THE AFRICAN JOURNAL OF INFORMATION AND COMMUNICATION (AJIC)

ISSUE 30, 2022

Published by the LINK Centre
University of the Witwatersrand (Wits)
Johannesburg, South Africa
https://www.wits.ac.za/linkcentre

ISSN 2077-7213 (online version)
ISSN 2077-7205 (print version)

The African Journal of Information and Communication (AJIC)

Issue 30, 2022

Published by the LINK Centre, School of Literature, Language and Media (SLLM), Faculty of Humanities, University of the Witwatersrand (Wits), Johannesburg, South Africa
https://ajic.wits.ac.za

The African Journal of Information and Communication (AJIC) is a peer-reviewed, interdisciplinary, open access academic journal focused on the myriad dimensions of electronic and digital ecosystems that facilitate information, communication, innovation and transformation in African economies and in the broader Global South. Accredited by the South African Department of Higher Education and Training (DHET), *AJIC* publishes online, free to the user, under a Creative Commons licence, and does not impose article processing charges. *AJIC* is indexed in Scientific Electronic Library Online (SciELO) SA, the Directory of Open Access Journals (DOAJ), Sabinet African Journals and Wits University WIReDSpace, and is hosted on the Academy of Science of South Africa (ASSAf) Khulisa Journals platform.

Editors

Managing Editor: Tawana Kupe, Vice-Chancellor, University of Pretoria, tawana.kupe@up.ac.za
Corresponding Editor: Lucienne Abrahams, Director, LINK Centre, University of the Witwatersrand, PO Box 601, Wits 2050, Johannesburg, South Africa, ajic.submissions@gmail.com
Publishing Editor: Chris Armstrong, Research Associate, LINK Centre, University of the Witwatersrand, Johannesburg, South Africa, chris.armstrong@wits.ac.za

Peer-Reviewing
AJIC acknowledges with gratitude the following peer reviewers of submissions published in this issue: Lucienne Abrahams, Colin Chibaya, Edmore Chindenga, Marcel Dunaiski, Barry Dwolatzky, Sebnem Er, Shafika Isaacs, Gillian Marcelle, Uche Mbanaso, Sunday Oladejo, Gregory Ugbo and Temple Uwalaka.

Production
Sub-editing: LINK Centre
Proofreading: Linda Van de Vijver
Desktop-publishing: LINK Centre

AJIC is published by the LINK Centre, School of Literature, Language and Media (SLLM), Faculty of Humanities, University of the Witwatersrand (Wits), PO Box 601, Wits 2050, Johannesburg, South Africa. The LINK Centre is based at the Wits Tshimologong Digital Innovation Precinct, 41 Juta Street, Braamfontein, Johannesburg, https://www.tshimologong.joburg

ISSN 2077-7213 (online version)
ISSN 2077-7205 (print version)

Past issues of *AJIC*, and its precursor *The Southern African Journal of Information and Communication (SAJIC)*, are available at https://ajic.wits.ac.za/issue/archive and https://www.wits.ac.za/linkcentre/sajic

CONTENTS

RESEARCH ARTICLES

Roles played by Nigerian YouTube micro-celebrities during the COVID-19 pandemic

Aje-Ori Agbese
Associate Professor, Department of Communication, University of Texas Rio Grande Valley, Edinburg, Texas
https://orcid.org/0000-0003-4783-9113

Abstract
In 2020, during the COVID-19 pandemic, Nigerian social media micro-celebrities were prominent players in the dissemination of information. This study examines the roles that one group of Nigerian micro-celebrities, YouTube video bloggers (vloggers)—also known as "YouTubers"—played during the pandemic. The research analysed the contents of COVID-19-themed videos that 15 popular Nigerian YouTubers posted on their channels between 29 February and 5 August 2020. The study was guided by the two-step flow of communication theory, in terms of which information first flows from mass media to opinion leaders, who then, in the second step, share the information with their audiences. The study found that all 15 YouTubers played positive roles as opinion leaders—by providing health and safety information on COVID-19, challenging myths, and educating audiences through entertainment. Only two of the YouTubers studied were found to have shared some information that misinformed their audiences about the virus and how to fight it. The study therefore concluded that Nigerian YouTubers, as opinion leaders, can be important allies to governments and organisations when health crises arise in the country.

Keywords
COVID-19, communication, social media, micro-celebrities, YouTubers, opinion leaders, two-step flow of communication theory, Nigeria

Acknowledgement
The author thanks Dr Matthew Heinz for invaluable feedback during the writing of this article.

DOI: https://doi.org/10.23962/ajic.i30.13910

Recommended citation
Agbese, A. (2022). Roles played by Nigerian YouTube micro-celebrities during the COVID-19 pandemic. *The African Journal of Information and Communication (AJIC), 30*, 1-20. https://doi.org/10.23962/ajic.i30.13910

1. Introduction

In times of crisis, social media channels are often the "initial source of information" when news breaks (Wohn & Bowe, 2016, p. 1). For example, when the Zika and Ebola pandemics started, YouTube saw a "tremendous surge in viewer traffic" (Bora et al., 2018, p. 321). Those who rush to social media in times of crisis do so believing that the information posted there is valid and trustworthy (Cuomo et al., 2020). This high level of trust that most social media users have in the information they receive through these channels makes it necessary to explore the roles that influential social media, led by their most prominent users, can have during a pandemic.

Such exploration is acutely necessary because social media spread large quantities of both false and true information (Tangwa & Munung, 2020). Moreover, there is evidence to suggest that false information spreads "significantly farther, faster, deeper and more broadly than truth" on social media (Vosoughi et al., 2018, p. 2). It has been found, for instance, that during the Zika epidemic from 2015 to 2016, misleading posts on Facebook were more popular than accurate ones (Bora et al., 2018). Within weeks of the outbreak of COVID-19, fearmongering, misinformation, and conspiracy theories regarding the virus were rife on social media (Depoux et al., 2020). Accordingly, social media can escalate public fear and undermine public health efforts because they have enormous influence on their audiences' actions, beliefs, and interests (Mookadam et al., 2019). Therefore, social media's influence during the COVID-19 pandemic requires close examination, especially because lockdowns increased the amount of time that people spent on social media.

A prominent element of social media's influence is its micro-celebrities—users who have achieved celebrity status through social media (Senft, 2008; Kostygina et al., 2020). Unlike traditional celebrities who achieve their fame through traditional media (e.g., movies, music, or television), micro-celebrities' fame comes from self-produced content and providing direct and frequent intimate access to their lives (Seo & Hyun, 2018). They reach many people quickly because followers are instantly notified about new posts. The more followers that a micro-celebrity has, the greater their influence (Chung & Cho, 2017).

For example, research has determined that popular micro-celebrities have considerable influence on their followers' choices and decisions (Abidin, 2015). Kirkpatrick et al. (2018) found that micro-celebrities' product recommendations yielded 11 times more profit than other forms of advertising. Schouten et al. (2020) found that people trusted micro-celebrities more than traditional celebrities when choosing celebrity-endorsed products. Therefore, companies regularly use micro-celebrities in their

marketing. Popular micro-celebrities also receive sponsorships and are paid for product endorsements or using products on their channels. However, there is a paucity of research on micro-celebrities' roles beyond advertising and marketing (Kostygina et al., 2020).

Accordingly, this study examined the roles played by micro-celebrities in a health context. Specifically, the research examined the roles that popular Nigerian YouTube video bloggers (vloggers)—also known as "YouTubers"—played during the early months of the COVID-19 pandemic in the country, between 29 February and 5 August 2020. The study focused on YouTubers because YouTube is "an important vehicle for sharing and disseminating timely health-related information, both in its function as a repository of videos and as a social networking interface where users can interact and socialize" (Madathill et al., 2015, p. 174). When COVID-19 caused global lockdowns, millions of people turned to YouTube to satisfy their need for quarantine information, self-care information, and entertainment (YouTube, 2020a). YouTube is also an important source because popular YouTubers develop relationships with their viewers and influence them (Senft, 2008).

Another reason for this study's focus on Nigerian YouTubers is the prominence of Nigeria's celebrity culture. In 2019, it was estimated that 62% of Nigerians were online and highly active on social media; that about 53% of Nigerian internet users visited YouTube daily; and that many were content creators (Udodiong, 2019). Some Nigerian YouTubers, such as Mark Angel, Dimma Umeh, and Taaooma Akpaogi, were found to have become celebrities and achieved global fame (Oludimu, 2019).

The research analysed the contents of 56 COVID-19-themed videos posted by 15 popular Nigerian YouTubers. The analysis of the findings was guided by the two-step flow of communication theory, in terms of which information first flows from mass media to opinion leaders, and then, in the second step, from the opinion leaders to their audiences.

2. Literature review

Understanding micro-celebrity on YouTube

The YouTube video-sharing platform, launched in 2005, is currently used by approximately 2.5 billion people worldwide (Kemp, 2022). Its popularity lies in "its user-generated content, which includes tutorials, reviews, reactions, pranks, confessionals, and much more" (Miller, 2017, p. 3). YouTubers, described by Jerslev (2016, p.5233) as "video bloggers (vloggers) who regularly post videos on their personal YouTube channels", speak directly to audiences on niche subjects through the camera, and broadcast from private environments such as kitchens, living rooms, and bedrooms.

According to Jerslev (2016, p. 5238), being a YouTuber requires "continuous and multiple uploads of performances of a private self" and the use of "access, immediacy, and instantaneity" to build intimacy. This means that, unlike traditional celebrities who guard their privacy and separate their private and public lives, YouTubers blur that boundary and give audiences constant access to their private lives (Marwick, 2015). In addition, while glamour and extraordinariness characterise traditional celebrities, ordinariness, closeness, and equality characterise YouTubers. YouTubers must also quickly respond to comments, sometimes in a video, to maintain a positive relationship with their followers (Song, 2018).

In this micro-celebrity world, the number of likes, comments, shares, and subscribers that a YouTuber gains determines their success. Consequently, successful YouTubers can become influential figures whom people consult for information, entertainment, and recommendations (Abidin, 2015). For example, Sobande (2017) found that Black women in Britain relied on popular natural hair YouTubers for hair tips and product recommendations. Coates et al. (2020) found that children who watched their favorite YouTubers eating unhealthy snacks increasingly ate unhealthy snacks.

Research suggests that a YouTuber's credibility is tied to their perceived authenticity and closeness (Jerslev, 2016). Authenticity can be tied to a perception that a YouTuber is real and free from corporate control (Salyer & Weiss, 2020). According to Baker and Rojek (2019), authenticity is a valuable tool on YouTube because the platform's identity as an uncommercialised do-it-yourself space where ordinary people can freely express themselves requires genuineness. Therefore, audiences expect authenticity and honesty from YouTubers.

YouTubers can express authenticity in several ways. These include saying and showing that they are accessible, spontaneous, ordinary, and always themselves. They can also document real issues, share intimate information, and suggest that they and their audiences are alike (Jerslev, 2016). Furthermore, YouTubers can build authenticity through intimate conversations. Salyer and Weiss (2019) and Tolbert and Drogos (2019) found that people regarded their favourite YouTubers as friends when the YouTubers were perceived as authentic. And it was found that such YouTubers were particularly influential among their subscribers.

However, Marwick and boyd (2011, p. 124) point out that authenticity does not have a universal definition because what people regard as authentic depends on "the person doing the judging". Therefore, a YouTuber must find a balance between "personal

authenticity and audience expectations" in order to appeal to, gain, and maintain subscribers (Marwick & boyd, 2011, p. 127). YouTubers can quickly lose followers when they are seen as inauthentic for any reason (Baker & Rojek, 2019).

Another important trait of successful YouTubers is closeness. According to Salyer and Weiss (2020), closeness means audiences feel connected to a YouTuber. Similar to how they build authenticity, YouTubers can create an "impression of connectedness" by providing continuous updates on their lives, being relatable, and seeking input from their viewers (Jerslev, 2016, p. 5241). Lifestyle YouTubers, for example, build closeness "by presenting themselves as friends and equals" (Baker & Rojek, 2019, p. 4).

A disadvantage of closeness, however, is that it can create "parasocial" relationships—in which audience members feel a false sense of connection or intimacy with the YouTuber—that make the audience members highly susceptible to doing what a YouTuber asks (Tolbert & Drogos, 2019). Social media "are especially potent in establishing parasocial relationships of trust and intimacy" because they are structured and presented as "a direct exchange between equals" (Baker & Rojek, 2019, p. 9). Niu et al. (2021) found that this parasocial structure was particularly strong for YouTube audiences during COVID-19 lockdowns because YouTubers met people's need for human connection.

In addition to authenticity and closeness, research suggests YouTubers are influential when they are perceived as relatable (in appearance and in the information that they provide), inspiring, sincere, attractive, informal, experienced yet ordinary (imperfect), and sharing similar demographic characteristics with subscribers (Djafarova & Trofimenko, 2019; Smith, 2017). YouTube audiences' choices of whom to watch or follow are also contextual, because people choose channels based on what they need at a particular moment (Marwick & boyd, 2011).

Popular YouTubers' extensive reach and influence make them potentially important sources of leadership roles in certain situations (Senft, 2008). For example, rates of loneliness and depression increased in the United States during COVID-19-related lockdowns (Rosenberg et al., 2021). Consequently, US YouTubers participated in social media's #StayHome #WithMe (SHWM) movement, helping people to cope and connect with others by posting entertaining and comforting content that reduced people's stress and diverted their attention from pandemic-related stressors (Niu et al., 2021). Sofian (2020) found that five popular Indonesian YouTubers raised public awareness about COVID-19 to counter false information when the Indonesian government did not.

Theoretical framework

This study applied Katz and Lazarsfeld's (1955) theory of a two-step flow of communication to the contemporary social media context. This theory holds that information first flows from the media to opinion leaders and then, in the second step, to a less involved public. Katz and Lazarsfeld (1955) argued that opinion leaders were casual but influential acquaintances, friends, and family who could shape their peers' attitudes and behaviours through interpersonal, face-to-face communication. They were also well-connected, strongly exposed to media, and associated with bringing new innovations to the community. For Hameed and Sawicka (2017), opinion leaders are people "who have a greater-than-average share of influence within their community" (2017, p. 36).

However, Bennett and Manheim (2006, p. 215) challenge the relevance of Katz and Lazarsfeld's conception of communication flow in the contemporary context, arguing that "the combination of social isolation, communication channel fragmentation, and message targeting technologies have produced a very different information recipient" from the 1950s. They argue that people are now less likely to congregate in groups to receive information, and that social media have made face-to-face communication less prevalent, creating a one-step flow of information (without opinion leaders).

However, many authors still see the relevance of two-step conceptions. Starbird and Palen (2012) argue that rather than removing opinion leaders, social media have provided new opportunities for opinion leaders to exert their influence—thus maintaining the existence of a two-step flow of information. Winter and Neubaum (2016) point to the power of social media in the hands of opinion leaders, stating that such media provide "an ideal venue for influencing others" (2016, p. 2). Schäfer and Taddicken's (2015) study on German internet users identifies pockets of opinion leaders and a framework resembling Katz and Lazarsfeld's conception of two-step communication flow.

Bergström and Jervelycke Belfrage (2018) found that opinion leaders on social media are those who bring attention to, and add context to, certain news items, and thus people perceive them as crucial news providers. Hansen et al. (2011, p. 23) find that bloggers are influential opinion leaders because they can "build audiences that rival pre-digital media and challenge more established information providers." Turcotte et al. (2015) find that people increasingly trust news outlets that opinion leaders endorse on social media. News-sharing, whereby people share "information that is already available elsewhere" and make it "personally relevant to their social network",

also suggests a two-step flow of communication (Oeldorf-Hirsch & Sundar, 2015, p. 241). In another study, Velasquez (2012) discovers that expertise cues from popular social media figures generate the greatest feedback in only public discussions. Zimmermann et al. (2020) find that YouTubers who cite sources gain greater perceived credibility.

At the same time, there is evidence that opinion leaders on social media can "amplify the effects of disinformation" when they do not verify information or simply echo what others have said (Dubois et al., 2020, p. 8). In the Nigerian context, social media influencers have been found to share conspiracy theories and misinformation to grow audiences (Hassan, 2020).

This study was grounded in the assumption that conceptions of a two-step flow of communication are still relevant today, and focused on the following research questions:

- What types of content did Nigerian YouTubers create and share concerning COVID-19 during the pandemic?
- What role(s) did Nigerian YouTubers play during the COVID-19 pandemic in Nigeria?
- To what extent did Nigerian YouTubers provide false or misleading information on COVID-19?

3. Research design

To find Nigerian YouTubers' videos on COVID-19, a search for "Nigerian YouTubers" (key term) was conducted on YouTube. This produced 1,020 people, whose number of subscribers ranged from just seven to over 6 million.

To be included in the study, a YouTuber had to: be a Nigerian in Nigeria (YouTubers indicate their nationality and locations in their profiles); be an individual (not a duo or group); have at least 30,000 subscribers; and have posted content on COVID-19. These criteria produced 15 people. Combined, the 15 YouTubers selected for the study had, as of 7 August 2020, more than 2.9 million subscribers and over 360 million views (see Table 1).

These 15 individuals' YouTube channels were then searched for videos they had posted on COVID-19 between 29 February 2020 (when Nigeria's first COVID-19 case was announced) and 5 August 2020. Altogether, 56 COVID-19-themed videos were found (see Table 1). The least viewed of these videos had, as of 7 August 2020, been watched 2,775 times, and the most viewed had been watched 1.2 million times.

Table 1: YouTubers' number of subscribers and number of COVID-themed videos

YouTuber	No. of subscribers	No. of COVID-themed videos posted
Tomi's Colour Pavilion	686,000	2
MC Shem Comedian	388,000	5
Dimma Umeh	355,000	3
SisiYemmieTV	348,000	5
Taaooma Apaokagi Maryam	212,000	3
Maraji's World	191,000	4
Layefa Beauty	171,000	2
Mr. Macaroni (Debo Adebayo)	142,000	9
Tayo Aina	141,000	1
Kelechi Mgbemena	81,500	1
Tolulope Solutions	54,300	2
Nelo Okeke	50,900	3
Dyna Ekwueme	50,300	8
Vivian Okezie	47,500	3
Em Etetim	31,100	5
Totals	2,949,600	56

Following the identification of COVID-19 videos, each video was reviewed twice and evaluated qualitatively, with a focus on the roles being played by the YouTubers.

4. Findings

The YouTubers' characteristics and styles

The 15 YouTubers all lived in three large Nigerian cities: 12 in Lagos, two in Port Harcourt, and one in Abuja. The YouTubers comprised three males (MC Shem, Tayo Aina, and Mr. Macaroni) and 12 females. In terms of content, four of the channels—those of two of the males (MC Shem and Mr. Macaroni) and two of the females (Maraji and Taaooma)—consisted primarily of comedy. Male YouTuber Tayo Aina's channel focused on travel and real estate; female YouTuber Tomi focused on lifestyle and natural health remedies, and female YouTuber SisiYemmie's channel focused on food and lifestyle. The other five channels, all run by females, were focused on lifestyle and beauty.

It was found that the YouTubers employed several authenticity and closeness techniques to build intimacy, including (in the case of the females) showing their faces without makeup, speaking directly to viewers as if they were friends and family, including family and friends in videos, and sharing private information.

The YouTubers also built intimacy through their locations, shooting their videos in personal spaces such as cars, bedrooms, kitchens, and living rooms. In several cases, the YouTubers recorded themselves while they engaged in an activity, such as running an errand, attending a party, visiting a hair salon, or speaking to friends.

The female YouTubers were found to be more likely than the males to use closeness techniques to connect with their audiences, including intimate conversations and using emotions to build closeness; crying when discussing personal problems in their relationships or health; frequent updates for viewers; and encouragement of feedback. One YouTuber, Ekwueme, identified her subscribers as "Dynamites" as a means to build closeness.

Four YouTubers (males MC Shem and Mr. Macaroni, and females Maraji and Taaooma) played fictional characters as part of their aforementioned emphasis on comedy. These characters also tended to use memorable catchphrases to build familiarity (Mr. Macaroni's "you are doing well", for example). MC Shem and Mr. Macaroni never switched out of their fictitious characters, while Maraji and Taaoma did, infrequently, post videos on their personal lives.

The YouTubers' roles

The qualitative analysis of the roles played by the 15 YouTubers across the 56 videos identified three main themes: (1) YouTubers as information providers, (2) YouTubers as myth-busters, and (3) YouTubers as entertainers.

YouTubers as information providers

The YouTubers informed audiences on coronavirus and its impact. The lifestyle YouTubers acted like reporters and provided perspective on what was happening. In a variety of ways, they shared news and discussed the lockdowns, coping strategies, stocking up on necessities, hygiene, masks, and COVID-19 symptoms. One way was through promotions, which they used to introduce new innovations to Nigerians. For example, while giving viewers coping tips in the "spirit of quarantine", Ekwueme (2020a) promoted *Naija Lyfe*, an entertainment app, and encouraged people to use it while staying at home. Macaroni, during his live shows, promoted VBank, a digital banking service for receiving or sending money—because people could not go to the bank.

The YouTubers also participated in news sharing. Five YouTubers shared information from CNN, Al Jazeera, the *Los Angeles Times*, Ghanaian YouTuber Wode Maya, and Twitter. Umeh's (2020) 19 April video entitled "Can we talk about this???!!!" used news screenshots from CNN, Al Jazeera, and the *Los Angeles Times* to highlight and contextualise the maltreatment of Black Africans in China because of the coronavirus. She discussed how upset she was that the pandemic had taken a racist turn that blamed Africans in China for COVID-19. She wanted more people talking

about it because "It makes no sense. It hurts too much, and just makes you question so many things" (Umeh, 2020). CNN also played in the background in four people's videos, and Shem incorporated CNN in three skits. Interestingly, only Ekwueme cited a local media source, Instablog Naija. However, she did this to counter some information the site provided on Rivers State's lockdown (she lives there). This suggests she used her position as a YouTuber to challenge media discourse by sharing her arguments and position. Health sources also appeared in the videos, including a medical doctor in Macaroni's #Luckdownmillionaire, the Nigerian Centre for Disease Control (NCDC), UNICEF, and the World Health Organisation.

Another way in which the YouTubers shared information was by documenting what they were doing. The lifestyle YouTubers, in particular, showed themselves social distancing, wearing masks, and using hand sanitiser when they went out. They described how unusual it was to see their typically busy cities looking quiet and empty. Several YouTubers also highlighted some challenges that the pandemic and lockdown created in Nigeria. These challenges included loneliness, people not wearing masks or social distancing because of conspiracy theories, panic buying, and robberies. In a video entitled "Lagos lockdown/A day in my life/Social distancing???", SisiYemmie painted a dire picture of Lagos and its lockdown's effect on people's access to food. She called Lagos' lockdown "pointless" because people moved freely from 6 a.m. to 8 p.m. but were required to stay home from 8 p.m. to 6 a.m. (SisiYemmieTV, 2020). She questioned the logic in letting people go out at all. "Are we going to say, Coro, wherever you dey [are], behave yourself, to avoid catching the virus?" she asked (SisiYemmieTV, 2020).

Etetim also shared the difficulty that organisations faced in enforcing social distancing rules. In a video entitled "What life under quarantine in Nigeria really looked like", posted on 26 June, she showed a bank where people crowded the doors without masks and were not social distancing (Etetim, 2020). Although she showed a place where people followed health guidelines, she said that, overall, Nigerians were not taking coronavirus seriously "because the number of deaths isn't so alarming" (Etetim, 2020). She concluded: "Nigeria is not really a place where people are very open to being careful or listening to what people are saying" (Etetim, 2020).

In an 18-minute self-described "rant" entitled "Nigerians are wicked", Ekwueme (2020b) also discussed several coronavirus-related issues. She talked about Nigerians exploiting the pandemic in Rivers State. For example, although the State Governor announced a lockdown starting at 6 p.m. on 26 March, police closed the state's borders early, before 10 a.m., and started charging people 1,000 to 2,000 Naira (approx. USD2.5 to USD5) to enter the state. She also pointed to the fact that the cost of food and essential commodities had soared, stating that the cost of vegetables rose sharply on 26 March, resulting in panic buying and hoarding.

Ekwueme then talked about the challenge of staying home without the kind of support that people in the United States, Canada, and several other countries received from their governments. Requiring Nigerians to stay at home daily without food or financial aid was especially difficult for those who relied on daily incomes and had no savings. Ekwueme (2020b) said: "As citizens, we are entitled to salaries every month at least until this thing is done and dusted." She said that hunger would otherwise undermine the lockdown's purpose because unless people had money and food at home while isolating, they would go out "to fend for themselves. Lockdown won't happen in Nigeria. It will never happen if you're not providing the essentials for the people" (Ekwueme, 2020b). She added that, unlike other leaders who addressed their citizens about the pandemic at least once a day, Nigeria's President Muhammadu Buhari was "nowhere to be found", because he had not addressed the country. She described the government as "paralyzed", with "no clue whatsoever" on the pandemic (Ekwueme, 2020b). Coincidentally, Buhari addressed Nigeria on 29 March 2020. Ekwueme's rant video generated 618 comments, with many supporting her opinions and observations.

In a video skit entitled "My mother has corona virus (COVID-19)", Shem highlighted how Nigerians were maltreating and shunning those with coronavirus-like symptoms, such as sneezing and coughing, without proof of a positive test. Ekwueme (2020b) said the fear of discrimination probably dissuaded people from getting tested when they had COVID-like symptoms, and this would negatively affect the government's efforts.

The YouTubers also discussed personal issues related to the pandemic and how they were coping. They said that they had learned new skills, decluttered closets, bonded with family, exercised, and cooked. Aina said that he found it difficult to create, and felt unmotivated and lonely. Etetim shared these sentiments. She said that she first viewed the lockdown as "a mini-vacay. A break the world needed" that would last a week (Etetim, 2020). Aside from highlighting challenges, two YouTubers provided information to help people too. For those who did not want to go shopping, Umeh and Okeke shared WhatsApp numbers that people in Lagos and Port Harcourt could use for home-delivered groceries or curbside shopping at grocery stores, which were innovations in Nigeria too.

However, the only prevention and safety behaviours that the YouTubers emphasised were using hand sanitiser, wearing masks, no touching or hugging, isolating, and taking vitamin C. Fourteen YouTubers did not address covering your face when you sneezed or coughed, cleaning and disinfecting surfaces, and how long people should wash their hands for with soap. Overall, the videos under this theme confirmed Bergström and Jervelycke Belfrage's (2018) finding that social media leaders can bring attention to news that others missed and can also add context.

YouTubers as myth-busters

The second observed role involved busting COVID-19 myths. The YouTubers tackled several myths, including the myths that spreading onions around the house would kill the virus; that COVID-19 was like Ebola and would be eradicated quickly; that Uber Eats, TikTok, and Disney+ created COVID-19; and that Christians and people who ate starchy foods could not contract the virus. They countered the myths with facts and satire. To prove that Black people could get COVID-19 because the first case in Nigeria was an Italian man, Tomi showed Idris Elba's announcement of his positive COVID-19 test. Ekwueme also addressed the myth that COVID-19 only affected wealthy people. She said the myth was common because test kits were not widely available and only prominent people's deaths were announced on the news and social media. As a medical doctor's wife, she emphasised that anyone could get it. Surprisingly, none of the YouTubers addressed the popular Nigerian conspiracy that blamed mobile 5G networks for COVID-19 (Adebayo, 2020; Wonodi et al, 2022).

YouTubers as entertainers

Finally, the YouTubers provided entertainment through comedy skits. Here, the YouTubers showed Nigerians' ability to find a "comic dimension" in any issue (Afolayan, 2013, p.164). As Nigerians also view social media as a "laughing space" where they can still highlight societal issues, it was not surprising that these videos were the most viewed (123,237 to 1.2 million) (Yékú, 2016, p. 249). These numbers matched Niu et al.'s (2021) finding that people turned to YouTube for entertainment and distraction during the pandemic. Johnston's (2017) finding that comedy can increase viewership and engagement is also supported. The lifestyle YouTubers also employed comedic strategies such as blundering to make their videos fun. A popular strategy that the comedy YouTubers used was satire, which refers to using humour, ridicule, or exaggeration to expose and criticise people's depravities.

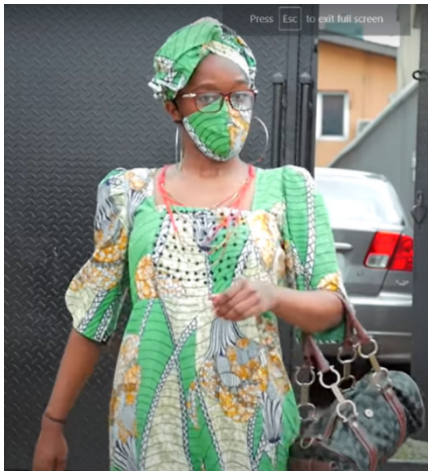
For example, Taaooma depicted a coronavirus-fighting soldier in a music video she posted on 18 April that probed Nigerians' resistance to compliance unless the government used force. In another skit, Macaroni (2020) satirically presented loneliness, robbery, fraud, and hunger as the "many children of coronavirus" in Nigeria because the government did not provide palliatives or prepare Nigerians. He also called COVID-19 the "hunger virus" (Macaroni, 2020). These were jabs at the economic and security problems that the pandemic created or heightened in Nigeria.

In two videos, Maraji also used satire. In a 28 March video, she satirically exposed the types of people (conspiracy theorists, newscasters, panicky, calm, indifferent, serious, and church lawbreakers) that the pandemic created. For example, the conspiracy theorists believed the virus was "planned work" and an "economic strategy" to raise prices and make money (Maraji's World, 2020). The church lawbreakers violated lockdown regulations and went to church because "coronavirus cannot hold us down. We are children of God. What will affect others cannot affect us" (Maraji's World,

2020). Her second video, entitled "Wearing masks in a pandemic", highlighted Nigerians' adaptability. The video started with frightening music and images of Chinese people wearing masks, and then switched to an upbeat Nigerian song that played in a fashion show where Maraji showed how Nigerians had made masks a fashion statement. Maraji sashayed out of the house in different clothes, for men and women of different ages, with matching face masks (Figure 1).

Shem also ridiculed the things that people used for masks in a skit entitled "Face mask". In it, his mother made a face mask using plastic bottles (Figure 2). The skit exposed and poked fun at Nigerians who used anything, including soap dishes, for face masks, an issue that memes also highlighted, and addressed the lack of information on correct face masks in Nigeria (Dynel, 2020).

Figure 1: Fashionable face mask



Source: Screenshot from Maraji's World (2020)

Figure 2: Mama Shem's face mask

Source: Screenshot from MC Shem Comedian (2020)

False or misleading information

To control misinformation, YouTube started using an automated system on 16 March 2020 to flag and remove misleading COVD-19 content (YouTube, 2020b). However, the system also removed safe COVID-19 videos and caused self-censorship on YouTube. To avoid getting flagged or removed, four of the YouTubers studied said that they could not say the word "coronavirus" and instead used the words "virus", "corona", or other nicknames such as "Coro", "rona", and "rororo". Three of the YouTubers spelt "coronavirus" as "corona virus" in video titles. But this self-censorship did not stop two YouTubers from sharing misleading information.

Tomi's vlog on 19 March 2020, entitled "Is This a Cure for Corona Virus? Find Out" promoted a COVID-19 cure. She shared a hairdryer method that she said cured COVID-19 patients in a London hospital. The method involved putting a hairdryer

on cool and blowing air around the face. She cited "reliable sources in London" and said the hairdryer method "has been helping a few patients get out of this" (Tomi's Colour, 2020). She said that she could not reveal the hospital or her sources for safety reasons and asked viewers to trust that the information was credible. However, the information that Tomi provided on the hairdryer method echoed a viral video that Facebook and YouTube removed in March 2020 for being false (Dunlop, 2020).

Ekwueme (2020b) also shared unsubstantiated information. She said poor Nigerians were "denied testing" and test kits were not available to the poor in Nigeria (Ekwueme, 2020b). However, she provided no sources to support the information. Ekwueme and Tomi's comments confirmed Dubois et al.'s (2020) finding that micro-celebrities can strengthen disinformation when they simply echo sentiments or do not verify information.

5. Discussion and conclusion

Guided by the two-step flow of communication theory, this study sought to examine the roles that 15 Nigerian YouTubers played during the COVID-19 pandemic and to fill the gap on Nigerian micro-celebrity research. The findings revealed that Nigerian YouTubers can be important sources of information. During the pandemic, they provided information, raised awareness, entertained, challenged myths, and acted as opinion leaders online. These YouTubers also participated in the #stayathome movement that encouraged people to remain indoors during lockdown.

This suggests that YouTubers played positive roles during the pandemic through social media. Several of this study's findings also matched previous studies on YouTubers. Like Western YouTubers, Nigerian YouTubers, particularly women, use authenticity and closeness to engage and interact with audiences (Miller, 2017; Salyer & Weiss, 2019). However, contrary to Jerslev's (2016) definition of a YouTuber as a vlogger, the data characterised Nigerian YouTubers as more than vloggers. A Nigerian YouTuber is more likely to be a content creator who can also attract an audience through fictional characters.

According to Marwick (2015), micro-celebrities sometimes adopt fake identities to hide their real identities in order to address the impossibility of maintaining a single identity and/or to target different audiences. This may explain why Nigerian YouTubers play fictional characters and combine content types. Therefore, anyone studying YouTubers must define them in ways that capture the unique characteristics and conditions that match their context.

This study also confirmed previous findings that opinion leaders exist in newer media (Choi, 2014). As opinion leaders, the 15 YouTubers understood that people would look to them for information, perspective, and entertainment during the lockdown. Therefore, they created content to meet those needs. They also interpreted and channelled information from news sources to their audiences through vlogs, comedic skits and more, opined on social issues, offered corrections without lecturing through comedy, shared innovations, and encouraged viewers to share their experiences and thoughts in the comments section. This supported Niu et al.'s (2021) finding that YouTubers helped people to cope and illustrated how YouTubers can lead public discussions on relevant issues in Nigeria (Grzywińska & Borden, 2012). The feedback that they generated from users could produce invaluable public perspective on social issues in Nigeria. As news sharers, they confirmed that news flows from the media to opinion leaders, who then share it with their followers (Oeldorf-Hirsch & Sundar, 2015).

However, two individuals shared false information, which confirmed Wonodi et al.'s (2022) finding that Nigerian social media was rife with falsehoods on COVID-19. Tomi's cure video was particularly risky because her channel, which had the most subscribers among the 15 YouTubers studied, shares natural health do-it-yourself remedies, which many Nigerians prefer over pharmaceuticals (Alabi et al., 2021). During the pandemic, 60% of Nigerians said that herbal medicine could successfully treat COVID-19, and 80% believed that they could not contract COVID-19 because they used herbal medicine diligently (Alabi et al., 2021).

Therefore, when YouTubers mislead their audiences they can negatively impact the medical advice and choices that people receive or make (Olapegba et al., 2020). When YouTubers do not verify information or rush to share what they find, they can become echo chambers for fake sources and can put people's lives at risk. Consequently, YouTubers must research and confirm the information and sources that they receive before sharing it.

Overall, despite using a small sample, the study found important information on Nigerian YouTubers. If another health crisis occurs in Nigeria, the Nigerian government and health organisations will benefit from including micro-celebrities in health campaigns to reach and educate people. Future studies could examine Nigerian YouTubers' influence from the audience's perspective. Studies could also include analysis of the content of the YouTube comments section, so as to better gauge audience engagement with the videos and messages.

References

Abidin, C. (2015). Communicative intimacies: Influencers and perceived interconnectedness. *Ada: Journal of Gender, New Media, and Technology*, *8*, 1–16. https://doi.org/10.7264/N3MW2FFG

Adebayo, B. (2020, May 19). UK regulator sanctions Nigerian Christian channel over 5G conspiracy theory claims. *CNN*. https://www.cnn.com/2020/05/19/africa/ofcom-sanctions-5g-conspiracy-theory-intl/index.html

Afolayan, A. (2013). Hilarity and the Nigerian condition. *The Journal of Pan African Studies*, *6*(5), 156–174.

Alabi, G.O., Dada, S. O., Adebodun, S. A., & Obi, O.C. (2021). Knowledge of COVID-19 and perception of Nigerians towards the use of herbal medicine in its treatment. *Nigerian Journal of Pharmaceutical Research*, *17*(2), 157–166. https://doi.org/10.4314/njpr.v17i2.2

Alexander, J. (2020, March 4). YouTube is demonetizing videos about coronavirus, and creators are mad. *The Verge*. https://www.theverge.com/2020/3/4/21164553/youtube-coronavirus-demonetization-sensitive-subjects-advertising-guidelines-revenue

Baker, S. A., & Rojek, C. (2019). The Belle Gibson scandal: The rise of lifestyle gurus as micro-celebrities in low-trust societies. *Journal of Sociology*, *56*(3), 388–404. https://doi.org/10.1177/1440783319846188

Bennett, W. L., & Manheim, J. B. (2006). The one-step flow of communication. *The Annals of the American Academy of Political and Social Science*, *608(1)*, 213–232. https://doi.org/10.1177/0002716206292266

Bergström, A. & Jervelycke Belfrage, M. (2018). News in social media: Incidental consumption and the role of opinion leaders. *Digital Journalism*, *6*(5), 583–598. https://doi.org/10.1080/21670811.2018.1423625

Bora, K., Das, D., Barman, B., & Borah, P. (2018). Are internet videos useful sources of information during global public health emergencies? A case study of YouTube videos during the 2015–16 Zika virus pandemic. *Pathogens and Global Health, 112*(6), 320–328. https://doi.org/10.1136/bmjgh-2020-002604

Choi, S. (2014). The two-step flow of communication in Twitter-based public forums. *Social Science Computer Review*, *33*(6), 696–711. https://doi.org/10.1177/0894439314556599

Chung, S., & Cho, H. (2017). Fostering parasocial relationships with celebrities on social media: Implications for celebrity endorsement. *Psychology & Marketing*, *34*(4), 481–495. https://doi.org/10.1002/mar.21001

Coates, A. E., Hardman, C. A., Halford, J. C. G., Christiansen, P., & Boyland, E. J. (2020). "It's just addictive people that make addictive videos": Children's understanding of and attitudes towards influencer marketing of food and beverages by YouTube video bloggers. *International Journal of Environmental Research and Public Health*, *17*, 449, 1–18. https://doi.org/10.3390/ijerph17020449

Cuomo, M. T., Tortora, D., Giodano, A., Festa, G., Metallo. G., & Martinelli, E. (2020). User-generated content in the era of digital well-being: A netnographic analysis in a healthcare marketing context. *Psychology & Marketing*, *37*, 578–587.

