and innovations in public and private entities, and the competitiveness of the AI products and services market in the country which can drive demand and attract further investment. The proposed AI Readiness Index for Africa is presented in Table 3.

Table 3. The proposed AI readiness index for Africa

Dimension	Description	Indicators	Comment
Government pillar			
Vision	Does the government have a vision for supporting the development and implementation of AI?	National AI Strategy	Adapt as is from the Government AI Readiness Index
Governance and ethics	Are there the right regulations and ethical frameworks including data protection and privacy legislation to support implementation of AI in a way that builds trust and legitimacy?	• Data protection and privacy legislation • Cybersecurity legislation • National Ethics Framework • Legal framework's adaptability to digital business models	Adapt as is from the Government AI Readiness Index
Digital capacity	What is the existing digital capacity within government?	• Government procurement of advanced technology • ICT use and Government efficiency • Online services	Adapt as is from the Government AI Readiness Index
Adaptability	Can the government change, adapt and innovate effectively?	• Effectiveness of Government • Government's responsiveness to change	Adapt as is from the Government AI Readiness Index
Technology sector pillar			
Size	How large is the technology sector that will supply governments with AI technologies?	• Number of technology unicorns • Market value of public technology companies • Value of trade in ICT services (per capita) • Value of trade in ICT goods (per capita) • Computer software spending	Adapt as is from the Government AI Readiness Index
Research & development	Percentage of Government expenditure on Research & Development for AI, quantity and quality of research output in AI and AI Researcher hire costs	Percentage of Government investment in Research & Investment in AI Quantity and quality of research output in AI AI Researcher hire costs	Rename innovation capacity to Research & Development & adopt the description in column two and merge the two sets of indicators

(*continued*)

Table 3. (*continued*)

Dimension	Description	Indicators	Comment
Education	Overall education level of the country defined by the UNDP education index, level of digital literacy, digital component of the education system covering among others introduction to programming, statistics, and analytical and critical thinking, and basics of data protection and intellectual property	Overall education level Level of digital literacy Digital component of the education system	Rename Human capital dimension to education dimension and add new indicators-general population literacy and digital component of non-STEM education system
Data and infrastructure			
Infrastructure	• Does the country have a good technological infrastructure to support AI technologies? • Availability of latest technologies • Internet bandwidth • Population connected to the internet particularly broadband (broad band subscriptions) obtainable from ITU • 3G or higher network coverage • Ownership of Internet enabled devices	• Telecommunications infrastructure • 5G infrastructure • Percentage of the population connected to 3G or higher network • 3G or higher network coverage in the country • Percentage of the population owning internet enabled devices	Add: • Population connected to the internet particularly broad band subscriptions • 3G or higher network coverage • Population owning internet enabled devices
Data	Is there good availability of data that could be used to train AI models?	• Open government data • Statistical capacity • Mobile-cellular telephone subscriptions • Internet users (% of adult population)	Add availability of open private data
Data representativeness	Is the data available likely to be representative of the population as a whole	• Gender gap in Internet usage • Socioeconomic gap in Internet usage	Adapt as is from the Government AI Readiness Index
Economy sector pillar (new pillar)			
Level of general employment	Number of employed persons in a country as a percentage of the total labour force as specified by the World Economic Forum	Level of employment	New dimension under new pillar
Level of employment in data science and AI roles	Level of employment in data and AI roles	Percentage of jobs advertised for Data Science & AI related roles	New dimension under new pillar

(*continued*)

Table 3. (*continued*)

Dimension	Description	Indicators	Comment
Gross Domestic Product-Per Capita Purchasing Power Parity (GDP-PPP)	An individual's ability to buy the same quantity of an item in different countries	Ability of the local population to buy the same quantity of an item in different countries	New dimension under new pillar

4 Discussion, Conclusion and Future Work

This paper has attempts to develop the African AI Readiness Index by examining the AI readiness needs of African countries against what is currently covered in existing Government AI Readiness Indices. The proposed index expands, elaborates, and zooms in on key modern technology development needs of Africa. This process led to suggestions to rename some dimensions to make them broader or more focused, suggestions for new dimensions that are very important for Africa's technological development, and suggestions for a new pillar about the performance of the economy that determines the demand for AI-based products, goods, and services which in turn can influence further investment by both locals and foreigners in the sector. The level of jobs in Data Science and AI roles is an indicator of the level of AI use and AI innovations in the country. On the other hand, the GDP-PPP of a country is an indicator of how competitive a country's market for AI products and services is compared to its neighbors and globally. This is an important driver for adoption.

Future work will extend the literature review to general drivers for technological development for Africa in order to capture more salient needs for Africa's AI readiness. We will also study more literature on AI readiness for other continents including case studies to further refine the reediness dimensions and indicators. Further, we will consider a field study on a sample of African countries in different regions to validate the proposed AI readiness Index for Africa.

References

1. ESCAP, U.: Artificial Intelligence in Asia and the Pacific (2017). https://www.unescap.org/resources/artificial-intelligence-asia-and-pacific. Accessed 10 Aug 2022
2. Arakpogun, E.O., Elsahn, Z., Olan, F., Elsahn, F.: Artificial intelligence in Africa: challenges and opportunities. In: Hamdan, A., Hassanien, A.E., Razzaque, A., Alareeni, B. (eds.) The fourth industrial revolution: implementation of artificial intelligence for growing business success. SCI, vol. 935, pp. 375–388. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-62796-6_22
3. Rossi, F.: Building trust in artificial intelligence. J. Int. Aff. **72**(1), 127–134 (2018)
4. Haseeb, M., Mihardjo, L.W., Gill, A.R., Jermsittiparsert, K.: Economic impact of artificial intelligence: new look for the macroeconomic assessment in Asia-Pacific region. Int. J. Comput. Intell. Syst. **12**(2), 1295 (2019)
5. Susar, D., Aquaro, V.: Artificial intelligence: opportunities and challenges for the public sector. In Proceedings of the 12th International Conference on Theory and Practice of Electronic Governance, pp. 418–426 (2019)

6. Oxford Insights: Government AI Readiness Index (2019). https://www.oxfordinsights.com/ai-readiness2019. Accessed 22 July 2022
7. Gillham, J.: The macroeconomic impact of artificial intelligence. Price water house Coopers (2018). https://doi.org/10.13140/RG.2.2.21506.38083
8. Jain, S.: US $320 billion by 2030. The potential impact of AI in the Middle East (2018). https://www.pwc.com/m1/en/publications/potential-impact-artificial-intelligence-middle-east.html. Accessed 28 July 2022
9. Chen, N., Christensen, L., Gallagher, K., Mate, R., Rafert, G.: Global economic impacts associated with artificial intelligence. Anal. Group 1–23 (2016)
10. APET: Africa's Socio-Economic Development AI for Africa (2018)
11. Vernon, D.: Robotics and artificial intelligence in Africa [regional]. IEEE Robot. Autom. Mag. **26**(4), 131–135 (2019)
12. ESCAP, U.: Artificial intelligence in the delivery of public services (2019). https://www.unescap.org/publications/artificial-intelligence-delivery-public-services. Accessed 10 Aug 2022
13. Oxford Insights: Government AI Readiness Index (2017). https://www.oxfordinsights.com/government-ai-readiness-index2021. Accessed 21 July 2022
14. Oxford Insights: Government AI Readiness Index (2021). https://www.oxfordinsights.com/ai-readiness2021. Accessed 22 July 2022
15. Daniel, Z., et al.: The AI index 2021 annual report (2021). https://hai.stanford.edu/ai-index-2021. Accessed 09 July 2022
16. Daniel, Z., et al.: The AI index 2022 annual report (2022). https://hai.stanford.edu/ai-index-2022. Accessed 09 July 2022
17. International Institute of Communications: Artificial intelligence in the Asia-Pacific region (2020). https://www.iicom.org/wp-content/uploads/IIC-AI-Report-2020.pdf. Accessed 13 Aug 2022
18. Moher, D., et al.: Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (PRISMA-P) statement. Syst. Rev. **4**(1) (2015). https://doi.org/10.1186/2046-4053-41
19. Dybå, T., Dingsøyr, T.: Strength of evidence in systematic reviews in software engineering. In Proceedings of the ACM-IEEE International Symposium on Empirical Software Engineering and Measurement, Kaiserslautern, Germany (2008)
20. Montoya, L., Rivas, P.: Government AI readiness meta-analysis for Latin America and the Caribbean. In: Proceedings of the 2019 International Symposium on Technology in Society (2019)
21. Oxford Insights: Government AI Readiness Index (2020). https://www.oxfordinsights.com/ai-readiness2020. Accessed 22 July 2022
22. Oxford Insights: Government AI Readiness Index (2021). https://www.oxfordinsights.com/ai-readiness2021. Accessed 15 Aug 2022
23. ITU: Measuring digital development: facts and figures (2019). https://www.itu.int/en/ITU-D/Statistics/Pages/facts/default.aspx. Accessed 15 June 2020
24. Oxford Insights: Government AI Readiness Index (2022). https://www.oxfordinsights.com/ai-readiness20222. Accessed 15 Aug 2022
25. Szoszkiewicz, L., Mickiewicz, A.: Open data: toward achieving and measuring the sustainable development goals (2021). https://doi.org/10.1007/978-3-319-71059-4_129-1
26. ECOSOC – Economic and Social Council: Special edition: progress towards the Sustainable Development Goals (2019). 9 May 2019. E/2019/68
27. ECA, UNDP, Open Data for Development Network & World Wide Web Foundation. The Africa data revolution report 2016. Addis Ababa (2017)