Depoux, A., Martin, S., Karafillakis, E., Preet, R., Wilder-Smith, A., & Larson, H. (2020). The pandemic of social media panic travels faster than the COVID-19 outbreak. *Journal of Travel Medicine*, *27*(3), 1–2. https://doi.org/10.1093/jtm/taaa031

Djafarova, E., & Trofimenko, O. (2019). Instafamous – credibility and self-presentation of micro-celebrities on social media. *Information, Communication & Society*, *22*(10), 1432–1446. https://doi.org/10.1080/1369118X.2018.1438491

Dubois, E., Minaeian, S., Paquet-Labelle, A., & Beaudry, S. (2020). Who to trust on social media: How opinion leaders and seekers avoid disinformation and echo chambers. *Social Media + Society*, *6*(2), 1–13. https://doi.org/10.1177/2056305120913993

Dunlop. W. G. (2020, March 19). *Hot air from saunas, hair dryers won't prevent or treat COVID-19*. AFP Fact Check. https://factcheck.afp.com/hot-air-saunas-hair-dryers-wont-prevent-or-treat-covid-19.

Dynel, M. (2020). COVID-19 memes going viral: On the multiple multimodal voices behind face masks. *Discourse & Society*, *32*(2), 1–21. https://doi.org/10.1177/0957926520970385

Ekwueme, D. (2020a, April 5). *Self-isolating in my lonely marriage*. [Video]. YouTube. https://www.youtube.com/watch?v=n6WiElLEJHs

Ekwueme, D. (2020b, March 29). *Nigerians are wicked*. [Video]. YouTube. https://www.youtube.com/watch?v=BebgbpMV10I

Etetim, E. (2020, June 26). *What life under quarantine in Nigeria really looked like*. [Video]. YouTube. https://www.youtube.com/watch?v=h2Hx70ZzQT4

Grzywińska, I., & Borden, J. (2012). The impact of social media on traditional media agenda setting theory: The case study of Occupy Wall Street movement in USA. In B. Dobek-Ostrowska, B. Lodzki, & W. Wanta (Eds.), *Agenda setting old and new problems in the old and new media* (pp. 133–155). University of Wroclaw Press.

Hameed, T., & Sawicka, B. (2017). The importance of opinion leaders in agricultural extension. *World Scientific News*, *76*, 35–41.

Hansen, D. L., Smith, M. A., & Schneiderman, B. (2011). *Analyzing social media networks with NodeXL: Insights from a connected world.* Morgan Kaufmann. https://doi.org/10.1016/B978-0-12-382229-1.00002-3

Hassan, I. (2020, March 27). COVID-19: The dual threat of a virus and a fake news epidemic. *Premium Times*. https://opinion.premiumtimesng.com/2020/03/27/covid-19-the-dual-threat-of-a-virus-and-a-fake-news-epidemic-by-idayat-hassan/?utm_source=headtopics&utm_medium=news&utm_campaign=2020-03-27

Jerslev, A. (2016). In the time of the microcelebrity: Celebrification and the YouTuber Zoella. *International Journal of Communication*, *10*, 5233–5251.

Johnson, K. F., Worth, A., & Brookover, D. (2019). Families facing the opioid crisis: Content and frame analysis of YouTube videos. *The Family Journal: Counseling and Therapy for Couples and Families*, *27*(2), 209–220. https://doi.org/10.1177/1066480719832507

Johnston, J. (2017). Subscribing to sex edutainment: Sex education, online video and YouTube star. *Television and New Media*, *18*(1), 76–92. https://doi.org/10.1177/1527476416644977

Katz, E., & Lazarsfeld, P. F. (1955). *Personal influence: The part played by people in the flow of mass communications*. Free Press.

Kemp, S. (2022, August 15). YouTube statistics and trends. *Datareportal*. https://datareportal.com/essential-youtube-stats

Keyton, J. (2011). *Communication research: Asking questions, finding answers* (3rd ed.). McGraw-Hill.

Kirkpatrick, N., Pederson J., & White, D. (2018). Sport business and marketing collaboration in higher education. *Journal of Hospitality, Leisure, Sport & Tourism Education*, *22*, 7–13. https://doi.org/10.1016/j.jhlste.2017.11.002

Kostygina, G., Tran, H., Binns, S., Szczypka, G., Emery, S., Vallone, D., & Hair, E. (2020). Boosting health campaign reach and engagement through use of social media influencers and memes. *Social Media + Society*, *6*(2), 1–12. https://doi.org/10.1177%2F2056305120912475

Madathill, K. C., Rivera-Rodriguez, A. J., Greenstein, J. S., & Gramopadhye, A. K. (2015). Healthcare information on YouTube: A systematic review. *Health Information Journal*, *21*(3), 173–194. https://doi.org/10.1177/1460458213512220

Maraji's World. (2020, March 28). *Different types of people now* [Video]. YouTube. https://www.youtube.com/watch?v=kRXUlFjImQ0

Marwick, A. E. (2015). You may know me from YouTube: (Micro-)celebrities in social media. In P. D. Marshall & S. Redmond (Eds.), *A companion to celebrity* (pp. 194–212). John Wiley & Sons. https://doi.org/10.1002/9781118475089.ch18

Marwick, A. E., & boyd, d. (2011). I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. *New Media & Society*, *13*(1), 114–133. https://doi.org/10.1177/1461444810365313

MC Shem Comedian. (2020, May 22). *Face mask* [Video]. YouTube. https://www.youtube.com/watch?v=fwFKqLj8D8g

Miller, B. (2017). YouTube as educator: A content analysis of issues, themes and educational value of transgender-created online videos. *Social Media + Society*, *3*(2), 1–12. https://doi.org/10.1177/2056305117716271

Mookadam, F., Oz, M., Siddiq, T. J., Almader-Douglas, D., Crupain, M., & Khan, M. S. (2019). Impact of unauthorized celebrity endorsements on cardiovascular healthcare. *Future Cardiology*, *15*(6), 387–390. https://doi.org/10.2217/fca-2019-0020

Mr. Macaroni. (2020, April 19). *Wahala in the society* [Video]. YouTube. https://www.youtube.com/watch?v=HyiEeUmMkTA.

Niu, S., Bartolome, A., Mai, C., & Ha, N. (2021). #StayHome #WithMe: How do YouTubers help with COVID-19 loneliness? In *CHI '21: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1–15). https://doi.org/10.1145/3411764.3445397

Oeldorf-Hirsch, A., & Sundar, S. S. (2015). Posting, commenting, and tagging: Effects of sharing news stories on Facebook. *Computers in Human Behavior*, *44*, 240–249. https://doi.org/10.1016/j.chb.2014.11.024

Olapegba, P. O., Ayandele, O., Kolawole, S. O., Oguntayo, R., Gandi, J. C., Dangiwa, A. L., Ottu, I. F. A., & Iorfa, S. K. (2020). A preliminary assessment of novel coronavirus (COVID-19) knowledge and perceptions in Nigeria. *medRxiv pre-print*. https://doi.org/10.1101/2020.04.11.20061408

Oludimu, T. O. (2019, March 27). Professional Nigerian YouTubers have to look outside YouTube to make money. *Techpoint Africa*. https://techpoint.africa/2019/03/27/ways-nigerian-youtubers-make-money/

Rosenberg, M., Luetke, M., Hensel, D., Kianersi, S., Fu, T., & Herbenick, D. (2021). Depression and loneliness during April 2020 COVID-19 restrictions in the United States, and their associations with frequency of social and sexual connections. *Social Psychiatry & Psychiatric Epidemiology, 56*, 1221–1232. https://doi.org/10.1007/s00127-020-02002-8

Salyer, A. M., & Weiss, J. K. (2020). Real close friends: The effects of perceived relationships with YouTube microcelebrities on compliance. *The Popular Culture Studies Journal*, *8*(1), 139–156.

Schäfer, M. S., & Taddicken, M. (2015). Mediatized opinion leaders: New patterns of opinion leadership in new media environments? *International Journal of Communication*, *9*, 960–981.

Schouten, A. P., Janssen, L., & Verspaget, M. (2020). Celebrity vs. influencer endorsements in advertising: The role of identification, credibility and product-endorser fit. *International Journal of Advertising*, *39*(2), 258–281. https://doi.org/10.1080/02650487.2019.1634898

Seo, M., & Hyun, K. D. (2018). The effects of following celebrities' lives via SNSs on life satisfaction: The palliative function of system justification and the moderating role of materialism. *New Media & Society*, *20*(9), 3479–3497. https://doi.org/10.1177/1461444817750002

Senft, T. M. (2008). *Camgirls: Celebrity and community in the age of social networks*. Peter Lang.

SisiYemmieTV. (2020, May 2). *Lagos lockdown/A day in my life/Social distancing???* [Video]. YouTube. https://www.youtube.com/watch?v=qIr1bSgs_2s

Smith, D. R. (2017). The tragedy of self in digitised popular culture: The existential consequences of digital fame on YouTube. *Qualitative Research*, *17*(6), 699–714. https://doi.org/10.1177/1468794117700709

Sobande, F. (2017). Watching me watching you: Black women in Britain on YouTube. *European Journal of Cultural Studies*, *20*(6), 655–671. https://doi.org/10.1177/1367549417733001

Sofian, F. A. (2020). YouTubers creativity in creating public awareness of COVID-19 in Indonesia: A YouTube content analysis. In *2020 International Conference on Information Management and Technology (ICIMTech)* (pp. 881–886). IEEE. https://doi.org/10.1109/ICIMTech50083.2020.9211149

Song, H. (2018). The making of microcelebrity: AfreecaTV and the younger generation in neoliberal South Korea. *Social Media + Society*, *4*(4), 1–10. https://doi.org/10.1177/2056305118814906

Starbird, K., & Palen, L. (2012). (How) will the revolution be retweeted? Information diffusion and the 2011 Egyptian uprising. In S. Poltrock, C. Simone, J. Grudin, G. Mark, & J. Riedl (Eds.), *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work* (pp. 7–16). ACM. https://doi.org/10.1145/2145204.2145212

Tangwa, G. B., & Munung, N. S. (2020). COVID-19: Africa's relation with epidemics and some imperative ethics considerations of the moment. *Research Ethics*, *16*((3-4), 1–11. https://doi.org/10.1177/1747016120937391

Tolbert, A. N., & Drogos, K. L. (2019). Tweens' wishful identification and parasocial relationships with YouTubers. *Frontiers in Psychology*, *10*, 1–15. https://doi.org/10.3389/fpsyg.2019.02781

Tomi's Colour Pavilion. (2020, March 19). *Is this a cure for corona virus? Find out* [Video]. YouTube. https://www.youtube.com/watch?v=z8gfCwr5eTI

Toromade, S. (2020, March 20). Timeline of coronavirus cases in Nigeria. *Pulse.* https://www.pulse.ng/news/local/coronavirus-timeline-and-profile-of-cases-in-nigeria/k9p6lbk

Turcotte, J., York, C., Irving, J., Scholl, R. M., & Pingree, R. J. (2015). News recommendations from social media opinion leaders: Effects on media trust and information seeking. *Journal of Computer-Mediated Communication*, *20(*5), 520–535. https://doi.org/10.1111/jcc4.12127

Udodiong, I. (2019, February 8). Here is how Nigerians are using the internet in 2019. *Pulse*. https://www.pulse.ng/bi/tech/how-nigerians-are-using-the-internet-in-2019/kz097rg

Umeh, D. (2020, April 19). *Can we talk about this???!!!* [Video]. YouTube. https://www.youtube.com/watch?v=GQ00RvYilNc.

Uzuegbunam, C. E. (2017). Between media celebrities and the youth: Exploring the impact the of emerging celebrity culture on the lifestyle of young Nigerians. *Mgbakoigba: Journal of African Studies*, *6*(2), 130–141.

Velasquez, A. (2012). Social media and online political discussion: The effect of cues and informational cascades on participation in online political communities. *New Media & Society*, *14*, 1286–1303. https://doi.org/10.1177/1461444812445877

Vosoughi, S., Roy, A., & Aral, D. (2018). The spread of true and false news online. *Science*, *359*(6380), 1146–1151. https://doi.org/10.1126/science.aap9559

Wegener, C., Prommer, E., & Linke, C. (2020). Gender representations on YouTube: The exclusion of female diversity. *M/C Journal, 23*(6). https://doi.org/10.5204/mcj.2728

Winter, S., & Neubaum, G. (2016). Examining characteristics of opinion leaders on social media: A motivational approach. *Social Media + Society*, *2*(3), 1–12. https://doi.org/10.1177/2056305116665858

Wohn, D. Y., & Bowe, B. J. (2016). Micro agenda setters: The effect of social media on young adults' exposure to and attitude toward news. *Social Media + Society*, *2*(1), 1–12. https://doi.org/10.1177/2056305115626750

Wonodi, C., Obi-Jeff, C., Adewumi, F., Keluo-Udeke, C., Gur-Aire, R., Krubiner, C., Jaffe, E. F., Bamiduro, T., Karron, R., & Faden, R. (2022). Conspiracy theories and misinformation about COVID-19 in Nigeria: Implications for vaccine demand generation communications. *Vaccine*, *40*, 2114–2121. https://doi.org/10.1016/j.vaccine.2022.02.005

Yékú, J. (2016). Akpos don come again: Nigerian cyberpop hero as trickster. *Journal of African Cultural Studies*, *28*(3), 245–261. https://doi.org/10.1080/13696815.2015.1069735

YouTube. (2020a, March 16). Protecting our extended workforce and community. https://blog.youtube/news-and-events/protecting-our-extended-workforce-and

YouTube. (2020b, June 25). YouTube during COVID-19. https://youtube.com/trends/articles/what-it-means-to-stayhome-on-youtube

Zimmermann, D., Noll, C., Gräßer, L., Hugger, K., Braun, L. M., Nowak, T., & Kaspar, K. (2020). Influencers on YouTube: A quantitative study on young people's use and perception of videos about political and societal topics. *Current Psychology*, *41,* 6808–6824. https://doi.org/10.1007/s12144-020-01164-7

A word embedding trained on South African news data

Martin Canaan Mafunda
Doctoral Candidate, Department of Physics, University of KwaZulu-Natal, Westville Campus, Durban
https://orcid.org/0000-0001-9008-5834

Maria Schuld
Senior Researcher and Software Developer, Xanadu Quantum Technologies, Toronto; and Researcher, University of KwaZulu-Natal, Westville Campus, Durban
https://orcid.org/0000-0001-8626-168X

Kevin Durrheim
Distinguished Professor, Department of Psychology, University of Johannesburg
https://orcid.org/0000-0003-2926-5953

Sindisiwe Mazibuko
Doctoral Candidate, Department of Psychology, University of KwaZulu-Natal, Pietermaritzburg Campus, South Africa
https://orcid.org/0000-0003-4376-4230

Abstract
This article presents results from a study that developed and tested a word embedding trained on a dataset of South African news articles. A word embedding is an algorithm-generated word representation that can be used to analyse the corpus of words that the embedding is trained on. The embedding on which this article is based was generated using the Word2Vec algorithm, which was trained on a dataset of 1.3 million African news articles published between January 2018 and March 2021, containing a vocabulary of approximately 124,000 unique words. The efficacy of this Word2Vec South African news embedding was then tested, and compared to the efficacy provided by the globally used GloVe algorithm. The testing of the local Word2Vec embedding showed that it performed well, with similar efficacy to that provided by GloVe. The South African news word embedding generated by this study is freely available for public use.

Keywords
natural language processing (NLP), word embedding, Word2Vec, GloVe, news data, South Africa

Acknowledgements
The authors thank Media Monitoring Africa (MMA) for making the data used for this project freely available. We are also grateful for support from the University of KwaZulu-Natal's Big Data and Informatics' Research Flagship, South Africa's National Research Foundation (NRF-Grant UID: 137755), and the South African Centre for Digital Language Resources (SADiLaR-Grant #OR-AAALV). SADiLaR is a national centre supported by the South African Department of Science and Innovation (DSI).

DOI: https://doi.org/10.23962/ajic.i30.13906

Recommended citation
Mafunda, M. C., Schuld, M., Durrheim, K., Mazibuko, S. (2022). A word embedding trained on South African news data. *The African Journal of Information and Communication (AJIC), 30*, 1-24. https://doi.org/10.23962/ajic.i30.13906

1. Introduction

Word embeddings are finding increasing use in the social sciences as tools to analyse social groups through the language they produce. They are computer models that use machine learning to develop representations of words as vectors or points in a high-dimensional space. The points are constructed so that relations between words, such as the use of two words in a similar semantic or grammatical context, can be measured as a distance between two points in the space. This gives rise to an "arithmetic of meaning".

The use of word embeddings as tools for studying culture and language is acknowledged as a new, emerging field of research (Arseniev-Koehler & Foster, 2020; Kozlowski et al., 2019). The ever-growing existence of large but "messy" pools of textual data harvested from social and traditional media is driving interest in word embeddings as key mechanisms for natural language processing (NLP). Word embeddings are used by companies such as Facebook to tag harmful posts, e.g., content written with the sole purpose of spreading false or misleading information on COVID-19 vaccination programmes. Badri et al. (2022) have demonstrated the role of word embeddings in text tagging or text detection. Their study uses fastText and GloVe word embeddings to detect offensive and hate speech in social media content.

The meaning captured by word embeddings is specific to the data that the machine-learning algorithm (model) is trained on. The development, training, and evaluation

of word embedding models must therefore be context-specific. Examples of word embeddings linked to a certain domain are: the NukeBERT model (Jain et al., 2020) that is trained on texts from the nuclear and atomic energy section; specialised embeddings for finance (Theil et al., 2020); and embeddings trained on certain languages, such as Setswana and Sepedi (Marivate et al., 2020) or Croatian (Svoboda & Beliga, 2017).

Even when a widely spoken language such as English is used in a dataset, geographic contexts will induce specific terms or relationships between words that are of critical importance to researchers in fields such as social and political sciences. For instance, political scientists have used word embeddings to recover rich knowledge, through semantic projections, about the behaviour of the main political parties in South Africa in respect of illegal foreign nationals (Grand et al., 2022). In addition, Durrheim et al. (2022) have demonstrated how word embeddings provide a useful tool to study cultural bias, showing that calculating the difference between two bipolar bias vectors (centroids) gives rise to another vector which represents a bias dimension. Other researchers have used the bias dimension to study stereotypes in word embeddings (Kozlowski et al., 2019).

The need to gather knowledge that is unique to a specific field or research area is what motivated us to carry out this study, which developed and evaluated a new word embedding trained on a large corpus of online South African news articles from outlets including *Daily Maverick*, *News24* and *Independent Online (IOL)*. The embedding was trained using Word2Vec's Skip-Gram algorithm (Mikolov et al., 2013), and the dataset used was provided by Media Monitoring Africa (MMA). The word embedding we generated is publicly available via a github repository.[1] It is, to the best of our knowledge, the first publicly available word embedding trained on South Africa news article data, and thus forms a valuable addition to the field of NLP in African contexts (Marivate et al., 2020). The embedding will allow researchers to investigate the meanings of numerous words from within a South African context and to seek answers to culturally or politically oriented South African research questions—such as, to give but one small example, how the African National Congress (ANC) and Democratic Alliance (DA) relate to terms such as "corruption" and "white monopoly capital".

This article introduces the word embedding and explains the choices we made in data preprocessing and in training of the Word2Vec algorithm that generated the embedding. We also present results from extensive validation testing of the embedding, and comparative testing between the performance of our locally generated Word2Vec

1 https://github.com/Mafunda/SouthAfricanNewsEmbeddings

embedding and an embedding generated by the internationally recognised GloVe algorithm. We conducted the comparison using 14 standard analogy benchmark tasks, and found that our local South African Word2Vec embedding scored very competitively with the GloVe embedding—and in some cases scored better.

Section 2 of this article describes the Word2Vec and GloVe algorithms; section 3 sets out approaches to evaluating word embeddings; section 4 describes the preparation and configuration of the dataset; section 5 describes the implementation, evaluation, and refinement of the word embedding; section 6 describes our work to maximise robustness of the embedding through determining variances and testing ensembles of embeddings; section 7 provides results from our comparative evaluation of the performance of our South African Word2Vec embedding against the performance of a GloVe embedding; section 8 provides findings from validation of our local embedding against South African benchmarks; and section 9 provides conclusions.

2. The Word2Vec and GloVe algorithms

Word2Vec

Word2Vec is a common algorithm for training word embeddings and is powered by the statistical power of neural network models. It was first introduced in 2013 by Tomas Mikolov and his research collaborators from Google. In our study, the Word2Vec algorithm was used to learn a word embedding from a South African news articles database.

This Word2Vec algorithm consists of two model architectures and two training methods. The two model architectures are Skip-Gram and CBOW (continuous bag of words), while the two training methods are the hierarchical softmax and negative sampling. The Skip-Gram model aims to predict context from a given word. Skip-Gram is slow, and good at learning infrequent words. On the other hand, the CBOW aims to predict a word from a given context of words. CBOW is fast, and is good at learning common words. The hierarchical softmax is good at training with infrequent words, and negative sampling is good at training with common words and low-dimension vectors.

Word2Vec is similar to other commonly used approaches for learning word embeddings such as GloVe (global vectors for word representation) (Pennington et al., 2014), BERT (bi-directional encoder representations from transformers) (Devlin et al., 2018), GPT (generative pre-trained transformer) (Radford et al., 2018), fastText (Bojanowski et al., 2017; Santos et al., 2017), and ELMo (embeddings from language model) (Peters et al., 2018), to name just a few. Since one goal of this study was to compare the performance of the Word2Vec and GloVe algorithms, we now briefly review the GloVe model.

GloVe (global vectors for word representation)

GloVe, like Word2Vec, is an unsupervised learning algorithm for generating word embeddings. According to the model's developers, Pennington et al. (2014), GloVe is a count-based, global log bilinear regression model that combines two embedding methods, namely global matrix factorisation and local context window. The model is based on the observation that the most appropriate starting point for word vector learning is the ratios of co-occurrence probabilities rather than the probabilities themselves. In other words, the GloVe model is built on the intuition that the ratios of co-occurrence probabilities among words potentially encode some kind of a relation among words.

3. Evaluation of word embeddings

To ensure that word embeddings are useful and can be deployed to solve downstream NLP tasks, the quality and reliability of a word embedding needs to be assured through validation tests. Several approaches to evaluating the quality of word embeddings have been reported. Bakarov (2018) divides the methods of evaluation into two categories, namely: (1) extrinsic; and (2) intrinsic.

According to Bakarov (2018), methods of *extrinsic evaluation* are anchored on the idea that every downstream NLP task is a form of word embedding evaluation. In other words, methods of extrinsic evaluation entail leveraging the potential of word embeddings to be used as feature or input vectors when training supervised machine-learning algorithms (like the maximum entropy model). Therefore, a rule of thumb for methods of extrinsic evaluation is that any downstream NLP task can be considered as an evaluation method, e.g., for the task of sentiment analysis, text classification, or part-of-speech tagging, to mention only a few (see Bakarov (2018) for more examples).

The methods of *intrinsic evaluation*, on the other hand, involve experiments which are designed to compare word embeddings with human judgments on word relations. This was of particular interest to our study because we made use of locally inspired analogy tasks—e.g., matching politicians to political parties—for model evaluation based on South African news article data. According to Bakarov (2018), methods of intrinsic evaluation are divided into four sub-categories: (1) methods of conscious evaluation; (2) methods of subconscious evaluation; (3) thesaurus-based methods; and (4) language-driven methods. In this study, we used methods of conscious evaluation to evaluate the South African news word embedding and therefore we now limit our discussion to describing those methods.

According to Bakarov (2018), the core methods of conscious evaluation are (1) word semantic similarity, (2) word analogy, (3) thematic fits, and (4) synonym detection. The *word semantic similarity* method is based on the idea that distances between words in an embedding space can be evaluated through the human heuristic judgments on the actual semantic distances between these words. For example, we would expect the distance between *cup* and *mug* defined by a number from the interval [0, 1] to be in the region of 0.8 since these words are nearly synonymous, that is, they are used similarly in language.

The *word analogy* method is the second most popular method for evaluating word embeddings (Bakarov, 2018). First introduced by Mikolov et al. (2013), word analogies are based on the idea that arithmetic operations in a word vector space can be predicted by humans. For instance, given a set of three words or word pairs—e.g., the two politicians "Julius Malema" and "Jacob Zuma", as well as the party "EFF" (Economic Freedom Fighters, founded by Malema)—the task would be to predict the word D such that the relation Julius_Malema : EFF is the same as the relation Jacob Zuma : D (Pereira et al., 2016; Turian et al., 2010). In this case, the target word would be "ANC" (African National Congress), which is the party of ex-President Jacob Zuma. Word analogies are also known as "analogical reasoning", "linguistic regularities", and/or "word semantic coherence". In this study, we used both word semantic similarity and word analogy methods to evaluate the quality of our South African news embedding.

4. Dataset preparation and configuration

Data

This study used a text corpus of 1,312,125 news articles, which were provided, upon request, by MMA from its news database. The text dataset consisted of news articles that were published between 1 January 2018 and 17 March 2021. It should be noted that the database was not in the public domain, and access was granted in response to our individual request.

Data preparation

Raw texts are by nature "noisy" and therefore require some text preprocessing before they can be used to train machine-learning algorithms such as the Word2Vec model. Text preprocessing for this study was done with the help of several open source

Python software packages, including the natural language toolkit (NLTK) (Loper & Bird, 2002), beautifulsoup (Richardson, 2007), and gensim (Řehůřek &, Sojka, 2011a). The sequence of preprocessing steps included: splitting documents (multi-sentences) into single sentences (also known as sentence tokenisation); removing all words containing single uppercase letters surrounded by lowercase letters in order to remove JavaScript; and converting all words to lowercase letters. Further, preprocessing included the removal of: html tags; expressions such as "\xad" and "displayad"; words that contained substrings ("windowtextcolor"), and punctuation and digits.

We did not remove stopwords, following a growing trend in the machine-learning literature. Rahimi and Homayounpour (2022) recommend the retention of stopwords when learning word representations for solving sentiment classification problems. This is because the removal of stopwords such as "no" and "don't" can potentially change the polarity of words in documents.

Data preparation also included the creation of n-grams (bigrams and trigrams) using the Phraser model of the gensim package. Bigrams are pairs of words that are repeatedly mentioned together in a given text corpus. For example, during our data preparation, words such as Jacob and Zuma were joined to produce a bigram Jacob_Zuma because they occurred together more than our determined minimum threshold of collocations. Similarly, we joined three words together into a trigram if they consecutively and consistently occurred together within the news articles corpus. For example, the word combination President Jacob Zuma was joined to produce a President_Jacob_Zuma trigram.

We conclude this section with an example of a "messy" text followed by its "clean" version after data preprocesing:

- *Before preprocessing:* PRESIDENT Jacob Zuma has declared a special official funeral for renowned author and poet, Prof. William Keorapetse Kgositsile, a renowned veteran activist and a giant of the liberation struggle who died on Wednesday.
- *After preprocessing:* president jacob zuma has declared a special official funeral for renowned author and poet prof william keorapetse kgositsile a renowned veteran activist and a giant of the liberation struggle who died on Wednesday.

Hyperparameter settings

Table 1 shows the hyperparameter names and values used to train the embedding (with Python's gensim package). As mentioned in section 2, the Word2Vec algorithm learns word embeddings using one of its two model architectures: Skip-Gram or CBOW. In our study, we used Skip-Gram. Also as mentioned above, Word2Vec uses two training methods to learn word embeddings: hierarchical softmax (Goodman, 2001) and negative sampling (Mikolov et al., 2013). We adopted negative sampling, and a hyperparameter negative value of 10. The role of the hyperparameter "negative" is to specify the number of "noise words" that the model is allowed to draw on during model training.

Table 1: Hyperparameter settings used to train the embedding

Parameter name	Value
minimum word count (m)	50
window size	10
architecture	Skip-Gram (s1)
training method	negative sampling (h0)
negative	10
vector dimension size (d)	250

For two of the hyperparameters—minimum word count and vector dimension—the hyperparameters seen in Table 1 (50 and 250, respectively)—were only finalised through experiments conducted on the initial embedding (see section 5).

5. Implementation, evaluation, and refinement of the word embedding

Implementation of the embedding

The gensim package (Řehůřek & Sojka, 2011b), implementable in the Python environment, was used to build and train the Word2Vec algorithm. The popularity and convenience of implementing the Word2Vec algorithm with gensim influenced our decision to select this implementation framework.

We used Google Colaboratory, an online environment for Python programming, to implement the Word2Vec algorithm with gensim. It took approximately eight hours to implement the embedding, starting from data preparation until the model finished training. Due to the large dataset size, we used a procedure in which data

was read in chunks of 10,000 sentences into a buffer holding 100,000 sentences, and after each read-in, the buffer was shuffled. This introduced a pseudo-randomness in which the first sentence in the corpus had a greater chance of being fed to the training procedure early on.

Performance evaluation measures

We measured the performance of the embedding using both "similarity" and "analogy" measures.

Similarity measure

The similarity measure probes the extent to which words are similar or dissimilar by measuring the distance between their respective vector representations in an embedding. More precisely, this measure typically uses the cosine similarity, or the size of the angle between two vectors belonging to any two given words, as a proxy for measuring the degree to which the two words are related. Given any two word vectors vec($word_1$) and vec($word_2$), where vec($word_i$) is the vector corresponding to a given word, the similarity value is computed as follows:

$$similarity(word_1, word_2) = \frac{vec(word_1)vec(word_2)}{||vec(word_1)||||vec(word_2)||} \quad \textbf{(1)}$$

Note that we normalised word vectors in the embedding, so that their norm is always 1. Equation (1) implies that highly similar words (or synonyms) have similarity values that are closer to 1, while highly dissimilar words have similarity values closer to –1. We used the WordSim353 dataset to evaluate our embedding. WordSim is a test dataset for measuring word similarity or relatedness (Agirre et al., 2009). The WordSim dataset consists of word pairs such as soccer and football, baseball and netball, etc. and their similarity scores. (The WordSim353 dataset is freely available for public use.[2])

Analogy measure

Analogy measurement tasks ask the embedding to predict the fourth word in a relational equation of the form "ANC is to Jacob_Zuma as EFF is to Julius_Malema". In a relational task, the model is given the first three words and asked to predict the fourth word that will solve the relational equation, i.e., Jacob_Zuma – ANC = [predict word] – EFF. A prediction is computed by retrieving the 10 nearest neighbours to the vector vec(Jacob_Zuma) – vec(ANC) + vec(EFF). If the correct word is found to be among these 10 nearest neighbours, the model is given a score of 1 (correct prediction). Otherwise, a score of 0 (incorrect prediction) is given. This

2 The WordSim353 dataset is at http://alfonseca.org/pubs/ws353simrel.tar.gz

method is commonly known as the accuracy@k method (Xu, 2018). It is called the accuracy@k method because the value of k is arbitrarily chosen and it measures the extent to which the model is penalised for producing k nearest neighbours.

The precision score for a set of analogies is then computed as follows:

$$\text{precision score} = \frac{\text{number of analogies correctly predicted by the model}}{\text{total number of tasks evaluated}} \times 100\% \quad \textbf{(2)}$$

For analogy measurement, we used the GloVe word analogy dataset. Publicly accessible via the GloVe website,[3] the dataset is made up of 14 analogy tasks, which are named as follows:

- capital-common-countries;
- capital-world;
- city-in-state;
- currency;
- family;
- adjective-to-adverb;
- opposite;
- comparative;
- superlative;
- present participle;
- nationality adjective;
- past tense;
- plural; and
- plural verbs.

Word analogies are relational equations of the form word1:word2::word3:word4 (translated verbally as word1 is to word2 as word3 is to word4). To restate, our goal in testing the South African news embedding's ability to solve analogy tasks was to measure how well the embedding predicted the fourth word ("word4").

In reporting the experiment results in this article, we use the following notations:

- "p" stands for %;
- "d" stands for dimension size; and
- "m" stands for "minimum word count".

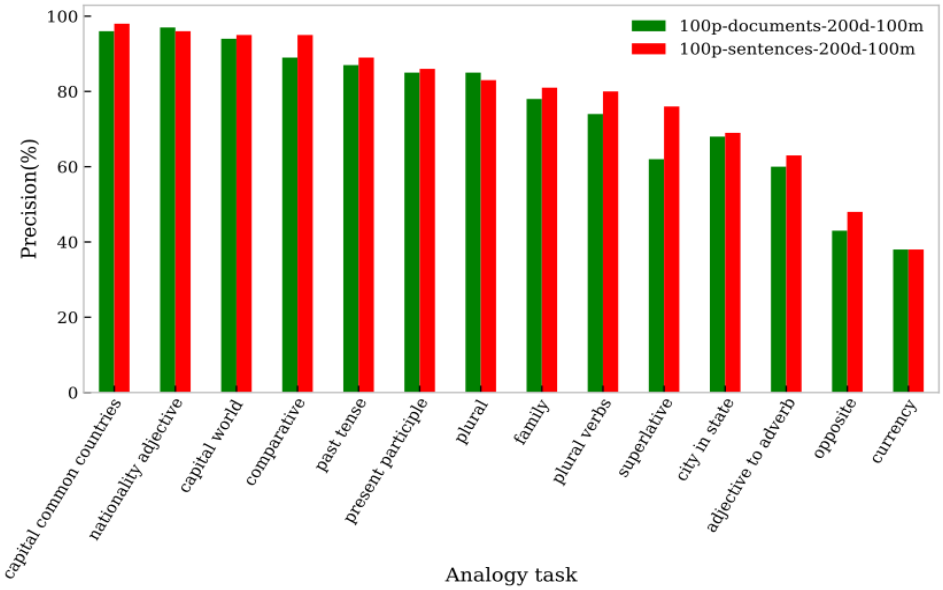
3 https://github.com/stanfordnlp/GloVe/tree/master/eval/question-data

For example, the notation "100p 100d 50m" denotes our word embedding trained on 100% (p) of the training dataset, with a word vector dimension size (d) of 100, and with words (tokens) with a minimum word count (m) of 50 (meaning that words not appearing 50 or more times were ignored during training). The reason for adopting 50 as the minimum word count is given below in the "determining a suitable minimum word count" sub-section.

Determining whether to train with sentences or documents

We conducted an experiment in order to determine whether the optimal training approach for our embedding was: (1) training based on data split into documents of news articles; or (2) training based on data split into sentences. As seen in Figure 1, we found that the precision of the word embedding trained on sentences was always the same or better than that of the word embedding trained on documents, with the sentence-contexts outperforming document-contexts in 11 (almost 80%) of the 14 analogy tasks. This finding was consistent with emerging best practices in the NLP literature (Gu et al., 2018).

Figure 1: Training with sentences versus documents (evaluated via 14 analogy tasks)



Refinement of hyperparameters through testing

As mentioned above, two of the hyperparameters could only be finalised once the embedding had been generated—allowing testing of the influence of different parameter settings on the embedding. We conducted experiments that measured the embedding's precision in conducting 14 analogy tasks when the value of a certain hyperparameter was varied.

Determining a suitable vector dimension size

We conducted a second experiment in order to determine the optimal dimension size, i.e., the size or dimensionality of the word vectors in the trained embedding. Identifying optimal dimensionality is important since it influences the space available to "encode meaning": a low dimension may result in under-fitting, a situation where there is not enough space to reflect the subtle levels of meaning, while a dimension that is too large may lead to model over-fitting, where all words are positioned far from each other and relational meaning is lost (see also Yin & Shen, 2018). To understand the impact of the vector dimension, we trained and compared four versions of our word embedding, with each version having the same training settings except for the dimension size, which was varied for the values 100, 200, 300 and 400.

Figure 2: Training with 4 different vector dimension sizes (evaluated via 14 analogy tasks)

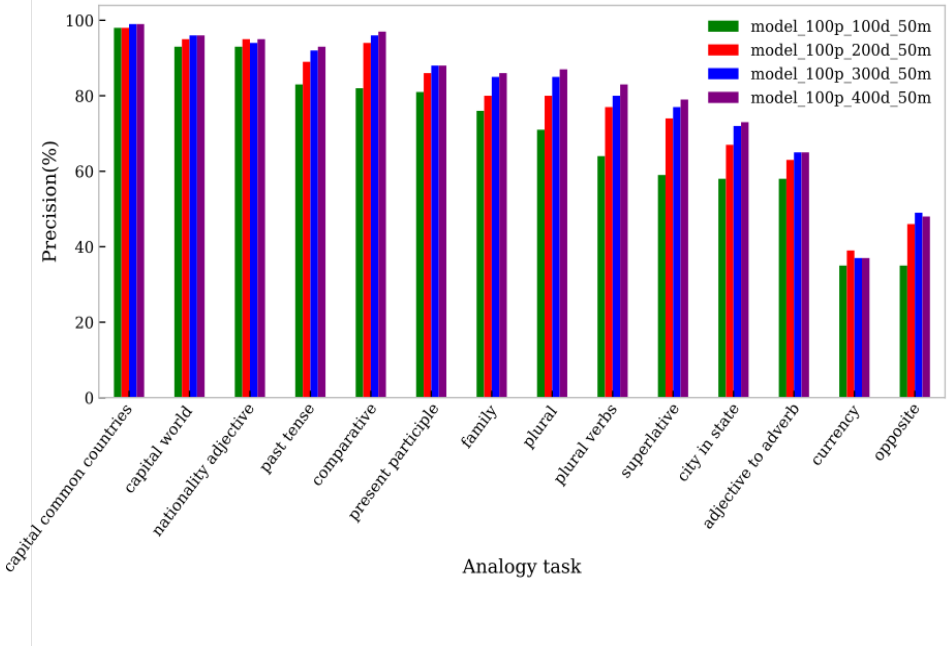


Figure 2 shows that while the precision consistently increased with higher dimensions, there was only a small improvement between 200 and 300, as well as a negligible improvement between 300 and 400. At the same time, due to the large vocabulary, 100 additional dimensions translated to 1.24×10^7 additional values (for a vocabulary of 124,000 words) that would have to be stored to describe the word vectors. We therefore decided to fix the vector dimension size at 250 in order to balance the physical size of the model with performance needs. This decision was justified on the grounds that we were reliant on the free version of Google's Colaboratory (Colab) platform to train and evaluate our models, and thus fixing the vector dimension size at 250 was necessary in order to reduce computational and time resources required

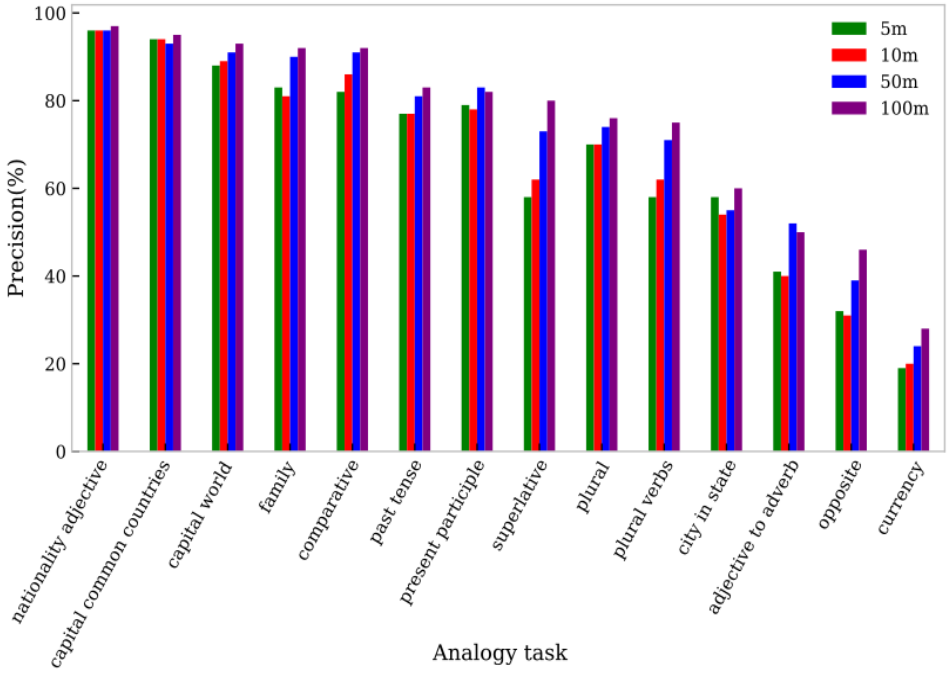
to train and evaluate embeddings. (The free version of Colab is limited in that one cannot leave the code running without being monitored, i.e., the code has to be constantly monitored to avoid premature termination of the task being executed.)