28. Youn Duermeijer, Ch., Amir, M., Schoombee, L.: Africa generates less than 1% of the world's research; data analytics can change that. Elsevier (2018). https://www.elsevier.com/connect/africa-generatesless-than-1-of-the-worlds-research-data-analytics-canchange-that. Accessed 15 June 2020
29. Coleman, C.: This region will be worth $5.6 trillion within 5 years – but only if it accelerates its policy reforms. In: World Economic Forum (2020). https://www.weforum.org/agenda/2020/02/africaglobal-growth-economics-worldwide-gdp/. Accessed 15 June 2020
30. Chaudron, S., Di Gioia, R., Gemo, M.: Children (0–8) and digital technology – a qualitative study across Europe. Publications Office of the European Union. Luxembourg (2017)
31. Awan, U., Sroufe, R., Kraslawski, A.: Creativity enables sustainable development: supplier engagement as a boundary condition for the positive effect on green innovation. J. Clean. Prod. **226**, 172–185 (2019). https://doi.org/10.1016/j.jclepro.2019.03.308
32. Shamim, S., Zeng, J., Shariq, S.M., Khan, Z.: Role of big data management in enhancing big data decisionmaking capability and quality among Chinese firms: adynamic capabilities view. Inf. Manag. **56**(6), 103135 (2019). https://doi.org/10.1016/j.im.2018.12.003
33. West, D.M., Allen, J.R.: How artificial intelligence is transforming the world. Report, 24 April 2018 (2018). https://www.bbc.com/news/business-44849492

An Adaptive and Dynamic Heterogeneous Ensemble Model for Credit Scoring

Tinofirei Museba(✉)

University of Johannesburg, Auckland Park, Johannesburg, South Africa
tmuseba@uj.ac.za

Abstract. The determination of the financial credibility of a person for a loan is a challenging task as many variables are taken into consideration. Recently, there has been a surge in the application of machine learning approaches in the design of robust and effective credit scoring models as part of the human social development agenda under the assumption that the variables will remain stable for a long time. However, in real-life, the behavior of customers changes over time and the variables used to quantify the financial credibility of a person for a loan such as past performances on debt obligations, profiling, main household, income and demographics tend to drift and evolve over time. This paper considers credit scoring as an ephemeral scenario as variables tend to drift over time and proposes the application of data stream learning techniques in credit scoring since they are tailored for incremental learning. This makes the scoring model to be able to detect and adapt to changes in the customer behavior.

We propose the Adaptive and Dynamic Heterogeneous Ensemble (ADHE) approach that is capable of learning incrementally and adapting to drifting variables and consists of models derived from different learning algorithms to exploit diversity. The prediction performance of ADHE is evaluated using datasets that are publicly available and we compared the accuracy and computational cost of ADHE with existing state of the art models. Our proposed approach performs significantly well when compared to existing state of the art benchmark models on prediction accuracy according to the non-parametric test.

Keywords: Credit scoring · Heterogeneous ensemble · Machine learning · Diversity

1 Introduction

Extending credit to customers is a great risk for most financial institutions and a robust and effective credit score has the potential to generate great profits to stakeholders [1]. The prevalence of imperfect information makes the possibility of risk occurring to be high in the credit market as a result of factors such as market failure, credit rationing and interest rates.

Credit risk poses a major threat for most financial institutions and the existence of an effective and robust credit risk management is a critical issue for financial institutions to

P. Ndayizigamiye et al. (Eds.): IDIA 2022, CCIS 1774, pp. 304–319, 2023.
https://doi.org/10.1007/978-3-031-28472-4_19

stay afloat in business. Credit scoring has emerged as one of the efficient and effective tool to properly mitigate information asymmetry and to manage credit risk [2].

The development of robust and effective credit scoring models to improve prediction accuracy has become an important task for financial institutions as it leads to significant boost in profits and avoids potential losses. Robust and efficient credit scoring models have large and economically significant influence on the profitability of financial institutions. Credit scoring can be considered as a binary classification problem. The rationale behind the development of robust and efficient models for credit scoring is to create a score and assist in the human social development agenda thereby improving the livelihoods of those living in deprivation. The generated score is used to discriminate loan applicants into two categories namely: credit worthy, one with good potential to pay the loan and not credit worthy, the risky applicants who are likely to default on their obligations to pay. The generated score hinges on the probability of the customer to likely default on his payment obligations and the task is converted to a classification problem [3]. Credit scoring models have been developed using either statistical techniques or machine learning techniques. Despite the simple implementation and interpretation advantages, statistical approaches suffer from insufficient accuracy and makes assumptions such as multivariate normality and are not consistent with real-life situations.

Machine learning techniques have been used to develop credit score models as they are capable of outperforming existing techniques based on statistical approaches in terms of prediction performance. Machine learning techniques handle nonlinear pattern classification problems well when compared to statistical techniques as they do not assume a stable distribution. Machine learning techniques provide an effective way of dealing with the big data in large, sparse and complex high dimensional samples [4] which makes them very suitable for developing credit scoring models in an online fashion, capable of detecting and adapting to change. Ensemble machine learning models work by pooling together the prediction outputs of several base learners and combining them to form the final decision and are able to generate complementary strengths and have been experimentally and analytically demonstrated to be accurate and stable when compared to single network models [5]. The superiority of ensemble models is due to the fact that different base models view the same data instance differently and this allows models to complement each other [6]. Although both accuracy and diversity of base learners are critical factors for constructing ensemble models, most existing models consider only one aspect and ignore the tradeoff between them [7]. In the case of credit score models, the variables may drift, requiring different amounts of diversity. The last decade has witnessed a surge in the application of machine learning in the design of credit score models as it forms part of the human social development agenda in improving the livelihoods of those living in deprivation.

To cover the research gap highlighted above, this paper proposes the development of robust, assertive and adaptive model for credit scoring. Our novel approach is based on a heterogeneous ensemble model capable of learning both batch and streaming data since variables of customers can drift anytime. The rest of the paper is organized as follows; Sect. 2 provides a review of related work on the application of ensembles of learning machines. Section 3 provides the methodology and base learners. Section 4 provides

data preprocessing approach, feature selection and parameter optimization. Section 5 looks at experimental design and results obtained. Section 6 provides the conclusion.

2 Related Work

Ensembles of machine learning improve their generalization performance by enhancing the accuracy of base learners, increasing the amount of diversity of base models together with the combination strategy of the prediction outputs. Ensemble of machines learning have demonstrated their superiority for classification tasks such as credit scoring. In credit scoring, the objective is to generate a model that is quantitative and exploits several attributes to be able to classify two categories of clients as creditworthy and non-creditworthy. Several ensemble contemporary approaches have been proposed in the literature. To enhance the predictive ability of the ensemble, the authors in [8] proposed an ensemble that combines the prediction output of learners generated from random forest and extreme gradient boosting algorithms. The base learners are optimized using particle swarm optimizers. Adjustable balanced sets are generated using the Balance Cascade approach. The approach incurs high computational cost. In [9], the authors improved the ensemble learning approach by creating a hybrid ensemble for credit scoring by incorporating the stacking approach to handle noise and weight assignment. The Bayesian optimization algorithm is used to optimize the hyper-parameters and assigns weights using the multi-stage backward cloud generator (MBCG) dynamic weighting strategy. The approach takes long to reveal new concepts due the process of assigning weights. The authors in [10] introduced a multi-level ensemble classifier to handle outliers in datasets and the approach uses PSO clustering. Weights are assigned to classifiers according to their prediction accuracy. If the parameters are not properly configured, the algorithm may converge prematurely. Wen-hui Hou [11] modified the ensemble approach by introducing the concept of classifier selection on imbalanced datasets to improve credit score prediction accuracy. The Synthetic Minority over-sampling technique is used to balance the training set. Classifier evaluation is performed using meta-training thereby slowing down the learning process. In [12], the authors proposed an ensemble learning approach for credit scoring that uses five feature selection algorithms and uses three types of voting and eight different ensemble models that are combined using soft voting. The approach is not resource efficient in terms of computing power and memory. An overfitting-cautious tree based dynamic heterogeneous was proposed by Yufei Xia [13]. Although empirical experiments conducted showed that the approach significantly outperforms homogeneous ensembles, the ensemble selection applied is not efficient and does not generate different amounts of diversity. In [14] the authors performed a comparative study of batch and data stream mining in credit scoring. Results obtained showed that the data stream learners are comparable to predictive models that are currently being used. Jian Xiao [15] suggested a not fully supervised learning, learning model that is cost sensitive, a group method that handles data and an ensemble technique to address the problem of classifying loan applicants. The proposed method introduces additional noise due to the fact that of the samples are not correctly labelled To improve the prediction accuracy of the ensemble, Xiaohong et al. [16] suggested an ensemble composed of models generated from different learning algorithms based on the generalized Shapley

value and the Choquet integral to create a credit score model. To test the feasibility of the Generalized Shapley Choquet Integral (GSCI), rigorous evaluations and comparative tests are performed to evaluate its ability to adapt and accurately classify applicants. The combination approach of using fuzzy measure and the aggregation operator may be ineffective and the evaluation metrics used are not enough for a sound conclusion of the performance of GSCI. Since in credit scoring, the variables used to quantify the financial credibility of a customer such as client information on past performances on debt obligations, profiling, income, demographics, possessions, income and census information may drift over time.

Inspired by previous works of the aforementioned authors, this paper proposes an adaptive and dynamic heterogeneous ensemble that combines models from extreme gradient boost and support vector machines to design a credit score to formulate a credit score model that is accurate and adaptive to drifting variables. Our proposal differs from existing models as it includes aspects such as dynamic model selection, parameter optimization, pool generation strategy. Regarding pool generation, XGBoost and support vector machines are used as base learning algorithms to generate models for the pool.

To adapt to drifting variables, the parameters of the two base learners are optimized using PSO.

3 Methodology

Credit risk assessment is a multidimensional and imbalanced problem mainly based on large volumes of variables such as job status, credit history and profiling. The proposed credit scoring model in this study uses a classifier ensemble to predict credit scoring datasets. The robust and assertive credit score model must adapt to different imbalance ratios and drifting variables. The datasets are split into training and testing sets. Training data is used as the training set and the test set is used as the validation set to evaluate model performance. The ensemble is constructed using models generated from two base classifiers namely: support vector machines and XGBoost and their parameters are optimized using the Particle Swarm Optimization (PSO) algorithm to improve the prediction result through the validation set. A robust and assertive ensemble model is then generated to perform the final prediction on the testing set.

3.1 Base Learners

To generate a pool of classifiers, two base learners, namely: support vector machines and XGBoost are employed. The two base learners are capable of striking a good balance between accuracy of the learning model and computational cost. A brief description of the two learning algorithms used in this study is provided in this subsection.

3.1.1 Support Vector Machines

Support vector machine is a learning algorithm that is powerful for both regression and classification problems in both stationary and time-varying domains. They have been widely applied used in domains such as credit scoring and have demonstrated

potential for robust credit scoring models. For classification problems, SVM generates an appropriate hyperplane that separates categories of data into proper classes regardless of whether the data is linear or not. This makes the distance of the hyperplane that separates the training instances to be large. In situations where the training instances are not linear, the SVM kernel function performs the mapping of the data to a high dimensional feature space. When data has been projected to a dimensional space that is higher, SVM can search for a proper hyperplane capable of separating two categories of data in a feature space of high dimension. Although SVM creates a maximum margin between training instances, scaling affects its prediction performance.

SVM is composed of two parameters that have to be tuned in order to approximate the loss function.

3.1.2 Extreme Gradient Boosting (XGBoost)

To solve real life classification problems, Chen and Guestrin [17] developed XGBoost. Due to its ability to mitigate model variances, it is mostly used by a number of winning solutions in Kaggle machine learning works for training models. XGBoost is based on the gradient boosting approach where the decision tree learners that are weak are combined into a strong model. To improve the efficiency and prediction performance of the learners, some modifications are performed. XGBoost employs a Taylor's expansion to quickly approximate the loss function and build a model capable of balancing bias and variance through the objective of decision trees with the objective of achieving superior generalization. In addition to that, a regularized term is attached to the loss function in order to addresss the complexity associated with the tree. As the level of complexities increases for the learning of trees, the normalization prevents the problems associated with over-fitness.

The XGBoost learning algorithm has the capacity to deal with data that is sparse, implementation of distributed and parallel computing flexibly [18]. XGBoost is based on both first and second derivatives The loss function generated is mostly like to be accurate and thus improving its capacity to precisely split the tree and generate the tree structure that is suitable to perform financial data analysis.

4 Data Preprocessing

The data is standardized, scaled within the 0–1 interval and missing values are approximated. The class imbalance problem of credit datasets is addressed. To standardize numeric features, we remove the mean and scale to unit variance. The 0–1 normalization process approach is adopted to scale the data properly. For a given feature x, the normalized $x^{'}$ is computed as follows:

$$\mathrm{x}' = \frac{\mathrm{x} - \min(\mathrm{x})}{\max(\mathrm{x}) - \min(\mathrm{x})}$$

where x' expresses the standardized value. Feature normalization greatly improves the accuracy of classifiers, especially those based on distance or edge calculations, rendering the model to be more assertive and accurate. Credit data is associated with missing values.

XGBoost incorporates a sparsity segmentation algorithm that estimates missing values accurately and performs well for modeling purposes in comparison with traditional methods designed to handle missing data. Standardization subdues the influence of outliers and extreme values are dealt with by the process of centralization. Datasets used for credit risk assessments are imbalanced. In most cases, clients defaulting on their payment obligations are fewer than those who are consistent with their payment obligations. Ignoring class imbalance may lead to the generation of a scoring model with high accuracy in determining non-defaulters but extremely low accuracy for the determination of defaults. To resolve the problem of class imbalance, this study employs the Synthetic Minority Oversampling (SMOTE) [19], an efficient and effective technique for handling class imbalance issues in both static and dynamic domains.

SMOTE works by synthesizing the underestimated minority sample based by exploiting feature space similarities that exist between the minority instances. Given a dataset T that is imbalanced, for each minority class instance $x_i \in T$, the SMOTE algorithm first searches the K nearest neighbor for x_i using Euclidean distance, after which one of the K nearest neighbors is randomly chosen and the feature vector difference between x_i and its corresponding nearest neighbor calculated. Finally, the feature vector difference multiplies a stochastic number and adds the new vector to x_i. The mathematical formulation for synthesizing a new minority sample is shown in Eq. (1).

$$x_{new} = x_i + \left(x_i^k - x_i\right) \times \delta, \tag{1}$$

where x_i^k is one of the nearest neighbors to x_i and δ is a random value which belongs to (0,1). Therefore, the synthesized minority instance x_{new} is a point along the line segment joining x_i and its nearest neighbor x_i^k.

4.1 Feature Selection

To perform feature selection, XGBoost utilizes scores that underline the importance of features to measure the average objective reduction when considering specific variables to split. A variable associated with a huge score is considered to be important when constructing a tree. Since this study employs XGBoost as one of the learning algorithms, the study carried out by Xia et al. (2017) [20] is followed to implement the ranks of scores for feature importance scores as a rule when adapting the sequential forward search (SFS) feature selection algorithm. SFS creates a number of candidate features by selecting features with higher scores and placing them into a subset and iteratively putting the rest of the features associated with highest scores into the subset. If a feature subset maximizes the cross validated accuracy, then it is chosen as the optimal feature set for model training in the steps that follow.

4.2 Parameter Optimization

The prediction performance of the credit scoring model is optimized by the SVM and XGBoost learning algorithms employed.

Several hyper-parameters optimization approaches exist and have their own merits and demerits.

Grid search has been widely used for optimization purposes and it performs an exhaustive search of specified parameters. The performance of the grid search algorithm is hindered by the prevalence of high dimensional data. As the number of parameters gets bigger and the space to be searched is extended, the computational cost of the grid search tends to increase dramatically. Bayesian hyper-parameter optimization has been used for most applications due to its superior performance and efficiency. Bayesian hyperparameter optimization creates a statistical model to map parameters to a specified objective which is normally carried out by performing cross-validation. Hyperparameters are iteratively selected, and their performance observed. The observations are collected, and a model is designed so that much information can be inferred concerning hyper-parameters. Hyper-parameter tuning is subjective and relies on empirical judgement and trial and error approaches [21].

In this study, Particle Swarm Optimization (PSO) is used to optimize hyper-parameters.

PSO is an extensively used heuristic algorithm. It is an evolutionary computational method developed by Kennedy and Eberhart [23].

For an appropriate optimization and generation of a proper subset of parameters and to optimize the efficacy of support vector machines and XGBoost, ADHE employs an adaptive heterogeneous Particle Swarm Optimizers.