Determining a suitable minimum word count

We conducted a third experiment in order to determine the most suitable word count for the embedding. The minimum word count sets the minimum word frequency for pruning the vocabulary available for model training. All words with minimum word count below a given specified threshold are ignored during training (Řehůřek & Sojka, 2011b). A high minimum word count leads to embeddings with a smaller but more robust vocabulary. However, such small models may not contain the words that researchers require in applications, and we found that even words used in the analogy test set (such as "policewoman") quickly became ignored if the minimum word count was too high.

Figure 3 shows the precision results of the four word embeddings, which were trained using the same hyperparameter settings except for the minimum word counts (which were set to 5, 10, 50 and 100, respectively).

Figure 3: Training with 4 different minimum word counts (evaluated via 14 analogy tasks)



It is evident in Figure 3 that the quality of the word embedding increased with increasing minimum word count. Nonetheless, since we intended to make the embedding publicly available for research, we decided to fix the minimum word count threshold at 50, which ensured that most of the words found in the vocabulary were associated with word vectors after building of the Word2Vec embedding. While the results show that a value of 100 was the optimal value for the minimum word count hyperparameter setting, the usefulness of our word embedding would be highly compromised if 100 were used—as only words whose frequency of occurrence reached 100 or more would appear in the vocabulary.

6. Maximising robustness: Determining variances and testing ensembles of embeddings

When developing and deploying a word embedding, one must seek to maximise its robustness. For example, one does not want the distance between word vectors in an embedding to significantly depend on random initialisation of weights in the training algorithm, or to depend on random permutations of the data when it is used during training of the algorithm. Likewise, the results should be robust against bootstrapping, or subsampling of the data (as long as the overall size or quality of the dataset—and therefore the information available—does not change). If large variances are produced by small changes in the training set, then this is evidence that the embedding does not generalise well (Antoniak & Mimno, 2018).

In order to maximise robustness of our embedding, we conducted tests, as described below, in order to:

- determine the variances produced by data shuffling, random initialisation and bootstrapping; and
- determine the degree to which generation of ensembles of embeddings would reduce variance and improve robustness.

Determining the variances

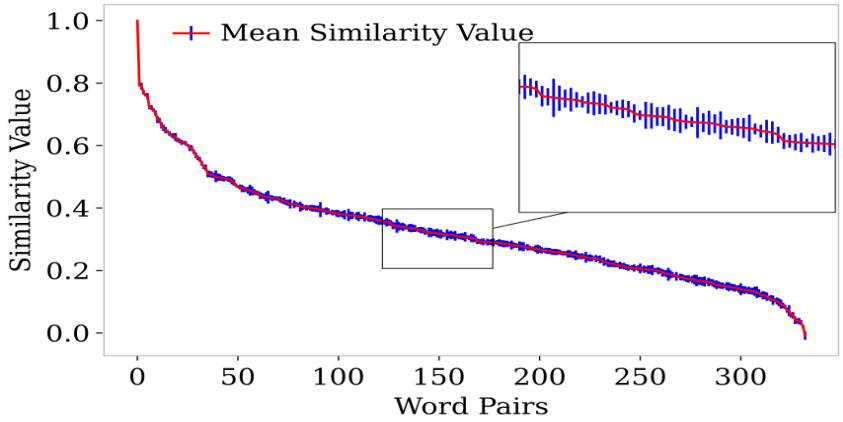
To understand how training stochasticity (data shuffling and random initialisation) and subsampling (bootstrapping) influenced the distances between word vectors in our embedding, we generated three ensemble word embeddings:

- 10p subsampled: This first ensemble consisted of 10 word embeddings trained on 10% of the sentences in the data corpus that were randomly subsampled for each word embedding. The resulting word embeddings therefore had different vocabularies.

- 10p shuffled: This second ensemble consisted of 10 word embeddings trained on the same subset of 10% randomly sampled sentences and the resulting embeddings therefore shared the same vocabulary. (The differences between the embeddings stemmed only from the differences in training procedures.)
- 100p shuffled: This third ensemble consisted of five word embeddings (a smaller number of embeddings, due to their size) that were trained on the entire training dataset. (Again, the difference in the embeddings stemmed only from the differences in training procedures.)

In addition to showing the variance of word similarities between different instances of the embedding, these three ensembles allowed us to study the effect of bootstrapping (when comparing 10p_subsampled vs 10p_shuffled), as well as the effect of the size of the training dataset (when comparing 10p_shuffled vs 100p_shuffled). It should be noted that the vocabulary of the smaller datasets was necessarily smaller as well, and we ignored analogies if one of the words (or the solution) was not part of the word embeddings' vocabulary. We calculated the 360 similarities of word pairs in the WordSim353 dataset (Agirre et al., 2009)[4] for all word embeddings in a set, and plotted the mean and variance of the results, as shown in Figures 4(a), 4(b), and 4(c).

Figure 4(a): Variance of embeddings based on a 10% subsampled set of 10 embeddings



4 http://alfonseca.org/pubs/ws353simrel.tar.gz

Figure 4(b): Variance of embeddings based on a set of 10 embeddings trained on shuffled 10% of the total training dataset

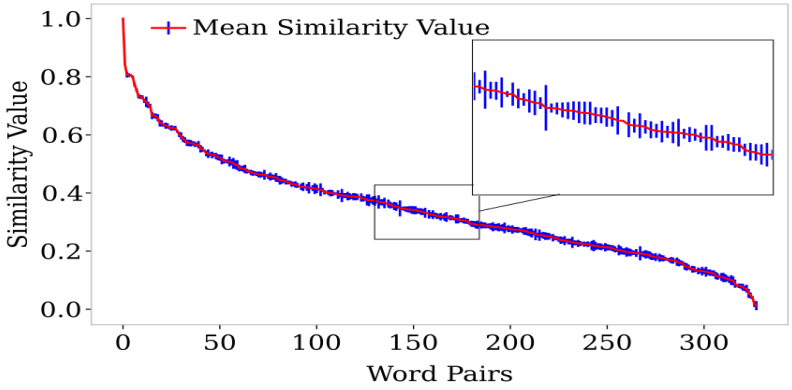
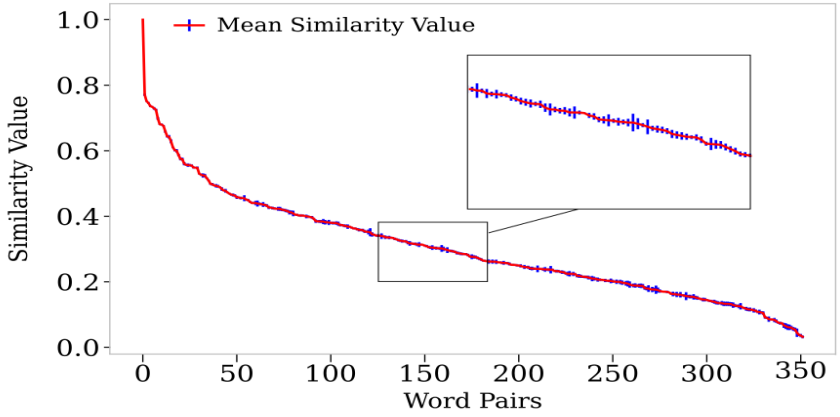


Figure 4(c): Variance of embeddings based on a set of five embeddings trained on shuffled full training dataset



The results confirmed our expectation that a variance in the training itself, as introduced by subsampling, would lead to an even larger variance of the word similarities. Furthermore, the results showed that a larger dataset led to a much lower variance. In the 100p model, for example, the variance was in fact low enough to reliably distinguish distances between words on an order of 0 to $2x10^{-4}$. These results suggested a strategy for how to make our word embedding more robust: build an ensemble model that united the prediction of several models (which is in fact standard practice to decrease variance (Antoniak & Mimno, 2018)).

Determining efficacy of ensemble embeddings

We investigated the efficacy of ensemble embeddings using the aforementioned 14 analogy benchmarks. For these benchmarks, averaging over models was not an option since the result of every analogy was a correct/incorrect answer. We therefore needed more elaborate rules for turning the decision of a single embedding into the ensemble decision. The rule we developed to combine analogy decisions into an ensemble was a rule that (1) computed the list of 10 closest neighbours to the vector in the analogy equation (such as vec(ANC) – vec(EFF) + vec(Zuma)) for every word embedding, and then (2) concatenated (joined together) the lists. We then computed the 10 words that appeared in the largest number of ensembles, and checked whether the desired result was in the final list (positive outcome) or not in the list (negative outcome).

Figure 5: Precision values for ensemble word embeddings evaluated on analogy tasks

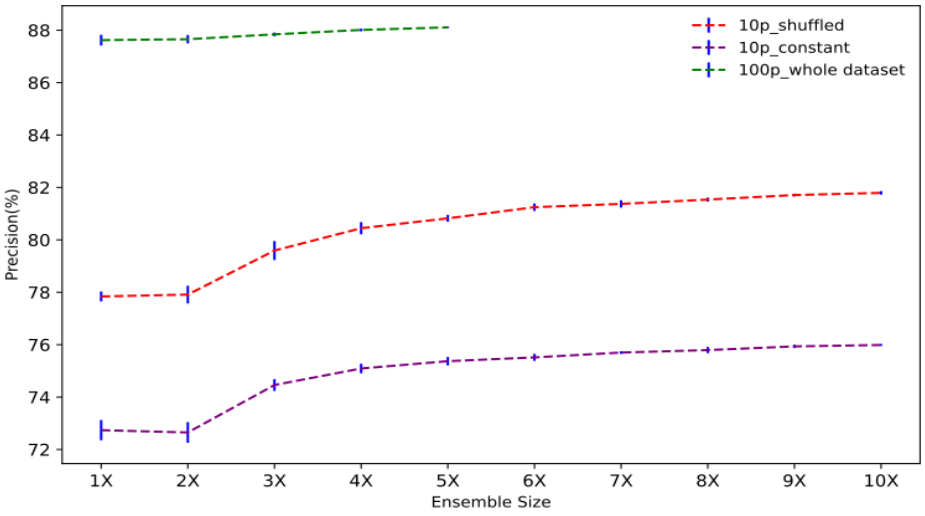


Figure 5 shows the average performance of ensembles of varying sizes on the analogy tasks, using the three different strategies mentioned above to create the training sets for each member of the ensemble. The benchmarks were calculated by randomly sampling, 10 times, ensemble members from the set of trained models. For example, the 2X ensemble word embedding in Figure 5 represents an ensemble consisting of two members, created by randomly sampling two models from a set of 10 word embeddings. This ensured that the results did not depend on the model used to make up the ensemble. This evaluation generated four important observations:

- Generation of ensemble embeddings improved the precision of the model's results.

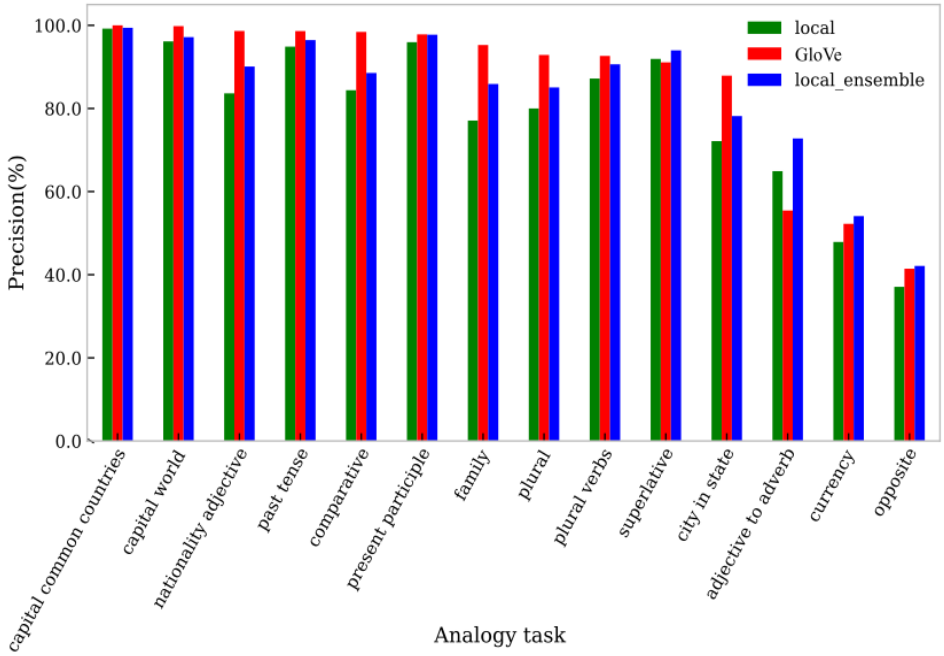
- The ensembles of embeddings performed better as they increased in size.
- This trend (improvement with increased ensemble size) was particularly strong for the 10p models (ensembles trained with smaller datasets), and not so prevalent for the ensembles trained on the full dataset.
- Significant improvements (over single models) occurred even in ensembles composed of three or more embeddings, an important finding given that smaller ensembles are easier to store and use.

As a result, we are able to recommend the use of ensembles of three or more embeddings when datasets are relatively small, which is common in geographic areas, such as on the African continent, that are under-represented in terms of data representation.

7. Performance comparison between South African Word2Vec embedding and a GloVe embedding

Figure 6 shows a comparison between our South African news Word2Vec embedding and the GloVe model in performing the 14 analogy tasks. We trained a single embedding (that is, the "100p_250d_50m" embedding) and a five-member ensemble of the South African news word embedding for comparison, and, consistent with the findings above, the ensemble showed a better performance.

Figure 6: Results from comparison between Word2Vec and GloVe embeddings



Overall, as seen in Figure 6, our Word2Vec word embedding was competitive on most tasks when compared to the GloVe embedding. And, interestingly, our Word2Vec ensemble of embeddings outperformed the GloVe model on four analogy tasks—remarkable given that the latter has the advantage of a significantly larger training dataset size and a global training set domain. This finding was surprising for two reasons. First, the GloVe algorithm is an international benchmark that has been acknowledged as one of the most stable models for production of word embeddings (Wendlandt et al., 2018). Second, the GloVe embedding we compared ours with was trained on a much larger corpus than our local word embedding. The GloVe embedding had approximately 300,000 unique word tokens in each embedding—300,000 for the GloVe embedding versus 124,000 for our embedding. The importance of vocabulary size in ensuring high quality word embeddings is well-documented (Rodman, 2020). The results are therefore encouraging evidence of the robustness and performance of our South African news word embedding.

We observed no differences of statistical significance, between the performance of our Word2Vec word embedding model and the performance of the GloVe model, on the largely universal grammar tasks among the 14 analogy tasks (such as the present participle and superlative tasks), or on semantic analogy tasks (such as capital-common-countries task). (Because we were not able to check the word frequencies for the corpus used to train the GloVe model, we suspect that the equivalence in performance for the two models showed that the words used in the evaluation tasks were sufficiently represented in the corpuses used to train both the GloVe embedding and our local Word2Vec embedding.)

However, we found that for the capital-world analogy task, the South African news Word2Vec embedding did not perform well in analogy examples that involved tokens such as Brussels and Belgium. The results suggested that the news embedding appeared to have known Brussels more in the context of the European Union (EU). This inference followed from noting that the EU was predicted with high likelihood amongst a list of top 10 nearest neighbours as predicted by the news embedding to solve given relational tasks. Also, the South African news embedding failed to solve the relational task which involved Madrid and Spain. We observed that the South African news embedding appeared to know Madrid in the context of Real Madrid, the Spanish football team. This observation points to the need to build situation-specific word embeddings whose learned word vectors fully capture and represent the original views and standpoints expressed by news content creators at the time of writing.

It should also be noted that, during the comparison, we found that there was a subtlety that may have influenced the results to a small extent—a subtlety introduced by our rule that excluded analogies in a task if the embedding (or all of the embeddings in an ensemble) did not contain a word in the analogy. For example, if the local

embedding did not contain the word "nursultan" (representing "Nur-Sultan"), the capital-common-countries analogy of guessing Kazakhstan's capital city did not have to be solved, while the global embedding had to solve it, even if the global embedding contained very few examples mentioning the city. Accordingly, we resolved this issue by evaluating the models using only those relational tasks with words that were common to both models.

8. Validation of the embedding against local benchmarks

In addition to testing our South African embedding against the international benchmarks provided by the GloVe algorithm, it was also necessary to validate the embedding against local benchmarks that represented local contexts. For this purpose, we created two local analogy tasks. The first local analogy task involved matching politicians to political parties—for example, EFF is to Julius_Malema as ANC is to Jacob_Zuma. This politician-to-political party relational task consisted of 398 analogy tasks, of which our local Word2Vec South African news embedding model successfully solved 212 (53%)—according to the performance evaluation metrics described above. The second local analogy task involved matching cities to provinces—for example, KZN is to Durban as Western_Cape is to Cape_Town. This city-to-province analogy consisted of 586 relational tasks, of which our local Word2Vec model successfully solved 582 (a model performance of 99%). Due to the fluidity of political affiliations, we were not surprised that the politician-to-political party task proved to be more difficult than the city-to-province task.

However, we observed that, in certain circumstances, our South African news embedding failed to solve certain analogy tasks simply because text preprocessing did not include stemming, a practice that reduces all related tokens to their root word. As a result, the embedding was penalised for predicting DAS (which stands for DA's) instead of DA. When we implemented stemming, that increased the predictive power of the word embedding for the task of matching politicians with their respective political parties (predictive accuracy increased from 53% to 59%). However, stemming proved to be counterproductive for the second task: matching provinces with their cities (predictive power dropped from 99% to 69%).

The negative effects from stemming arose from the fact that a word stemmer is a model that is trained to reduce English words to original root words, and therefore any token that is given to the model will be reduced to the root word that is known to the model. The work of Al-Shammari and Lin (2008) supports our finding regarding the drop in model performance, for certain analogy tasks, following word stemmatisation. Consequently, we recommend development of a full range of localised NLP tools, including word stemmers and lemmatisers, that are optimised to handle local contexts, so as to facilitate training of word embeddings whose learned word vectors truly resemble local contexts.

9. Conclusions

In this article, we have presented a word embedding that was trained, using the Word2Vec algorithm, on South African news article data collated and stored by MMA. The full corpus consisted of news articles that were published between 1 January 2018 and 17 March 2021. We have presented results from testing of the impact of varied hyperparameters, changes in the training set, and ensemble-building, on the performance of the embedding. We have also presented results from comparison of the performance of our local Word2Vec embedding against the performance of the GloVe algorithm—results which showed competitive performance, and even superior performance in some instances, by our local South African embedding. Furthermore, we have provided results from two tests used to check the performance of our embedding against South African benchmarks.

We hope that the embedding and benchmarks we have presented promote further research in South African social sciences, and will help researchers who lack the resources required to train vast machine learning models for NLP. In particular, the word embedding contributed by this study presents researchers with an opportunity to use the word vectors as text encoders. For instance, researchers can use our word embedding in ways similar to how pre-trained word vectors produced by algorithms such as Word2Vec, GloVe, BERT and fastText are used to vectorise texts without the need to train word embeddings from scratch. Consequently, we hope and anticipate that our word embedding will play a significant role as the "embedding layer" in similar South African text analysis studies. We believe that contributions such as these are crucial to unlocking the potential of big data analysis in localised African contexts.

Disclosure statement

The authors report that there are no competing interests to declare.

Data availability statement

The word embeddings that support the findings of this study are openly available in figshare.[5] The code for reproducing the results is available in a github repository.[6]

References

Agirre, E., Alfonseca, E., Hall, K., Kravalova, J., Pasca, M., & Soroa, A. (2009). A study on similarity and relatedness using distributional and WordNet-based approaches. https://doi.org/10.3115/1620754.1620758

Al-Shammari, E. T., & Lin, J. (2008). Towards an error-free Arabic stemming. In *Proceedings of the 2nd ACM Workshop on Improving Non English Web Searching* (pp. 9–16). https://doi.org/10.1145/1460027.1460030

5 https://doi.org/10.6084/m9.figshare.18933728

6 https://github.com/Mafunda/SouthAfricanNewsEmbeddings

Antoniak, M., & Mimno, D., 2018. Evaluating the stability of embedding-based word similarities. *Transactions of the Association for Computational Linguistics, 6*, 107–119. https://doi.org/10.1162/tacl_a_00008

Arseniev-Koehler, A., & Foster, J. G. (2020). Sociolinguistic properties of word embeddings. https://doi.org/10.31235/osf.io/b8kud

Badri, N., Kboubi, F., & Chaibi, A. H. (2022). Combining FastText and Glove word embedding for offensive and hate speech text detection. *Procedia Computer Science*, *207*, 769–778. https://doi.org/10.1016/j.procs.2022.09.132

Bakarov, A. (2018). A survey of word embeddings evaluation methods. arXiv preprint arXiv:1801.09536.

Berardi, G., Esuli, A., & Marcheggiani, D. (2015). Word embeddings go to Italy: A comparison of models and training datasets. In *Proceedings of 6th Italian Information Retrieval Workshop.*

Bojanowski, P., Grave, E., Joulin, A., & Mikolov, T. (2017). Enriching word vectors with subword information. *Transactions of the Association for Computational Linguistics*, *5*, 135–146. https://doi.org/10.1162/tacl_a_00051

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805.

Durrheim, K., Schuld, M., Mafunda, M., & Mazibuko, S. (2022). Using word embeddings to investigate cultural biases. *British Journal of Social Psychology*, *00*, 1–13. https://doi.org/10.1111/bjso.12560

Goodman, J. (2001). Classes for fast maximum entropy training. In *2001 IEEE International Conference on Acoustics, Speech, and Signal Processing: Proceedings, 1* (pp. 561–564).

Grand, G., Blank, I. A., Pereira, F., & Fedorenko, E. (2022). Semantic projection recovers rich human knowledge of multiple object features from word embeddings. *Nature Human Behaviour*, *6*, 975–987. https://doi.org/10.1038/s41562-022-01316-8

Gu, Y., Leroy, G., Pettygrove, S., Galindo, M. K., & Kurzius-Spencer, M. (2018). Optimizing corpus creation for training word embedding in low resource domains: A case study in autism spectrum disorder (ASD). In *AMIA Annual Symposium Proceedings, 2018* (pp. 508–517).

Hunt, E., Janamsetty, R., Kinares, C., Koh, C., Sanchez, A., Zhan, F., Ozdemir, M., Waseem, S., Yolcu, O., Dahal, B., & Zhan, J. (2019). Machine learning models for paraphrase identification and its applications on plagiarism detection. In *2019 IEEE International Conference on Big Knowledge (ICBK)* (pp. 97–104). https://doi.org/10.1109/ICBK.2019.00021

Jain, A., Meenachi, D. N., & Venkatraman, D. B. (2020). NukeBERT: A pre-trained language model for low resource nuclear domain. arXiv preprint arXiv:2003.13821.

Kozlowski, A. C., Taddy, M., & Evans, J. A. (2019). The geometry of culture: Analyzing the meanings of class through word embeddings. *American Sociological Review*, *84*(5), 905–949. https://doi.org/10.1177/0003122419877135

Loper, E., & Bird, S. (2002). NTLK: The natural language toolkit. arXiv preprint cs/0205028.

Marivate, V., Sefara, T., Chabalala, V., Makhaya, K., Mokgonyane, T., Mokoena, R., & Modupe, A. (2020). Investigating an approach for low resource language dataset creation, curation and classification: Setswana and Sepedi. arXiv preprint arXiv:2003.04986.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. *Advances in Neural Information Processing Systems*, *26*. https://doi.org/10.48550/arXiv.1310.4546

Pennington, J., Socher, R., & Manning, C. D. (2014). GloVe: Global vectors for word representation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 1532–1543). https://doi.org/10.3115/v1/D14-1162

Pereira, F., Gershman, S., Ritter, S., & Botvinick, M. (2016). A comparative evaluation of off-the-shelf distributed semantic representations for modelling behavioural data. *Cognitive Neuropsychology*, *33*(3–4), 175–190. https://doi.org/10.1080/02643294.2016.1176907

Peters, M. E., Neumann, M., Iyyer, M., Gardner, M., Clark, C., Lee, K., & Zettlemoyer, L. (2018). Deep contextualized word representations. arXiv preprint arXiv:1802.05365.

Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). Improving language understanding by generative pre-training. https://s3-us-west-2.amazonaws.com/openai-assets/research-covers/language-unsupervised/language_understanding_paper.pdf

Rahimi, Z., & Homayounpour, M. M. (2022). The impact of preprocessing on word embedding quality: A comparative study. *Language Resources and Evaluation*, 1–35. https://doi.org/10.1007/s10579-022-09620-5

Řehůřek, R., & Sojka, P. (2011a). Gensim – statistical semantics in Python. NLP Centre, Faculty of Informatics, Masaryk University, Brno, Czech Republic. https://www.fi.muni.cz/usr/sojka/posters/rehurek-sojka-scipy2011.pdf

Řehůřek, R., & Sojka, P. (2011b). Gensim – Python framework for vector space modelling. NLP Centre, Faculty of Informatics, Masaryk University, Brno, Czech Republic.

Rezaeinia, S. M., Rahmani, R., Ghodsi, A., & Veisi, H. (2019). Sentiment analysis based on improved pre-trained word embeddings. *Expert Systems with Applications*, *117*, 139–147. https://doi.org/10.1016/j.eswa.2018.08.044

Richardson, L. (2007). Beautiful soup documentation. https://www.crummy.com/software/BeautifulSoup/bs4/doc/

Rodman, E. (2020). A timely intervention: Tracking the changing meanings of political concepts with word vectors. *Political Analysis*, *28*(1), 87–111. https://doi.org/10.1017/pan.2019.23

Santos, I., Nedjah, N., & de Macedo Mourelle, L. (2017). Sentiment analysis using convolutional neural network with fastText embeddings. In *Proceedings of the 2017 IEEE Latin American Conference on Computational Intelligence (LA-CCI)* (pp. 1–5). https://doi.org/10.1109/LA-CCI.2017.8285683

Svoboda, L., & Beliga, S. (2017). Evaluation of Croatian word embeddings. arXiv preprint arXiv:1711.01804.

Theil, C. K., Štajner, S., & Stuckenschmidt, H. (2020). Explaining financial uncertainty through specialized word embeddings. *ACM Transactions on Data Science*, *1*(1), 1–19. https://doi.org/10.1145/3343039

Turian, J., Ratinov, L., & Bengio, Y. (2010). Word representations: A simple and general method for semi-supervised learning. In *Proceedings of the 48th annual Meeting of the Association for Computational Linguistics* (pp. 384–394).

Wendlandt, L., Kummerfeld, J. K., & Mihalcea, R. (2018). Factors influencing the surprising instability of word embeddings. arXiv preprint arXiv:1804.09692

Xu, R., Yang, Y., Otani, N., & Wu, Y. (2018). Unsupervised cross-lingual transfer of word embedding spaces. arXiv preprint arXiv:1809.03633.

Yin, Z., & Shen, Y. (2018). On the dimensionality of word embedding. *Advances in Neural Information Processing Systems*, *31*. https://doi.org/10.48550/arXiv.1812.04224

Using machine learning to predict low academic performance at a Nigerian university

Ebiemi Allen Ekubo
Lecturer, Department of Computer Science, Niger Delta University, Amassoma, Nigeria
https://orcid.org/0000-0001-9348-5630

Bukohwo Michael Esiefarienrhe
Associate Professor, Department of Computer Science and Information Systems, and Researcher, Material Science Innovation and Modelling (MaSIM), Faculty of Natural and Agricultural Sciences (FNAS), North-West University, Potchefstroom, South Africa
https://orcid.org/0000-0002-1631-9759

Abstract
This study evaluates the ability of various machine-learning techniques to predict low academic performance among Nigerian tertiary students. Using data collected from undergraduate student records at Niger Delta University in Bayelsa State, the research applies the cross-industry standard process for data mining (CRISP-DM) research methodology and the Waikato Environment for Knowledge Analysis (WEKA) tool for modelling. Five machine-learning classifier algorithms are tested—J48 decision tree, logistic regression (LR), multilayer perceptron (MLP), naïve Bayes (NB), and sequential minimal optimisation (SMO)—and it is found that MLP is the best classifier for the dataset. The study then develops a predictive software application, using PHP and Python, for implementation of the MLP model, and the software achieves 98% accuracy.

Keywords
machine learning, educational data mining, student academic performance, university, cross-industry standard process for data mining (CRISP-DM), Waikato Environment for Knowledge Analysis (WEKA), classifier algorithms, J48 decision tree, logistic regression (LR), multilayer perceptron (MLP), naïve Bayes (NB), sequential minimal optimisation (SMO), Nigeria, Niger Delta University

Acknowledgements
The authors thank the Material Science Simulation and Modelling (MaSIM) research focus area in the Faculty of Natural and Agricultural Sciences (FNAS), the School of Postgraduate Studies, and the Department of Computer Science and Information Systems, all at North-West University, for their assistance throughout this research. We also thank Niger Delta University for making its data available for research purposes. This paper draws on the contents of the first-listed author's PhD thesis at North-West University (Ekubo, 2020).

DOI: https://doi.org/10.23962/ajic.i30.14839

Recommended citation
Ekubo, E. A., & Esiefarienrhe, B. M. (2022). Using machine learning to predict low academic performance at a Nigerian university. *The African Journal of Information and Communication (AJIC), 30*, 1-33. https://doi.org/10.23962/ajic.i30.14839

1. Introduction

Low academic performance is a challenge faced by many higher education institutions worldwide (Romero et al., 2010). The challenge is more serious in developing countries, where the phenomena undermining academic success are more numerous and more pronounced than in developed-world settings (Al-Zoubi & Younes, 2015). The UN has identified the challenges present in developing countries that ultimately result in low academic performance as poverty, unstable electricity, lack of funding, diverse health challenges, environmental degradation, level of social development (low school enrolment rate), vulnerability in employment, maternal mortality, and quality of life of slum dwellers (UN DESA, 2013). With regard to low academic performance of undergraduate students, the research shows that the effects on students can be long-term, resulting in low self-esteem, unease, and fear of failure (Aryana, 2010; Nsiah, 2017; Nurmi et al., 2003). For tertiary institutions, low academic performance limits their growth in terms of, inter alia, execution of operations and rankings in global indices (Serdyukov, 2017). Thus, the challenge needs continuous monitoring and the creation of incentives to help boost students' progress.

The issue of low academic performance is prevalent in Nigeria, where many universities record a high number of low-performing undergraduate students (Oyebade & Dike, 2013). Educational research in Nigeria concerned with low academic performance has considered numerous factors, including those related to the student, the environment, the institution, and the relevant level(s) of government (Farooq et al., 2011).

The discipline of data mining seeks to discover useful patterns in large sets of data, so as to predict future outcomes (Gullo, 2015). In education, the use of data mining has led to improvements in predicting factors that will render certain students prone to struggling with their studies, to failure, and to dropping out (Hughes & Dobbins, 2015). Results from these predictions can assist stakeholders in addressing the factors identified as contributing to poor student performance. Therefore, it is necessary for tertiary institutions to make use of this useful prediction method.

Some research carried out in Nigeria using data-mining techniques with educational data has already looked at predicting students' performance using selected attributes (Adeyemo & Kuye, 2006; Oyerinde & Chia, 2017). These studies predicted students' performance using various attributes, but did not develop software based on their prediction to monitor and track the progress of students' performance. The absence of a dedicated software tool that enables Nigerian tertiary institutions to track their students' performance, despite the abundance of packages for data mining, justified this study's efforts to identify the highest-performing data-mining algorithm and to implement it in the form of customised software. This study used a case study of a single university.

Universities need to document the attributes of low performers, classify the low performers using data-mining techniques, and develop a model that can identify likely low performers. The model developed can serve as a foundation for the design and implementation of systems that can provide potential low performers with the assistance that they need to perform better. The goal of this study was to develop such a model, based on data from a single Nigerian university—Niger Delta University in Bayelsa State—and with potential applicability to multiple universities.

This research sought to: (1) identify factors that are causing the poor performance of undergraduates in Nigeria; (2) collect and represent these factors as features in machine-readable format for data mining; (3) identify the best set of features from the total features collected for predicting low academic performance; (4) identify which machine-learning technique could best classify low-performing students based on the selected features; and (5) develop a customised software system that uses the identified best machine-learning algorithm to identify low-performing students and automatically recommend various interventions.

Using data collected from the university's undergraduate student records and applying the cross-industry standard process for data mining (CRISP-DM) research methodology and the Waikato Environment for Knowledge Analysis (WEKA) tool for modelling, five machine-learning classifier algorithms were used to determine which one was the most accurate in predicting poor student performance.

Section 2 of this article reviews related literature; section 3 discusses the methodology followed; section 4 presents the testing of the five classifier algorithms; section 5 presents the feature selection process; section 6 presents the development and evaluation of the predictive software application; and section 7 provides conclusions and a recommendation.

2. Literature review

Educational data mining is the multidisciplinary research area that applies data-mining techniques to educational data (Romero & Ventura, 2013). The educational environment regularly produces large amounts of data (Romero & Ventura, 2013), and with the use of data-mining techniques, stakeholders can gain knowledge to help them understand learners and improve their learning process (Algarni, 2016). Using educational data-mining techniques can also benefit society in general, as every society thrives socially and economically when its education system performs optimally (Mitra, 2011).

Yağcı (2022) proposes a model that predicts the final exam grades of students using their midterm exam grades. Six algorithms—random forest (RF), nearest neighbour, support vector machines (SVMs), logistic regression (LR), naïve Bayes (NB), and k-nearest neighbours (K-NN)—are used, and their accuracy falls within the range of 70 to 75%. Thus, the Yağcı (2022) study contributes to identifying high-risk students who are likely to fail their exams. Our study was different from that that of Yağcı (2022), because we used the cumulative grade point average (CGPA), which is technically a student's average of all courses, instead of a single exam grade for a subject. Dhilipan et al. (2021) use four data-mining algorithms to predict the performance of students. The four algorithms used, and their levels of accuracy, are: LR (97.05%), decision tree (88.23%), entropy (91.19%), and K-NN (93.71%). While the Dhilipan et al. (2021) focuses on prediction of student performance in general, our study was focused more on identification of low-performing students. In our study, not all students' data was used—only the data related to low-performing students. Also, the algorithms we used differed from those used by Dhilipan et al. (2021).

Vergaray et al. (2022) present a stacking multi-classification model for mining academic performance of students, using five classification algorithms, with their output as a feeder to the stacking model. The five classification algorithms with their accuracies are: extra trees (ET) (57.41%), RF (61.96%), decision tree (91.44%), adaptive boosting (AdaBoost) (59.65%), and extreme gradient boosting (XGBoost) (83.3%). The proposed stacking model is then applied, combining the five algorithms, enabling accuracy of 92.86%.

So as to predict student performance at a very early stage, Li et al. (2022) propose an end-to-end deep-learning model that can automatically (with no manual intervention) extract features from datasets to form a two-dimensional convolutional neural network (CNN) of behaviours reflective of the dataset. Their experiment is conducted using university students' data from Beijing, and the results show improvement over traditional data-mining algorithms. Ofori et al. (2020) present findings from a review of literature related to using machine-learning algorithms

in the prediction of students' academic performance and their learning outcomes. The reviews show various algorithms used in the prediction of students' academic performance and the highest levels of accuracy achieved.

Abu Zohair (2019) mines data on postgraduate students' performance using 50 students' datasets obtained from a university. The authors split their data into dataset1 and dataset2, and use five data-mining algorithms to train their datasets, with the following results in terms of accuracy: multilayer perceptron (MLP) (60.5%), NB (71.1%), SVMs (76.3%), K-NN (65.8%), and linear discriminant analysis (LDA) (71.1%). Flanagan et al. (2022) use the linear kernel SVMs model to predict the performance and engagement of students' behaviour using digital textbooks during an open book test. A good level of accuracy is achieved in identifying, at an early stage, students who are low performers.

Factors that influence low academic performance generally vary from society to society, and from individual to individual. In many developed countries, low academic performance often relates to personal issues such as a lack of inspiration to study or emotional trauma resulting from certain situations (Banerjee, 2016). However, in many developing countries, factors emanating from institutions and the government, in addition to individual factors, also contribute to poor academic performance. Thus, the causes of poor academic performance in developing countries can broadly be divided into individual, institutional, and governmental factors.

The individual factors relate to students' capacity to focus and concentrate on their academic work. Some individual factors that affect the performance of students in the Arab countries of Jordan and Oman, as outlined by Al-Zoubi and Younes (2015) and Alami (2016), are: poor motivation, lack of planning, low self-confidence, fretfulness about exams, poor examination practices, low opinion of the course, lack of interest in the course, laziness, and lack of future plans. Other individual factors identified as affecting low student performance are medical or mental problems, and the family's financial background (Al-Zoubi & Younes, 2015). Factors associated with low tertiary performance in Nigeria include students' uncertainty about their future because of high levels of unemployment; the large number of low-salary earners; lack of funding (e.g. bursaries); a lot of competition for scholarships; self-sponsoring students' distractions due to their engagement in temporary jobs; engaging in small-scale business activities such as buying and selling of goods and services, which then occupies most of their time, leaving little for studies; lack of planning; poor study habits; negative peer influence; family crises; lack of support from guardians or parents; students' view of a course as difficult; and the family's financial or psychological background (Oyebade & Dike, 2013).

Contributing factors associated with tertiary institutions, according to Frimpong et al. (2016), writing in the Ghanaian context, include poor conditions in the learning environment and educators' inadequate knowledge of educational and psychological matters.

Government factors include unstable electricity within and outside the school environment; insufficient security in learning environments; failure to adequately address high levels of poverty in the country; insufficient budgets for the education sector; and pay disputes between the academic staff of higher education institutions and government, which result in regular strike activities (Adeyemi & Adeyemi, 2014; Longe, 2017; Ugar, 2018).