Particle Swarm Optimizers (PSO) is a heuristic population based iterative, global and stochastic optimization algorithm inspired by the social behavior of bird flocking or fish schooling to conduct an intelligent search for the optimal solution [22]. PSO is derivative free as it does not require the optimization problem to be differentiable, therefore not requiring gradients making it applicable to a variety of problems such as those with discontinuous or non-convex and multimodal problems. In this study, particles of the swarm are instantiated at individual level thereby introducing heterogeneity among the swarm particles. Individual particles instantiation allows particles in a swarm to assume different search behaviors by randomly selecting velocity and position update rules from a behavior pool thereby creating a swarm that consists of explorative particles and exploitative particles, giving the optimization algorithm the ability to explore and exploit throughout the search process thereby avoiding premature convergence.

4.3 SVM and XGBoost Parameters

Support Vector Machines and XGBoost are the base learners used in this study. SVM is selected since its learning rate is independent of the amount of data and enables proper investigation of the role of diversity and adaptation to new concepts or drifting variables without confounding factors. Moreover, its intuitive to separate the two classes by a margin that can be controlled. The separating margin guards best against small errors in selecting the perfect decision boundary. The approach is robust since only a fraction of all data points are support vectors. This study uses the RBF kernel. Two parameters are considered, namely C and gamma. The parameter C trades off misclassification of training examples. The gamma parameter defines how far the influence of a single training example reaches.

The derived parameters of XGBoost are as follows (Table 1):

Table 1. Derivation of hyperparameters of XGBoost

Hyperparameter	Derivation
Learning rate	Weight value reduced and node robustness optimized
Tree depth maximum	Maximum tree depth generates a greater value, increases tree complexity and evades overfitting
Subsample ratio	Sample rate
Column subsample ratio	Features sampled to construct tree
Maximum child weight	Overfitting reduction via sum determination of minimum leaf nodes sample weights
Maximum delta step	Tree weight changes limit factor for unbalanced data
Gamma	Minimum loss function reduction necessary to evade over fitting
Number of estimators	The number of iterations of XGBoost

4.4 Dynamic Classifier Selection

After the pool of models has been generated, the Dynamic Ensemble Selection (DES) is applied to select the most accurate and diverse models. The Dynamic Ensemble Selection selects classifiers based on their competence on some fitness function. For this reason, the ensemble of classifiers is determined in a dynamic fashion [24]. Diversity and accuracy are two important features of ensemble learning. Intuitively speaking, the key to the success of an ensemble of classifiers is that the base classifiers perform diversely [25]. For this study, classifiers are selected based on their accuracy on the validation set and their diversity to accommodate both batch and incremental learning since variables tend to drift over time. To select classifiers based on accuracy and diversity, we apply the Selection by Accuracy and Diversity (SAD) algorithm [26]. The SAD algorithm works as follows:

1. Train a set of heterogeneous classifiers
2. Determine each classifier's accuracy on validation set
3. Accurate classifiers are selected or most top accurate classifiers
4. Measure the diversity between the most accurate classifiers
5. Classifiers with strong diversity are selected into the ensemble and selection process is repeated until predetermined size of the ensemble
6. Use majority voting to combine classifiers into an ensemble
7. Evaluate the generalization performance of the ensemble.

Due to its simplicity and ease of interpretation, the Q static [27] is used as a diversity measure in this study. To generate a pool of classifiers, the training dataset is used to learn classifiers. Given a training dataset $D_{train} = \{x, y\}$, where x is an $M \times N$ dimensional feature matrix and $y \in \{0,1^N$ denote the label. If y yields a value of 1, it is an indication of a default application and a value of 0 implies creditworthiness. Two base learners are used in an attempt to create a heterogeneous ensemble architecture. The base models used are

SVM and XGBoost. The two base learners are popularly used in credit scoring studies, generate high performance and are highly efficient as credit scoring models need to be resource efficient as they are updated frequently to adapt to changes as variables may drift [28]. For both batch and incremental learning, accuracy is a measure that is frequently used. Since server class imbalance may exist and may lead to misleading information, the use of a diversity measure compensates the absence of posterior probability prediction. Accuracy is an effective metric for both batch and incremental learning as it provides coherent information regarding changes in the data especially in credit scoring models where variables can drift.

5 Experimental Design

To validate the effectiveness of our proposed approach, we employ four real world credit datasets. Two popular datasets, the Australian and German datasets are publicly available on the UCI Machine Learning Repository [29]. The two datasets are widely used in related literature for credit scoring. This makes it feasible to perform a detailed comparison of the generalization performance of our proposed approach with other studies. The Australian and German datasets are relatively too small to cover all aspects of credit scoring. We therefore introduce two more large datasets with more instances. From the peer to peer (P2P), we employ the RRDai data from the Chinese Internet finance enterprise called RenRenDai that consists of loan data for the year 2017. The dataset is publicly available on https://www.renrendai.com. The second real-world big credit dataset is called the LendingClubLoan and is sourced from a Peer to Peer (P2P) lending company called LendingClub. The dataset consists of complete loan data from the year 2007 to 2015 and includes the current loan status and recent information regarding payments. A summary of the four datasets is presented in Table 2.

Table2. Formulation of credit score data

Dataset	Instances	Features	Training set	Test set
German	1000	24	800	200
Australia	690	14	552	138
RRDai Ren	1421	17	1137	284
LendingClubLoan	2642	11	2114	528

5.1 Evaluation Measures

The commonly used performance metrics is the average accuracy in credit scoring. The overall generalization performance of the learning model is reflected there.in exploring the effectiveness of the model to differentiate nondefault and default applications, six evaluation metrics are empirically employed to predict the generalization performance

of our proposed approach. These metrics are Accuracy (ACC), Brier score (BS), Area Under the rock curve, F1 score, Type 1 and Type 11 errors. The later two performance metrics evaluate performances models with more detail than others. When a type 1 error occurs, a default loan application will have been wrongly classified and a type 11 error occurs when a nondefault is classified as default. In a confusion matrix the numbers are accurately classified as TP for good borrowers and TN for bad borrowers. Numbers of misclassified loan applications are represented by two variables FP and FN. The following is a representation of performance metrics.

$$\text{The average accuracy (ACC)} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} \tag{2}$$

$$\text{The type 1 errors} = \frac{\text{FP}}{\text{TN} + \text{FP}} \tag{3}$$

$$\text{The type 11 errors} = \frac{\text{FN}}{\text{TN} + \text{FP}} \tag{4}$$

The Brier score is a performance metric used to measure the accuracy of the predicted probability and the calibration of the prediction performance. The Brier score is within the interval 0 to 1 and the value of the interval represents probabilistic predictions from perfect to poor. A lower Brier score reflects better predictions.

The Brier score can be expressed as:

$$\text{BS} = \frac{1}{\text{N}} \sum_{i=1}^{N} (p_i - y_i), \tag{5}$$

where N is the number of samples and p_i and y_i denote the probability prediction and the true label respectively of sample i.

The F1-score considers both precision and recall of classification models. It is the harmonic average of precision and recall and ranges within the interval of 0 to 1. F1 can be expressed as:

$$\text{F1} = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}, \tag{6}$$

where precision is the proportion of positive samples in positive cases and is defined as:

$$\text{Precision} = \frac{\text{FP}}{\text{TN} + \text{FP}} \tag{7}$$

Recall is expressed as the proportion of predicted positive cases in the total positive cases and is expressed as

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{8}$$

The Area Under the Roc Curve (AUC) performs a global assessment by computing the area under the Receiver Operating Characteristics Curve (ROC) according to the predicted scores which is plotted against the True Positive Rate (TPR) and False positive Rate (FPR).

5.2 Experimental Results

In this section, a comprehensive comparison between our proposed model and the benchmark models is performed. All models are validated over four real world credit datasets in the six evaluation metrics. in administering these experiments, Python 2.7on a desktop with 36 GHz intel 7 CPU, 8 GB RAM and Microsoft Windows 10 operating system were used.

To demonstrate the prediction power of the Adaptive and Dynamic Heterogeneous Ensemble (ADHE), we evaluated the benchmark models on four datasets. Ten classifiers, namely, XGBoost, SVM, ANN, GBDT, LightGBM, CatBoost, Decision Tree, Leveraging Bagging, Adaptive Random Forest and ADHE are validated across the datasets to serve as baselines. Six performance metrics were applied to evaluate the efficacy of the proposed ADHE approach. Table 3 shows the average prediction performance of each benchmark model on six performance metrics. The Adaptive and Dynamic Heterogeneous Ensemble (ADHE) performs the best among the four datasets and takes the first place in most cases among the evaluation metrics. Adaptive Random Forest model also demonstrated good performance. Leveraging Bagging and Decision tree present acceptable results over the datasets. Therefore they were selected as benchmarks models in most literature related to credit scoring. The ACC measure provides comparative results in the four datasets. The performance on CatBoost and LightGBM measure shows patterns that differ for most trends, which inversely demonstrates the importance of evaluating models from different aspects such as discriminatory ability and probability.

Table 3. Performance of Benchmark models across datasets

Dataset	Classifier	ACC	F1	AUC	BS	TYPE 1	TYPE 11
Australia	XGBoost	72.43	53.65	62.05	0.2268	56.49	12.65
	SVM	71.85	43.85	58.85	0.1678	53.66	16.46
	ANN	73.57	56.43	67.35	0.2465	50.42	13.68
	GBDT	70.58	59.35	63.48	0.1874	65.48	10.48
	LightGBM	71.35	61.67	61.35	0.2247	56.34	11.56
	CatBoost	74.45	54.32	59.85	0.1778	62.23	14.28
	Decision Tree	72.85	62.28	60.59	0.1764	58.45	15.09
	Leverage Bagging	74.33	57.76	55.79	0.1743	63.54	17.36
	Adaptive Random Forest	72.85	54.34	61.75	0.1659	54.35	11.49
	ADHE	**75.43**	**72.48**	**70.48**	**0.1604**	**53.05**	**10.43**
German	XGBoost	66.34	58.58	63.46	0.2213	62.53	12.87
	SVM	69.34	62.87	67.82	0.2174	61.27	13.48
	ANN	73.47	59.42	63.28	0.1783	59.74	15.85

(*continued*)

Table 3. (*continued*)

Dataset	Classifier	ACC	F1	AUC	BS	TYPE 1	TYPE 11
	GBDT	68.76	64.37	59.86	0.1674	56.46	11.78
	LightGBM	71.26	69.72	60.47	0.2034	60.74	12.35
	CatBoost	74.64	67.48	58.73	0.2167	58.79	14.56
	Decision Tree	72,58	65.39	65.48	0.1784	64.72	16.86
	Leverage Bagging	74.78	70.84	69.87	0.1674	68.43	11.38
	Adaptive Random Forest	68.85	69.87	64.89	0.1689	64.58	12.69
	ADHE	**76.83**	**73.74**	**71.84**	**0.1567**	**69.79**	**11.21**
RRDai	XGBoost	74.47	59.47	72.38	0.2217	0.1874	14.23
	SVM	69.83	61.35	69.85	0.2145	0.2167	16.78
	ANN	71.68	58.67	64.58	0.1673	0.2214	13.29
	GBDT	68.73	63.48	62.87	0.1784	0.1894	15.63
	LightGBM	70.39	60.79	67.89	0.2242	0.1787	12.84
	CatBoost	66.58	57.45	68.89	0.1783	0.2231	17.62
	Decision Tree	68.38	70.34	70.38	0.1687	0.1875	12.31
	Leveraging Bagging	72.68	72.84	73.49	0.1602	0.2262	11.87
	Adaptive Random Forest	74.86	76.58	71.29	0.1687	0.1876	13.56
	ADHE	**72.35**	**70.35**	**74.67**	**0.1583**	**0.1671**	**11.05**
LendingClubLoan	XGBoost	69.47	66.89	71.86	0.1873	0.2217	16.81
	SVM	65.67	63.47	68.47	0.2241	0.1764	13.48
	ANN	63.28	59.83	66.89	0.1784	0.2137	12.75
	GBDT	66.87	68.42	63.12	0.1843	0.1785	14.39
	LightGBM	69.84	71.89	67.84	0.1798	0.2236	13.89
	CatBoost	67.38	69.84	69.54	0.2345	0.1706	12.34
	Decision Tree	70.83	71.73	73.42	0.1865	0.1689	13.26
	Leveraging Bagging	73.48	74.86	74.89	0.1721	0.1654	12.19
	Adaptive Random Forest	71.89	72.49	72.36	0.1672	0.1702	11.87
	ADHE	**77.58**	**76.69**	**74.85**	**0.1612**	**0.1603**	**10.89**

5.3 Comparison of Computational Cost

In credit scoring, a resource efficient model is important since it stores a large number of clients. Computational cost must be taken into account when designing credit score models. An efficient credit score provides quick responses to loan applicants. Since variables can drift, overheads in training often hinder the process of frequently updating the credit scoring model. The selection of models from a pool is expected to decrease the computational time significantly. In this section, a comparison of the computational cost of benchmark models and our proposed approach, ADHE, is performed. Computational cost is measured by using a single training time [30], which is easily computed as a single cross-validation. Table 4 shows the single training time of the benchmark models on the four datasets.

Table4. Computational cost of single training time of models across datasets

	Australian	German	RRDai	LendingClubLoan
XGBoost	0.36	0.51	1.37	1.36
SVM	0.27	0.46	1.14	0.98
ANN	0.68	1.65	2.36	3.16
GBDT	1.98	3.48	2.94	2.45
LightGBM	2.86	6.35	4.35	3.83
CatBoost	4.83	5.78	6.45	3.37
Decision Tree	0.74	1.26	2.34	1.83
Leveraging Bagging	3.75	4.25	3.28	2.94
Adaptive Random Forest	2.68	1.76	2.46	3.43
ADHE	3.18	3.48	3.25	5.68

The result of assessing the computational cost of the benchmark models reveals a tradeoff between computational cost and model performance. Models based on algorithm use a few computational resources even though their generalization performance is poor when compared to benchmark models. For datasets with more features and are large, the training process is slowed down. The dynamic selective heterogeneous process employed by ADHE credit scoring model shows a comparative performance.

5.4 Statistical Tests of Significance

The non-parametric Friedman ranking test is employed for the ten baseline models for the determination of significant statistical differences. Reliable comparisons can only be demonstrated through the integration of different evaluation measure, given that they exhibit different merits and limitations. To properly conduct the statistical differences of prediction performance [31, the author suggested the application of non-parametric tests in place of $k + 1$ parametric tests due to the fact that the parametric tests assumptions

are often violated in the process of comparing learning models [32]. The Friedman test has a null hypothesis which highlights those differences existing among rankings of the ten baseline models is neither accidental nor influenced by their significant levels. When the hypothesis of the test is rejected, the post hoc Nemenyi test [32] can be administered for paired comparisons.

The Friedman statistics is calculated as follows:

$$\chi_F^2 = \frac{12N}{K(K+1)}\left[\sum_j AR_J^2 - \frac{k(k+k+1)2}{4}\right] \tag{9}$$

The variable k represents the number of models while N represents the number of datasets. AR_j^2 represents average rankings of the j-th model over all the datasets. When a significant average difference among the models is identified, it results in the rejection of the null hypothesis of the Friedman test for a given evaluation measure.

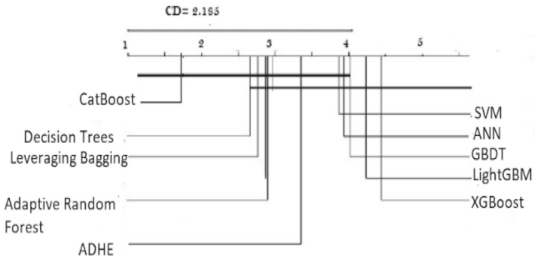


Fig. 1. The CD diagram with the results of the Nemenyi Test

Figure 1 shows the results of the significance test. The groups of classification models connected by a line segment do not significantly differ. The ADHE model's performance is significantly superior to Adaptive Random Forest, Leveraging Bagging, Decision Trees and CatBoost.

6 Conclusion

Credit scoring models capable of accurately distinguishing between creditworthy and non-creditworthy customers for emerging economies. Ensembles of classifiers have gained a lot of popularity in the domain of credit scoring due to their superior performance. A few studies in the literature have considered an adaptive and dynamic heterogeneous ensemble credit score model since customer variables tend to drift. This paper proposed a novel adaptive and dynamic heterogeneous ensemble composed of base models generated from two base learners. To validate the efficacy and performance of ADHE, empirical experiments were conducted to assess the generalization performance. Four real world datasets are used across six metrics to evaluate ADHE and the benchmarks. Prediction results obtained indicate that our approach notably outperforms the majority of homogeneous and heterogeneous ensemble benchmark models.

Regarding computational costs, our approach was ranked among the best a few resource-efficient benchmark models compared. The approach proposed can be

improved in future by including models generated from more base learners, In addition, different ensemble selection strategies can be explored to optimize prediction performance of credit scoring models.

Acknowledgement. All resources used mobilized by the author. This work was not funded by an individual or organization.