3. Methodology

The process model followed in this study was the widely used CRISP-DM. It involves six steps (Wirth & Hipp, 2000):

- *Business understanding:* Scouting relevant stakeholders and assembling vital information to make sure the goals of the research are achievable.
- *Data understanding:* Data gathering and investigation to ensure that the data has quality and shows useful patterns.
- *Data preparation*: Cleaning and converting gathered data into a suitable format for mining and ensuring the dataset fits the selected modelling tool.
- *Modelling:* Use of the cleaned dataset and application of selected algorithms for knowledge mining.
- *Evaluation:* Examination of the output of the modelling step to ensure the knowledge gained is in line with the proposed goals of the project.
- *Deployment:* Display of the discovered knowledge by either integrating it into an existing system or developing a new system to assist stakeholders in benefiting from the knowledge gained (Wirth & Hipp, 2000).

In line with the CRISP-DM process model, the study consisted of the following steps: a survey of literature (as summarised above in section 2) on data-mining of educational data and on the causes of low performance in Nigerian universities; the gathering of student details from Niger Delta University and the storage of the details in Microsoft Excel; the testing of five machine-learning algorithms to determine the best algorithm for identifying low student performance based on collected data; using four feature selection techniques to choose the optimal features from the dataset; and designing and testing a predictive system, using the PHP programming language, for the identification of potential low performers.

The CRISP-DM methodology was used in this study because it provided all the phases necessary for our research endeavour, including a phase for data collection and pre-processing (necessary for us as our data was mostly in the form of physical

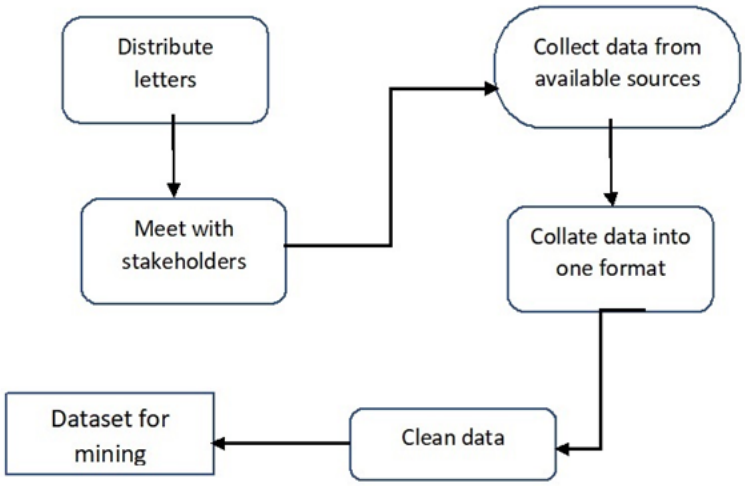
files obtained from the university administrator). The methodology also provided a phase for data mining and then construction of a solution for implementation. Using the CRISP-DM methodology saved us from having to combine numerous methodologies for the different parts of the research, particularly the software development (we relied on the CRISP-DM deployment phase to develop the software). The WEKA tool was used for modelling because of its simplicity and its ability to present results in a concise form.

The study used data stored in the Niger Delta University repository, and sampled 2,348 low-performing students' data, i.e., data on students with a CGPA of less than 3.00 (CGPA < 3.00).

Data collection

Figure 1 shows the complete data collection process followed in this study. The process involved letters being distributed to key stakeholders for permission to obtain data, which led to a meeting with stakeholders to agree on terms regarding the data-collection process and privacy issues. The next steps were the collection and collation of existing data from the available sources into one format, cleaning data to rid it of inaccurate or incomplete information, and preparation of the data for mining.

Figure 1: The data collection process



Source: Ekubo (2020, p. 36)

The lead researcher (the first-listed author, Ekubo) had a meeting with the Deputy Vice-Chancellor (Academic) of the Niger Delta University, who oversees the handling of all information related to students' academic results. The DVC assisted the researcher by sending out memos to all faculty deans, heads of department,

faculty officers, and faculty examination officers to assist the researcher in data-gathering from the university repositories. The faculty officers stored the students' details in spreadsheet files and gave the researcher the data on USB flash drives. The collection of students' results data involved meeting with examination officers within departments. Where the examination officers were unavailable, the heads of department provided the data, which was in either pdf or hard copy files. The researcher transferred the files on USB flash drives immediately to his laptop and stored all hard copy files in a file jacket.

Data preparation

The data collected was incomplete and inaccurate. The students' details collected from the university comprised 10,472 records. After manually inserting CGPAs from the pdfs and hard copies, the total number of students with assigned CGPAs comprised 5,631 records, of which students with CGPAs less than 3.00 comprised 3,481 records. This formed the entire population of low-performing undergraduate students used in this research.

Attribute selection

At the end of the data preparation process, one class attribute (CGPA, the core attribute used for the prediction) and 24 dependent attributes were selected for mining. Table 1 shows the class attribute and 24 dependent attributes, the attributes' variable codes, and the codes' corresponding values. In the table, "JAMB score" refers to the student's score on the exam provided by the Joint Admissions and Matriculation Board (JAMB)—the exam called the "UTME" (the Unified Tertiary Matriculation Examination). The "course from JAMB" attribute indicates whether the course the student was admitted to study was the actual course the student chose to study from JAMB. The "post-UTME score" is the score the student obtained on the university's internal entrance examination.

Also in Table 1, "SSCE" refers to the Senior Secondary School Certificate Examination. The student's "average SSCE score" was calculated as the average of the student's pre-university academic performance. Nigerian high school students write exams in a minimum of seven subjects and a maximum of nine subjects, and earn grades A, B, C, D, E, or F, based on their performance. For the purpose of the research, the values 6, 5, 4, 3, 2 or 1 replaced the respective grades, and the sum for each student divided by the number of subjects the student wrote gave the value of the average SSCE score.

Table 1: The attributes used in the data mining

Attribute	Variable codes	Values
Class attribute: CGPA	HL LL	2.50 – 2.99 0.01 – 2.49
Sex	M F	Male Female
Age	B30 30A	Below 30 years 30 years and above
Marital status	S M	Single Married
Attended primary school	NO YES	No Yes
Secondary school type	PRI PUB	Private Public
Secondary school area	URB RUR	Urban Rural
Sponsor type	GUAD PAR SELF	Guardian Parents Self-sponsor
Sponsor qualification	DEG NODEG NOEDU	Educated with degree Educated without degree No formal education
Sponsor income (per month)	LOW MED HIGH	Below N50,000 N50,000 – N100,000 Above N100,000
Sponsor support	LOW MED HIGH	Little support Average support Great support
Family size	SMALL MED LAR	1 – 4 5 – 9 Above 9
Work and study	YES NO	Yes No
University accommodation	CMPS OFFCMPS	Campus Off-campus
Years before admission (gap between completing secondary school and commencing university)	NONE B5 5A	None Below 5 years 5 years and above
Course from JAMB	YES NO	Yes No
Course interest	LOW AVE HIGH	Little interest Average interest High interest

Attribute	Variable codes	Values
Weekly study time	LOW AVE HIGH	Less than 10hrs 10 – 20hrs Above 20hrs
Postgraduate degree	NO YES NS	No Yes Not sure
Own smart phone	YES NO	Yes No
Smart phone assistance	ASGMT STUDY NONE	For assignment For studying None
Sports activeness	LOW HIGH	A little active Very active
JAMB score	LOW AVE HIGH	Below 180 180 – 250 Above 250
Post-UTME score	LOW AVE HIGH	Below 180 180 – 250 Above 250
Average SSCE score	LOW AVE HIGH	Less than 4.00 4.00 – 4.99 5.00 and above

The dataset was split into two parts: one for training and the other for testing of the models to be built with the five machine-learning classifier algorithms chosen for testing:

- J48 decision tree;
- logistic regression (LR);
- multilayer perceptron (MLP);
- naïve Bayes (NB); and
- sequential minimal optimisation (SMO).

These algorithms were selected for testing based on their frequent use in existing similar studies (see literature reviewed in section 2) and their high rates of accuracy in such studies. The dataset also underwent a feature selection process, to retrieve the optimal features from the dataset for prediction of low academic performance, using the following four feature selection techniques:

- correlation;
- gain ratio;
- information gain; and
- ReliefF.

The next phase of the study was development of a predictive software application, using the best-performing machine-learning model (among the five tested) and the optimal features (as identified in the feature selection exercise). Finally, the study evaluated this predictive application with the test dataset to determine its effectiveness.

4. Testing of the five classifier models

In line with the objectives of this research, the study developed five machine-learning models. It started by splitting the entire dataset into two parts, using a 70:30 ratio. The first part, with 70% of the dataset, was for training, while the other 30% was for testing, i.e., testing the model built to ensure that the model performed well with unseen data. This study employed the resample filter available in WEKA for the splitting purpose. This filter ensured that the dataset split has no duplicates.

For the modelling process, the study used the WEKA tool to build models with the training dataset for the five selected machine-learning algorithms, namely J48 decision tree, LR, MLP, NB and SMO. Figures 2 to 6 below provide snapshots of the results from the modelling processes.

Figure 2: J48 decision tree model for training dataset

```
Time taken to build model: 0.07 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances        1566               95.3135 %
Incorrectly Classified Instances        77                4.6865 %
Kappa statistic                          0.8979
Mean absolute error                      0.0643
Root mean squared error                  0.2085
Relative absolute error                 13.8717 %
Root relative squared error             43.299  %
Total Number of Instances             1643

=== Detailed Accuracy By Class ===

               TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
               0.912    0.023    0.958      0.912    0.934      0.899    0.961     0.941     HL
               0.977    0.088    0.951      0.977    0.964      0.899    0.961     0.960     LL
Weighted Avg.  0.953    0.064    0.953      0.953    0.953      0.899    0.961     0.953

=== Confusion Matrix ===

   a    b   <-- classified as
 547   53 |    a = HL
  24 1019 |    b = LL
```

Figure 3: Logistic regression (LR) model for training dataset

```
Time taken to build model: 0.32 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances        1578               96.0438 %
Incorrectly Classified Instances        65                3.9562 %
Kappa statistic                          0.9143
Mean absolute error                      0.063
Root mean squared error                  0.1835
Relative absolute error                 13.5776 %
Root relative squared error             38.1076 %
Total Number of Instances             1643

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.935    0.025    0.956      0.935    0.945      0.914    0.982     0.979     HL
                 0.975    0.065    0.963      0.975    0.969      0.914    0.982     0.986     LL
Weighted Avg.    0.960    0.050    0.960      0.960    0.960      0.914    0.982     0.984

=== Confusion Matrix ===

    a    b   <-- classified as
  561   39 |    a = HL
   26 1017 |    b = LL
```

Figure 4: Multilayer perceptron (MLP) model for training dataset

```
Time taken to build model: 51.48 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances        1596               97.1394 %
Incorrectly Classified Instances        47                2.8606 %
Kappa statistic                          0.9381
Mean absolute error                      0.0302
Root mean squared error                  0.156
Relative absolute error                  6.5135 %
Root relative squared error             32.3991 %
Total Number of Instances             1643

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.952    0.017    0.969      0.952    0.960      0.938    0.992     0.991     HL
                 0.983    0.048    0.972      0.983    0.978      0.938    0.992     0.993     LL
Weighted Avg.    0.971    0.037    0.971      0.971    0.971      0.938    0.992     0.992

=== Confusion Matrix ===

    a    b   <-- classified as
  571   29 |    a = HL
   18 1025 |    b = LL
```

Figure 5: Naïve Bayes (NB) model for training dataset

```
Time taken to build model: 0 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances        1461               88.9227 %
Incorrectly Classified Instances       182               11.0773 %
Kappa statistic                          0.7601
Mean absolute error                      0.1164
Root mean squared error                  0.2901
Relative absolute error                 25.0931 %
Root relative squared error             60.2594 %
Total Number of Instances             1643

=== Detailed Accuracy By Class ===

               TP Rate   FP Rate   Precision   Recall  F-Measure  MCC      ROC Area  PRC Area  Class
               0.838     0.081     0.855       0.838   0.847      0.760    0.945     0.939     HL
               0.919     0.162     0.908       0.919   0.913      0.760    0.945     0.958     LL
Weighted Avg.  0.889     0.132     0.889       0.889   0.889      0.760    0.945     0.951

=== Confusion Matrix ===

   a   b   <-- classified as
 503  97 |   a = HL
  85 958 |   b = LL
```

Figure 6: Sequential minimal optimisation (SMO) model for training dataset

```
Time taken to build model: 1.19 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances        1580               96.1656 %
Incorrectly Classified Instances        63                3.8344 %
Kappa statistic                          0.9164
Mean absolute error                      0.0383
Root mean squared error                  0.1958
Relative absolute error                  8.2693 %
Root relative squared error             40.6697 %
Total Number of Instances             1643

=== Detailed Accuracy By Class ===

               TP Rate   FP Rate   Precision   Recall  F-Measure  MCC      ROC Area  PRC Area  Class
               0.923     0.016     0.970       0.923   0.946      0.917    0.954     0.924     HL
               0.984     0.077     0.957       0.984   0.970      0.917    0.954     0.952     LL
Weighted Avg.  0.962     0.055     0.962       0.962   0.961      0.917    0.954     0.942

=== Confusion Matrix ===

    a    b   <-- classified as
  554   46 |    a = HL
   17 1026 |    b = LL
```

After successful modelling of the training dataset, the study used the testing dataset to test the models built to confirm how well each built model performed with the unseen dataset. Figures 7 to 11 are snapshots of the results from this process.

Figure 7: J48 decision tree model for testing dataset

```
=== Re-evaluation on test set ===

User supplied test set
Relation:     NewDataset2-weka.filters.unsupervised.instance.Resample-S1-Z70.0-no-replacement-V
Instances:     unknown (yet). Reading incrementally
Attributes:   25

=== Summary ===

Correctly Classified Instances         684               97.0213 %
Incorrectly Classified Instances        21                2.9787 %
Kappa statistic                          0.9349
Mean absolute error                      0.0443
Root mean squared error                  0.1685
Total Number of Instances              705

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.934    0.009    0.984      0.934    0.958      0.936    0.976     0.967     HL
                 0.991    0.066    0.963      0.991    0.977      0.936    0.976     0.977     LL
Weighted Avg.    0.970    0.046    0.971      0.970    0.970      0.936    0.976     0.973

=== Confusion Matrix ===

   a   b   <-- classified as
 239  17 |   a = HL
   4 445 |   b = LL
```

Figure 8: Logistic regression (LR) model for testing dataset

```
=== Re-evaluation on test set ===

User supplied test set
Relation:     NewDataset2-weka.filters.unsupervised.instance.Resample-S1-Z70.0-no-replacement-V
Instances:     unknown (yet). Reading incrementally
Attributes:   25

=== Summary ===

Correctly Classified Instances         680               96.4539 %
Incorrectly Classified Instances        25                3.5461 %
Kappa statistic                          0.9227
Mean absolute error                      0.0608
Root mean squared error                  0.1724
Total Number of Instances              705

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.934    0.018    0.968      0.934    0.950      0.923    0.988     0.986     HL
                 0.982    0.066    0.963      0.982    0.972      0.923    0.988     0.992     LL
Weighted Avg.    0.965    0.049    0.965      0.965    0.964      0.923    0.988     0.990

=== Confusion Matrix ===

   a   b   <-- classified as
 239  17 |   a = HL
   8 441 |   b = LL
```

Figure 9: Multilayer perceptron (MLP) model for testing dataset

```
=== Re-evaluation on test set ===

User supplied test set
Relation:     NewDataset2-weka.filters.unsupervised.instance.Resample-S1-Z70.0-no-replacement-V
Instances:     unknown (yet). Reading incrementally
Attributes:   25

=== Summary ===

Correctly Classified Instances         693               98.2979 %
Incorrectly Classified Instances        12                1.7021 %
Kappa statistic                          0.9631
Mean absolute error                      0.0195
Root mean squared error                  0.1205
Total Number of Instances              705

=== Detailed Accuracy By Class ===

               TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
               0.969    0.009    0.984      0.969    0.976      0.963    0.998     0.997     HL
               0.991    0.031    0.982      0.991    0.987      0.963    0.998     0.999     LL
Weighted Avg.  0.983    0.023    0.983      0.983    0.983      0.963    0.998     0.998

=== Confusion Matrix ===

   a   b   <-- classified as
 248   8 |   a = HL
   4 445 |   b = LL
```

Figure 10: Naïve Bayes (NB) model for testing dataset

```
=== Re-evaluation on test set ===

User supplied test set
Relation:     NewDataset2-weka.filters.unsupervised.instance.Resample-S1-Z70.0-no-replacem
Instances:     unknown (yet). Reading incrementally
Attributes:   25

=== Summary ===

Correctly Classified Instances         629               89.2199 %
Incorrectly Classified Instances        76               10.7801 %
Kappa statistic                          0.7673
Mean absolute error                      0.1168
Root mean squared error                  0.2918
Total Number of Instances              705

=== Detailed Accuracy By Class ===

               TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC A
               0.855    0.087    0.849      0.855    0.852      0.767    0.944     0.938
               0.913    0.145    0.917      0.913    0.915      0.767    0.944     0.956
Weighted Avg.  0.892    0.124    0.892      0.892    0.892      0.767    0.944     0.950

=== Confusion Matrix ===
```

Figure 11: Sequential minimal optimisation (SMO) model for testing dataset

```
=== Re-evaluation on test set ===

User supplied test set
Relation:     NewDataset2-weka.filters.unsupervised.instance.Resample-S1-Z70.0-no-replacement-V
Instances:     unknown (yet). Reading incrementally
Attributes:   25

=== Summary ===

Correctly Classified Instances         681               96.5957 %
Incorrectly Classified Instances        24                3.4043 %
Kappa statistic                          0.9256
Mean absolute error                      0.034
Root mean squared error                  0.1845
Total Number of Instances              705

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.930    0.013    0.975      0.930    0.952      0.926    0.958     0.932     HL
                 0.987    0.070    0.961      0.987    0.974      0.926    0.958     0.957     LL
Weighted Avg.    0.966    0.050    0.966      0.966    0.966      0.926    0.958     0.948

=== Confusion Matrix ===

   a   b   <-- classified as
 238  18 |   a = HL
   6 443 |   b = LL
```

Model performance

This sub-section presents the performance of the trained and tested models by exploring the number of correctly classified datasets and looking at performance in terms of six metrics.

Table 2: Comparison of models' performance on training dataset, based on correctly and incorrectly classified student data

Metric	J48	LR	MLP	NB	SMO
Correctly classified students	1566	1578	**1596**	1461	1580
Incorrectly classified students	77	65	**47**	182	63
Correctly classified HL students	547	561	**571**	503	554
Incorrectly classified HL students	53	39	**29**	97	46
Correctly classified LL students	1019	1017	1025	958	**1026**
Incorrectly classified LL students	24	26	18	85	**17**

Table 2 compares the performance of the classifier models based on correctly and incorrectly classified student data for the training dataset. The acronyms "LL" and "HL" in the table stand for "LowLow" and "HighLow", which were this study's two hyperparameters, i.e., parameters whose values controlled the learning process. Each machine-learning model had to deliver either an LL or HL classification for each student record. The focus in this study was on classifying low-performing students

into two groups—HL and LL—with the idea being that students in the LL group would be the ones to receive urgent assistance to improve their chances of performing well. As seen in the table, the MLP algorithm correctly classified the highest number of students (1,596) for the entire training dataset and misclassified the lowest number of students (47). SMO had the next best performance, with 1,580 records classified correctly and 63 misclassifications. LR followed SMO with a classification margin of two less than SMO: 1,578 records were correctly classified, and 65 records were incorrectly classified. J48 correctly classified 1,566 student records and misclassified 77 records. NB showed the poorest performance, with 1,461 correctly classified records and 182 misclassified records.

For the performance of the algorithms in correctly classifying students in the HL class, MLP outperformed the other algorithms, while for the performance of the algorithms in correctly classifying students in the LL class, SMO performed the best. However, the difference in performance between SMO and MLP in correctly classifying students in the LL class was just one record.

Table 3: Comparison of the models' performance on the training dataset, using six selected metrics

Model	recall	specificity	ROC curve	F-measure	kappa	RMSE
J48	0.977	0.912	0.961	0.964	0.8979	0.2085
LR	0.975	0.935	0.982	0.969	0.9143	0.1835
MLP	0.983	**0.952**	**0.992**	**0.978**	**0.9381**	**0.1560**
NB	0.919	0.838	0.945	0.913	0.7601	0.2901
SMO	**0.984**	0.923	0.954	0.970	0.9164	0.1958

Table 3 shows the five models' performance on the training dataset in terms of six selected metrics:

Recall is the proportions of positive instances that are correctly classified. It is often called sensitivity.

$$\text{Recall} = \frac{TP}{TP + FN}$$

Specificity is the proportion of true negatives that are correctly predicted by the model.

$$\text{Specificity} = = \frac{TN}{TN + FP}$$

The **receiver operating characteristic (ROC) curve** gives the performance of a classification model at all classification thresholds.

True positive rate

False positive rate

True positive rate (TPR) is a synonym for recall and is therefore defined as follows:

$$\text{True positive rate (TPR)} = \frac{TP}{TP + FN}$$

False positive rate (FPR) is defined as follows:

$$\text{False positive rate (FPR)} = \frac{FP}{TF + TN}$$

Therefore, the ROC curve plots TPR vs. FPR at different classification thresholds.

F-measure is a measure that conveys the balance between recall and precision and can be said to be the harmonic mean of recall and precision.

$$\text{F-measure} = \frac{2 * Precision * Recall}{Precision + Recall}$$

The **kappa value** is a statistic used to control only those instances that may have been correctly classified by chance.

$$\text{Kappa can be calculated as: Kappa} = \frac{total\ accuracy - random\ accuracy}{1 - random\ accuracy}$$

The **root mean square error (RMSE)** is a standard way to measure the error of a model in predicting quantitative data. RMSE is calculated as:

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(\hat{y}_i - y_i)^2}{n}}$$

where,

$\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_n$ are predicted values

$y_1, y_2, \ldots, y_n$ are observed values

n is the number of observations

As seen in the results in Table 3, SMO had the best performance (98.4%) in terms of the recall metric, followed closely by MLP (98.3%). For all of the other five metrics—specificity, ROC curve, F-measure, kappa and RMSE—MLP performed best. For all six metrics, the weakest performer was NB.

After modelling the training dataset, the study compared the performance of the five algorithms on the testing dataset. Table 4 shows the results of the comparison of the performance based on correctly and incorrectly classified data, and Table 5 shows the performance in terms of the six metrics.

Table 4: Comparison of models' performance on test dataset, based on correctly and incorrectly classified student data

	J48	LR	MLP	NB	SMO
Right All	684	680	**693**	629	681
Wrong All	21	25	**12**	76	24
Right HL	239	239	**248**	219	238
Wrong HL	17	17	**8**	37	18
Right LL	**445**	441	**445**	410	443
Wrong LL	**4**	8	**4**	39	6

As seen in Table 4, the MLP algorithm performed best, correctly classifying the entire dataset, and achieved the best performance for classifying students in the HL and LL groups. Overall across the metrics, as seen in the findings presented up to this point, MLP was the best performer.

Table 5: Comparison of the models: Performance on the test dataset, using the six selected metrics

Model	recall	specificity	ROC curve	F-measure	kappa	RMSE
J48	**0.991**	0.934	0.976	0.977	0.9349	0.1685
LR	0.982	0.934	0.988	0.972	0.9227	0.1724
MLP	**0.991**	**0.969**	**0.998**	**0.987**	**0.9631**	**0.1205**
NB	0.913	0.855	0.944	0.915	0.7673	0.2918
SMO	0.987	0.930	0.958	0.974	0.9256	0.1845

As seen in Table 5, for the recall metric, MLP and J48 both achieved the highest value (99.1%). For the other five metrics—specificity, ROC curve, F-measure, kappa and RMSE—MLP was the best performer. For all six metrics, NB scored worst.

5. Feature selection

Feature selection in data mining helps models by looking out for the most important features in the dataset, reducing complexities and increasing the accuracy of the model (Neumann et al., 2016). This study ranked the attributes in order of importance using four feature selection techniques available in WEKA, namely: correlation, gain ratio, information gain, and ReliefF. Figures 12 to 15 are snapshots of the results.

Figure 12: Results from correlation feature selection technique

```
=== Attribute selection 10 fold cross-validation (stratified), seed: 1 ===

average merit      average rank  attribute
 0.51  +- 0.007      1   +- 0        8 SponQual
 0.426 +- 0.005      2.4 +- 0.49     5 SecType
 0.424 +- 0.009      2.6 +- 0.49    12 WorkStudy
 0.381 +- 0.007      4.4 +- 0.49    13 UniAcc
 0.382 +- 0.007      4.6 +- 0.49     6 SecArea
 0.321 +- 0.006      6   +- 0       21 SptAc
 0.293 +- 0.006      7.4 +- 0.49    14 BeAdmYrs
 0.293 +- 0.006      7.6 +- 0.49     7 SponType
 0.273 +- 0.004      9.6 +- 0.49    10 SponSup
 0.27  +- 0.008      9.8 +- 0.87    15 JambCou
 0.261 +- 0.009     10.8 +- 0.87     2 Age
 0.252 +- 0.004     11.8 +- 0.4     17 WkStud
 0.222 +- 0.004     13.4 +- 0.49    24 AveSc
 0.219 +- 0.007     13.7 +- 0.64     9 SponInc
 0.203 +- 0.005     15.3 +- 0.46    22 JambSc
 0.199 +- 0.008     15.7 +- 0.78    19 SmPhn
 0.188 +- 0.008     16.9 +- 0.3     18 PgDeg
 0.162 +- 0.004     18   +- 0       23 PumeSc
 0.151 +- 0.004     19   +- 0        1 Sex
 0.125 +- 0.002     20.1 +- 0.3     16 CouInt
 0.119 +- 0.004     20.9 +- 0.3     11 FamSize
 0.088 +- 0.004     22   +- 0       20 SmPhnAss
 0.045 +- 0.007     23   +- 0        4 AtPri
 0.01  +- 0.007     24   +- 0        3 MarStat
```

Figure 12 gives the ranking of the features (attributes) in order of contribution to the low performance of students, ranked from highest to lowest, as determined by the correlation feature selection technique. The four most-correlating attributes were: sponsor qualification, secondary school type, work and study, and university accommodation. The four least-correlating attributes were: family size, smart phone assistance, attended primary school, and marital status.

Figure 13: Results from gain ratio feature selection technique

```
=== Attribute selection 10 fold cross-validation (stratified), seed: 1 ===

average merit      average rank  attribute
 0.194 +- 0.006      1   +- 0        8 SponQual
 0.142 +- 0.006      2.4 +- 0.49    12 WorkStudy
 0.136 +- 0.003      3.2 +- 0.98     5 SecType
 0.132 +- 0.002      3.5 +- 0.67    24 AveSc
 0.126 +- 0.003      5   +- 0.45    17 WkStud
 0.119 +- 0.005      6   +- 0.45    13 UniAcc
 0.109 +- 0.004      6.9 +- 0.3      6 SecArea
 0.097 +- 0.003      8.2 +- 0.4      7 SponType
 0.092 +- 0.003      8.9 +- 0.54    22 JambSc
 0.085 +- 0.003      9.9 +- 0.3     10 SponSup
 0.077 +- 0.003     11.2 +- 0.4     21 SptAc
 0.075 +- 0.003     11.9 +- 0.54     9 SponInc
 0.069 +- 0.004     12.9 +- 0.3     18 PgDeg
 0.061 +- 0.002     14.2 +- 0.4     14 BeAdmYrs
 0.054 +- 0.004     15.4 +- 0.66     2 Age
 0.053 +- 0.003     15.5 +- 0.81    15 JambCou
 0.047 +- 0.004     17.4 +- 0.8     19 SmPhn
 0.044 +- 0.002     18   +- 0.63    23 PumeSc
 0.042 +- 0.002     18.5 +- 0.81    16 CouInt
 0.019 +- 0.002     20.1 +- 0.3     20 SmPhnAss
 0.016 +- 0.001     20.9 +- 0.3      1 Sex
 0.012 +- 0.001     22.1 +- 0.3     11 FamSize
 0.009 +- 0.003     22.9 +- 0.3      4 AtPri
 0     +- 0.001     24   +- 0        3 MarStat
```

Figure 13 provides the feature rankings generated by the gain ratio technique. The strongest (most-correlating) four attributes were: sponsor qualification, work and study, secondary school type, and average SSCE score. The weakest (least-correlating) four attributes were: sex, family size, attended primary school, and marital status.

Figure 14: Results from information gain feature selection technique

```
=== Attribute selection 10 fold cross-validation (stratified), seed: 1 ===

average merit      average rank  attribute
 0.223 +- 0.006      1   +- 0        8 SponQual
 0.194 +- 0.005      2   +- 0       17 WkStud
 0.18  +- 0.004      3   +- 0       24 AveSc
 0.145 +- 0.004      4.1 +- 0.3      7 SponType
 0.135 +- 0.005      5.8 +- 1.17    22 JambSc
 0.135 +- 0.005      5.8 +- 0.75    10 SponSup
 0.132 +- 0.003      6.8 +- 0.98     5 SecType
 0.129 +- 0.006      7.5 +- 0.67    12 WorkStudy
 0.113 +- 0.004      9.1 +- 0.3      9 SponInc
 0.106 +- 0.004     10.3 +- 0.46     6 SecArea
 0.103 +- 0.004     10.8 +- 0.87    13 UniAcc
 0.092 +- 0.006     11.8 +- 0.4     18 PgDeg
 0.078 +- 0.003     13.3 +- 0.46    14 BeAdmYrs
 0.077 +- 0.003     13.7 +- 0.46    21 SptAc
 0.067 +- 0.003     15.2 +- 0.4     23 PumeSc
 0.062 +- 0.003     15.8 +- 0.4     16 CouInt
 0.053 +- 0.003     17.2 +- 0.4     15 JambCou
 0.048 +- 0.003     17.8 +- 0.4      2 Age
 0.028 +- 0.002     19   +- 0       20 SmPhnAss
 0.027 +- 0.002     20   +- 0       19 SmPhn
 0.019 +- 0.001     21.1 +- 0.3     11 FamSize
 0.016 +- 0.001     21.9 +- 0.3      1 Sex
 0.002 +- 0         23   +- 0        4 AtPri
 0     +- 0         24   +- 0        3 MarStat
```

Figure 14 presents the features ranked by the information gain technique. The four strongest attributes were found to be: sponsor qualification, weekly study time, sverage SSCE score, and sponsor type. The weakest attributes were: family size, sex, attended primary school, and marital status.

Figure 15: Results from ReliefF feature selection technique

```
=== Attribute selection 10 fold cross-validation (stratified), seed: 1 ===

average merit      average rank  attribute
 0.375 +- 0.009      1   +- 0       11 FamSize
 0.353 +- 0.008      2   +- 0        7 SponType
 0.309 +- 0.008      3.6 +- 0.66    17 WkStud
 0.3   +- 0.01       4.1 +- 1.04     8 SponQual
 0.3   +- 0.005      4.4 +- 0.66    10 SponSup
 0.283 +- 0.009      6.5 +- 0.92     2 Age
 0.278 +- 0.006      6.9 +- 0.7     14 BeAdmYrs
 0.274 +- 0.007      7.5 +- 0.67    22 JambSc
 0.26  +- 0.006      9.2 +- 0.6      9 SponInc
 0.251 +- 0.008     10.7 +- 1.1     24 AveSc
 0.246 +- 0.006     11.1 +- 0.94    23 PumeSc
 0.241 +- 0.008     11.8 +- 1.08    16 CouInt
 0.234 +- 0.007     12.9 +- 1.45    18 PgDeg
 0.228 +- 0.008     14.2 +- 1.54    21 SptAc
 0.219 +- 0.009     15.8 +- 1.33    20 SmPhnAss
 0.218 +- 0.01      15.8 +- 1.25     5 SecType
 0.219 +- 0.007     15.9 +- 1.3      6 SecArea
 0.199 +- 0.007     17.9 +- 0.94    13 UniAcc
 0.193 +- 0.006     18.8 +- 0.6      1 Sex
 0.181 +- 0.005     20.1 +- 0.54    12 WorkStudy
 0.168 +- 0.009     20.8 +- 0.4     15 JambCou
 0.075 +- 0.003     22   +- 0       19 SmPhn
 0.057 +- 0.006     23   +- 0        3 MarStat
 0.017 +- 0.003     24   +- 0        4 AtPri
```

As shown in Figure 15, use of the ReliefF technique generated these four strongest attributes: family size, sponsor type, seekly study time, and sponsor qualification. The weakest features were: course from JAMB, own smart phone, marital status, and attended primary school.

The results as set out in Figures 11 to 14 indicated that, across all the techniques used, certain attributes were consistently identified as most relevant, meaning they were likely to contribute greatly to predicting a student's classification as HL or LL. At the same time, there were certain attributes that consistently emerged as the least relevant, suggesting that these features contribute least to predicting a student's classification.

Accordingly, we tested the performance of the five classifier algorithms through successive modelling of the features in each feature selection technique starting from the top four attributes. The ROC curve and RMSE values, which are two widely used metrics (Caruana & Niculescu-Mizil, 2004), were used to evaluate the performance of the models. Table 6 provides a summary of the results obtained.

Table 6: Performance summary of feature selection algorithms

Algorithm	ROC curve	RMSE	Best range
correlation	0.993	0.1410	15–21
gain ratio	0.994	0.1471	14–17
information gain	0.997	0.1382	15–16
ReliefF	0.995	0.1416	18–20

Table 6 shows that all algorithms achieved a higher level of performance on the reduced dataset than on the complete dataset. The information gain feature selection algorithm had the highest performance, with an ROC curve value of 99.7% and an RMSE value of 13.82%, and with the features identified as most relevant ranging between 15 and 16 of the features in Table 1.

Performance of MLP classifier using the 16 most relevant attributes

Since MLP had proven to be the best classifier, we built a new MLP model, now using only the 16 attributes determined to be most relevant by the information gain feature selection technique. The modelling process began by extracting the selected attributes in WEKA and then using the MLP classifier to build the model. Figure 16 shows the model built, and Table 7 presents the performance of the model using the 16 most relevant attributes identified.

Figure 16: MLP model built with 16 most relevant attributes, using the training dataset

```
Time taken to build model: 6.92 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances        1606               97.748  %
Incorrectly Classified Instances        37                2.252  %
Kappa statistic                          0.9513
Mean absolute error                      0.0269
Root mean squared error                  0.1382
Relative absolute error                  5.7956 %
Root relative squared error             28.7039 %
Total Number of Instances             1643

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.965    0.015    0.973      0.965    0.969      0.951    0.997     0.995     HL
                 0.985    0.035    0.980      0.985    0.982      0.951    0.997     0.998     LL
Weighted Avg.    0.977    0.028    0.977      0.977    0.977      0.951    0.997     0.997

=== Confusion Matrix ===

    a    b   <-- classified as
  579   21 |    a = HL
   16 1027 |    b = LL
```

Table 7: Performance of MLP using 16 most relevant attributes dataset

Item	Value	Metric	Value
Right All	1606	recall	0.985
Wrong All	37	specificity	0.965
Right HL	579	ROC curve	0.997
Wrong HL	21	F-Measure	0.982
Right LL	1027	kappa	0.9513
Wrong LL	16	RMSE	0.1382

In Table 7, all the metric values for the 16 most relevant attributes dataset show good performance. A comparison of these metrics values with those for the complete database of all attributes revealed the following: recall was 98.5% for the 16 most relevant attributes dataset, while it was 98.3% for the complete database; specificity was 96.5% for the most relevant attributes and 95.2% for the complete database; the ROC curve was 99.7% for the most relevant attributes and 99.2% for the complete database; the F-measure was 98.2% for the most relevant attributes and 97.8% for the complete database; the kappa value was 95.13% for the most relevant attributes and 93.81% for the complete database; and the RMSE was 13.82% for the most relevant attributes and 15.6% for the entire database.

The results showed improvement for every metric value using the 16 most relevant attributes when compared with using all 24 attributes. It was thus concluded that the MLP classifier, using the 16 most relevant features ranked with the information gain algorithm, was the best classifier for the dataset generated by this study. Accordingly, it was determined that it would be beneficial to design and implement the prediction application with the MLP classifier and the 16 identified attributes.

The 16 selected attributes

Table 8 shows the variable codes and values, and descriptions, of the 16 most relevant attributes identified from the modelling process. These were the features used for development of the predictive application.

Table 8: Selected attributes

Attribute	Variable codes	Values
Sponsor qualification	DEG NODEG NOEDU	Degree No degree No education
Weekly study time	LOW AVE HIGH	Below 10hrs 10 – 20hrs Above 20hrs
Average SSCE score	LOW AVE HIGH	Below 4.00 4.00 – 4.99 Above 5.00
Sponsor type	GUAD PAR SELF	Guardian Parents Self-sponsor
JAMB score	LOW AVE HIGH	Below 180 180 – 250 Above 250
Sponsor support	LOW MED HIGH	Little Average Great
Secondary school type	PRI PUB	Private Public
Work and study	YES NO	Yes No
Sponsor income (per month)	LOW MED HIGH	Below 50k 50k – 100k Above 100k
Secondary school area	URB RUR	Urban Rural
University accommodation	CMPS OFFCMPS	Campus Off-campus

Attribute	Variable codes	Values
Postgraduate degree	NO YES NS	No Yes Not sure
Years before admission (gap between completing secondary school and commencing university)	NONE B5 5A	None Below 5 5 above
Sports activeness	LOW HIGH	Little active Very active
Post-UTME score	LOW AVE HIGH	Below 180 180 – 250 Above 250
Course interest	LOW AVE HIGH	Little Average High

6. Development and evaluation of predictive software application

The predictive application was developed with the goal of creating a tool that could help Nigerian universities to identify potentially low-performing students—so as to be able to provide targeted intervention measures to address these students' needs.

Application development

The design process followed the following four steps: input students' features, which were the 16 most relevant attributes identified; use the best classifier algorithm identified, which was the MLP model; predict students' failure risk level as HL or LL; and output an intervention summary for students, with HL students requiring low intervention and LL students requiring high intervention. The main users of the software would be faculty officers in charge of gathering and storing students' information. Their role would be to use the application to obtain predictions and to forward results to the relevant authorities in the institution for the purpose of intervention in the form of one or more of: study load rescheduling, counselling, financial intervention in the form of bursary, and/or government assistance through various funding agencies for indigent students.

A 17th attribute was added (as seen below in Figure 17), called "matric number", to identify specific students whose information had been processed. This was the ID number of the student, unique to each student and eliminating the need to record student names in the application. The software interface was developed with the PHP programming language, while the part of the software that would handle the prediction used the Python programming language—incorporating the MLP model built in WEKA. Snapshots of the predictive application follow in Figures 17 to 19.