Declaration of Competing Interests. The author declares that he has no known competing financial interest or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Caldarelli, A., Fiondella, C., Maffei, M., Zagaria, C.: Managing risk in credit cooperative banks: lessons from a case study. Manage. Account. Res. **32**, 1–15 (2016)
2. Frame, W.S., Srinivasan, A., Woosley, L.: The effect of credit scoring on small business lending. J. Money Credit Banking 813–825 (2001)
3. Crook, J.N., Edelman, D.B., Thomas, L.C.: Recent developments in consumer credit risk assessment. Eur. J. Oper. Res. **183**, 1447–1465 (2007)
4. Rudin, C., Wagstaff, K.L.: Machine learning for science and society. Mach. Learn. **95**(1), 1–9 (2014)
5. Wang, G., Hao, J., Ma, J., Jiang, H.: A comparative assessment of ensemble learning for credit scoring. Expert syst. Appl. **38**(1), 223–230 (2011)
6. Tsai, C.F.: Combining cluster analysis with cluster ensembles to predict financial distress. Inf. Fusion **16**, 46–58 (2014)
7. Xiao, H., Xiao, Z., Wang, Y.: Ensemble classification based on supervised clustering for credit scoring. Appl. Soft Comput. **43**, 73–86 (2016)
8. He, H., Zhang, W., Zhang, S.: A novel ensemble method for credit scoring: adaption of different imbalance ratios. Expert Syst. Appl. **98**, 105–117 (2018)
9. Yao, J., Zhongyi Wang, L., Wang, M.L., Jiang, H., Chen, Y.: A novel hybrid ensemble credit scoring model with stacking based noise detection and weight assignment. Expert Syst. Appl. **198**, 15 (2022)
10. Singh, I., Kumar, N., Srinivasa, K.G., Maini, S., Ahuja, U., Jain, S.: A multi-level classification and modified PSO clustering based ensemble approach for credit scoring. Appl. Soft Comput. **111**, 107687 (2021)
11. Hou, W.H., Wang, X.K., Zhang, H.Y., Wang, J.Q., Li, L.: A novel dynamic ensemble selection classifier for an imbalanced dataset: an application for credit risk assessment. Knowl. Based Syst. **208**, 106462 (2020)
12. Nalic, J., Martinovia, G., Zagar, D.: New hybrid data mining model for credit scoring based on feature selection algorithm and ensemble classifiers. Adv. Eng. Inform. **45**, 101130 (2020)
13. Xia, Y., Zhao, J., He, L., Li, Y., Niu, M.: A novel tree-based dynamic heterogeneous ensemble method for credit scoring. Expert Syst. Appl. **159**, 113615 (2020)
14. Barddal, J.P., Loezer, L., Enembreck, F., Lanzuolo, R.: Lessons learned from data stream classification applied to credit scoring. Expert Syst. Appl. **162**, 113899 (2020)
15. Jin Xiao, X., Zhong, Z.Y., Xie, L., Xin, G., Liu, D.: Cost-Sensitive semi-supervised selective ensemble model for customer credit scoring. Knowl. Based Syst. **189**, 15 (2020)
16. Chen, X., Li, S., Xu, X., Meng, F., Cao, W.: A novel GSCI-based ensemble approach for credit scoring. IEEE Access **8**, 222449–222465 (2020)

17. Chen, T., Guestrin, C.: XGBoost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 785–794. ACM (2016)
18. Wang, S.X., Dong, P.F., Tian, Y.J.: A novel method of statistical line loss estimation for distribution feeders based on feeder cluster and modified XGBoost. Energies **10**(12), 2067 (2017)
19. Chawla, N.V., Bowyer, K.W., Hall, L.O., Kegelmeyer, W.P.: SMOTE: synthetic minority over-sampling technique. J. Artif. Intell. Res. **16**, 321–357 (2002)
20. Xia, Y., Li, C., Liu, N.: A boosted decision tree approach using Bayesian hyper-parameter optimization for credit scoring. Expert Syst. Appl. **78**, 225–241 (2017)
21. Chang, K.H., Chu, H.H., Tong, L.I.: Establish decision tree-based short term default credit risk assessment models. Commun. Stat. Theory Methods **45**(23), 6803–6815 (2016)
22. Engelbrecht, A.P.: Computational Intelligence: An introduction. Wiley, Chichester (2002)
23. Kennedy, J., Eberhart, R.: Particle swarm optimization. In: Proceedings of IEEE International Conference on Neural Networks, Perth, Australia, 27 November 27–1 December 1995, vol. 4 pp. 1942–1948 (1995)
24. Cruz, R.M., Sabourin, R., Cavalcanti, G.D.: META-DES oracle: meta-learning and feature selection for dynamic ensemble selection. Inf. Fusion **38**, 84–103 (2017)
25. Minku, L.L., Yao, X.: DDD: a new ensemble approach for dealing with concept drift. IEEE Trans. Knowl. Data Eng. **24**(4), 619–633 (2012)
26. Yang, L.: Classifier selection for ensemble learning based on accuracy and diversity. Proc. Eng. **15**, 4266–4270 (2011)
27. Yule, G.: On the association of attributes in statistics. Philos. Trans. Roy. Soc. London Ser. A **194**, 257–319 (1900)
28. Avery, R.B., Calem, P.S., Canner, G.B.: Consumer credit scoring: do situational circumstances matter? J. Bank. Finance **28**, 835–856 (2004)
29. Asuncion, A., Newman, D.: UCI Machine Learning Repository. Publishing (2007)
30. Zhang, W., He, H., Zhang, S.: A novel multi-stage hybrid model with enhanced multi-population niche genetic algorithm: an application in credit scoring. Expert Syst. Appl. **121**, 221–232 (2019)
31. Lessmann, S., Baesens, B., Seow, H.V., Thomas, L.C.: Benchmarking state-of-the-art classification algorithms for credit scoring: an update of research. Eur. J. Oper. Res. **247**, 124–136 (2015)
32. Demsar, J.: Statistical comparison of classifiers over multiple data sets. J. Mach. Learn. Res. **7**, 1–30 (2006)

An Ensemble Model Based on Learning Vector Quantization Algorithms for Early Detection of Cassava Diseases Using Spectral Data

Emmanuel Ahishakiye[1,2(✉)], Waweru Mwangi[2], Petronilla Murithi[2], Ruth Wario[3], Fredrick Kanobe[4], and Taremwa Danison[4]

[1] Department of Networks, Data Science and Artificial Intelligence, Kyambogo University, Kampala, Uganda
ahishema@gmail.com
[2] Department of Computing, Jomo Kenyatta University of Agriculture and Technology, Nairobi, Kenya
[3] Department of Informatics, University of Free State, Bloemfontein, South Africa
[4] Department of Computer Science, Kyambogo University, Kampala, Uganda

Abstract. In Sub-Saharan Africa, cassava is the second most significant food crop after maize. Cassava brown streak disease (CBSD) and cassava mosaic virus disease (CMD) combined account for nearly 90% of productivity losses. Automating the detection and classification of crop diseases could help professionals diagnose diseases more accurately and allow farmers in remote locations to monitor their crops without the help of specialists. Machine learning algorithms have been used in the early detection and classification of crop diseases. Previous research has used plant image data captured with smartphones. However, disease symptoms must be observable to use this strategy (using image data). Unfortunately, once symptoms appear on the aerial part of the plant, the root, which is the edible part of the plant, is destroyed. In this study, we used spectral data in a three-class classification challenge for diagnosing cassava diseases. We propose an ensemble model based on Generalized Learning Vector Quantization (GLVQ), Generalized Matrix LVQ (GMLVQ), and Local Generalized Matrix LVQ (LGMLVQ). Experimental results revealed that the LGMLVQ model had the best overall performance on the precision, recall, and F1-score followed by our proposed ensemble model, the GMLVQ model performed third, and finally GLVQ model. Also, using an accuracy performance metric, LGMLVQ had overfitting issues even though it had the highest accuracy of 100%, followed by our proposed ensemble model with an accuracy of 82%, and then the third in performance was the GMLVQ model with an accuracy of 74% and the least performed model on accuracy was GLVQ model with an accuracy of 56%.

Keywords: Ensemble model · Crop diseases · Machine learning · LVQ

1 Introduction

Early disease identification in crops is critical for food security, especially in Sub-Saharan Africa. In Sub-Saharan Africa, cassava is the second most significant food crop after

P. Ndayizigamiye et al. (Eds.): IDIA 2022, CCIS 1774, pp. 320–328, 2023.
https://doi.org/10.1007/978-3-031-28472-4_20

maize [1]. Because of its persistence in severe locations, as well as its tolerance for extreme environmental stress conditions and poor soils, cassava continues to acquire importance in Africa as a staple food eaten by more than 500 million people every day. However, various pests and diseases, including Cassava Bacterial Blight (CBB), Cassava Brown Steak Disease (CBSD), Cassava Green Mite (CGM), and Cassava Mosaic Disease (CMD), pose a significant danger to crop productivity [1]. However, the two major diseases that affect cassava production are cassava brown streak disease (CBSD) and cassava mosaic virus disease (CMD). In Sub-Saharan Africa, CBSD and CMD combined account for nearly 90% of productivity losses [2]. To identify cassava disease, government professionals visit various sections of the country and visually score the plants by looking for disease indicators on the leaves [3] which is notoriously subjective and it is not uncommon for specialists to differ on a plant's diagnosis. Automating the detection and classification of crop diseases could help professionals diagnose diseases more accurately and allow farmers in remote locations to monitor their crops without the help of specialists [3].

Machine learning algorithms have been used in the early detection and classification of crop diseases [4–7]. In this study, we used spectral data [8] in a three-class classification challenge for diagnosing cassava diseases. Previous research has used plant image data captured with smartphones [9–12]. However, disease symptoms must be observable to use this strategy (using image data) [8]. Unfortunately, once symptoms appear on the aerial part of the plant, the root, which is the edible part of the plant, is destroyed. By nature, this spectral data used is exceptionally high-dimensional, with more than 3600 characteristics or dimensions [8]. Conventional principal component analysis (PCA) [13] was used as a dimension reduction strategy. Motivated by the current developments and many influential studies in automated disease detection and classification using machine learning techniques, this study proposes an Ensemble Model based on Learning Vector Quantization Algorithms in the Classification of Cassava Diseases using Spectral Data. By developing an ensemble model that may improve the crop disease detection accuracy that is critical for decision-making by farmers, this study is formulated within the research area of Artificial Intelligence for Development (AI4D) and is aimed at contributing to United Nations Sustainable Development Goals (SDGs) number two, "End hunger, achieve food security and improved nutrition and promote sustainable agriculture".

1.1 Objectives of This Study

The novelty of this work is in the development of an ensemble model. An ensemble model is necessary for two reasons. The first step is to lessen bias. A base model's restricted hypothesis space could result in a significant bias. A more expressive ensemble may be less biased than the base models. More functions can be realized by an ensemble of n linear classifiers at a cost of only n times as much computation, which is frequently preferable to enabling a completely generic hypothesis space that might necessitate exponentially more computation. The second justification is to reduce variance. Think about a group of $K = 5$ classifiers that were combined by majority voting. At least three out of the five classifiers must misclassify a new example for the ensemble to do so.

This is supposedly less frequent than a single incorrect classification made by a single classifier.

2 Overview of Learning Vector Quantization

Learning Vector Quantization (LVQ) is a family of statistical pattern classification techniques that tries to learn prototypes (codebook vectors) that represent class regions. Hyperplanes between prototypes are used to define the class regions, resulting in Voronoi partitions [14]. Initially introduced in 1986, Learning Vector Quantization (LVQ) is a family of prototype-based supervised classification methods [15]. The main benefit of LVQ is the simplicity of interpreting the trained model, and it has been used in many practical applications. To achieve faster convergence or better generalization behavior, several modifications to the original LVQ method have been proposed in the literature. Prototypes that represent the classes in the given data are determined by the algorithms during the training phase. Based on the prototypes' receptive fields, predictions regarding novel samples are produced. To put it another way, a novel sample is classified by calculating its distance from all prototypes and labeling it with the closest prototype. Different approaches can be used to compute the prototypes and define their receptive fields [16]. The prototypes' ability to be interpreted in feature space is a significant characteristic of the LVQ family. The prototypes can, however, be converted into the original space under specific circumstances even if the feature space is challenging and impossible to comprehend. Along with the prototypes, the GMLVQ and LGMLVQ variations utilize and build a "relevance matrix" that can be used to produce low-dimensional discriminant visualizations [16]. In Generalized Learning Vector Quantization (GLVQ), the cost function is chosen in a manner that ensures the learning rule that is produced meets the convergence requirement [17]. The first LVQ learning rules were heuristic and vulnerable to startup issues, sluggish convergence issues, and instability. However, two primary strategies have been put out for creating explicit cost functions from which learning rules can be derived using the steepest descent or ascent, as well as for addressing the issue of convergence of the original LVQ algorithms [14].

2.1 Related Work

To identify cassava diseases, the study [8] classified spectral data using Generalized Matrix Relevance Learning Vector Quantization (GMLVQ) and compared it to traditional classification methods in the space of expansion coefficients. Experimental results revealed that although GMLVQ and Linear Discriminant Analysis (LDA) are conceptually similar, the LVQ technique offers better flexibility in terms of extensions concerning the number of prototypes per class and the usage of local relevance matrices. The study [18] proposed a novel framework for distance-based classification known as divergence-based learning vector quantization (DLVQ). The study also evaluated the use of DLVQ with conventional Euclidean distances for various parameter values. The Wisconsin Breast Cancer dataset and cassava image data showed that DLVQ can produce improved classification accuracies and Receiver Operating Characteristics. The study [19] used a Convolutional Neural Network (CNN) model and Learning Vector Quantization (LVQ)

algorithm-based method for tomato leaf disease detection and classification. Experimental results on healthy and diseased leaf images revealed that the proposed method successfully categorizes four different types of tomato leaf diseases.

3 Materials and Methods

3.1 Dataset Used

Spectral data [8] was obtained from the leaves of cassava plants in two conditions: when they were healthy and when they were infected with the two diseases CBSD and CMD, both of which had visible symptoms. A CI-710 miniature leaf spectrometer was used to collect this data (CID BioScience Inc 2010). The device is powered by USB from a device (such as a tablet or laptop), allowing the setup to be mobile and collect data in the field. The device is clamped onto a leaf of a certain plant to gather data, and the profile of the amount of light absorbed or reflected is captured on the device as a spectrogram for each position of the clamped leaf. The intensity and structure of the spectra are influenced by a variety of environmental conditions, with illumination being particularly relevant. As a result, data was gathered in the field under identical lighting conditions. Data was collected for plants aged 6 to 9 months from Nase 3, Nase 4, Nase 14, Nase 19, Alado Alado, Magana, Orera, and NAROCass 2 cassava types. Three plants were selected for each variety, and three leaves were chosen for each plant. On each leaf lobe, two spectral measurements were taken: one on the best (least affected/non-symptomatic) section and one on the worst (most affected/symptomatic) section (see Fig. 1). Because the spectrometer only measures a small area of the plant, around 2cm in diameter, readings for each leaf lobe were taken to ensure a representative and reliable sample. The dataset was collected in Uganda with the guidance of agricultural and bio-chemical experts from the National Crops Resources Research Institute (NaCRRI)[1]. The National Agricultural Research and Development Research Institute (NaCRRI) is the Ugandan government's entity in charge of agricultural research. A total of 1656 data points were collected

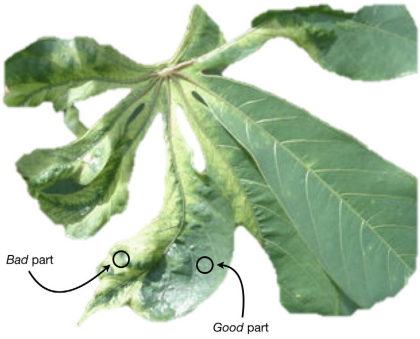


Fig. 1. Depiction of asymptomatic (good) and symptomatic (bad) parts of the plant. Adapted from [8].

[1] www.naro.go.ug

across three classes: healthy, CMD, and CBSD. For details, refer to [8]. Figure 2 shows an example of a dataset for each class of cassava leaves.

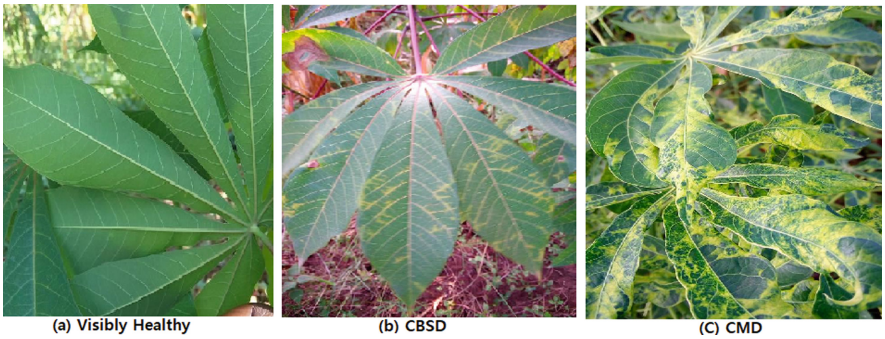


Fig. 2. Cassava image data with 3 categories used in this study, Adopted from [8].

3.2 The Proposed Model

In this study, we propose an ensemble model by aggregating three Learning Vector Quantization (LVQ) algorithms as the base models. The LVQ algorithms considered are Generalized Learning Vector Quantization (GLVQ) [17], Generalized Matrix LVQ (GMLVQ) [20], and Local Generalized Matrix LVQ (LGMLVQ) [20]. The goal of ensemble learning is to choose a collection, or ensemble, of hypotheses, such as h_1, h_2, and h_n, and aggregate their predictions using methods like voting, averaging, or a higher level of machine learning [21]. In this study, we used majority voting [22] to build our ensemble. It works by first building two or more separate models from your training data. When required to make predictions for new data, a Voting Classifier can be used to wrap your models and average the predictions of the sub-models. We did several experiments with different parameter values and the optimal results were obtained with default parameters. We maintained the default parameters and their values for GLVQ, GMLVQ, and LGMLVQ in all our experiments.