Figure 17 shows the welcome screen, which offers instructions to users on the relevant student details to gather, and informs users that only authorised persons can access the system.

Figure 17: Welcome screen

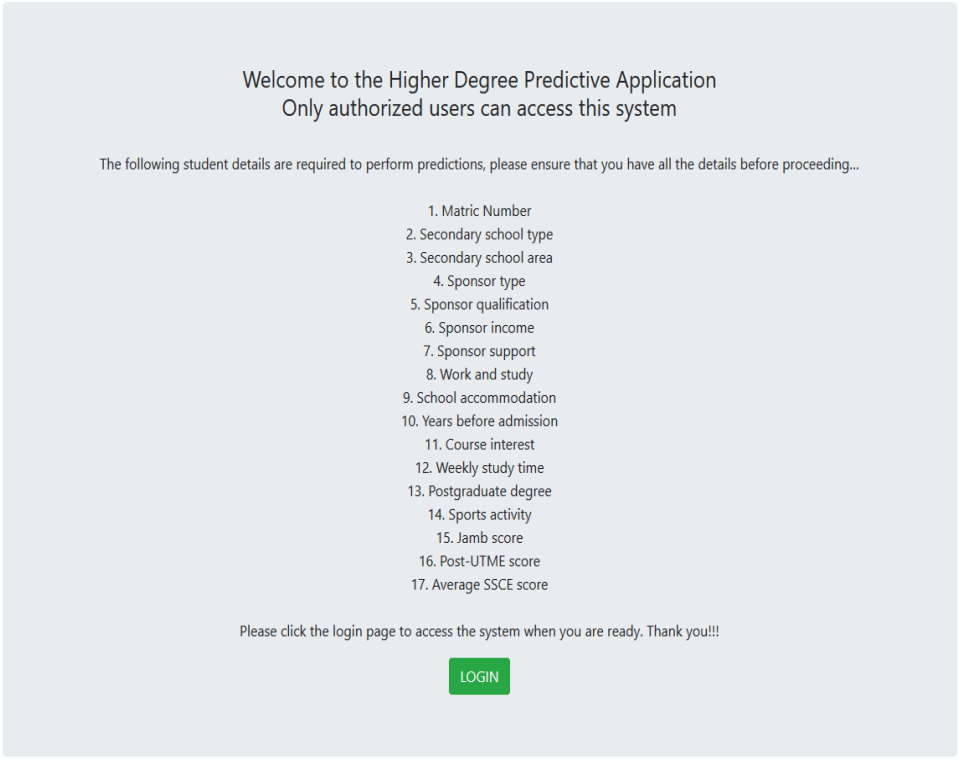


Figure 18 shows the user input page for the 16 most relevant attributes and the added "Matric number". Users select options from the dropdown menu for each feature. The page also has a button for viewing the predicted results, and another button for resetting the page.

Figure 18: User input page

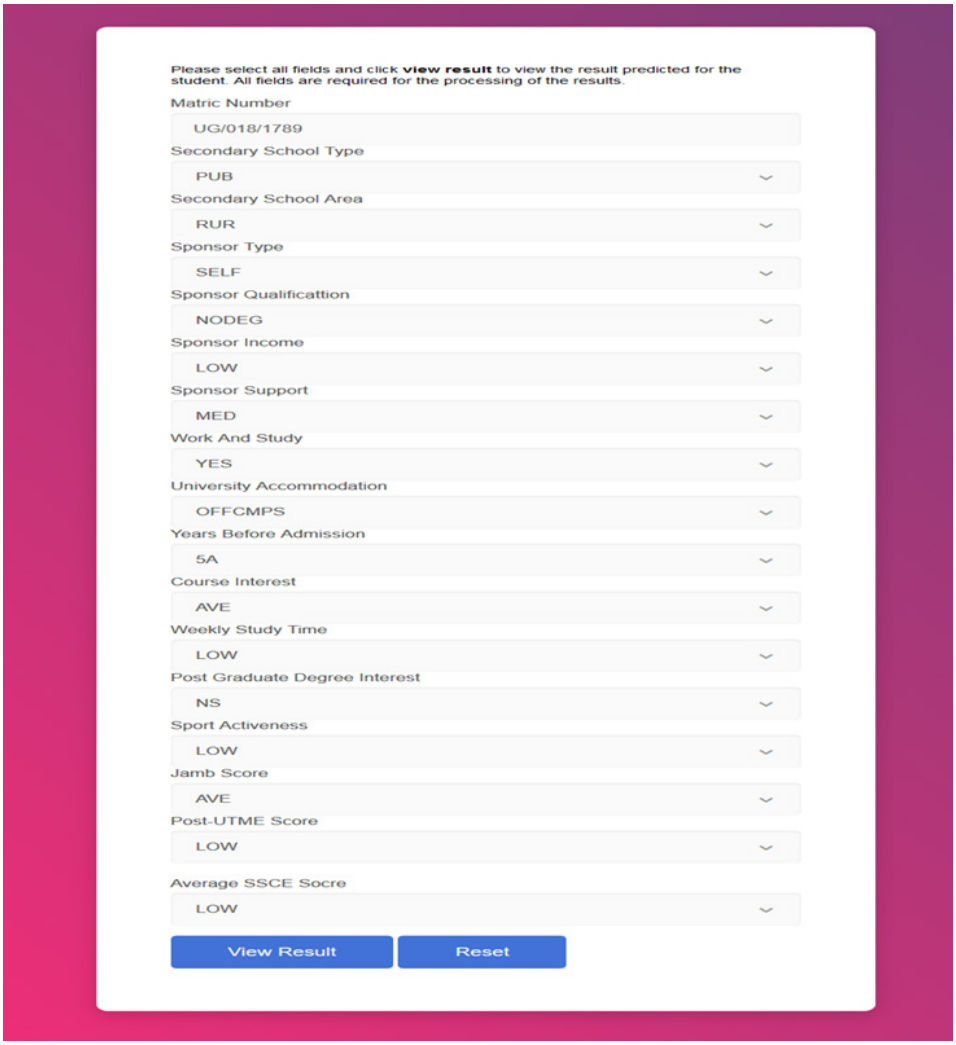
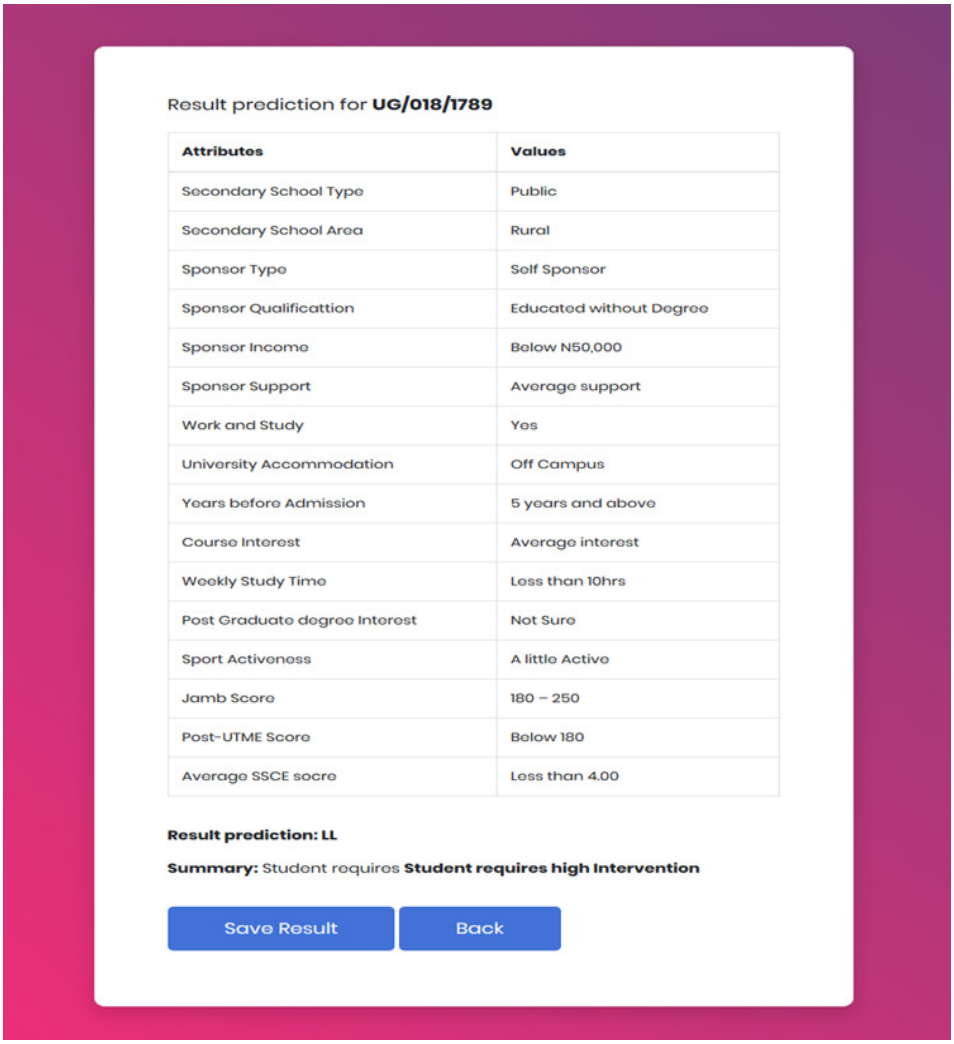


Figure 19 displays a "Result prediction" for a student, indicating (in this example case) an LL classification and specifying the need for "high intervention".

Figure 19: Prediction results page



Application evaluation

The software was evaluated in two phases. The first phase evaluated the software based on its requirements, and the second phase evaluated the software based on its performance. The first phase ensured that the application met the following requirements:

- Simple and interactive interface: Easy navigation between pages and straightforward instructions for users

- Secured system: Login interface to ensure that only authorised persons can access the software
- Avoidance of prediction error: Notifications to users of empty fields to ensure the application has received values for all 17 features before allowing the generation of a prediction
- Rapid results: Provision of results with a single click
- Provision of risk level: Offering a result prediction value for each student, and a risk intervention summary based on predicted value

The second phase of the evaluation involved testing the software using the 30% testing dataset, which contained 705 records. The process of evaluation began by entering the details of each record into the predictive application to obtain a prediction. Figure 20 shows a cross-section of the stored results in Excel, including "Actual" and "Predicted" values in terms of the "LowLow" and "HighLow" classifications.

Figure 20: Cross-section of results from testing dataset

MatNum	SecType	SecArea	SponType	SponQual	SponInc	Spon3_up	WorkStud	UniAcc	BeAdmYrs	CourIrt	WkStud	PgDeg	SptAc	JambSc	PumeSc	AveSc	Actual	Predicted	LowxLow	HighLow
UG/010/0301	P_b	Rur	Self	NoDeg	Med	Med	Yes	OffCmps	5A	Ave	Ave	Nis	Low	Low	Ave	Ave	LowLow	LL	A	
UG/010/0302	P_b	Rur	Self	NoDeg	Med	Low	Yes	OffCmps	5A	Ave	Ave	Nis	Low	Low	Low	Low	LowLow	LL	A	
UG/010/0303	P_b	Rur	Par	NoDeg	Med	High	No	OffCmps	B5	Ave	Ave	Nis	Low	Ave	High	Ave	LowLow	LL	A	
UG/010/0304	Pri	Urb	G_ad	Deg	High	High	No	OffCmps	B5	High	Ave	Yes	Low	High	Ave	High	HighLow	HL		A
UG/010/0305	Pri	Rur	Self	NoDeg	Med	Low	Yes	OffCmps	5A	Ave	Low	Nis	Low	Low	Ave	Ave	LowLow	LL	A	
UG/010/0306	Pri	Rur	G_ad	Deg	High	Med	Yes	Cmps	5A	Ave	High	Nis	High	Low	Low	Low	HighLow	HL		A
UG/010/0307	Pri	Rur	Self	NoDeg	Med	Low	Yes	OffCmps	B5	Ave	Low	Nis	High	Low	Ave	Low	LowLow	HL	D	
UG/010/0308	P_b	Rur	Self	NoDeg	Med	Low	Yes	Cmps	5A	Ave	Ave	Nis	High	Ave	Ave	Ave	LowLow	LL	A	
UG/010/0309	P_b	Rur	Self	NoDeg	Low	Low	Yes	OffCmps	5A	Ave	Low	Nis	Low	Low	Low	Low	LowLow	LL	A	
UG/010/0310	P_b	Urb	G_ad	Deg	High	High	No	OffCmps	B5	Ave	Ave	Nis	High	Ave	High	High	HighLow	HL		A
UG/010/0311	P_b	Urb	G_ad	Deg	High	High	No	OffCmps	B5	Ave	Ave	Nis	High	Ave	High	High	HighLow	HL		A
UG/010/0312	Pri	Rur	Self	NoDeg	Med	Med	Yes	OffCmps	5A	Low	Ave	Nis	Low	High	Ave	Ave	LowLow	LL	A	
UG/010/0313	P_b	Urb	G_ad	Deg	High	High	No	Cmps	B5	High	High	Yes	High	Ave	High	High	HighLow	HL		A
UG/010/0314	Pri	Urb	G_ad	Deg	High	High	No	Cmps	B5	High	High	Yes	High	Ave	High	High	HighLow	HL		A
UG/010/0315	Pri	Rur	G_ad	Deg	High	Med	Yes	Cmps	5A	Ave	High	Nis	High	Low	Low	Ave	HighLow	HL		A
UG/010/0316	Pri	Urb	G_ad	Deg	High	High	No	OffCmps	B5	High	Low	Yes	Low	Ave	Low	Ave	HighLow	HL		A
UG/010/0317	P_b	Rur	Self	NoDeg	Med	Med	Yes	OffCmps	5A	Ave	Ave	Nis	Low	High	Ave	Ave	LowLow	LL	A	
UG/010/0318	P_b	Rur	Self	NoDeg	Med	Med	Yes	Cmps	5A	Ave	High	Nis	High	Low	Low	Ave	LowLow	LL	A	
UG/010/0319	P_b	Rur	G_ad	NoDeg	Low	Low	No	OffCmps	B5	Ave	Low	Nis	Low	Low	Low	Low	LowLow	LL	A	
UG/010/0320	Pri	Urb	G_ad	Deg	High	Med	No	Cmps	B5	High	High	Yes	High	Ave	High	Ave	HighLow	HL		A
UG/010/0321	P_b	Urb	Self	NoDeg	Med	Low	Yes	OffCmps	B5	High	High	Nis	Low	Ave	Ave	Ave	LowLow	LL	A	
UG/010/0322	P_b	Rur	Par	NoDeg	Med	High	No	OffCmps	B5	Ave	Low	Nis	Low	Ave	Low	Ave	LowLow	LL	A	
UG/010/0323	Pri	Urb	G_ad	Deg	High	High	No	Cmps	None	High	High	Yes	High	High	Ave	Ave	HighLow	HL		A
UG/010/0324	P_b	Urb	G_ad	Deg	High	High	No	OffCmps	B5	High	Ave	Yes	Low	High	Ave	Ave	HighLow	HL		A
UG/010/0325	P_b	Rur	G_ad	Deg	Med	Med	Yes	OffCmps	B5	Ave	High	Nis	Low	Ave	Low	Ave	HighLow	HL		A
UG/010/0326	P_b	Rur	Par	NoEdu	Med	Med	No	Cmps	5A	Low	Low	Nis	High	Ave	Low	Low	LowLow	LL	A	
UG/010/0327	Pri	Urb	G_ad	Deg	Med	High	No	Cmps	B5	High	High	Nis	High	High	High	Ave	HighLow	HL		A
UG/010/0328	P_b	Rur	Par	NoDeg	Med	High	Yes	OffCmps	B5	Ave	Ave	Nis	Low	Ave	High	Ave	LowLow	LL	A	
UG/010/0329	P_b	Rur	Self	NoDeg	Low	Low	Yes	OffCmps	5A	Ave	Ave	Nis	Low	Low	Low	Ave	LowLow	LL	A	

From the data in the spreadsheet shown in Figure 20, a confusion matrix was constructed, as represented in Table 9, to determine the accuracy of the predictive application when engaging with the testing dataset.

Table 9: Confusion matrix

	Actual LL	Actual HL
Predicted LL	446	11
Predicted HL	3	245

The confusion matrix in Table 9 shows that 446 out of 449 records were correctly predicted for the LL group, and 245 out of 256 records were correctly predicted for the HL group. Furthermore, using the values obtained from the confusion matrix, the values obtained for recall and sensitivity were both 99.3%, and the value obtained for specificity was 95.7%. For prevalence, the software obtained 63.7% for LL students and 36.3% for HL students, which meant that high-risk students were found to be twice the number of low-risk students in the institution. The software achieved accuracy of 98%, which demonstrated a high level of dependability. For precision, the software achieved 99.3%, and its F-measure was 98.5%. Thus, the performance obtained from an evaluation of the software showed that the application could accomplish its objectives.

7. Conclusion and recommendation

This study was grounded in an understanding that low student performance is a global challenge that affects every tertiary academic institution in the world—and that the challenge is particularly prominent in developing countries, including Nigeria. In looking for ways to monitor and improve the performance of students, the study explored machine learning as a technique to assist Nigerian universities in identifying the potential for low academic performance among undergraduate students. With a dataset collected from Niger Delta University, which comprised 24 attributes, this study identified the 16 features (student attributes) that give optimal prediction results when subjected to machine learning. The study also identified the MLP classifier algorithm as the best model for generating such predictions. Furthermore, the study developed a predictive software application for potential deployment by Nigerian universities, using the 16 attributes and MLP model identified.

It is recommended that future research should study the performance of the MLP model, and of predictive applications, using datasets from other Nigerian higher institutions—so as to work towards enabling a robust and unified system for all institutions in the country.

References

Abu Zohair, L. M. (2019). Prediction of student's performance by modelling small dataset size. *Journal of Educational Technology in Higher Education*, *16*, 27. https://doi.org/10.1186/s41239-019-0160-3

Adeyemi, A.M., & Adeyemi, S.B. (2014). Institutional factors as predictors of students' academic achievement in colleges of education in South Western Nigeria. *International Journal of Educational Administration and Policy Studies*, *6*(8), 141–153. https://doi.org/10.5897/IJEAPS2014.0342

Adeyemo, A. B., & Kuye, G. (2006). Mining students' academic performance using decision tree algorithms. *Journal of Information Technology Impact*, *6*(3), 161–170.

Alami, M. (2016). Causes of poor academic performance among Omani students. *International Journal of Social Science Research*, *4*(1), 126–136. https://doi.org/10.5296/ijssr.v4i1.8948

Algarni, A. (2016). Data mining in education. *International Journal of Advanced Computer Science and Applications, 7*(6), 456–461. https://doi.org/10.14569/IJACSA.2016.070659

Al-Zoubi, S. M., & Younes, M. A. B. (2015). Low academic achievement: Causes and results. *Theory and Practice in Language Studies*, *5*(11), 2262–2268. https://doi.org/10.17507/tpls.0511.09

Aryana, M. (2010). Relationship between self-esteem and academic achievement amongst pre-university students. *Journal of Applied Sciences*, *10*(20), 2474–2477. https://doi.org/10.3923/jas.2010.2474.2477

Banerjee, P. A. (2016). A systematic review of factors linked to poor academic performance of disadvantaged students in science and maths in schools. *Cogent Education*, *3*(1). https://doi.org/10.1080/2331186X.2016.1178441

Caruana, R., & Niculescu-Mizil, A. (2004). Data mining in metric space: An empirical analysis of supervised learning performance criteria. In *Proceedings of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 69–78). https://doi.org/10.1145/1014052.1014063

Dhilipan J., Vijayalakshmi N., Suriya S., & Arockiya, C. (2021). Prediction of students performance using machine learning. *IOP Conference Series: Materials Science and Engineering, 1055*, 012122. https://doi.org/10.1088/1757-899X/1055/1/012122

Ekubo, A. E. (2020). *Predictive system for characterizing low performance of undergraduate students using machine-learning techniques*. PhD thesis, North-West University, South Africa.

Farooq, M. S., Chaudhry, A. H., Shafiq, M., & Berhanu, G. (2011). Factors affecting students' quality of academic performance: A case of secondary school level. *Journal of Quality and Technology Management*, *7*(2), 1–14.

Flanagan, B., Majumdar, R., & Ogata, H. (2022). Early-warning prediction of student performance and engagement in open book assessment by reading behavior analysis. *International Journal of Educational Technology in Higher Education*, *19*(41), 1–23. https://doi.org/10.1186/s41239-022-00348-4

Frimpong, E. A., Agyeman, G. A., & Ofosu, F. F. (2016). Institutional factors affecting the academic performance of polytechnic students in Ghana. *International Journal of Humanities & Social Science Studies (IJHSSS)*, *2*(5), 102–109.

Glewwe, P., & Kremer, M. (2006). Schools, teachers, and education outcomes in developing countries. In E. Hanushek, & F. Welch (Eds.), *Handbook of the economics of education, vol. 2* (pp. 945–1017). https://doi.org/10.1016/S1574-0692(06)02016-2

Gullo, F. (2015). From patterns in data to knowledge discovery: What data mining can do. *Physics Procedia*, *62*, 18–22. https://doi.org/10.1016/j.phpro.2015.02.005

Hughes, G., & Dobbins, C. (2015). The utilization of data analysis techniques in predicting student performance in massive open online courses (MOOCs). *Research and Practice in Technology Enhanced Learning*, *10*(1), 1–18. https://doi.org/10.1186/s41039-015-0007-z

Li, X., Zhang, Y., Cheng, H., Li, M., & Yin, B. (2022). Student achievement prediction using deep neural network from multi-source campus data. *Complex and Intelligent Systems*, *8*, 5143–5156. https://doi.org/10.1007/s40747-022-00731-8

Longe, O. (2017). Graduate unemployment in Nigeria: Causes, consequences and remediable approaches. *American International Journal of Contemporary Research*, *7*(4), 63–73.

Mitra, D. (2011). Pennsylvania's best investment: The social and economic benefits of public education. Pennsylvania State University.

Mohammed, A. (2022). Data mining classification simplified: Key types, steps & 6 best classifiers. Hevo. https://hevodata.com/learn/data-mining-classification

Neumann, U., Riemenschneider, M., Sowa, J. P., Baars, T., Kälsch, J., Canbay, A., & Heider, D. (2016). Compensation of feature selection biases accompanied with improved predictive performance for binary classification by using a novel ensemble feature selection approach. *BioData Mining*, *9*(36), 1–14. https://doi.org/10.1186/s13040-016-0114-4

Nsiah, H. (2017). Fear of failure and the academic performance of students from low-income families. *International Journal of Education and Social Science, 4*(10), 19–26.

Nurmi, J. E., Aunola, K., Salmela-Aro, K., & Lindroos, M. (2003). The role of success expectation and task-avoidance in academic performance and satisfaction: Three studies on antecedents, consequences and correlates. *Contemporary Educational Psychology*, *28*(1), 59–90. https://doi.org/10.1016/S0361-476X(02)00014-0

Ofori, F., Maina, E., & Gitonga, R. (2020). Using machine learning algorithms to predict students' performance and improve learning outcome: A literature based review. *Journal of Information and Technology*, *4*(1), 33–55.

Ololube, N. P. (2013). The problems and approaches to educational planning in Nigeria: A theoretical observation. *Mediterranean Journal of Social Sciences*, *4*(12). https://doi.org/10.5901/mjss.2013.v4n12p37

Oyebade, S. A., & Dike, C. (2013). Restructuring Nigerian tertiary (university) education for better performance. Paper presented at the 11th Annual Meeting of the Bulgarian Comparative Education Society, Plovdiv, Bulgaria, 14-17 May.

Oyerinde, O. D., & Chia, P. A. (2017). Predicting students' academic performances – A learning analytics approach using multiple linear regression. *International Journal of Computer Applications*, *157*(4), 37–44. https://doi.org/10.5120/ijca2017912671

Romero, C., & Ventura, S. (2013). Data mining in education. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *3*(1), 12–27. https://doi.org/10.1002/widm.1075

Romero, C. B., Ventura, S. N., Pechenizkiy, M., & Baker, R. S. (Eds.) (2010). *Handbook of educational data mining*. CRC Press. https://doi.org/10.1201/b10274

Serdyukov, P. (2017). Innovation in education: What works, what doesn't, and what to do about it? *Journal of Research in Innovative Teaching & Learning*, *10*(1), 4–33. https://doi.org/10.1108/JRIT-10-2016-0007

Ugar, A. A. (2018). ASUU strike: The Federal Government and Nigerian educational system. *International Journal of Education and Research*, 6(5), 19–32.

UN Department of Economic and Social Affairs (UN DESA). (2013). *World economic and social survey 2013: Sustainable development challenges*.

Vergaray, A. D., Guerra C., Cervera, N., & Burgos, E. (2022). Predicting academic performance using a multiclassification model: Case study. *International Journal of Advanced Computer Science and Applications (IJACSA)*, *13*(9), 881–889. https://doi.org/10.14569/IJACSA.2022.01309102

Wirth, R., & Hipp, J., (2000). CRISP-DM: Towards a standard process model for data mining. In *Proceedings of the 4th International Conference on the Practical Applications of Knowledge Discovery and Data Mining* (pp. 29–39). Springer-Verlag.

Yağcı, M. (2022). Educational data mining: Prediction of students' academic performance using machine learning algorithms. *Smart Learning Environments*, *9*(1), 1–9. https://doi.org/10.1186/s40561-022-00192-z

Radio, mobile communications, and women's empowerment: Experiences in Mathare, Nairobi

Grace Githaiga
PhD Candidate, Institute for Development Studies, University of Nairobi
https://orcid.org/0000-0002-1383-3795

Norbert Wildermuth
Associate Professor, Department of Communication and Arts, Roskilde University, Denmark
https://orcid.org/0000-0003-0564-5214

Abstract
This article presents findings from a study of young women's empowerment through the use of information and communication technology (ICT), specifically the use of radio and mobile communications, in the Mathare informal settlements of Nairobi. Data was collected through focus group discussions and in-depth interviews with young women. The findings were extrapolated through the lens of Longwe's five-stage women's empowerment framework (Longwe, 1991). Longwe posits that the empowerment of women can be viewed through a prism of ascension, the initial stage being *welfare*, followed by *access*, *conscientisation*, *participation*, and then *control*, the highest level of empowerment. The study finds that the use of radio and mobile communications has empowered many young women to engage in conversations on issues affecting them, and with useful information on how to improve their means of livelihood. At the same time, mobile communications have enabled many young women to achieve financial inclusion by engaging in income-generating activities. In addition, mobile communications have empowered young women by providing them with access to financial services and the ability to manage their finances. Such empowerment, when viewed through the Longwe (1991) framework, is a remarkable progression on empowerment by young women through the initial stages up to the highest levels of *participation* and *control*. At the same time, the study finds that the empowerment of young women through ICTs is held back to some extent by socioeconomic and cultural factors that are the result of patriarchal traditions and mindsets.

Keywords
information and communication technology (ICT), radio, mobile communications, empowerment, young women, informal settlements, Mathare, Nairobi

Acknowledgement
This article draws on elements of Githaiga's PhD research. The authors thank the young women of Mathare Pioneer Youth Group who generously shared information on their empowerment through ICTs.

DOI: https://doi.org/10.23962/ajic.i30.13753

Recommended citation
Githaiga, G., & Wildermuth, N. (2022). Radio, mobile communications, and women's empowerment: Experiences in Mathare, Nairobi. *The African Journal of Information and Communication (AJIC), 30*, 1-22. https://doi.org/10.23962/ajic.i30.13753

1. Introduction and context

Harnessing the transformative potential of information and communication technologies (ICTs) is widely acknowledged as essential to the process of women's empowerment (Nangooba, 2020; Niyonzima & Bhuju, 2021; Varriale et al., 2022). ICTs are considered in this study as a broad category of technological tools used to create, preserve, manage, transmit, and receive data with the intention of meeting a variety of human needs (Rahman, 2014). This study is focused on the use of two particular ICTs: radio and mobile communications.

Batliwala (1994) views women's empowerment as the ability of women to influence external actions relating to their welfare and well-being. In this case, empowerment means that women gain control, and the capacity to make choices individually or collectively, and then transform these choices into desired outcomes. Women's empowerment is manifest in interpersonal socio-economic, cultural, and political dimensions (Singh & Vanka, 2020). The scope of empowerment varies, depending on the realities of the women in specific contexts. For instance, empowerment is dependent on whether women have access to ICTs. This implies that access is dictated by prevailing social, economic, and cultural circumstances. These realities underpinned the decision to use, in this study, the Longwe (1991) gender empowerment framework, which views women's empowerment as progressing through five stages, from the base levels of *welfare* and *access* to the optimum levels of *conscientisation*, *participation*, and *control*.

The context of this study is the Mathare informal settlements, the second largest cluster of informal settlements, after Kibera, in Kenya's capital city of Nairobi. It has a population of about 206,564 people, according to the Kenya Population and Housing Census (Kenya National Bureau of Statistics, 2019). The Mathare informal settlements are characterised by hundreds of structures, densely packed and laid out without adhering to essential construction guidelines. Many residents work in the informal business sector inside the settlements and also in the surrounding estates (UN Habitat, 2020). Unemployment, abuse of drugs and alcohol, gang culture, lack of essential services, and other forms of deprivations are rampant in the settlements, making life difficult, especially for young, vulnerable women (Kovačič & Lundine, 2014).

Research by Wamuyu (2017) in Mathare found that the settlements had poor access to digital infrastructure. A digital gender audit of Nairobi slums, including Mathare, by the World Wide Web Foundation (2015) established that only 20% of women could access the internet, compared to 57% of men. Komen and Ling (2021) found that young women in Mathare face difficulties in accessing certain ICTs, mainly because of traditional gender roles and stereotypes in Kenya—based on a patriarchal culture that does not promote gender equity. Against the background of rapid digitalisation by the Government of Kenya, this study sought to uncover the extent to which young women in the Mathare informal settlements are harnessing the power of ICTs, specifically radio and mobile communications.

2. Literature review and theoretical framework

Empowerment through radio

Radio plays an essential role in the dissemination of information. Radio communication can also be a powerful tool for amplifying the voices of women by providing them a platform to speak out about the issues that matter to them (Rimmer, 2021; Tijani-Adenle, 2022). Debates on women's empowerment through radio often focus on how the medium can be used to improve access by women to information and education, as well as their participation in public spheres (Varriale et al., 2022; Wei et al., 2021). Ekwok (2018) and Simmons (2019) argue that radio communication, in particular, is well-suited to reach out to women in rural and marginalised areas, especially those who may be illiterate or have limited access to other forms of media.

Community radio stations, in particular, have been instrumental in enhancing gender equality and women's empowerment. Niyonzima and Bhuju (2021) observe that in the rural populations of Rwanda, community radio stations have consistently raised awareness on gender issues by encouraging conversations geared towards empowering women and promoting gender equality.

Heywood (2020) has found that women-themed radio programmes empower women politically and economically. In marginalised communities and conflict-prone countries, radio is a key source of information for women. In Mali, for instance, radio programmes have enabled women to share their stories widely, thus challenging traditional gender roles and promoting gender equality (Heywood & Ivey, 2021). This suggests that radio programmes can be used to empower women, but only when they are inclusive of women's voices and varied perspectives.

Solomon (2019) has found that in South Sudan, women's participation in radio shows as hosts, guests, or callers allows them to easily share their ideas on development and to access information that has improved their involvement in development issues. Thus, radio can be a vital tool in the empowerment of women in marginalised communities, as it gives them both voice and opportunities to participate actively in development processes.

Even with all the advancements in ICTs, radio, one of the oldest ICTs, is still an essential tool for accessing information, particularly in marginalised and impoverished settings. In Kenya, community radio stations such as Realist FM in Kiambu County are embracing ICTs and providing greater opportunities to their audiences to access and participate in their radio programmes (Mwangi, 2021).

Crider (2022) and Fajula et al. (2021) caution that the use of radio for women's empowerment is not without its risks, and that careful attention must be paid to the content of programmes, the target audience, and the overall context in which they are broadcast. For instance, although the radio sector in Kenya is well advanced and thriving, there is still a lack of gender policies in many radio stations to address imbalances in organisational structures and programming outputs (Ntshangase, 2021).

Empowerment through mobile communication technologies

In this article, the term "mobile communication technologies" refers to communication platforms and services accessible through smartphones. The GSM Association's *Mobile Gender Gap Report 2021* (GSMA, 2021) shows that women in low- and middle-income countries are 7% less likely than men to own a mobile phone, and 15% are less likely to use their phones to access the internet. In Sub-Saharan Africa, women are 37% less likely to use mobile internet compared to men, indicating that the region is among those with the largest mobile gender gap globally. Rotondi et al. (2020) find that with an improvement in access to mobile phone communications, coverage, and the narrowing of digital divides, women's economic empowerment is enhanced in Sub-Saharan Africa.

The findings by Rotondi et al. (2020) are complemented by those of Hussien and Khedr (2019), who established in the context of Egypt that mobile phone ownership among women contributes to their economic advancement. Hussien and Khedr (2019) found that women who had access to mobile phones were more likely to participate in decision-making, to have more control over their finances, and to be more confident about their ability to communicate with others. This is an important finding which suggests that mobile phones can be a powerful tool for empowering women in developing countries. Although it is not clear from the findings how long-term these effects are, they are particularly relevant in the context of Egypt, where women have historically been disfranchised by patriarchal traditions and mindsets.

Though Rotondi et al. (2020) and Hussien and Khedr (2019) offer optimistic perspectives on the potential of mobile technology in the empowerment of women, their focus is only on economic empowerment. Other dimensions of women's empowerment, such as socio-cultural, interpersonal, and political dimensions, are also important considerations in order to show holistically how mobile technology can enhance women's ability to determine the direction of their lives. For instance, a study by Hussain and Amin (2018) found that although mobile technology made it possible for women in Afghanistan to improve their incomes, they could not go against the existing patriarchal power relations. Due to the likelihood of social repercussions, the Afghani women could not use the internet to acquaint themselves with information, for example on their right to divorce, or to share, privately on social media platforms, incidents of abuse. Only a few bold women rejected this control, thereby advancing their empowerment and weakening oppressive patriarchal relations (Hussain & Amin, 2018).

These studies show that that there is more to women's empowerment than economic empowerment. There is therefore a need to understand the extent to which mobile communication technologies can improve other aspects of women's lives. In Nigeria, for example, Abubakar et al. (2017) found that the use of the WhatsApp social media platform contributed to women's empowerment by expanding their freedom of expression and enabling their active engagement in economic, social, and political activities. The study provides valuable insights into the potential of mobile communications to empower women. The women could not, however, benefit from the platform fully because of the high costs of accessing the internet.

Bailur and Masiero (2017) have also established that mobile phones were instrumental in empowering women in resource-constrained areas of Kenya, Uganda, and Ghana. Bailur and Masiero (2017) further show that mobile communications can be used

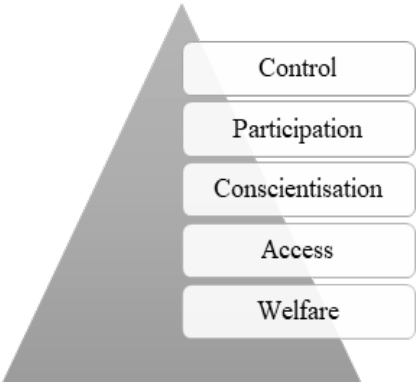
to close the gender gaps and thus improve incomes for women. This is especially important in developing countries where women often have less access than men to financial resources. With regard to the use of mobile phones by young people, Porter et al. (2020) provide accounts of a study of young people in South Africa, Malawi, and Ghana. This study found that young people primarily used mobile phones for romantic and sexual relationships, rather than for entrepreneurial and educational advancements, thus constraining trajectories on women's empowerment.

On the whole, these studies show that mobile communications are not a panacea for women's lack of empowerment; they must be used in conjunction with other interventions, such as financial literacy, to improve the gains attained. Additionally, it is important to consider the contextual factors that shape women's access to and use of mobile communications. For example, in societies that are still steeped in patriarchal traditions, women may face cultural barriers to using mobile communications, or may lack the necessary skills and knowledge. Therefore, efforts to promote women's empowerment through mobile communications need to be sensitive to constraints in the local context and must take these factors into account.

Theoretical framework

The study used Longwe's women's empowerment framework (1991). The framework conceptualises women's empowerment as progressing through five levels of measurable actions. These five levels, in ascending order (see Figure 1), are: 1) *welfare*, 2) *access*, 3) *conscientisation*, 4) *participation*, and 5) *control*. These levels are not viewed as linear but as mutually reinforcing, with every progression symbolising advancement in women's empowerment.

Figure 1: Longwe (1991) women's empowerment framework



Source: Adapted from Longwe (1991, p. 151)

Welfare is at the foundational level of this framework. This level encompasses women's empowerment in respect of factors such as an increase in income and socio-economic status. Focus at this level is on the material welfare of women in aspects such as income levels, the supply of food, and medical care. The key critique at this level is that women are regarded as passive recipients of welfare opportunities, not as active and creative creators of their material needs, hence empowerment in the actual sense does not occur.

Access is the second level in this framework. This level is focused on access as a process towards the empowerment of women. This level views gender gaps as emanating from the inequalities in access to opportunities, information, and other resources. In this regard, empowerment entails putting measures in place that would raise awareness among women of the existing gaps. Awareness is regarded as enabling women to pursue actions towards accessing an equal and fair share of the available resources within the household and in their communities.

The third level is *conscientisation*. This level involves recognition of structural forces contributing to discrimination against women. At this level, women engage in collective efforts to ensure that they are not disadvantaged. This includes ensuring that the roles of the different genders are relevant, fairness in the division of labour, and that there is no domination of one gender by the other. Empowerment at this stage also encompasses women's access to production and marketing facilities. These are complemented by legal reforms to ensure that there is equality of access to opportunities.

The next stage is *participation*, where the focus is on implementing actions arising from the conscientisation of women. At this level, women engage in collective actions aimed at challenging the widely held beliefs that have disadvantaged women. The primary concern is re-evaluating participation at all levels of decision-making, since the enhanced participation of a mobilised group is likely to facilitate the further empowerment of women. Women thus gain more power to influence decisions through mobilisation and networking. This level is crucial to the process of empowerment since there is a collective commitment to addressing structural issues that give rise to gender imbalances.

The last level in Longwe's gender empowerment framework is *control*. The control stage enables women to have control of the resources gained from their collective actions in the preceding stages. Empowerment occurs in the process of social change, thereby balancing power relations between men and women. Accordingly, neither gender dominates the other. Harnessing control enables women to influence their overall well-being, together with the well-being of their communities and society at large.

However, as previously noted, the Longwe (1991) gender empowerment framework is not a linear model. Women going through the process of empowerment may not necessarily move from one stage to the next in linear ascendancy. Rather, empowerment moves back and forth between the stages, sometimes horizontally, at other times vertically, as circumstances and experiences change. In terms of the Longwe framework, the empowerment of women is ascertained through tracking visible outcomes that are deduced through strategic data collection methods and analysis. This study was based on a qualitative approach that used in-depth interviews and focus group discussions for data collection and analysed the data through thematic content analysis. The aim of using a qualitative approach was to identify patterns and themes in the process of the empowerment of young women through personal and collective experiences, and to examine how these have changed over time as a result of their use of radio and mobile communications.

3. Research design and approach

The research adopted a qualitative design. A qualitative design was regarded as the most suitable to draw out the perceptions, experiences, and practices of young women in informal settlements on the extent to which radio and mobile communications were empowering them. As observed by Denzin and Lincoln (2011, p. 3), a qualitative design "is a situated activity that locates the observer in the world". The interviews with young women were meant to turn their perceptions, experiences, and practices into a series of representations and conversations (Jaeger, 1997). The research approach was a case study of the experiences of women in the Mathare Pioneer Youth Group.

Framework of research analysis

Longwe's (1991) framework is based on the principle that women's empowerment is a process of change that is relational, contextual, and dynamic. The framework was used to assess the experiences of young women with radio and mobile communications, and how these experiences had influenced their empowerment. The framework enabled conversations with young women about their current situation in terms of their access to resources, levels of conscientisation, participation in decision-making, and control over their lives. Based on this assessment, the study was able to identify patterns and trends in the experiences of young women and to determine the level/s at which the women were located on Longwe's framework.

In the welfare stage, the framework determined whether young women were able to access basic needs such as food, water, shelter, and healthcare. In addition, the framework analysed the ability of young women to access information, education, and economic opportunities. In the conscientisation stage, the framework assessed the awareness of young women about their rights and other issues affecting them. In the participation stage, the framework analysed evidence on the involvement of young women in decision-making and in the implementation of policies and

programmes affecting them. The final stage, control, analysed the ability of young women to exert control over their own lives, and their influence on the resources and institutions affecting them.

Sampling

The study population comprised young women from a local group known as the Mathare Pioneer Youth Group based in the Mathare informal settlements. The group accommodates vulnerable community members facing varieties of hardships in the slums. Both men and women belong to the group, although there are more young women than men. Shared challenges as a result of their vulnerable circumstances brought together members of this group to engage in activities that would improve their lives. The Mathare Pioneer Youth Group comprises 53 members, from which 36 young women were purposively selected to participate in focus group discussions. Further, 13 of the young women who participated in the focus group discussions were selected for in-depth interviews. The criterion for selection for the focus group or individual interviews was ownership of a smartphone. The average age of the young women who participated in the study was 23 years.

Data collection and analysis

This study used two methods to collect data: focus group discussions and individual interviews.

Focus groups were used to generate conversations and an understanding of issues among the participants. The discussions sought to bring out ways and means of empowerment among the young women through their use of radio and mobile communications—as well as barriers to their use of these ICTs.

Individual interviews enabled participants to express freely their opinions and experiences, devoid of bias and influence from other participants. In-depth interviews delved into the young women's use of radio and mobile communications, and barriers to their use of these ICTs for self-improvement.

The focus groups and individual interviews were based on informed consent, which was obtained from all the participants prior to data collection. The collected data was thereafter transcribed, synthesized, analysed, and interpreted in accordance to emerging themes.

Limitations of Longwe (1991) in data collection and analysis

In spite of the successes of Longwe's (1991) women's empowerment framework, it has some methodological limitations that need to be considered when analysing the process of empowerment of young women in informal settlements such as Mathare through radio and mobile communications.

First, the model does not include a practical mechanism for measuring or analysing the economic empowerment of women. This is a significant limitation, as economic empowerment is often regarded as one of the most important indicators of women's

empowerment. This study overcame this limitation by interviewing young women about their economic activities and by collecting data on women-owned businesses in the informal settlements.

In addition, the framework does not take into account the agency of women in overcoming structural disadvantages without external support. To overcome this limitation, the study took into consideration the agency of women in informal settlements, and how they may be using radio and mobile communications to empower themselves. This included looking at the content that they are sharing, as well as the participation rates and impact of radio and mobile communications initiatives. Also, the framework does not account for the diversity of experiences of women living in informal settlements.

4. Findings

Empowerment through radio

During the focus group discussions, the young women noted that they are loyal to certain radio stations. Specifically, 21 out of 36 women said that they listen to and regularly interact with Ghetto Radio 89.5 FM[1] programmes because they air content that is relevant to their lives. They said that Ghetto Radio 89.5 FM prioritises the promotion of local businesses by women entrepreneurs in informal settlements. Ghetto Radio uses Facebook and Twitter to interact with local entrepreneurs and invites them to participate in sessions and programmes on the radio station. The participants said that they rely on advice and information from the radio sessions, such as information about how to start and run successful businesses. In addition to offering a platform to women to showcase their businesses, the radio station airs advertisements for local enterprises at a discounted rate, especially for women entrepreneurs. A participant in the focus group discussions noted:

> Ghetto Radio has a show named "Hustle Mtaani", [whose focus is on promoting local businesses in the informal settlements and advising local entrepreneurs]. The chairperson of our group, Mathare Pioneer Youth group, was previously a guest on the show. She told the story of the group and of its successes on rearing chickens in Mathare. Through her story on the radio show, the demand for chicks increased substantially. As a result, we procured chicken and eggs incubators and expanded our chicken-rearing business. We are now supplying day-old chicks in Mathare and Korogocho slums. Also, we occasionally pay for advertising on Ghetto Radio because of its broad reach in Mathare and affordable advertising rates.

In addition, two of the participants in the interviews had been beneficiaries of skills training, through internships, offered by Ghetto Radio. The station provides skills

1 http://www.ghettoradio.co.ke

training programmes to build and strengthen the capacity of youth in informal settlements. One of the beneficiaries said:

> At one time, I made a video imitating Mbusi, a popular radio presenter then working at Ghetto Radio, that went viral on Twitter. The radio station got in touch and offered me [a] one-year internship programme with prospects for future employment. I was surprised because I only had a high school certificate, and most internships required at least an undergraduate degree. While attached to the radio station, I later found out that their focus was on developing skills and promoting talents, especially youth from the informal settlements. As an intern at the station, I enrolled for a journalism diploma at Rias Technical College and improved my writing and reporting skills. (Interview respondent 4)

Another beneficiary of an internship at Ghetto Radio said:

> It was always my dream to be a DJ though I never got the opportunity [to pursue it]. In 2018, Ghetto Radio posted on their Facebook page that they were offering an opportunity to nine female individuals within Mathare interested in becoming DJs. I was among the few who were selected. Through this opportunity, I am currently among well-known female DJs within Mathare and am always invited whenever Ghetto Radio hosts events within Nairobi County. The training and exposure have increased my client base and improved my standard of living. (Interview respondent 2)

The research also found that by frequently playing "Gengetone", a relatively new genre of music in the evolution of Kenyan hip hop, Ghetto Radio is in effect facilitating the empowerment of young women. Frequent airplay by Ghetto Radio of this emergent genre of music has increased the revenue streams of local artists and led to the growth of the creative economy in Mathare. In the words of one focus group participant:

> Gengetone music has created immense opportunities for women in Mathare. Some of us earn an income as music producers, musicians, DJs and some [as] "video vixens". If it were not for radio stations such as Ghetto Radio [giving] airplay to Gengetone music, probably the creative economy in Mathare would not have survived. Besides, by playing this music, the artists from Mathare can raise awareness of life in the ghetto and prompt the government to tackle issues such as rampant crime in slums and poor standards of sanitation.

Another focus group participant noted:

> We have taken the initiative as Mathare Pioneer Youth group to organise cleanups on Saturdays. This has created awareness in our locality, as people are [now] more responsible when disposing of waste. Some group members have been able to earn incomes by collecting garbage from different areas in Mathare. Without this radio station we would not have had the idea of cleaning our environment, and at the same time, creating an opportunity for generating incomes for some of our members.

Empowerment through mobile communications

The research found that mobile phones are an indispensable form of communication for young women in Mathare. By using mobile phones, young women can interact with radio programmes, use mobile financial services, and generate an income, among other engagements for their well-being.

Interaction with radio

With respect to the women's frequent use of mobile communications to interact with radio programming, one of the participants in the focus group discussions said:

> Whenever the radio station airs content relevant to us, such as measures to conserve the environment, most of us send text messages or calls to participate in the discussion.

Mobile financial services

The study found that mobile phones have made it easier for the young women in Mathare to engage in financial transactions. One advantage is that mobile money services do not require advanced ICT knowledge. As a result, the service addresses young women's needs without requiring much effort on their part. M-PESA, the mobile money service provided by Kenya's largest mobile operator Safaricom,[2] was found to be the mobile money service that is most used by the research participants. And among the M-PESA services, the Lipa na M-PESA ("Pay with M-PESA") offering, allowing secure cashless payments for goods and services, was found to be particularly valuable. As a participant in one of the focus groups observed:

> The Lipa na M-PESA function is a game-changer for the young women engaged in various businesses. Women entrepreneurs running grocery shops and other small businesses have fully adopted Lipa na M-PESA services. This is because it relieves us of the problem [of] having to carry money home from our business premises. Considering the high rate of

2 http://www.safaricom.co.ke/personal/mpesa

> crime in Mathare, Lipa na M-PESA ensures that money is safe [in the mobile wallet]. Also, most of the women entrepreneurs in the group have previously experienced a reversal of mobile money transactions by clients once payment is done, leading to losses for the businesses. However, with Lipa na M-PESA there is no theft through the reversal of transactions as it does not allow that function.

Lipa na M-PESA was developed by Safaricom to be a more efficient and secure way to pay for goods and services when compared to the M-PESA mobile money service. With Lipa na M-PESA, once a client initiates a transaction, only the receiving merchant can reverse that transaction. Therefore, there are no instances of theft through the reversal of transactions. Interviewee 3, a grocery retailer, observed:

> The use of mobile money has contributed immensely to saving costs and convenience in my business. Through Lipa na M-PESA, I can pay my water and electricity bills in my comfort without visiting physical offices. Also, the "Buy Goods" [feature within Lipa na M-PESA] option offered by M-PESA allows my clients to pay for their goods without incurring transaction charges. Thus, clients who prefer mobile money have the option of paying for their goods without incurring extra costs.

The Lipa na M-PESA function has two options for payment: a "pay bill" option and a "buy goods and services" option. The "buy goods and services" option allows for paying for goods and services at a participating merchant, where a buyer enters the merchant's Till Number into their M-PESA menu. With the "pay bill" option, a client incurs transaction costs, while the "buy goods and services" option does not attract any transaction charges. Interviewee 1 said:

> The mobile money platform facilitates financial transactions and is also instrumental in paying my medical insurance cover. I learned about M-TIBA, [an embedded health scheme within the mobile money payment platform], through an advertisement on Ghetto Radio. I subscribed for the service because I can pay for my medical insurance using my mobile phone every month. Besides, several hospitals within Mathare cover patients using M-TIBA, making it easier to seek medical services in the slums.

Interviewee 5, a hairdresser, observed:

> Banking services via mobile phone are convenient for conducting financial transactions from any location […]. The only exception is that you need a mobile money agent to withdraw your money. My phone allows me to operate a banking application that I can use to pay for electricity, school fees, and shopping. The transaction costs are [relatively] low, thereby affordable for most women in Mathare.

Interviewee 8, a freelance writer, said:

> As an online writer, I depend on mobile banking services to receive payments for my services. In the past, I had difficulties withdrawing money from my bank account, [but] once my bank connected my mobile wallet with my bank account, I had full access to my earnings as a freelancer […]. I can now transfer funds from my bank account to my mobile wallet, and I receive notifications from the bank about the latest financial offers.

Another participant in a focus group discussion said:

> Mobile banking services are vital in the growth of our chicken-rearing business, in that we can deposit profits from the businesses and later on invest in other ventures. We are all signatories to our group's bank account. Hence, we can keep track of our bank balances through our mobile phones. If cash withdrawals are to be made, each member must approve it. As such, there is no misuse of the group's funds. The result is that the group's savings have accumulated over the years, and we will soon be opening a bakery to serve the residents of Mathare. We believe that the convenience of mobile banking services has contributed to the financial success of Mathare Pioneer Youth group.

The study also found that some women are using the M-PESA M-Shwari[3] loan service. Some have borrowed from M-Shwari to start businesses. In the words of interviewee 7:

> For the past two years, I have been borrowing money via M-Shwari. I had initially borrowed money to pay bills, but with time, my limit grew, and I managed to borrow KES5,000 [approx. USD41[4]] to start a fruits business within Mathare. I borrowed money to buy a blender, a table, a few chairs, and fruits for the business. Within five weeks, I had made enough money to repay the loan and interest. As my business expanded, the loan limit also grew to KES13,000 [USD107]. However, I am usually cautious not to borrow money that I will not be in a position to repay on time. I am proud to say that M-Shwari has made it possible for me to start a business and earn an income to fend for myself and my siblings.

3 http://www.safaricom.co.ke/personal/m-pesa/m-pesa-services-tariffs/m-shwari
4 Using exchange rate on 30 November 2022.

According to interviewee 6:

> I sell second-hand clothes within Mathare to earn a living. I used to depend on money borrowed from M-Shwari to buy stock of second-hand clothing until Safaricom introduced Fuliza loans. The service allows me to restock clothes even when I do not have sufficient cash in my M-PESA account. The interest rates are low, and I only have to pay for the loan when I receive money on my M-PESA account. Fuliza has made it convenient for me to conduct my business because, at times, it might take weeks before I sell all the clothes and repay the loan. I have grown my Fuliza limit from a low of KES1,000 [approx. USD8] to KES4,300 [USD35] in the past three months.

M-Shwari is a mobile banking service offered by Safaricom that allows customers to save, lend, and borrow money using their mobile phones. If a customer needs to borrow money, they can take out an M-Shwari loan, which comes as credit in their account that they can withdraw or transfer. If a customer does not have enough money to complete an M-PESA transaction, they can use Fuliza, which is an overdraft service, to cover the shortfall. A focus group participant noted:

> Most of us have different loan applications on our mobile phones. The money borrowed from M-Shwari, Fuliza and other mobile loan applications is essential to running our businesses. It is difficult for us to borrow money from the bank because of the inability to meet bank requirements. Therefore, we rely on M-Shwari to facilitate our businesses and buy food whenever we experience financial challenges. The M-Shwari platform has taught us financial discipline because the loan limit is reduced whenever there is a late repayment of the loan. We are, therefore, encouraged to pay on time so that we can increase the limit and use the extra amount to expand our businesses. Raising capital for business is no longer an issue for the group.

Income generation

The study also found that the use of mobile communications has provided the young women with opportunities to earn an income. Interview respondent 9 indicated that she markets her hair and beauty products to clients within and beyond Mathare through social media, with some of her clients ordering products via her social media handles. Some of her clients also book appointments through text messaging and calls.

In the words of interviewee 11:

> I am a content creator and work with different brands in Kenya. I use my mobile phone to tweet about their products and services. On average, I spend KES50 [approx. 40 cents in USD] daily on data bundles which translates to KES350 [USD2.90] weekly. I get paid on average KES300 [USD2.46] daily to promote a product. On a good month, I make close to KES40,000 [USD328] for just promoting products online with my mobile phone. I can comfortably pay my siblings' school fees [and] house rent, and fend for the family.

According to Interviewee 12:

> I learned how to bake cakes by watching cooking tutorials on YouTube. I was surprised that my first attempt at making cupcakes turned out perfect. I discovered that few individuals were baking and decorating cakes within Mathare. I, therefore, saved up some money and started a business as a baker. I started with cupcakes and then proceeded to birthday and wedding cakes. We have a WhatsApp group for residents in our Flat [apartment building] that has helped me to get clients. Our Flat […] has 135 active WhatsApp members who share my cakes with potential clients. I receive countless orders for birthday cakes. There are instances whereby I no longer receive new orders until I am done delivering the ones I [have]. In future, I intend to expand my business and employ more youth in the slums.

Interviewee 10, an online shoe retailer, noted that her income level has increased by virtue of her social media platforms. In most cases, clients order shoes on her social media page.

The study found that one of the focus group participants runs a daycare centre, which makes it possible for the women in Mathare to engage in other businesses with the assurance that their children are safe. The woman created a WhatsApp group that enables her to coordinate her daycare services. Parents receive information about when to drop and pick up their children, and on the children's general welfare. She also takes into consideration that some parents own feature phones, and therefore provides the information through text messages. The efficient management of the daycare via the use of mobile communications has, according to his participant, attracted more parents and resulted in the expansion of the business. To supplement her income, this woman also sells *mandazis* (sweet fried bread), sourcing clients via word-of-mouth and her social media handles.

Barriers to empowerment through mobile phones

The study also identified several barriers to women's empowerment through use of mobile communications.

Socioeconomic barriers

The research participants indicated that their socio-economic status prevents their full utilisation of mobile communications. They indicated that they struggle to take care of their families and siblings, in addition to paying their school fees. Due to limited time and a scarcity of financial resources, they are reluctant to register for training that would make them more effective in using mobile communication technologies and in exploring the opportunities thereof. Nevertheless, many said that they are interested in learning new things. Most said that if they were offered affordable ICT training and were assured of opportunities in the ICT sector, they would be willing to dedicate time to learning more about ICT (focus group respondents).

However, some women who have been offered training opportunities in ICTs have been unable to participate. They said that this is because their priority is taking care of their families, and participating would mean losing income on the days when they are engaged in the training. A lack of transport to the training venues is also a barrier to their participation.

Cultural barriers

It was also found that cultural barriers against women have negatively impacted their uptake of mobile communications. With their socio-economic status already limited, cultural biases make the situation even worse. Interviewee 13 noted that some parents forbid their girls from owning mobile phones, as they believe that the phones will encourage them to have sexual relationships with men.

In a focus group discussion, some participants noted that some women are afraid to use mobile phones because of the belief that it will lead to the break-up of their relationships. Several women revealed that they have been accused of infidelity when their phone rings frequently (focus group respondents).

5. Analysis in terms of the Longwe (1991) theoretical framework

Empowerment through radio

Young women in the Mathare informal settlements can be said to have initiated the *welfare* stage in the Longwe (1991) framework through their engagements with Ghetto Radio, whose content has contributed to improvement of their socioeconomic status. The young women have also progressed to the next level, *access*, where they gain access to opportunities for skills development, career development, and support for their businesses—through the various programmes and initiatives offered by Ghetto Radio.

The third level, *conscientisation*, is being realised by some women, as they are slowly getting into jobs and careers that were previously dominated by men. This is evidence that the young women are progressing from merely having knowledge of existing inequalities to taking concrete steps to address these inequalities.

The final levels in the Longwe (1991) gender empowerment framework, *participation* and *control*, are being realised by some of the young women through their participation in initiatives as equals to men. Such young women are not merely passive recipients of development initiatives, but are actively shaping the conversations and content towards personal and social transformation.

Empowerment through mobile communications

The use of mobile communications has enabled many of the young women to progress from the level of *welfare* up to the *control* stage in Longwe's (1991) women's empowerment framework.

To start with, many of the young women initiated the *welfare* stage through their engagements with mobile money services. These financial services have provided them with a convenient way to access money and manage their finances, as well as with opportunities to start and grow their own businesses. The welfare stage has since progressed into the *access* stage for many of the women, who are now able to access financial services and healthcare, among other services, through mobile communications. Many of the young women have also advanced to the *conscientisation* stage through mobile communications, as they are now more aware of their financial options and are able to make more informed choices about how to manage their finances.

The *participation* stage has also been realised by many of the women, wherein they have amplified their voices on how their money and other resources are being used. The ability of many of the young women to generate income through mobile phones suggests that these women have even risen to the *control* stage of Longwe's (1991) women's empowerment framework, wherein they are not content with merely listening and speaking, and are controlling their personal and collective resources.

6. Conclusions

The study established that many young women in the Mathare informal settlements have been able to use ICTs, specifically radio and mobile communications, as important elements of their journey towards personal and collective empowerment—as viewed through the prism of Longwe's (1991) women's empowerment framework. It was found that ICTs have provided the young women with relevant information to access resources, opportunities to generate income, and opportunities to improve their general well-being.

The study provides important insights into how ICTs can be used to empower women, particularly in developing countries. It also highlights the potential of radio and mobile communications to reach and engage women in the Mathare informal settlements, for whom access to communications and other resources is often impeded by socioeconomic and cultural barriers grounded in patriarchal mindsets. Finally, the study provides a contribution to the literature on women's empowerment through ICTs, and can be used to inform policy and practice in this area.

References

Abubakar, N., Dasuki, S. I., & Quaye, A. (2017). The contribution of ICTs to women empowerment: A capability perspective. Paper presented to the Twenty-Third Americas Conference on Information Systems, Boston.

Bailur, S., & Masiero, S. (2017). Women's income generation through mobile internet: A study of focus group data from Ghana, Kenya, and Uganda. *Technology and Development*, *21*(1–2), 77–98. https://doi.org/10.1080/09718524.2017.1385312

Batliwala, S. (1994). The meaning of women's empowerment: New concepts from action. In G. Sen, A. Germain, & L. C. Chen (Eds.), *Population policies reconsidered: Health, empowerment, and rights*. Harvard University Press.

Booi, S. L., Chigona, W., Maliwichi, P., & Kunene, K. (2019). The influence of telecentres on the economic empowerment of the youth in disadvantaged communities of South Africa. Paper presented to the 15th IFIP WG 9.4 International Conference on Social Implications of Computers in Developing Countries, Dar es Salaam, 1–3 May. https://doi.org/10.1007/978-3-030-18400-1_13

Crider, D. (2022). Listening, but not being heard: Young women, popular music, streaming, and radio. *Popular Music and Society*, *45*(5), 600–616.

Denzin, N. K., & Lincoln, Y. S. (2011). *The SAGE handbook of qualitative research*. SAGE.

Ekwok, L. (2018). The effects of public service broadcasting on rural development: A study of the impact of selected programmes of Cross River Radio, Ikom and four other rural areas. *LWATI: A Journal of Contemporary Research*, *15*(3), 216–237.

Fajula, A., Barbeito, M., Barrio, E., Enrique, A. M., & Perona, J. J. (2021). Feminist stereotypes and women's roles in Spanish radio ads. *Media and Communication*, *9*(2), 39–51. https://doi.org/10.17645/mac.v9i2.3762

Golzard, V. (2019). Economic empowerment of Iranian women through the internet. *Gender in Management*, *35*(1), 1–18. https://doi.org/10.1108/GM-11-2017-0145

GSMA. (2021). *Connected women: The mobile gender gap report 2021*. https://www.gsma.com/r/wp-content/uploads/2021/06/The-Mobile-Gender-Gap-Report-2021.pdf

Hancock, D. R., & Algozzine, B. (2017). *Doing case study research: A practical guide for beginning researchers*. Teachers College Press.

Heywood, E. (2020). Radio journalism and women's empowerment in Niger. *Journalism Studies*, *21*(10), 1344–1362. https://doi.org/10.1080/1461670X.2020.1745668

Heywood, E., & Ivey, B. (2021). Radio as an empowering environment: How does radio broadcasting in Mali represent women's "web of relations"? *Feminist Media Studies*, 1–17. https://doi.org/10.1080/14680777.2021.1877768

Hussain, F., & Amin, S. N. (2018). 'I don't care about their reactions': Agency and ICTs in women's empowerment in Afghanistan. *Gender & Development*, *26*(2), 249–265. https://doi.org/10.1080/13552074.2018.1475924

Hussien, S., & Khedr, W. (2019). Do mobile phones empower women? Evidence from rural Egypt. *Journal of the Faculty of Commerce for Scientific Research*, *56*(2), 193–223. https://doi.org/10.1080/13552074.2018.1475924

Jaeger, R. M. (1997). Complementary research methods for research in education (2nd ed.). American Educational Research Association.

Kenya National Bureau of Statistics (2019). 2019 Kenya population and housing census.

Komen, L. J., & Ling, R. (2021). 'NO! We don't have a joint account': Mobile telephony, mBanking, and gender inequality in the lives of married women in western rural Kenya. *Information, Communication & Society*, *24*(16), 1–18. https://doi.org/10.1080/1369118X.2021.1927137

Kovačič, P., & Lundine, J. (2014). Mapping Kibera: Empowering slum residents by ICT. In S. Livingston & G. Walter-Drop (Eds.), *Bits and atoms: Information and communication technology in areas of limited statehood.* Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199941599.003.0008

Longwe, S. H. (1991). Gender awareness: The missing element in the Third World development project. In T. Wallace & C. March (Eds.), *Changing perceptions: Writings on gender and development* (pp. 149–157). Oxfam.

Lundine, J., Kovačič, P., & Poggiali, L. (2012). Youth and digital mapping in urban informal settlements: Lessons learned from participatory mapping processes in Mathare in Nairobi, Kenya. *Children, Youth and Environments*, *22*(2), 214–233. https://doi.org/10.7721/chilyoutenvi.22.2.0214

Macwan, U. (2017, April 24). Mobile technology, its importance, present and future trends. *Finextra*.

March, C., Smyth, I., & Mukhopadhyay, M. (1999) Women's empowerment (Longwe) framework. In *A guide to gender-analysis frameworks* (pp. 92–101). Oxfam. https://doi.org/10.3362/9780855987602.000

Mwangi, S. (2021). Appropriating new digital technology in community radio broadcasting in Kenya: A case study of Realist FM. *African Multidisciplinary Journal of Research*, *6*(1), 1–18.

Nangooba, M. (2020). Role of radio in improving livelihoods: The case of Central Broadcasting Services' project to empower women in savings and loan associations in Nsangi Sub-County, Wakiso District, Uganda. *Journal of Science and Sustainable Development*, *7*(1), 31–40. https://doi.org/10.4314/jssd.v7i1.3

Nayyar, T., Aggarwal, S., Khatter, D., Kumar, K., Goswami, S., & Saini, L. (2019). *Opportunities and challenges in digital literacy: Assessing the impact of digital literacy training for empowering urban poor women*. University of Delhi.

Niyonzima, D., & Bhuju, K. (2021). The role of community radio in promoting gender equality in Rwanda. In C. Simon, F. Elsebeth, G. Mariateresa, H. Shabbir, K. Iqbal, A. L. C. Lopez, J. Lugo-Ocando, & L. D. Pant (Eds.), *Handbook of research on discrimination, gender disparity, and safety risks in journalism* (pp. 343–365). IGI Global. https://doi.org/10.4018/978-1-7998-6686-2.ch018

Ntshangase, J. S. (2021). Mapping community radio in Sub-Saharan Africa. Journalism.co.za.

Porter, G., Hampshire, K., Abane, A., Munthali, A., Robson, E., De Lannoy, A., Tanle, A., & Owusu, S. (2020). Mobile phones, gender, and female empowerment in sub-Saharan Africa: Studies with African youth. *Information Technology for Development*, *26*(1), 180–193. https://doi.org/10.1080/02681102.2019.1622500

Rahman, H. (2014). The role of ICT in open and distance education. *Turkish Online Journal of Distance Education*, *15*(4), 162–169. https://doi.org/10.17718/tojde.47700

Rimmer, A. (2021). Breaking the silence: Community radio, women, and empowerment. *Community Development Journal*, *56*(2), 338–355. https://doi.org/10.1093/cdj/bsz030

Rotondi, V., Kashyap, R., Pesando, L. M., Spinelli, S., & Billari, F. (2020). Leveraging mobile phones to attain sustainable development. *Proceedings of the National Academy of Sciences*, *117*(24), 13413–13420. https://doi.org/10.1073/pnas.1909326117

Singh, S., & Vanka, S. (2020). Achieving women empowerment through ICT: Case of a government initiative in India. In P. Rekha, S. Birte, & S. Nilima (Eds.), *Handbook of research on new dimensions of gender mainstreaming and women empowerment* (pp. 87–104). IGI Global. https://doi.org/10.4018/978-1-7998-2819-8.ch006

Sinha, S. (2017). Multiple roles of working women and psychological well-being. *Industrial Psychiatry Journal*, *26*(2), 171–177. https://doi.org/10.4103/ipj.ipj_70_16

Simmons, J. (2019). The political economy of education for village development. In P.C. Reining & B. Lenkerd (Eds.), *Village viability in contemporary society* (pp. 287-314). Routledge. https://doi.org/10.4324/9780429267000-12

Solomon, D. R. G. (2019). *Women's empowerment through community radio: A case study of Bakitha Radio, South Sudan.* Master's dissertation, Makerere University, Uganda.

Tijani-Adenle, G. (2022). Women FM (W.FM): The women-focused radio station amplifying the voices of Nigerian women. In *M. Lindgren & J. Loviglio* (Eds.), *The Routledge companion to radio and podcast studies* (pp. 327–338).

UN Habitat (2020). *Informal settlements' vulnerability mapping in Kenya: The case of Mathare*.

Varriale, C., Pesando, L. M., Kashyap, R., & Rotondi, V. (2022). Mobile phones and attitudes toward women's participation in politics: Evidence from Africa. *Sociology of Development*, *8*(1), 1–37. https://doi.org/10.1525/sod.2020.0039

Wamuyu, P. K. (2017). Closing the digital divide in low-income urban communities: A domestication approach. *Interdisciplinary Journal of e-Skills and Lifelong Learning*, *13*, 117–142. https://doi.org/10.28945/3885

Wei, W., Sarker, T., Żukiewicz-Sobczak, W., Roy, R., Alam, G. M., Rabbany, M. G., Hossain, M. S., & Aziz, N. (2021). The influence of women's empowerment on poverty reduction in the rural areas of Bangladesh: Focus on health, education and living standard. *International Journal of Environmental Research and Public Health*, *18*(13), 1–18. https://doi.org/10.3390/ijerph18136909

World Wide Web Foundation. (2015). *Women's rights online: Translating access into empowerment*.

An assessment of website quality at Nigerian polytechnics and colleges of education

Oseikhuemen Ibhadode
UI/UX Designer, Midland Digital Solutions, Lagos
https://orcid.org/0000-0001-6297-532X

Adeola Opesade
Senior Lecturer, Department of Data and Information Science, University of Ibadan, Nigeria
https://orcid.org/0000-0001-9241-2798

Abstract
This study assessed the quality of the websites of Nigerian polytechnics and colleges of education. Using two web diagnostic tools, SEOptimer and W3C Markup Validation Service, a total of 213 sites were evaluated in terms of six performance indicators: search engine optimisation (SEO), usability, operational performance, social media integration, security, and HTML validation. The weakest performance across the sites of both polytechnics and colleges of education was found to be with respect to social media integration, with the vast majority of sites in both categories making no use of social media. The other two indicators against which sites in both categories of institutions tended to perform poorly were SEO and operational performance. The two areas where the sites in both institutional categories generally scored well were usability and HTML validation. Meanwhile, for the security indicator, the performance was highly variable among both polytechnic and college of education sites, with significant numbers of both strong and weak scores for sites. It was also found that polytechnic sites outperformed college of education sites on SEO to a statistically significant degree (bearing in mind that both categories of sites were weak in this area). It was also found that, among the polytechnic sites, the sites of the state government-owned institutions outperformed, to a statistically significant degree, the sites of the privately owned institutions in their operational performance measure.

Keywords
institutional websites, polytechnics, colleges of education, website quality, Nigeria, search engine optimisation (SEO), usability, operational performance, social media integration, security, HTML validation

DOI: https://doi.org/10.23962/ajic.i30.13987

Recommended citation
Ibhadode, O. & Opesade, A. (2022). An assessment of website quality at Nigerian polytechnics and colleges of education. *The African Journal of Information and Communication (AJIC), 30*, 1-20. https://doi.org/10.23962/ajic.i30.13987

1. Introduction

Websites play a vital role as the image maker of institutions because of their ability to boost institutional reputation and attract the desired audience. This makes it imperative for any educational institution to ensure optimised design and functionality of its website. According to Greenfield (2015), if a tertiary institution's web presence fails, the business operations will likely fail too because a non-optimised website can negatively impact on the school's operations, finance, enrolment, reputation, and credibility. Furthermore, a poorly designed school website can generate legal risks in terms of privacy and copyright issues, and credibility risks with respect to broken links, misspelt words, and outdated content (Greenfield, 2015). Tate (2017) states that visitors will often leave an institution's website if it has weaknesses in its site load time and navigation.

Geron (2018) states that institutional websites give site visitors such as prospective and current students, prospective and current lecturers, researchers, parents, alumni, employers, and companies information about the institution and reflect the institution's style, activities, and reputation. An outstanding tertiary institution website will also attract students and research fellows alike, who will have information about the school right at their fingertips due to a powerful web presence. A number of studies have investigated the website quality of academic institutions in different parts of the world, such as Iran (Andalib & Danaee, 2013), the Czech Republic (Kincl & Strach, 2012), Jordan (Almahamid et al., 2016), Ukraine (Kravchenko et al., 2021), Turkey (Akgül, 2021), India (Kaur et al., 2016), and Nigeria (Olaleye et al., 2018; Pechnikov & Nwohiri, 2012), amongst many others. Despite the fact that it should be the desire of any institution of higher learning to have their website perform optimally, most research efforts on the assessment of academic institutional websites have focused on university websites, leaving other institutions like polytechnics and colleges of education unattended to.

Nigeria's polytechnics and colleges of education are important institutions of tertiary education, and it is essential that their websites are of high quality. It therefore becomes pertinent to investigate the current state of these institutions' websites, in order to provide research-based evidence of their performance. In order to fill this

research gap, the present study focused on assessing the institutional websites of all accredited Nigerian polytechnics and colleges of education based on six performance indicators, namely: (1) search engine optimisation (SEO), (2) usability, (3) operational performance, (4) social media integration, (5) security, and (6) HTML validation. The study also sought to determine if there were correlations between institutional type and ownership type and the performance measures of their institutional websites.

Research questions

- How do Nigerian polytechnic and college of education websites perform in terms of search engine optimisation (SEO)?
- How do Nigerian polytechnic and college of education websites perform in terms of usability?
- How do Nigerian polytechnic and college of education websites perform in terms of operational performance?
- How do Nigerian polytechnic and college of education websites perform in terms of social media integration?
- How do Nigerian polytechnic and college of education websites perform in terms of security?
- How do Nigerian polytechnic and college of education websites perform in terms of their HTML validations?

Research hypotheses

- There are no significant differences in website performance between Nigeria's polytechnics and colleges of education.
- There are no significant differences in website performance among Nigeria's polytechnics based on institutional ownership type (federal, state, and private).
- There are no significant differences in website performance among Nigeria's colleges of education based on institutional ownership type (federal, state, and private).

2. Literature review

Previous empirical studies

Many studies have been conducted to evaluate the quality of websites. Some have adopted the subjective website quality evaluation paradigm, using questionnaires to collect data relevant to research questions, while others have adopted the objective website quality evaluation paradigm, using qualitative models or diagnostic tools.

To evaluate the quality of websites from users' perspectives, a questionnaire research instrument is often used as a data-collection tool to ask users for their opinions on the websites. Jundillah et al. (2019) evaluated e-learning websites based on the results of the Webqual questionnaire and calculations using the importance performance analysis method. The questionnaire was distributed to 95 students in various study

programmes at Stikubank University in Indonesia. The parameters measured were the Webqual 4.0 measures for standard of usability, quality information, and service interaction. The study found that the average student was satisfied with the quality of e-learning websites with which they interacted.

Andalib and Danaee (2013) measured the quality of the website of Payame Noor University in Iran. The study considered the effects on customer loyalty of four quality parameters—efficiency, accessibility, achievement, and security—via two variables: trust and satisfaction. A questionnaire was used as the research instrument, administered to 387 active website users. A website quality parameter that seemed to be a major determinant of website quality was user satisfaction. The authors found that trust, efficiency, and achievement play an essential role in customer loyalty.

Almahamid et al. (2016) studied the website quality of Middle East University in Jordan. They found, through a quantitative study, that perceived usefulness and perceived ease of use are key determinants of behavioural intention in the use of the university website. Kravchenko et al. (2021) used a quantitative, model-based methodology called website QEM (quality evaluation method) to assess the quality of two higher education websites in Ukraine. The authors focused on the following attributes: quality of content, quality of design, ease of use of the site for applicants, and students' reputation. Based on the parameters evaluated, Taras Shevchenko National University of Kyiv was found to be performing significantly better than Kyiv National University of Trade and Economics.

Kincl and Strach (2012) studied user satisfaction with 44 business school websites in the Czech Republic. They reported that user satisfaction is multidimensional: different features of the websites impact differently on the level of user satisfaction with the websites. They also documented evidence that users' inability to complete a task when on a website resulted in a significant negative effect on user satisfaction. Another study that investigated user satisfaction is that of Søruma et al. (2012), which found no positive correlation between website quality and user satisfaction. Elling et al. (2012) developed a website evaluation questionnaire (WEQ) specifically for governmental websites and focused on seven website quality parameters: ease of use, hyperlinks, structure, relevance, comprehension, completeness, and layout.

Furthermore, Akgül (2021) evaluated the accessibility, usability, quality performance, and readability aspects of all Turkish state and private university websites. The results from the study indicated low levels of accessibility, usability, quality performance, and readability. A variety of web diagnostic tools were used in the evaluation process, including, but not limited to, Google Mobile-friendly Test for mobile responsiveness,

Web Page Analyzer for quality performance test, Alexa Traffic Rank for website popularity ranking, CSS Validation Service and HTML Markup Validation Service. Numerous automated evaluations of website quality, using web diagnostic tools, have also been conducted. One of the earliest studies, by Choudrie et al. (2004), evaluated global e-government sites using diagnostic tools, and discovered that significant work was needed in order to make the portals examples of "best practice" e-government services. Shi (2006) found that significant work was needed to improve the quality of World Expo websites in order to provide quality information and services to their users. Jati and Dominic (2009) considered the quality of five Asian e-government websites, using web diagnostic tools that checked for website performance, quality of links, markup validation, and accessibility. Their study revealed that the five e-government websites were not performing optimally.

Suwawi et al. (2015) evaluated the website of Telkom University, Indonesia, using an internationally recognised software evaluation standard, ISO/IEC 9126. Their findings revealed that the three website quality parameters that needed to be improved were reliability, usability, and functionality. Kaur et al. (2016) evaluated the website quality of five Indian university websites using web diagnostic tools to determine their performance, speed, number of requests, load time, page size, SEO, mobile rendering, and security. In an evaluation of 51 Turkish e-government websites using web diagnostic tools, Akgül (2016) discovered that the e-government websites were not meeting the standards as regards response time, download time, page size, number of items, and markup validation.

Perçin (2021) used two decision-making models, namely the fuzzy decision making trial and evaluation laboratory (fuzzy DEMATEL) and the generalized Choquet fuzzy integral (GCFI), to evaluate the website quality of Turkish hospitals. The results of the study showed that hospitals should focus more on reliability, visual content, and assurance issues to improve websites. Dominic et al. (2011) used decision-making models—linear weightage model (LWM), analytical hierarchy process (AHP), fuzzy analytical hierarchy process (FAHP), and new hybrid model (NHM)—to determine the quality of five Asian e-government websites. They reported that the five e-government websites did not fulfil the criteria of a high-quality website. Other studies conducted by Chou and Cheng (2012), Dominic and Khan (2013), Dominic et al. (2013) and Faustina and Balaji (2016) also used decision-making models. Some other studies conducted by Abdel-Basset et al. (2018), Dani and Agrawal (2021), and Gong et al. (2021) have also used decision-making models.