3.3 Model Implementation

During this study, experiments were done in google colab[2]. Also, a windows 10 laptop, Intel Core i7 processor of 2.50 GHz, and 16 GB RAM with 2 GB of NVIDIA GeForce MX150 graphical processing unit (GPU) was used. The models were implemented using the python sklvq package [16], an open-source Python implementation of a set of learning vector quantization (LVQ) algorithms.

[2] https://colab.research.google.com/?utm_source=scs-index.

3.4 Performance Metrics Used

In this study, we evaluated the performance of our proposed model using accuracy, precision, recall, and F1-score as shown in Eqs. 1–4.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

$$F1 - Score = \frac{2 * Recall * Precision}{Recall + Precision} \tag{4}$$

4 Results and Discussion

In this section, we discuss the performance of our proposed ensemble model together with Generalized Learning Vector Quantization (GLVQ), Generalized Matrix LVQ (GMLVQ), and Local Generalized Matrix LVQ (LGMLVQ). The experimental results revealed that our proposed model performed well on precision, recall, and F1-score as shown in Table 1 below. However, the LGMLVQ model had the overall performance of all the algorithms considered during this study as shown in Table 2. Also, the GMLVQ model performed third in relation to the algorithms under study (Table 3) and finally GLVQ model had the least performance (Table 4). On accuracy, LGMLVQ had the highest accuracy of 100%, followed by our proposed ensemble model with an accuracy of 82% (Table 5). The third in performance was the GMLVQ model with an accuracy of 74% and the least performed model on accuracy was the GLVQ model with an accuracy of 56% (Table 5). The next step of this study is to compare the proposed model with deep learning models so that we can get a final model for implementation on a mobile device so that it may help farmers in the early detection of cassava diseases.

Table 1. Performance of the Ensemble Model

	Precision	Recall	F1-score
Healthy	0.88	0.82	0.85
CBSD	0.73	0.84	0.78
CMD	0.85	0.78	0.81

Table 2. Performance of the LGMLVQ Model

	Precision	Recall	F1-score
Healthy	1.00	1.00	1.00
CBSD	1.00	1.00	1.00
CMD	1.00	1.00	1.00

Table 3. Performance of the GMLVQ Model

	Precision	Recall	F1-score
Healthy	0.83	0.78	0.80
CBSD	0.66	0.74	0.69
CMD	0.75	0.70	0.73

Table 4. Performance of the GLVQ Model

	Precision	Recall	F1-score
Healthy	0.80	0.33	0.47
CBSD	0.49	0.73	0.58
CMD	0.56	0.62	0.59

Table 5. Performance of the Models on Accuracy

	GMLVQ	GLVQ	LGMLVQ	PROPOSED
Accuracy	74.00%	56.00%	100.00%	82.00%

5 Conclusion

In this study, we used spectral data in a three-class classification challenge for diagnosing cassava diseases. Previous research has used plant image data captured with smartphones. However, disease symptoms must be observable to use this strategy (using image data). Unfortunately, once symptoms appear on the aerial part of the plant, the root, which is the edible part of the plant, is destroyed. We propose an ensemble model based on Learning Vector Quantization (LVQ) algorithms. The proposed model is a hybrid of Generalized Learning Vector Quantization (GLVQ), Generalized Matrix LVQ (GMLVQ), and Local Generalized Matrix LVQ (LGMLVQ). We maintained the default parameters and their values for GLVQ, GMLVQ, and LGMLVQ in all our experiments. Experimental results revealed that the LGMLVQ model had the best overall performance on the accuracy, precision, recall, and F1-score followed by our proposed ensemble model, and the GMLVQ

model performed third and finally GLVQ model. However, more studies need to be done to determine why the LGMLVQ algorithm had overfitting issues and how to address them. The next step of this study is to compare the proposed model with deep learning models so that we can get a final model for implementation on a mobile device so that it may help farmers in the early detection of cassava diseases.

Acknowledgments. This work was supported by Kyambogo University. We are grateful for their financial support.

References

1. Owomugisha, G.: Computational intelligence & modeling crop disease data in Africa. University of Groningen (2020)
2. Zeyimo, B., et al.: Attempts to identify cassava brown streak virus in western democratic Republic of Congo. J. Agric. Sci. **11**(2), 31 (2019). https://doi.org/10.5539/jas.v11n2p31
3. Mwebaze, E., Gebru, T., Frome, A., Nsumba, S., Tusubira, J.: iCassava 2019 Fine-Grained Visual Categorization Challenge (2019). arXiv:1908.02900v2, http://arxiv.org/abs/1908.02900
4. Liu, J., Wang, X.: Plant diseases and pests detection based on deep learning: a review. Plant Methods **17**(1), 1–18 (2021). https://doi.org/10.1186/s13007-021-00722-9
5. Goyal, L., Sharma, C.M., Singh, A., Singh, P.K.: Leaf and spike wheat disease detection & classification using an improved deep convolutional architecture. Inform. Med. Unlocked **25**(April), 100642 (2021). https://doi.org/10.1016/j.imu.2021.100642
6. Rangarajan, A.K., Purushothaman, R., Ramesh, A.: Tomato crop disease classification using pre-trained deep learning algorithm. Procedia Comput. Sci. **133**, 1040–1047 (2018). https://doi.org/10.1016/j.procs.2018.07.070
7. Mohanty, S.P., Hughes, D.P., Salathé, M.: Using deep learning for image-based plant disease detection. Front. Plant Sci. **7**(September), 1–10 (2016). https://doi.org/10.3389/fpls.2016.01419
8. Owomugisha, G., Melchert, F., Mwebaze, E., Quinn, J.A., Biehl, M.: Matrix relevance learning from spectral data for diagnosing cassava diseases. IEEE Access **9**, 83355–83363 (2021). https://doi.org/10.1109/ACCESS.2021.3087231
9. Abayomi-Alli, O.O., Damaševičius, R., Misra, S., Maskeliūnas, R.: Cassava disease recognition from low-quality images using enhanced data augmentation model and deep learning. Expert Syst. **38**(7), 1–21 (2021). https://doi.org/10.1111/exsy.12746
10. Lilhore, U.K., et al.: Enhanced convolutional neural network model for cassava leaf disease identification and classification. Mathematics **10**(4) (2022). https://doi.org/10.3390/math10040580
11. Ramcharan, A., Baranowski, K., McCloskey, P., Ahmed, B., Legg, J., Hughes, D.P.: Deep learning for image-based cassava disease detection. Front. Plant Sci. **8**(October), 1–7 (2017). https://doi.org/10.3389/fpls.2017.01852
12. Ramcharan, A., et al.: A mobile-based deep learning model for cassava disease diagnosis. Front. Plant Sci. **10**(March), 1–8 (2019). https://doi.org/10.3389/fpls.2019.00272
13. Reddy, G.T., et al.: Analysis of dimensionality reduction techniques on big data. IEEE Access **8**, 54776–54788 (2020). https://doi.org/10.1109/ACCESS.2020.2980942
14. Nova, D., Estévez, P.A.: A review of learning vector quantization classifiers. Neural Comput. Appl. **25**(3–4), 511–524 (2013). https://doi.org/10.1007/s00521-013-1535-3

15. Kohonen, T.: Learning Vector Quantization for Pattern Recognition (ReportTKK-F-A601), Helsinki, Finland (1986)
16. van Veen, R., Biehl, M., de Vries, G.J.: sklvq: scikit learning vector quantization. J. Mach. Learn. Res. **22**, 1–6 (2021)
17. Sato, A., Yamada, K.: Generalized learning vector quantization. Adv. Neural Inf. Process. Syst. 423–429 (1996)
18. Mwebaze, E., et al.: Divergence-based classification in learning vector quantization. Neurocomputing **74**(9), 1429–1435 (2011). https://doi.org/10.1016/j.neucom.2010.10.016
19. Sardogan, M., Tuncer, A., Ozen, Y.: Plant leaf disease detection and classification based on CNN with LVQ algorithm. In: UBMK 2018–3rd International Conference on Computer Science and Engineering, pp. 382–385 (2018). https://doi.org/10.1109/UBMK.2018.8566635
20. Schneider, P., Biehl, M., Hammer, B.: Adaptive relevance matrices in learning vector quantization. Neural Comput. **21**(12), 3532–3561 (2009). https://doi.org/10.1162/neco.2009.11-08-908
21. Russell, S., Norvig, P.: Artificial Intelligence: A Modern Approach, 4th ed., vol. 48. Pearson Education Limited, London (2022)
22. Brownlee, J.: Machine Learning Mastery With Python: Understand Your Data, Create Accurate Models and Work Projects End-To-End (2016)

Author Index

P. Ndayizigamiye et al. (Eds.): IDIA 2022, CCIS 1774, pp. 329–330, 2023.
https://doi.org/10.1007/978-3-031-28472-4