In Nigeria, Olaleye et al. (2018) used web diagnostic tools such as WooRank, CheckMyColors, GTmatrix, Link Polularity, and W3C Markup Validator to perform a comparative analysis on 141 university websites across federal, state, and private universities. The study found that the private Nigerian universities generally had websites of better quality than the federal and state university websites.

The study reported on in this article is unique in the Nigerian context because it not only evaluated the quality of all 213 websites of Nigerian polytechnics and colleges of education, but also compared quality between the two types of institutions and among the three types of institutional ownership.

3. Research design and methodology

The research framework adopted for the study belongs to the objective website quality evaluation paradigm, and used two web diagnostic tools: SEOptimer and W3C's HTML Markup Validator. Advances in technology have provided web diagnostic tools, such as the two used in this study, which can rapidly generate in-depth insights and robust figures on a number of website quality parameters. This allows for a large number of websites—in this case, more than 200—to be rapidly evaluated and analysed.

SEOptimer provides a comprehensive report on a website based on website quality parameters when the website's URL is inputted into the analysis field of SEOptimer. It reviews a website's SEO, usability, operational performance, social presence, security, and technology used in developing the website. These parameters are the broad categories that form the general report for the website's quality and its tendency to be found faster by search engines. Some intricate factors that are reported on are backlinks, broken links, page speed, page size, number of resources, malware check, email privacy, and more.

HTML Markup Validator checks a website's HTML to validate the extent to which the website meets technical web quality. When the URL of a website is inputted into the analysis field, Markup Validator scans through the website's HTML structure for errors and warnings. The number of errors and warnings, and exactly what they are, are posted on the platform. A detailed report is given on each error and warning.

SEOptimer produces grades: A+, A, A–, B+, B, B–, C+, C, C–, D+, D, D– and F. We coded these grades into five categorical weights: Excellent (A+, A, A–), Good (B+, B, B–), Fair (C+, C, C–), Poor (D+, D, D–), or Very Poor (F). Markup Validator reports numeric values based on the number of HTML errors. Again we coded into five categorical weights: Excellent (1–25 errors), Good (26–50 errors), Fair (51–75 errors), Poor (76–100 errors), and Very Poor (more than 100 errors). Then each of the five weights, for the results from both tools, was assigned a numerical weight—5, 4, 3, 2, and 1—for Excellent, Good, Fair, Poor, and Very Poor respectively.

According to Burns and Grove (2009), a research design is a plan that is adopted for conducting a study in a way that there would be an adequate control over such factors that could hinder the credibility of the findings. The present study adopted the content analysis method of research.

The population of the study comprised 213 accredited Nigerian polytechnics and colleges of education. This study adopted a complete enumeration strategy, because the total population was of a manageable size (Glen, 2018). The disaggregation of the population for Nigerian polytechnics is 28 federal, 45 state, and 50 private, for a total of 123. In the case of Nigerian colleges of education, the disaggregation is 21 federal, 44 state, and 25 private, for a total of 90. The entities whose websites were evaluated for this research are approved by the respective regulatory bodies, which are the National Board for Technical Education (NBTE) for polytechnics and the National Commission for Colleges of Education (NCCE) for colleges of education. The study adopted a primary data collection method, with data collected in March 2020 through use of SEOptimer and W3C's Markup Validator.

To collect the necessary data, SEOptimer's URL (https://www.seoptimer.com) was inputted in a Google Chrome browser. The URL for each institution's website was entered in SEOptimer for auditing. The audit button was then selected in order to generate the result of the analysis based on the aforementioned website quality parameters. SEOptimer then returned the website quality of the audited institution on an A+ to F– grade scale, which, as explained above, were assigned categorical weights of 5, 4, 3, 2, or 1. The individual grades of the parameters were aggregated to give the website quality grade of the institution. There was in-depth analysis for each parameter, and intricate factors such as backlinks, broken links, page speed, page size, number of resources, malware check, and email privacy were reported back. A screen capture plugin on the Chrome browser was used to save the results of each website analysis.

The URL of each institution's website was also inputted into the analysis field of the W3C Markup Validator, which scanned through the website's HTML structure for errors and warnings. The number of errors and warnings, and exactly what they were, were posted on the platform. A detailed report was provided on each error and warning. The researchers could then view both the summary report and the detailed report. Screen capture was again used to save the result of each analysis.

The data collected from the web diagnostic tools were coded and analysed using Statistical Package for the Social Sciences (SPSS) version 20.

4. Findings

Search engine optimisation (SEO)

As shown in Table 1, the only two "Excellent" websites, in terms of SEO, were indentified at colleges of education. However, polytechnic websites were more strongly represented in the "Good" and "Fair" categories, and in terms of mode value, the polytechnics performed better ("Fair") than the colleges of education ("Poor"). At the same time, the polytechnic sites had roughly the same percentage (33.3%) of "Poor" ratings as the college of education sites (34.3%), and it was only because of the polytechnic sites' slightly higher percentage (35.5%) of "Fair" sites that the polytechnic sites achieved a mode value of "Fair". One of the 123 polytechnic websites had no SEO ranking reported for it by the SEOptimer. This might have been due to it not having adequate features upon which the tool could assess its SEO state.

Table 1: SEO

		Polytechnics	Colleges of education
	Excellent (%)	0 (0%)	2 (2.2%)
	Good (%)	20 (16.3%)	10 (11.1%)
	Fair (%)	43 (35.0%)	22 (24.4%)
	Poor (%)	41 (33.3%)	31 (34.4%)
	Very Poor (%)	18 (14.6%)	25 (27.8%)
	Unreported ranking (%)	1 (0.8%)	0 (0%)
Mode		Fair	Poor

Usability

As shown in Table 2, the mode value for usability was "Excellent" for both the polytechnic websites and the college of education websites, with the polytechnic sites slightly better represented in the "Excellent" and "Good" categories. At the same time, the percentages of "Very Poor" sites—15.4% of polytechnic sites, 17.8% of college of education sites—were notable. Also, one of the 123 polytechnic websites had no usability ranking reported by the SEOptimer. This was likely due to it not having adequate features upon which the tool could assess its usability state.

Table 2: Usability

		Polytechnics	Colleges of education
	Excellent (%)	84 (68.3%)	58 (64.4%)
	Good (%)	16 (13.0%)	11 (12.2%)
	Fair (%)	3 (2.4%)	4 (4.4%)
	Poor (%)	0 (0%)	1 (1.1%)
	Very Poor (%)	19 (15.4%)	16 (17.8%)
	Unreported ranking (%)	1 (0.8%)	0 (0%)
Mode		Excellent	Excellent

Operational performance

As shown in Table 3, the mode value for the operational performance of polytechnic websites was "Very Poor", while the mode value for the college of education websites was "Fair". The stronger mode value for the college of education sites was driven by their relatively high percentage (18.9%) of "Excellent" ratings, compared to only 9.8% of polytechnic sites. At the same time, the percentage of college of education sites (24%) in the "Very Poor" category was identical to the percentage of "Very Poor" polytechnic sites, showing that nearly one-quarter of sites in each category were in serious need of improvement in terms of their operational performance. Also, one of the 123 polytechnic websites had no operational performance ranking reported by the SEOptimer tool.

Table 3: Operational performance

		Polytechnics	**Colleges of education**
	Excellent (%)	12 (9.8%)	17 (18.9%)
	Good (%)	26 (21.1%)	13 (14. 4%)
	Fair (%)	26 (21.1%)	25 (27.8%)
	Poor (%)	28 (22.8%)	13 (14. 4%)
	Very poor (%)	30 (24.4%)	22 (24.4%)
	Unreported ranking (%)	1 (0.8%)	0 (0%)
Mode		Very Poor	Fair

Social media integration

As shown in Table 4, the sites of the polytechnics and colleges of education both had mode values of "Very Poor" in respect of social media integration, with 92.2% of college sites and 86.2% of polytechnic sites receiving a "Very Poor" score. Testing of the social media dimension uncovered the worst website performance in each category.

Table 4: Social media integration

		Polytechnics	**Colleges of education**
	Excellent (%)	1 (0.8%)	1 (1.1%)
	Good (%)	0 (0%)	0 (0.0%)
	Fair (%)	5 (4.1%)	1 (1.1%)
	Poor (%)	9 (7.3%)	4 (4. 4%)
	Very Poor (%)	106 (86.2%)	83 (92.2%)
	Unreported ranking (%)	2 (1.6%)	1 (1.1%)
Mode		Very Poor	Very Poor

The extremely large number of websites found to be "Very Poor" in their social media integration all tended to lack the use of one or more of Facebook, Twitter, Instagram,

YouTube, and LinkedIn. Thus, the vast majority of Nigerian polytechnics and colleges of education were not, at the time of the research, harnessing the opportunities afforded by integration of social media into their websites, that is, opportunities to drive traffic to their websites, to build their brands, and to attract potential future students. Two of the 123 polytechnic websites and one of the 90 college of education websites had no value reported for their social media integration.

Security

As shown in Table 5, only 35% of polytechnic websites and 28.9% of college of education websites were found to have "Excellent" security. While the mode value for both categories of institution was "Excellent", the majority of sites in both categories had scores of "Fair" or worse, and there were a large number of sites with "Poor" and "Very Poor" scores. Roughly one-quarter of sites in each category were found to have "Poor" security, and roughly one-fifth of sites in each category had "Very Poor" security.

Table 5: Security

		Polytechnics	Colleges of education
	Excellent (%)	43 (35%)	26 (28.9%)
	Good (%)	0 (0.0%)	0 (0.0%)
	Fair (%)	23 (18.7%)	21 (23.3%)
	Poor (%)	30 (24.4%)	23 (25.6%)
	Very Poor (%)	27 (22%)	20 (22.2%)
	Unreported ranking (%)	0 (0%)	0 (0%)
Mode		Excellent	Excellent

Websites with excellent security are SSL-enabled, integrate HTTPS redirection of their pages, are flagged as safe by popular malware scanners, and ensure that no email addresses are found in plain text on webpages. The results suggest that insufficient number of Nigerian polytechnics and colleges of education are showing the necessary level of seriousness in their efforts to ensure the security of their websites.

HTML validation

As shown in Table 6, large percentages of the websites of both polytechnics (66.7%) and colleges of education (72.2%) were found to be "Excellent" in their HTML validation, meaning that they had between 1 and 25 HTML errors. Accordingly, the mode value for HTML validation was "Excellent" for the sites of both types of institution. The fewer the HTML errors, the better the chances of these websites having good indexing on search engines. Two of the 123 three polytechnic websites and one of the 90 college of education websites had no HTML validation ranking reports.

Table 6: HTML validation

		Polytechnics	Colleges of education
	Excellent (%)	82 (66.7%)	65 (72.2%)
	Good (%)	21 (17.1%)	13 (14.4%)
	Fair (%)	6 (4.9%)	4 (4.4%)
	Poor (%)	4 (3.3%)	2 (2.2%)
	Very Poor (%)	8 (6.5%)	5 (5.6%)
	Unreported (%)	2 (1.6%)	1 (1.1%)
Mode		Excellent	Excellent

Polytechnic websites versus colleges of education websites

As seen in Table 7, which provides the results from a non-parametric Mann-Whitney U test, there were no significant differences in the usability, operational performance, social media integration, security, or HTML validation measures between the websites of the polytechnics and those of the colleges of education. There was, however, a statistically significant difference in SEO performance between the polytechnic websites and those of the colleges of education ($U = 4553.5$, $P = 0.027$). The mean rank of polytechnic websites was higher than that of college of education websites, meaning that, overall, the polytechnic websites outperformed the college of education websites, on SEO performance, to a statistically significant degree.

Table 7: Mann-Whitney U test results

Performance parameter	Type of institution	N	Mean rank	Sum of ranks	Mann-Whitney U	Wilcoxon W	Z	Asymp. Sig. (2-tailed)
SEO	Polytechnic	122	114.18	13929.5	4553.5	8648.5	–2.211	0.027
	College of education	90	96.09	8648.5				
	Total	212						
Usability	Polytechnic	122	108.93	13289	5194	9289	–0.802	0.423
	College of education	90	103.21	9289				
	Total	212						
Operational performance	Polytechnic	122	103.06	12573	5070	12573	–0.973	0.33
	College of education	90	111.17	10005				
	Total	212						

Performance parameter	Type of institution	N	Mean rank	Sum of ranks	Mann-Whitney U	Wilcoxon W	Z	Asymp. Sig. (2-tailed)
Social media integration	Polytechnic	121	108.02	13071	5079	9084	–1.349	0.177
	College of education	89	102.07	9084				
	Total	210						
Security	Polytechnic	123	108.8	13382.5	5313.5	9408.5	–0.517	0.605
	College of education	90	104.54	9408.5				
	Total	213						
HTML validation	Polytechnic	121	103.46	12518.5	5137.5	12518.5	–0.706	0.48
	College of education	89	108.28	9636.5				
	Total	210						

Polytechnic website performance and institutional ownership (federal, state, and private)

Table 8 shows the results from the Kruskal-Wallis test that was carried out to examine whether there were differences in website performance that correlated with institutional ownership (federal government, state government, and private). As shown in the table, there were no statistically significant differences in the SEO, usability, social media integration, security, and HTML validation measures of Nigeria's polytechnic websites based on their ownership type (federal, state, and private). There was, however, a statistically significant difference (Chi-Square = 6.263, df = 2, P = 0.044) in the operational performance of Nigeria's polytechnic websites based on institutional ownership type (federal, state, and private).

Further testing found that the statistically significant difference in the operational performance lay between sites of the state government-owned and privately owned polytechnics. The mean rank of state government-owned polytechnic websites was higher than that of privately owned polytechnic websites. This shows that state government-owned Nigerian polytechnic websites had, overall, a statistically significant higher operational performance measure than the sites of the privately owned Nigerian polytechnics. There was, however, no statistically significant difference between federal government-owned and state government-owned polytechnic websites, or between federal government-owned and privately owned polytechnic websites.

Table 8: Kruskal-Wallis test results (polytechnic site performance and ownership type)

Performance parameter	Ownership type	N	Mean rank	Chi-Square	df	Asymp. Sig.
SEO	Federal	28	59.54	0.176	2	0.916
	State	45	62.92			
	Private	49	61.32			
	Total	122				
Usability	Federal	28	64.7	1.418	2	0.492
	State	44	57.38			
	Private	50	63.34			
	Total	122				
Operational performance	Federal	28	63.54	6.263	2	0.044
	State	44	70.3			
	Private	50	52.62			
	Total	122				
Social media integration	Federal	28	66.93	3.736	2	0.154
	State	44	57.58			
	Private	49	60.68			
	Total	121				
Security	Federal	28	61.13	0.033	2	0.983
	State	45	61.89			
	Private	50	62.59			
	Total	123				
HTML validation	Federal	28	67	1.592	2	0.451
	State	45	59.43			
	Private	48	58.97			
	Total	121				

College of education website performance and institutional ownership (federal, state, and private)

Table 9 provides the results of a Kruskal-Wallis test looking for correlations between college of education website performance and institutional ownership (federal government, state government, and private). As seen in Table 9, there were no significant differences in any of the six quality measures based on institutional ownership type.

Table 9: Kruskal-Wallis test results (college of education site performance and ownership type)

Performance parameter	Ownership type	N	Mean rank	Chi-Square	df	Asymp. Sig.
SEO	Federal	21	56.31	5.505	2	0.064
	State	44	43.68			
	Private	25	39.62			
	Total	90				
Usability	Federal	21	49.17	0.8	2	0.67
	State	44	43.91			
	Private	25	45.22			
	Total	90				
Operational performance	Federal	21	50.69	1.165	2	0.559
	State	44	44.31			
	Private	25	43.24			
	Total	90				
Social media integration	Federal	20	48.65	2.728	2	0.256
	State	44	43.98			
	Private	25	43.88			
	Total	89				
Security	Federal	21	46.24	0.131	2	0.937
	State	44	44.52			
	Private	25	46.6			
	Total	90				
HTML validation	Federal	21	51.43	3.574	2	0.167
	State	44	41.43			
	Private	24	45.92			
	Total	89				

5. Analysis of findings

The present study investigated six research questions to determine the performances of Nigerian polytechnic and colleges of education websites in six website parameters, pertaining to search engine optimisation, usability, social media integration, operational performance, security, and HTML validation. The study reported different levels of performance across the institutional types and website parameters. As reported in the findings, the majority of polytechnic websites were fair in search engine optimisation, while the majority of college of education websites were poor in search engine optimisation. The implication of this is that most of Nigeria's polytechnic and college of education websites are missing many critical SEO elements that can ensure improved web presence, ranking in search engines, visibility, and discoverability in the internet space (Lemos & Joshi, 2017; MintTwist, 2016).

This finding is in line with that of Mbanaso et al. (2015), who found in 2015 that 88% of Nigeria's polytechnics and 68% of Nigeria's colleges of education that they studied had a basic web presence. This could, therefore, mean that the developers of these institutional websites have not regarded search engine optimisation as an important parameter for attention over the past several years. However, it was at the same time found that most of the institutional websites in Nigeria's polytechnics and colleges of education were excellent in terms of their usability. This shows that these websites were able to render properly on mobile and hand-held devices, and were easily used with mobile viewports, favicons, legible font sizes, and tap target sizing.

This study found that the majority of polytechnic websites had very poor operational performance measures, while the majority of college of education websites had fair operational performance measures. According to Marszałkowski et al. (2014), some of the issues that are associated with websites that have very poor operational performance are slow server response time, large page file size, JavaScript errors, poor optimisation of images, and the use of inline styles, while fair operational performance implies fewer issues such as non-optimisation of images, deprecated HTML tags, and the use of inline styles. It is important that Nigerian educational institutions continuously review these and other operational performance measures. Furthermore, most of the institutional websites in Nigeria's polytechnics and colleges of education were found to be very poor in their social media integration performance measures. This shows that these websites lacked adequate connectivity to large social media networks like Facebook, Twitter, Instagram, YouTube, and LinkedIn. They had also not taken advantage of social media resources like Facebook Pixel and Twitter Cards. This presents opportunities to harness social media platforms in the Nigerian post-school education sector.

A minority of institutional websites in Nigeria's polytechnics and colleges of education were found to be "Excellent" in terms of their security. Websites with excellent security are SSL-enabled, integrate HTTPS redirection of their pages, are flagged

as safe by popular malware scanners, and ensure that no email addresses are found in plain text on webpages. This finding shows that the developers of a minority of the websites regarded security as a very important parameter. This is alarming given the challenge of cyber insecurity that is pervasive in the country (Ubabukoh, 2016).

At the same time, most of the institutional websites in Nigeria's polytechnics and colleges of education were found to be excellent in terms of their HTML validation. The fewer the HTML errors, the better the chances of these websites having good indexing, which improves their rankings on search engines. Markup validation is important for two reasons. First, it helps in cross-browser, cross-platform, and future compatibility. Second, when search engines encounter websites that have markup errors, they can take a variety of possible decisions in response to those errors—with many of these decisions resulting in some keywords not being parsed correctly and parts of the webpage not being indexed (Heng, n.d.).

Three inferential statistical tests were carried out to investigate if institutional types (polytechnic and college of education) or institution ownership types (federal, state, and private) had any significant effect on the website performance measures. The first test revealed that institutional type had no significant effect on the websites' performances in usability, operational performance, social media integration, security, or HTML validation. The search engine optimisation of polytechnic websites was, however, significantly better than that of the college of education websites. Ownership type had no significant effect on each of the six website parameters of colleges of education, while state-owned polytechnic websites performed better than privately owned polytechnic websites only in the operational performance parameter.

These findings show that the performance of the websites was mostly not affected by either the institutional type or the ownership type. This then suggests that the attention of the developers of these institutions' websites and of the stakeholders in these institutions needs to be shifted towards improving their weak points while not relenting on maintaining their strong points in website quality. The comparative approach taken in this present study is in line with some existing studies such as those of Kravchenko et al. (2021) and Akgül (2021), where institutional websites were assessed based on different website parameter types and different institutional ownership types.

6. Conclusions

The most noteworthy findings from this study are, in our analysis, (1) the sites' weak performance with respect to social media integration, SEO, and operational performance; and (2) their highly variable scores on security. It is clear that the vast majority of Nigerian polytechnics and colleges of education were not, at the time of the data collection, taking social media sufficiently seriously—in spite of ample evidence of the great amount of time that Nigerian youth spend on social

media, as noted by Adaugo et al. (2015). The youth must certainly be a core target audience for Nigerian polytechnics and colleges of education. Deficiencies in SEO and operational performance also call into question the seriousness of these higher education institutions about the task of building their brands and attracting new students.

Meanwhile, the highly variable scores on website security must, of course, be cause for concern. While it is laudable that a sizeable number of polytechnic websites (35%) and college of education websites (28.9%) were found to have "Excellent" security, it is unacceptable that the majority of sites in each category scored "Fair" or worse in this crucial performance area—and that about one-quarter of sites in each category were found to have "Poor" security, and roughly one-fifth of sites in each category were found to have "Very Poor" security.

References

Abdel-Basset, M., Zhou, Y., Mohamed, M., & Chang, V. (2018). A group decision making framework based on neutrosophic VIKOR approach for e-government website evaluation. *Journal of Intelligent and Fuzzy Systems*, *34*(6), 4213–4224. https://doi.org/10.3233/JIFS-171952

Adaugo, C. Q., Ovute A. O., & Obochi, C. I. (2015). The influence of the social media on the Nigerian youths: Aba residents experience. *Journal of Research in Humanities and Social Science*, *3*(3), 12–20.http://www.questjournals.org/jrhss/papers/vol3-issue3/C331220.pdf

Akgül, Y. (2016). Quality evaluation of e-government websites of Turkey. In *Proceedings of 11th Iberian Conference on Information Systems and Technologies (CISTI).* https://doi.org/10.1109/CISTI.2016.7521567

Akgül, Y. (2021). Accessibility, usability, quality performance, and readability evaluation of university websites of Turkey: A comparative study of state and private universities. *Universal Access in the Information Society*, *20*(1), 157–170. https://doi.org/10.1007/s10209-020-00715-w

Almahamid, S. M., Tweiqat, A. F., & Almanaseer, M. S. (2016). University website quality characteristics and success: Lecturers' perspective. *International Journal of Business Information Systems*, *22*(1), 41–61. https://doi.org/10.1504/IJBIS.2016.075717

Andalib, Z. & Danaee, H. (2013). A study on measuring the quality of university website. *Management Science Letters*, *3*(7), 1955–1960. https://doi.org/10.5267/j.msl.2013.06.031

Bukhari, S. A. H. (2011). What is comparative study? SSRN. https://doi.org/10.2139/ssrn.1962328

Burns, N., & Grove, S. (2009). *The practice of nursing research: Appraisal, synthesis and generation of evidence* (6th ed.). Saunders Elsevier.

Cartay, L. (2017, January 24). How to integrate social media into your website to increase your ecommerce sales. *Digital Doughnut.* https://www.digitaldoughnut.com/articles/2017/january/how-to-integrate-social-media-into-your-website-to

Chou, W., & Cheng, Y. (2012). A hybrid fuzzy MCDM approach for evaluating website quality of professional accounting firms. *Expert Systems with Applications*, *39*(3), 2783–2793. https://doi.org/10.1016/j.eswa.2011.08.138

Choudrie, J., Ghinea, G., & Weerakkody, V. (2004). Evaluating global e-government sites: A view using web diagnostic tools. *Electronic Journal of e-Government*, *2*(2), 105–114. https://bura.brunel.ac.uk/handle/2438/6995

Dani, D., & Agrawal, G. (2021). Evaluating the quality of Indian school education boards' websites using multi criteria decision making models. *International Journal of Information Technology*, *13*, 1–9. https://doi.org/10.1007/s41870-018-0119-y

Dominic, P. D. D., Jati, H., & Hanim, S. (2013). University website quality comparison by using non-parametric statistical test: A case study from Malaysia. *International Journal of Operational Research*, *16*(3), 349–374.

Dominic, P. D. D., Jati, H., Sellapan, P., & Nee, G. K. (2011). A comparison of Asian e-government websites quality: Using a non-parametric test. *International Journal Business Information Systems*, *7*(2), 220–246. https://doi.org/10.1504/IJBIS.2011.038513

Dominic, P. D. D., & Khan, H. (2013). Measuring quality of Asian airline websites using analytical hierarchy process: A future customer satisfaction approach. In *Proceedings of 2013 Information Systems International Conference (ISICO)*.

Elling, S., Lentz, L., De Jong, M., & Van den Bergh, H. (2012). Measuring the quality of governmental websites in a controlled versus an online setting with the "website evaluation questionnaire". *Government Information Quarterly*, *29*(3), 383–393. https://doi.org/10.1016/j.giq.2011.11.004

Faustina, F., & Balaji, T. (2016). Evaluation of universities websites in Chennai city, India using analytical hierarchy process. In *Proceedings of 2016 IEEE International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT)*. https://doi.org/10.1109/ICEEOT.2016.7754850

Geron, G. (2018, February 26). Best practices in institutional websites for international visibility. *Educause.* https://er.educause.edu/articles/2018/2/best-practices-in-institutional-websites-for-international-visibility

Glen, S. (2018, May 10). Total population sampling. *Statistics How To.* https://www.statisticshowto.com/total-population-sampling

Gong, J. W., Liu, H. C., You, X. Y., & Yin, L. (2021). An integrated multi-criteria decision making approach with linguistic hesitant fuzzy sets for e-learning website evaluation and selection. *Applied Soft Computing*, *102*, 107118. https://doi.org/10.1016/j.asoc.2021.107118

Greenfield, M. (2015). Why the web matters to higher education. *Site Improve.* https://siteimprove.com/en/blog/why-the-web-matters-to-higher-education

Hanover Research. (2014, July 22). Why marketing strategy matters in higher education. https://www.hanoverresearch.com/insights-blog/why-marketing-strategy-matters-in-higher-education/

Heng, C. (n.d.). What is HTML and CSS validation? Should you validate your web page? *thesitewizard*. https://www.thesitewizard.com/webdesign/htmlvalidation.shtml

Jati, H., & Dominic, D. D. (2009). Quality evaluation of e-government website using web diagnostic tools: Asian case. *IEEE Computer Society*, *1*, 85–89. https://doi.org/10.1109/ICIME.2009.147

Jundillah, M. L., Suseno, J. E., & Surarso, B. (2019). Evaluation of e-learning websites using the Webqual method and importance performance analysis. *E3S Web of Conferences*, *125*, 24001, 1–5. https://doi.org/10.1051/e3sconf/201912524001

Kaur, S., Kaur, K., & Kaur, P. (2016). An empirical performance evaluation of universities website. *International Journal of Computer Applications (0975–8887)*, *146*(15), 10–16. https://doi.org/10.5120/ijca2016910922

Khraim, H. S. (2015). The impact of search engine optimization dimensions on companies using online advertisement in Jordan. *American Journal of Business and Management*, *4*(2), 76–84. https://pdfs.semanticscholar.org/66fe/22cb0c4077235b9e9a5523396d5eb28f4c62pdf

Kincl, T., & Strach, P. (2012). Measuring website quality: Asymmetric effect of user satisfaction. *Behaviour and Information Technology*, *31*(7), 647–657. https://doi.org/10.1080/0144929X.2010.526150

Kravchenko, Y., Leshchenko, O., Dakhno, N., & Radko, M. (2021). Comparative evaluation of universities websites quality. *CEUR Workshop Proceedings*, *3132*, 166–175.

Lawal, I. (2019, March 28). Making colleges of education relevant in 21st century. *The Guardian.*https://guardian.ng/features/making-colleges-of-education-relevant-in-21st-century/

Lemos, Y. J., & Joshi, A. R. (2017). Search engine optimization to enhance user interaction. In *Proceedings of 2017 International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud).* https://doi.org/10.1109/I-SMAC.2017.8058379

Maksimava, M. (2022, July 19). Site quality: Google's major ranking signal explained. *SEO PowerSuite Blog.* https://www.link-assistant.com/news/site-quality.html

Marszałkowski, J., Marszałkowski, J. M., & Drozdowski, M. (2014). Empirical study of load time factor in search engine ranking. *Journal of Web Engineering*, *13*(1), 114–128.

Massey, A. P., Khatri, V., & Montoya-Weiss, M. M. (2007). Usability of online services: The role of technology readiness and context. *Decision Sciences Institute*, *38*(2), 277–308.

Mbanaso, U., Chukwudebe, G., & Atimati, E. (2015). Nigeria's evolving presence in cyberspace. *The African Journal of Information and Communication (AJIC)*, *15*, 106–113. https://doi.org/10.23962/10539/20334

MintTwist (2016, May 20). The importance of SEO for universities. https://www.minttwist.com/blog/seo-universities/

Nadeema, W., Andreini, D., Salo, J., & Laukkanen, T. (2015). Engaging customers online through websites and social media: A gender study of Italian Generation Y clothing consumers. *International Journal of Information Management*, *35*(4), 432–442. https://doi.org/10.1016/j.ijinfomgt.2015.04.008

National Board for Technical Education (NBTE). (n.d.). https://net.nbte.gov.ng

National Commission for Colleges of Education (NCCE) (n.d.). http://www.ncceonline.edu.ng

Napier, S. (n.d.). What makes a website a high quality website? *SEO Expert.* https://www.seoexpert.co.uk/high-quality-websites

Nwagwu, W. E., & Agarin, O. (2008). Nigerian university websites: A webometric analysis. *Webology*, *5*(4) http://www.webology.org/2008/v5n4/a65.html

Ogungbe, V. (n.d.). The importance of polytechnic education in Nigeria. *InfoGuide Nigeria.* https://infoguidenigeria.com/importance-polytechnic-education-nigeria/

Olaleye, S. A., Sanusi, I. T., Ukpabi, D. C., & Okunoye, A. (2018). Evaluation of Nigeria universities websites quality: A comparative analysis. *Library Philosophy and Practice*, 1717. https://digitalcommons.unl.edu/libphilprac/1717

Park, H., & Kim, Y. (2014) The role of social network websites in the consumer–brand relationship. *Journal of Retailing and Consumer Services*, *21*(4), 460–467. https://doi.org/10.1016/j.jretconser.2014.03.011

Pechnikov, A. A., & Nwohiri, A. M. (2012) Webometric analysis of Nigerian university websites. *Webology 9*(1). http://www.webology.org/2012/v9n1/a95.html

Perçin, S. (2019). A combined fuzzy multicriteria decision-making approach for evaluating hospital website quality. *Journal of Multi-Criteria Decision Analysis*, *26*(3–4), 129–144. https://doi.org/10.1002/mcda.1671

Rocha, A. (2012). Framework for a global quality evaluation of a website. *Online Information Review*, *36*(3), 374–382. https://doi.org/10.1108/14684521211241404

Shi, Y. (2006). Quality services need quality websites: An evaluation of three official world expo home pages using web diagnostic tools. In *Proceedings of 2006 IEEE International Conference on Service Operations and Logistics, and Informatics*. https://doi.org/10.1109/SOLI.2006.328962

Slawski, B. (2011, April 19). Just what is website quality? *SEO by the sea.* http://www.seobythesea.com/2011/04/just-what-is-web-site-quality/

Søruma, H., Andersen, K. N., & Vatrapu, R. (2012). Public websites and human–computer interaction: An empirical study of measurement of website quality and user satisfaction. *Behaviour and Information Technology*, *31*(7), 697–706. https://doi.org/10.1080/0144929X.2011.577191

Suwawi, J. D. D., Darwiyanto, E., & Rochmani, M. (2015). Evaluation of academic website using ISO/1EC 9126. In *Proceedings of 2015 3rd International Conference on Information and Communication Technology (ICoICT)*. https://doi.org/10.1109/ICoICT.2015.7231426

Tate, E. (2017, April 5). Prospective students seek information, but colleges don't always provide. *Inside Higher Ed.* https://www.insidehighered.com/digital-learning/article/2017/04/05/colleges-need-make-websites-easier-find-info

Torkildson, A. (2018, November 8). 13 new ways to incorporate social media in your website. *Social Media Explorer.* https://socialmediaexplorer.com/content-sections/news-and-noise/13-new-ways-to-incorporate-social-media-into-your-website/

Ubabukoh, O. (2016, June 7). Nigeria ranks 16th in cyberattacks vulnerability index. *Punch*. https://punchng.com/nigeria-ranks-16th-cyberattacks-vulnerability-index/

Evaluation of ABET accreditation path for a representative African undergraduate computer science programme

Raymond Greenlaw
Professor, Department of Computing, Mathematical, and Statistical Sciences, University of Namibia, Windhoek
https://orcid.org/0000-0003-0690-2112

Kauna Mufeti
Associate Professor, Department of Computing, Mathematical, and Statistical Sciences, University of Namibia, Windhoek
https://orcid.org/0000-0003-4864-1387

Abstract
As of the beginning of 2022, only four of 54 countries in Africa had programmes accredited by ABET—the widely used international accreditation body for tertiary education programmes in computing, engineering, applied and natural sciences, and engineering technology. This article provides results from an evaluation of a representative non-ABET-accredited African undergraduate computer science programme—the University of Namibia (UNAM) Bachelor of Science (BSc) in Computer Science—in terms of its potential for ABET accreditation. The study evaluates the UNAM programme against ABET's General Criteria (GC), and also against ABET's computer science Program Criteria (PrCr), in order to determine the steps UNAM would need to take were it to seek ABET accreditation for the programme. The authors also evaluate the level of difficulty that each of the steps would likely involve. The authors' evaluation aims to be replicable, so as to provide a potential tool for use by any African higher education entity seeking ABET accreditation for a learning programme.

Keywords
computer science, programme accreditation, ABET, curriculum development, quality assurance

Acknowledgement
First-listed author Greenlaw's participation in this research was supported by a Fulbright Fellowship.

DOI: https://doi.org/10.23962/ajic.i30.13987

Recommended citation
Greenlaw, R., & Mufeti, K. (2022). Evaluation of ABET accreditation path for a representative African undergraduate computer science programme. *The African Journal of Information and Communication (AJIC), 30*, 1-20.
https://doi.org/10.23962/ajic.i30.13987

1. Introduction

Accreditation is fast becoming a vital component of quality assurance and an indicator of competent university graduates worldwide. Defined as a "recognition accorded to an institution that has met the standards or criteria established by a competent agency or association" (Blauch, 1959), accreditation was initially mainly thought of as a US phenomenon, designed to regulate education in the absence of centralised government control (Mullen, 2001). With numerous notable benefits, such as providing standards and programme criteria, facilitating inter-institutional transfers, raising educational quality and standards, and providing guidance on the prestige of institutions (Blauch, 1959), accreditation is now a widely accepted practice worldwide (Stensaker, 2011). Many professions, including auditing, accounting, and engineering, require students to graduate from accredited programmes before being able to practise as professionals. Programmes in science, technology, engineering and mathematics (STEM) areas often seek accreditation as a way to distinguish themselves and to inform the general public about their quality.

ABET accredits programmes in four main areas: computing, engineering, applied and natural sciences, and engineering technology.[1] ABET's accreditation criteria are developed by experts, member societies, and by ABET commissions. There is a separate commission for each of the four main areas in which ABET accredits programmes. When criteria are being created or updated, the public is able to provide input and suggest modifications. At the time of this writing in mid-2022, ABET accredits 4,361 programmes at 850 institutions in 41 countries (ABET, n.d.-a). Each year, more than 175,000 students graduate from ABET-accredited programmes (ABET, n.d.-a).

1 "ABET" originally stood for "Accreditation Board for Engineering and Technology", but the wording behind the acronym was discontinued when ABET expanded to include computing and applied sciences.

To qualify for ABET accreditation, a programme needs to satisfy both ABET's General Criteria (GC) and the particular programme area's specific Program Criteria (PrCr).[2] PrCr are informed by the latest curriculum developments in a specific discipline, and hence specify additional requirements beyond the GC. The PrCr typically focus on student outcomes, curriculum, and faculty, with the bulk of the requirements being in the curriculum area.

The questions addressed in this study are: How near to meeting ABET accreditation criteria is a typical Sub-Saharan African programme in computer science? And, by extension, what would it take for such a programme to become ABET accreditable? We evaluated and analysed a representative programme critically, to see how well it measured up to ABET's requirements. The programme we chose to focus on is the Computer Science programme at the University of Namibia (UNAM), because it is the one that we are most familiar with (as we both work at the UNAM) and because we found that it is a representative programme in Sub-Saharan Africa. Our aim was to develop an approach which is general enough that it will highlight the kinds of challenges that many programmes, including those not in computer science, could expect to encounter while seeking ABET accreditation.

Also, so as to be fair to UNAM and to the UNAM programme that we evaluated in this study, it must be pointed out that it was fully understandable and expected that we were able to find some shortcomings in respect of the programme's compliance with ABET criteria, as the programme was not designed with specific attempts to meet ABET's GC and PrCr.

There is very little research literature relating directly to the method of analysis and evaluation we followed. However, of some relevance is the work of Atchison et al. (1968), Austing et al. (1979), and Tucker (1991), who laid early groundwork for the curriculum of computing programmes. The curriculum evolution for computing continued in Cassel et al. (2008); The Joint Task Force on Computing Curricula (2001); and the ACM/IEEE-CS Joint Task Force on Computing Curricula (2013). One of the most recent computing curriculum models is presented by ACM and IEEE-CS (2020). In Besterfield-Sacre et al. (2000) and Engel et al. (2009), foundations for the accreditation standards of computing programmes are presented

2 Technically, it is possible for a programme to receive accreditation under just the GC, but this is not the typical case. Most programmes also satisfy PrCr.

and their history outlined. Raj et al. (2022) is recent work discussing how to interpret ABET's computer science PrCr using competencies is presented. The ABET computing curriculum and accreditation criteria continue to be refined, as new directions in the computing discipline emerge.

2. ABET accreditation criteria

For the purposes of this study, we focus on the Computing Accreditation Commission (CAC), although much of the discussion also applies to the other ABET commissions because ABET harmonised its GC in 2009 (ABET, n.d.-b). ABET's Computing GC consist of eight individual criteria that programmes must satisfy:

- Students;
- Program Educational Objectives (PEOs);
- Student Outcomes (SOs);
- Continuous Improvement (CI);
- Curriculum;
- Faculty;
- Facilities; and
- Institutional Support.

The Students criterion requires that performance is evaluated and progress monitored so that students can attain the SOs. This in turn allows graduates to attain the PEOs. Advising is required in curriculum and career matters. Policies must be established and followed regarding transfer students. Programmes must enforce graduation requirements. For the PEOs criterion, programmes must establish PEOs, publish them, involve constituents in their revision, and make sure that the PEOs are in line with the institution's mission, the needs of its constituents, and the ABET GC.

For the SOs criterion, ABET has five SOs: (1) analysing a problem and identifying its solution; (2) designing, implementing, and evaluating a solution to a problem; (3) communication skills; (4) professional responsibility; and (5) teamwork. The CI criterion requires that programmes have a robust system in place that is used regularly to determine the extent to which the SOs are being attained. The system must lend itself to programme improvements. The Curriculum criterion requires mathematics appropriate to the programme, and specifies the necessary breadth and depth of topics covered. Computing topics must include the following: "1. Techniques, skills,

and tools necessary for the computing practice. 2. Principles and practices for secure computing. 3. Local and global impacts of computing solutions on individuals, organizations, and society" (ABET, n.d.-c, p. 6).

The Faculty criterion requires that the faculty members are competent and sufficient to develop and offer the programme in a manner that enables students to graduate on time. The Facilities criterion requires that the facilities and library are sufficient for students and faculty members to carry out their work. The Institutional Support criterion requires that the institution provide sufficient resources for the programme to operate in an appropriate and effective manner.

The ABET computer science PrCr, as is typical of ABET PrCr, focus on three GC areas: SOs, Curriculum, and Faculty. In the SOs for computer science, a sixth SO, in addition to the five listed above in respect of the GC, is added. It reads: "Apply computer science theory and software development fundamentals to produce computing-based solutions" (ABET, n.d.-c, p. 10). Clearly, the requirements in this statement apply only to computer science and not broadly to all fields involving computing, e.g., not information technology (IT) or information systems (IS). This is why the requirement is included in the PrCr, and not the GC. The Faculty criterion requires that some faculty members have PhDs in the field of computer science.

With respect to the Curriculum criterion in the PrCr for computer science, at least 40 hours of instruction are required, including:

- instruction on algorithms and complexity, computer science theory, concepts of programming languages, and software development;
- instruction on a general-purpose programming language;
- instruction on computer architecture and organisation, information management, networking and communication, operating systems, and parallel and distributed computing;
- instruction on computing-based systems at varying levels of abstraction;
- a project requiring integration and application of knowledge and skills acquired in earlier course work;
- at least 15 hours of rigorous mathematics, including discrete mathematics; and
- at least six hours in rigorous natural science courses, where the scientific method and laboratory work are included (ABET, n.d.-c).

It is our view that the UNAM BSc in Computer Science is representative of undergraduate computer science programmes in Sub-Saharan Africa, because we have found that many schools in the region follow similar models, e.g.,

- Bahir Dar University (n.d.), Ethiopia;
- Osun State University of Nigeria (n.d.);
- School of Advanced Level Studies (n.d.), Seychelles;
- Uganda Technology and Management University (n.d.);
- Unicaf University (n.d.), operating in several African countries;
- University of Dar es Salaam (n.d.);
- University of Ghana (n.d.);
- University of Nairobi (n.d.);
- University of Pretoria (n.d.); and
- University of Zimbabwe (n.d.).

3. Evaluation in terms of ABET GC Curriculum criterion and computer science PrCr

We now turn to the evaluation of UNAM's BSc in Computer Science programme against the curriculum requirements of ABET's GC Curriculum criterion and computer science PrCr. The course information presented in this section is available in the UNAM *School of Science Prospectus 2022* (UNAM, 2022a) and the UNAM *School of Engineering and the Built Environment Prospectus 2022* (UNAM, 2022b).

Mapping of UNAM programme in ABET format

A key part of ABET's Self-Study Report procedures is "Table 5-1", which is where programmes must describe their curricula (ABET, 2022d). Accordingly, Tables 1 to 4 below represent our mapping of UNAM's programme in the ABET Table 5-1 format, with each table corresponding to a year of study. We scaled UNAM's courses from 16 credits to four credits, which we regard as equivalents across the Namibian and US credit frameworks. And it should be noted that in the UNAM programme each full course has four contact hours per week and an associated practical session of three additional hours per week. Half-courses have four contact hours per week and practical sessions of 1.5 hours per week. The numbers in the columns represent the number of credits assigned to a given course, in the ABET accounting system. For example, Introduction to Digital Electronics (CIT 3511) has two credits of science and two credits of "other".

Table 1: First year of UNAM's programme (inserted into ABET's Table 5-1)

Course	Required (R), Elective (E), or Selected Elective (SE)	Math	Science	Computing Topics: Fundamental (F) or Advanced (A)	General Education (GE)	Other
English Communication & Study Skills (LCE 3419)	R				4	
Computer Literacy (CLC 3509)	R					2
Basic Mathematics (MAT 3511)	R					4
Programming Fundamentals I (CMP 3511)	R			4F		
Introduction to Digital Electronics (CIT 3511)	R		2			2
Fundamentals of Information Technology I (CIT 3512)	R					2
English for Academic Purposes (LEA 3519)	R				4	
Contemporary Social Issues (CSI 3580)	R				2	
Programming Fundamentals II (CMP 3512)	R			4F		
Precalculus (MAT 3512)	R					4
Fundamentals of Information Technology II (CIT 3521)	R					2
Introduction to Statistics (STS 3522)	R					4

Table 2: Second year of UNAM's programme (inserted into ABET's Table 5-1)

Course	Required (R), Elective (E), or Selected Elective (SE)	Math	Science	Computing Topics: Fundamental (F) or Advanced (A)	General Education (GE)	Other
Introduction to Database Systems (CMP 3611)	R			4F		
Object Oriented Programming I (COS 3611)	R			4F		
Mathematics for Computer Science (CMP 3651)	R	4				
Computer Networks I (CIT 3611)	R			4F		
Advanced Databases (COS 3632)	R			2F, 2A		
Object Oriented Programming II (COS 3612)	R			4F		
Computer Networks II (CIT 3612)	R			4F		
Computer Organisation & Architecture (CIT 3652)	R			2F, 2A		

Table 3: Third year of UNAM's programme (inserted into ABET's Table 5-1)

Course	Required (R), Elective (E), or Selected Elective (SE)	Math	Science	Computing Topics: Fundamental (F) or Advanced (A)	General Education (GE)	Other
Computer Theory (CMP 3711)	R			2F, 2A		
Software Engineering (CMP 3731)	R			2F, 2A		
Emerging Technologies (COS 3731)	R	4		4F		
Data Structures and Algorithms (COS 3711)	R			2F, 2A		
Operating Systems (COS 3732)	R			2F, 2A		
Human Computer Interaction (COS 3712)	R			2F, 2A		
Research Methodology (CMP 3752)	R			3F, 1A		
Web Design & Programming (CMP 3772)	R			2F, 2A		

Table 4: Fourth year of UNAM's programme (inserted into ABET's Table 5-1)

Course	Required (R), Elective (E), or Selected Elective (SE)	Math	Science	Computing Topics: Fundamental (F) or Advanced (A)	General Education (GE)	Other
Research Project (CMP 3810)	R			4F, 4A		
Network System Security (CMP 3821)	R			2F, 2A		
Wireless and Mobile Computing (CIT 3711)	R			2F, 2A		
Numerical Methods and Operations Research (CMP 3811)	SE			2F, 2A		
Distributed Systems (CMP 3851)	SE			2F, 2A		
Artificial Intelligence (COS 3871)	SE			2F, 2A		
Data Warehousing & Data Mining (COS 3812)	R			2F, 2A		
Entrepreneurship and Management of IT Systems (CMP 3832)	SE			2F, 2A		
Real Time Multimedia (CMP 3812)	SE			2F, 2A		
Database Programming (CMP 3872)	SE			2F, 2A		
Totals		4	2	53F, 31A 84	10	20
Total Credits: 120						

Identification of shortcomings

As can be observed in Table 1, UNAM's curriculum does combine technical and professional components to prepare students for a career, further study, and lifelong professional development in computer science. However, it is lacking in the General Education (GE) area (shortcoming 1, denoted s1). There are only 10 credits of general education, where eight of them are in English. The rest are coming from a half-semester course discussing social issues.

In Table 2 we see the programme's only required mathematics course beyond the precalculus level, Mathematics for Computer Science (CMP 3651), which is essentially a discrete mathematics course. Thus, the programme is lacking appropriate mathematics for computer science (s2). In Table 4, we see the programme has a total of 84 hours of up-to-date computing (53 fundamental and 31 advanced) topics, which provides both the breadth and depth required by ABET and far exceeds the 30 hours required in the GC for computing. This number also exceeds the 40 hours of required computer science topics as specified in the PrCr.

The programme includes techniques, skills, and tools necessary for computer science, and includes some discussion of the principles and practices for secure computing, but these security concepts could be stressed more throughout (s3). There is coverage of the local and global impacts of computing solutions on organisations, but more material could be added to focus on impacts on individuals (s4) and society (s5). Two other shortcomings identified are in respect of coverage of parallel computing (s6) and coverage of distributed computing (s7). Neither is included in a required course. Distributed Systems (CMP 3851), where students who choose this elective get exposure to distributing computing, is merely a selective elective, and parallel computing is entirely absent from the curriculum.

ABET requires at least 15 credit hours of mathematics that must include discrete mathematics and must have mathematical rigor at least equivalent to introductory calculus. Table 4 shows that the UNAM programme only has four credits of mathematics—i.e., the four-credit Mathematics for Computer Science (CMP 3651) course—with a rigor equivalent to introductory calculus (s8).

ABET also requires at least six credit hours in natural science course work, where students must develop an understanding of the scientific method and perform laboratory work. UNAM's curriculum does not meet this requirement (s9), as Table 1 shows that Introduction to Digital Electronics (CIT 3511) contains just two credits of science, and there are no other required science courses. The scientific method and appropriate laboratory work are missing, and the number of credits is deficient.

ABET's computer science PrCr include an additional student outcome (SO 6), which requires the application of computer science theory and software development fundamentals to produce computing-based solutions. UNAM's Computer Theory (CMP 3711) course is only a pre-requisite for a single course: Artificial Intelligence (COS 3871), which is a selective elective. Thus, UNAM's coverage of the first half of SO 6—application of computer science theory—is weak (s10). The second half of SO 6—software development to produce computing-based solutions—is satisfied by the programme's inclusion of Object Oriented Programming II (COS 3612), Software Engineering (CMP 3731), Human Computer Interaction (COS 3712), and Research Project (CMP 3810).

The only remaining item in the ABET computer science PrCr is in the Faculty criterion, which states: "In addition to the General Criteria Faculty requirements, some full-time faculty members must have a Ph.D. in computer science " ABET (n.d.-c, p. 11). In UNAM's BSc in Computer Science programme, three of the six faculty members have (at the time of writing this article) PhDs in computer science, thus satisfying the ABET computer science PrCr Faculty criterion.

In the next section, we examine the steps that would need to be taken to bring UNAM's BSc in Computer Science curriculum into compliance with ABET Curriculum criteria and with ABET computer science SO 6 (s10).

4. Required changes for compliance with ABET GC Curriculum criterion and computer science PrCr

In section 3, we identified 10 shortcomings (s1–s10) in UNAM's BSc in Computer Science in relation to ABET's GC Curriculum criterion and ABET's computer science PrCr (SO 6). Table 5 summarises the shortcomings and proposes solutions, assigning each solution a degree of difficulty ("easy" or "medium").

Table 5: Shortcomings and solutions

	Shortcoming	Solution	Degree of difficulty
s1	Lack of General Education (GE) courses.	Reduce computing requirements and introduce additional GE courses.	Easy
s2	Lack of appropriate mathematics for computer science in terms of the GC.	See s8.	Medium
s3	Insufficient coverage of principles and practices for secure computing.	Include some modules in existing courses.	Easy
s4	Lack of coverage of the local and global impacts of computing solutions on individuals.	Where the impact on organisations is discussed, include discussions on individuals.	Easy
s5	Lack of coverage of the local and global impacts of computing solutions on society.	Where the impact on organisations is discussed, include discussions on society.	Easy
s6	Lack of compulsory coverage of distributed computing.	Make Distributed Systems (CMP 3851) a required course rather than a selective elective.	Easy
s7	Lack of coverage of parallel computing.	In the Distributed Systems course, add fundamental coverage of parallel computing.	Easy
s8	Lacking 11 credit hours of mathematics that must, in terms of the computer science PrCr, have mathematical rigor at least equivalent to introductory calculus.	Reduce computing requirements and add three appropriate four-credit mathematics courses.	Medium
s9	Missing six credit hours of natural science course work intended for science and engineering majors, where students must develop an understanding of the scientific method and perform lab work.	Reduce computing requirements and add in two appropriate science courses.	Medium
s10	Lack of application of computer science theory to produce computing-based solutions.	Computer Theory (CMP 3711) can be made a pre-requisite for several other courses, and theory integrated more throughout the programme.	Medium

In Table 5, we have classified solutions to six of the 10 shortcomings detected in UNAM's programme as "easy" solutions. These are solutions that would require minimal effort—e.g., an administrative change, or a small addition of material to a course.

With respect to solutions to three of the shortcomings (s2, s8, and s9), computing requirements would need to be traded off for mathematics and science courses. We ranked these solutions as being of "medium" difficulty, requiring substantially more effort to execute than easy solutions. Although it is not too difficult to rearrange the curriculum, the students might struggle with the additional mathematics—and students might struggle financially if UNAM had to increase fees in order to cover the costs of additional mathematics courses. Shortcoming s10 is ranked as being of medium difficulty because it involves changes to the pre-requisite structure throughout the curriculum, and also to the content of a number of courses.

5. Evaluation in terms of remaining ABET GC (excluding Curriculum GC)

Now that we have addressed the ABET computer science PrCr and the Curriculum requirements of the ABET GC, we must now evaluate UNAM's BSc in Computer Science with respect to the remaining seven GC (excluding the Curriculum GC), which are:

- Students;
- Program Educational Objectives (PEOs);
- Student Outcomes (SOs);
- Continuous Improvement (CI);
- Faculty;
- Facilities; and
- Institutional Support.

Students

UNAM computer science faculty members evaluate student performance in their courses, and student progress is monitored by advisors, so that students are aware of programme requirements and can graduate in a timely manner. The programme has not established SOs or PEOs in an ABET style. The curriculum leads to students attaining SOs 1 and 2: analysing, designing, implementing, and evaluating a computing-based solution to a problem. SOs 3, 4, and 5 deal with communication, professional and ethical responsibilities, and teamwork. These are not a specific focus of the programme and material would need to be added to bring the curriculum into compliance with these SOs. Only the Research Project module (CMP 3810) partially addresses SO 3. Students are

advised regarding curriculum and career matters. The programme has and enforces policies for transfers. The programme does not allow academic credit to be granted for work. There is a system of checks and balances to make sure that all students who graduate meet all graduation requirements. The programme needs to establish PEOs.

PEOs

As noted, the programme does not have ABET-style PEOs. However, the programme has clear goals for its graduates in the form of a set of its own nine SOs. The programme focuses on what students need to be able to attain by graduation rather than on what the students will be doing three to five years after graduation. These SOs are not publicly documented, as ABET requires of its SOs and PEOs. The programme would need to define PEOs and their constituents. A documented, systematically utilised, and effective process for the periodic review of the PEOs, involving programme constituencies, would need to be implemented to ensure they meet ABET's requirements.

SOs

The programme would need to adopt ABET's GC SOs (1–5) and the computer science PrCr's SO 6. Its internal nine SOs do in fact map to a number of ABET's SOs. Along with the PEOs, the programme would need to publish the SOs on its website, as well as enrolment data.

CI

Because the programme has not adopted ABET's SOs, there is no formal assessment process in place to assess these items. The programme would need to develop and regularly use an appropriate, documented processes for assessing and evaluating the extent to which the SOs are being attained. Those results would need to drive programme improvements. Many programmes struggle with the CI criterion. It can take a considerable length of time to develop and implement a good CI process. The programme's faculty members would need to develop additional expertise in assessment and evaluation.

Faculty

Each faculty member has the expertise and educational background necessary to contribute to the programme as expected. Faculty members have the education, professional credentials and certifications, and professional experience necessary. They are sufficient in number to maintain continuity, stability, oversight, student interaction, and advising. However, the faculty members need more professional development in order to remain current. The faculty has the responsibility and authority to improve the programme. When PEOs and SOs are defined and adopted, the faculty will need to modify the curriculum so that the SOs can be attained and, in turn, the PEOs reached.

Facilities

Faculty members have good offices, classrooms are sufficient, and there is sufficient work space for students. Internet connectivity is reasonable, as is the software provided to faculty members for teaching their courses. However, it is a challenge for UNAM to maintain high-quality and up-to-date laboratory and library services. Some of the machines in the labs are outdated, as are some of the materials in the library. Through additional online subscriptions, the library's collection can become sufficient without too much added cost. The institution and programme do, however, need to work with the Namibian government and industry to improve the programme's facilities. The main difficulty is securing an adequate budget for replacement of aging equipment, on a rotational basis.

Institutional Support

The programme is attracting and retaining faculty members. However, during a 2021 restructuring exercise at the institution, the School of Computing (where the BSc in Computer Science was housed) was merged with the Department of Mathematics and Statistics. This move seems to be contrary to the trend with computing departments elsewhere in world, which tend to be expanding rather than being absorbed into other departments. The programme has a technical support specialist, but, as stated above, the resources to maintain and upgrade faculty equipment are lacking. Additional resources are needed to provide the students and faculty members with up-to-date equipment. In addition, the budget for continued professional development may need to be augmented.

We now turn to consideration of ABET's fees. For programmes outside the US, ABET has an annual maintenance fee (at the time of writing in mid-2022) of USD1,530 per campus per commission, plus the same fee for each accredited programme at a campus. This means that were UNAM's BSc in Computer Science to become ABET-accredited, there would be an annual fee of USD3,060 (comprising the USD1,530 campus fee and the USD1,530 programme fee). Additionally, when a site visit occurs, which is normally every six years for a programme that meets ABET's requirements, there is a basic fee of USD8,235 and an additional fee of the same amount for each of the evaluators (usually there are two, thus costing USD16,470), in addition to the evaluators' travel expenses. Other fees may apply. For a complete listing of fees, see ABET (n.d.-e).

Table 6 summarises the non-Curriculum GC shortcomings and proposes solutions, with each solution given a difficulty ranking of "easy", "medium", or "difficult".

Table 6: Shortcomings and solutions with respect to remaining ABET GC (excluding Curriculum GC)

Shortcoming	Solution	Degree of difficulty
SOs are defined, but they are not ABET's SOs.	Adopt ABET's SOs. Modify the curriculum to make sure SOs 3, 4, and 5 can be attained by the students. SO 6 is addressed elsewhere.	Difficult
No ABET-style PEOs are defined.	Formalise the goals that the programme is trying to have its students attain three to five years after graduation. Formalise an ABET process for maintaining the PEOs and keeping the constituents involved.	Difficult
PEOs and SOs are not published.	Make the PEOs and SOs publicly known. Publish enrolment data.	Easy
Lack of CI process.	Develop and implement a complete assessment and evaluation process, and use it to improve the programme.	Difficult
Some faculty members not remaining professionally current.	Provide professional development opportunities and funds so that faculty members can remain current in computer science.	Medium
Lack of support to ensure SOs are attained and PEOs achieved.	Modify the curriculum to make sure the SOs can be attained and that they support the PEOs.	Medium
Out-of-date equipment in computing labs and library.	The library needs some additional funding in order to be able to subscribe to addutional electronic resources, and significant funds need to be budgeted for upgrading/ replacing equipment.	Difficult
Insufficient institutional support.	The 2021 reorganisation merged the School of Computing into the Department of Mathematics and Statistics. This move has reduced the focus on, and some of the resources for, computer science. The administration must be encouraged to work with and support the programme. Computer science programmes sprang out of math programmes 40 years ago, and thus, moving such a programme back inside a mathematics department was a step in the wrong direction.	Difficult
Paying ABET's annual dues and visit fees.	Budget needs to be allocated to pay for ABET's annual dues and fees. These are significant costs for UNAM.	Difficult

6. Conclusions

In this article we have sought to chart the way forward for a typical African undergraduate computer science programme—UNAM's BSc in Computer Science—if it were to seek ABET accreditation. We have also sought to conduct our evaluation of the UNAM programme in such a manner that other institutions would be able to make use of the methodology if they wish to consider ABET accreditation. At the same time, however, our approach could not be applied blindly and directly to all programmes. Adjustments would need to be made to cater to the specifics of the existing programme. And there is little doubt that pursuit of ABET accreditation of a programme requires strong buy-in from both faculty members teaching the programme and decision-makers in the broader institutional administration.

References

ABET. (n.d.-a). Accreditation. https://www.abet.org/accreditation

ABET. (n.d.-b). Rationale for revising criteria 3 and 5. https://www.abet.org/rationale-for-revising-criteria-3-and-5/

ABET. (n.d.-c). Criteria for accrediting computing programs, 2022-23. Computing Accreditation Commission, ABET. https://www.abet.org/accreditation/accreditation-criteria/criteria-for-accrediting-computing-programs-2022-2023

ABET. (n.d.-d). Self-study templates. Computing Accreditation Commission, ABET. https://www.abet.org/accreditation/accreditation-criteria/self-study-templates

ABET. (n.d.-e). Fees for programs outside the U.S. https://www.abet.org/accreditation/cost-of-accreditation/fees-for-programs-outside-the-u-s/

ACM/IEEE-CS Joint Task Force on Computing Curricula. (2013). *Computer science curricula 2013 (CS2013).* Association for Computing Machinery (ACM) and IEEE Computer Society (IEEE-CS).

Association for Computing Machinery (ACM), & IEEE Computer Society (IEEE-CS). (2020). *Computing curricula 2020 (CC2020): Paradigms for global computing education.* https://www.acm.org/binaries/content/assets/education/curricula-recommendations/cc2020.pdf

Atchison, W. F., Conte, S. D., Hamblen, J. W., Hull, T. E., Keenan, T. A., Kehl, W. B., … Young, D. M. (1968). Curriculum 68: Recommendations for academic programs in computer science: A report of the ACM Curriculum Committee on Computer Science. *Communications of the ACM, 11*(3), 151-197. https://doi.org/10.1145/362929.362976

Austing, R. H., Barnes, B. H., Bonnette, D. T., Engel, G. L., & Stokes, G. (1979). Curriculum '78: Recommendations for the undergraduate program in computer science—a report of the ACM Curriculum Committee on Computer Science. *Communications of the ACM*, *22*(3), 147-166. https://doi.org/10.1145/359080.359083

Bahir Dar University. (n.d.). Bachelor of Science in Computer Engineering. https://bit.bdu.edu.et/content/computer-engineering

Besterfield-Sacre, M., Shuman, L. J., Wolfe, H., Atman, C. J., McGourty, J., Miller, R. L., … Rogers, G. M. (2000). Defining the outcomes: A framework for EC-2000. *IEEE Transactions on education*, 100-110. https://doi.org/10.1109/13.848060

Blauch, L. E. (1959). *Accreditation in higher education.* US Government Printing Office.

Cassel, L., Clements, A., Davies, G., Guzdial, M., McCauley, R., McGettrick, A., … Weide, B. W. (2008). *Computer science curriculum 2008: An interim revision of CS 2001.* Association for Computing Machinery (ACM) and IEEE Computer Society (IEEE-CS). https://www.acm.org/binaries/content/assets/education/curricula-recommendations/computerscience2008.pdf

Engel, G., Impagliazzo, J., & LaMalva, P. (2009). A brief history of the computing sciences accreditation board (CSAB): Promoting quality education in the computing fields. Paper presented at 2009 IEEE Conference on the History of Technical Societies. https://doi.org/10.1109/HTS.2009.5337839

Greenlaw, R., & Mufeti, T. K. (2022). Reducing cyber crime in Africa through education. Paper presented at World Engineering Education Forum and The Global Engineering Deans Council (WEEF&GEDC) Conference, 28 November to 1 December, Cape Town.

Osun State University of Nigeria. (n.d.). Curriculum for B.Sc. Degree in Computer Science. https://uniosun.edu.ng/Curriculum/SET/ICT/CSC.pdf

Raj, R. K., Kumar, A. N., Sabin, M., & Impagliazzo, J. (2022). Interpreting the ABET computer science criteria using competencies. In *SIGCSE 2022: Proceedings of the 53rd ACM Technical Symposium on Computer Science Education V. 1* (pp. 906-912). https://doi.org/10.1145/3478431.3499293

School of Advanced Level Studies. (n.d.). http://www.sals.edu.gov.sc

Stensaker, B. (2011). Accreditation of higher education in Europe – moving towards the US model? *Journal of Education Policy, 26*(6), 757-769. https://doi.org/10.1080/02680939.2010.551785

The Joint Task Force on Computing Curricula. (2001). Computing curricula 2001. *Journal on Educational Resources in Computing (JERIC), 1*(3es). https://doi.org/10.1145/384274.384275

Tucker, A. B. (1991). Computing curricula 1991–report of the ACM/IEE-CS Joint Curriculum Task Force. *Communications of the ACM*, 34(6). 68–84. https://doi.org/10.1145/103701.103710

Uganda Technology and Management University. (n.d.). Bachelor of Science in Computer Science. http://utamu.ac.ug/academics/bachelors-degree-programmes/bachelor-of-science-in-computer-science

Unicaf University. (n.d.). BSc in Computer Science. https://unicafuniversity.com/course/bsc-computer-science

University of Dar es Salaam. (n.d.). Bachelor of Science in Computer Science. https://cse.udsm.ac.tz/programs/program-details/bsc-in-computer-science-%28bsc-cs%29

University of Ghana. (n.d.). Department of Computer Science: Undergraduate courses. https://dcs.ug.edu.gh/undergrad-course

University of Nairobi. (n.d.). Bachelor of Science in Computer Science. https://computerscience.uonbi.ac.ke/admission-content-type/bachelor-science-computer-science

University of Namibia (UNAM). (2022a). *School of Science prospectus 2022*. https://www.unam.edu.na/wp-content/uploads/2022/01/Prospectus-2022-School-of-Science.pdf

UNAM. (2022b). *School of Engineering & the Built Environment prospectus 2022*. https://www.unam.edu.na/wp-content/uploads/2022/01/Prospectus-2022-School-of-Engineering.pdf

University of Pretoria. (n.d.). BSc (Computer Science). https://www.universityofpretoria-international.com/programmes/bsc-computer-science

University of Zimbabwe. (n.d.). Bachelor of Business Studies & Computing Science.

CRITICAL INTERVENTIONS

Understanding state-level variations in India's digital transformation

Smriti Parsheera
Fellow, CyberBRICS project, Fundação Getulio Vargas (FGV) Law School, Rio de Janeiro; and PhD Candidate, Indian Institute of Technology Delhi, New Delhi
https://orcid.org/0000-0001-8985-0594

Abstract
India's digital transformation is often described in terms of the country's vast internet penetration, growing mobile connectivity, and widespread uptake of digital payments and other online services. Undoubtedly, India has made tremendous progress on all these fronts over the last few years. But digitalisation narratives founded upon aggregate national statistics bear the risk of assuming a homogeneity of digital access and experience in the country. This article highlights some of the state-level differences in digital access, skills, and infrastructure across India—as a basis for dispelling assumptions about the homogeneity and universality of India's digital transformation. The article draws attention to the varying levels of digital readiness within India, and to the need to account for these variations in the design and implementation of the country's digital initiatives.

Keywords
India, digital transformation, digital inclusion, digital readiness, state level

Acknowledgements
This article draws upon elements of a forthcoming CyberBRICS[1] project report. The author is grateful to Larissa Galdino de Magalhães Santos at the CyberBRICS project, and the editorial team at *The African Journal of Information and Communication (AJIC),* for valuable inputs.

DOI: https://doi.org/10.23962/ajic.i30.15082

Recommended citation
Parsheera, S. (2022). Understanding state-level variations in India's digital transformation. *The African Journal of Information and Communication (AJIC), 30*, 1-9. https://doi.org/10.23962/ajic.i30.15082

1 See https://cyberbrics.info. The BRICS countries are Brazil, Russia, India, China and South Africa.

1. Introduction

The Constitution of India describes the country, in Article 1, as "a Union of States" (Republic of India, 1950). The country is divided into 28 states and eight union territories, with the responsibility to administer different functional areas in these territories divided between the central and state governments. At the central level, India's digital transformation strategy is built on the foundational blocks of strengthening access, the adoption of e-governance initiatives, and the empowerment of citizens. These key objectives are articulated in the government's Digital India programme, which has been in operation since 2015 (Government of India, n.d.a). Digital India brings together a number of different schemes and initiatives that include projects with a pan-Indian application, as well as those developed or implemented at the level of state governments.

The success of the Digital India programme is often reported in terms of statistics about the country's growing internet access, mobile penetration, adoption of digital services, and booming start-up culture (Kaka et al., 2019; Prasad, 2022). India has indeed seen tremendous developments on all these fronts. Yet a study of the country's digital trajectory is incomplete without the inclusion of a more localised understanding of how the digital story is playing out in the country's different states (Singh et al., 2013).

Accordingly, this article focuses on the differences in digital adoption and implementation across India's states, and how those differences may contribute to differentiated outcomes for citizens. Based on a review of the available data and literature, this article notes that, in addition to administrative variations, digitisation outcomes are also influenced by factors like the country's rural–urban divide, varying levels of digital literacy and skills, and other socio-economic considerations. All of these play a role in determining how a diverse spectrum of user groups across the country experience, in varying ways, India's digital transformation.

2. A "mobile first" approach to digital transformation

Being the second most populous country in the world, it is unsurprising that many of India's digital achievements are highlighted in terms of the size of its user base. According to official estimates, India currently has more than 1.15 billion mobile connections and almost 830 million internet subscribers (TRAI, 2022). These developments have spurred a "mobile first" strategy that can be seen in the design of services offered by the government, as well as in private offerings. For instance, the government's UMANG (Unified Mobile Application for New-age Governance) app emphasises a "mobile first" approach while seeking to bring together all e-governance initiatives on its platform (MeitY, n.d.). The Jan Dhan-Aadhaar-Mobile (JAM) scheme offers another important example. JAM focuses on leveraging India's mobile phone strength (*Press Trust of India*, 2021) and enrollments in its digital identity programme, Aadhaar (UIDAI, 2022), to push for the adoption of banking services in

the country (PM India, n.d.), on the assumption that most individuals already have a mobile phone and digital identity, thus paving the way for the adoption of other digitally enabled services.

The central government's Direct Benefit Transfer (DBT) scheme is an example of such a service. Under the DBT scheme, disbursements of various welfare benefits are made directly to the bank accounts of individuals, with Aadhaar being the "preferred" mode of authentication (Government of India, n.d.b). In August 2021, the government also introduced a new digital payment solution called e-RUPI, which is a pre-paid digital voucher that can be received by a beneficiary on their phone as an SMS or QR code. It is meant to be used only for certain designated purposes, such as an e-RUPI transfer by the Health Ministry to avail health services at designated hospitals that have agreed to redeem the e-RUPI vouchers. The government's long-term plan is to connect e-RUPI with several other disbursements under the DBT scheme (Ministry of Finance, 2021), meaning that the mobile phone will play an even more central role in an individual's ability to access their welfare entitlements.

The assumption about near-ubiquitous mobile adoption in the country is also visible in the widespread use of one-time passwords (OTPs), received via SMS, for accessing various products and services. In some cases, this electronic means of verification coexists with offline alternatives. For instance, the Income Tax Department currently allows individuals to verify their tax filings through various electronic means, including an Aadhaar-based OTP, as well as by using traditional postal services.

But an offline option is not always provided. When the government launched its CoWin portal to manage the administration of COVID-19 vaccinations, it was mandatory for all users to complete an OTP check in order to log onto the system. Access to vaccination bookings was, therefore, predicated on a user having either their own mobile phone or the ability to seek assistance from others who had a phone and were willing to help. The rush for limited vaccination slots in the early days of the vaccination rollout made it clear that the digital-only strategy inevitably favoured mobile-savvy citizens over others (Santuka, 2021).

The assumption about ubiquitous access has, therefore, been challenged in the light of available evidence on the inequities that exist in India's digital ecosystem. The country's digital divide has most notably been demonstrated with respect to both gender (Barboni et al., 2018) and the rural–urban divide (Pandey, 2020; Singh et al., 2013). The latest Mobile Gender Gap Report by GSMA found a 14% gap between Indian men and women for mobile ownership and a 41% gap for mobile internet usage (GSMA, 2022). With respect to the rural–urban divide, 2021 data released by the Telecom Regulatory Authority of India (TRAI) showed rural and urban internet densities (subscriptions per 100 people in the population) standing, respectively, at 37% and 104% (TRAI, 2022). This gap is significant in itself, given that the majority

of India's population (approximately 65%) resides in its rural areas (World Bank, n.d.). The differences become all the more stark when the data is disaggregated at the level of different states.

3. Understanding state-level variations

For the purposes of telecommunications licensing, India is divided into 22 service areas. This demarcation broadly corresponds to the boundaries of India's states, subject to some variations. For instance, states such as Madhya Pradesh, Andhra Pradesh, and Bihar, which have each been bifurcated into two separate states, continue to exist as one service area. Further, a service area may consist of multiple smaller states, as seen in the case of the North Eastern service area, which includes six states from that region (Arunachal Pradesh, Manipur, Meghalaya, Mizoram, Nagaland, and Tripura).

The data points to significant variations in the state of access across different parts of the country. According to TRAI's reports, internet access density—the number of internet subscribers per 100 people—ranges widely, from the highest figure of 186% in Delhi[2] to the lowest figure of 36% in the Bihar service area (TRAI, 2022). Further, with the exception of one service area (Kerala, which has a rural internet density of 149 per 100 people and an urban internet density of 64 per 100 people), the figures for rural internet density are lower than for urban density in all service areas.

Intra-service area variations can also be seen in cases where two or more states are part of the same service area. The service area of Bihar, which, as noted above, has the poorest internet density in the country, consists of two states—Bihar and Jharkhand—that have widely different internet densities, of 43% and 10% respectively. Similarly, Chhattisgarh, which forms part of the Madhya Pradesh service area, performs significantly worse than its counterpart state, Madhya Pradesh, in the same service area. Chhattisgarh has the country's lowest internet density, at 7%, while internet density in Madhya Pradesh stands at 63% (TRAI, 2022). Finally, data released by the government in 2021, about villages that still remain outside the grid of mobile connectivity, found that there were 25,000 unconnected villages, and the majority of these villages were concentrated in the states of Odisha, Madhya Pradesh, and Maharashtra (Srivastava, 2021).

Another general trend that emerges is that the regions that fare worst in terms of inclusiveness in internet access also perform relatively poorly on other human development indicators (Bhardwaj, 2021; Parsheera, 2019b). A more granular sub-division, down to the levels of districts, blocks, and villages—which are smaller units

2 One of the possible reasons could be that many users have two or more internet subscriptions.

of administration within a state—can be expected to reflect a similar trend. TRAI's data on internet subscriptions does not go down to the level of these units, but other sources offer relevant indications. Intra-state variations have, for instance, been observed in the rollout of the government's BharatNet project, which seeks to provide broadband connectivity to most of India's villages. In the state of Himachal Pradesh, all the Gram Panchayats (village-level units) selected for coverage in the first phase of the project were situated "in the relatively better-off[3] districts of Hamirpur, Mandi and Solan" (Parsheera, 2019a)—thus excluding large portions of the state's tribal groups, which are concentrated in the districts of Lahaul-Spiti, Kinnaur, and Chamba.

In addition, the network experience of users in less developed regions has been found to be poorer, compared to service areas with a higher revenue-earning potential. India's North Eastern Region, which has historically been among the country's most neglected and underdeveloped areas (Gokhale, 2022; Barua, 2011), offers an example. According to an analysis by Open Signal, the average data download and upload speeds in the North Eastern Region were, respectively, 23.4% and 43.3% lower than the national average (Khatri, 2022). The neighbouring state of Assam, which also lies in India's North Eastern Region (although it is a separate telecommunications service area), displayed similar results. This may partially be explained by the unwillingness on the part of telecommunication providers to invest in infrastructure in remote and inaccessible areas with comparatively low revenue potential. In addition to commercial considerations, state-level policies (as discussed in the next section) also influence the quality of digital access available to their residents.

It has also been found that the ability of different user groups to benefit from e-governance solutions is highly dependent on their level of digital skills. In this context, researchers have found that only 38% of households in India are digitally literate—61% in urban areas and 25% in rural areas (Mothkoor & Mumtaz, 2021). (A household is regarded as digitally literate if at least one person can operate a computer and use the internet.) In certain regions—such as Uttar Pradesh, Madhya Pradesh, Chhattisgarh, Bihar, and parts of Jammu and Kashmir—the rural digital literacy rates are less than 20%. In contrast, states such as Goa and Kerala have achieved more than 70% digital literacy in both rural and urban areas. The researchers attribute this to specific digital literacy initiatives undertaken by those states, such as a computer literacy programme in Kerala and a memorandum of understanding to promote digital literacy entered into by the state of Goa with Google (Mothkoor & Mumtaz, 2021).

3 In terms of economic development and annual rate of growth of the district's income (Government of Himachal Pradesh, 2018).

4. Right-of-way policies

Ensuring "broadband for all" is contingent upon the establishment of the necessary infrastructure in the form of optical fibre, copper cables, telecommunication towers, and other apparatus required for the operation of networks. The establishment of such infrastructure requires permissions for use of public land and public rights-of-way. In terms of India's constitutional dispensation, the legislative responsibilities are divided between the central government and the states, with some sectors driven by the former and others by the latter. The power to make laws relating to posts, telegraphs, and communications is vested in the central government (Entry 31, List 1, Seventh Schedule, Constitution of India, 1950). But responsibility for land use policy, law, and regulation, which have an important bearing on the setting up of telecommunications infrastructure, falls under the domain of state governments (Entry 18, List 2, Seventh Schedule, Constitution of India, 1950). This allows the states to determine the policies, costs, and timelines for land-related permissions for laying down telecommunications infrastructure.

The inconsistent implementation of right-of-way policies in India has been found to delay the expansion of India's telecommunications infrastructure (GSMA, 2020; Marwah, 2019). This has been a problem not only for private operators but also for the roll-out of the government's own BharatNet project, which, as noted above, aims to achieve universal broadband access (TRAI, 2016). In addition to permissions from state governments (and local authorities within them), right-of-way permissions are sometimes also required from central government bodies like the National Highway Authority of India and the Indian Railways (TRAI, 2016). As a result, the requirement of authorisations from multiple authorities often results in extensive delays and exorbitant fees.

In some cases, government functionaries have observed a direct correlation between low broadband penetration and non-alignment of right-of-way rules by the state governments (Aulakh, 2021). Right-of-way challenges also contribute to the low rate of fiberisation of telecommunications towers in the country, i.e., the low proportion of towers that are connected to a fibre-based backhaul. At present, only about 30% of India's towers (and fewer than 7% of households) are connected to fibre (Surya, 2021). The rates of fiberisation are lowest in mountainous states such as Himachal Pradesh, and in remote parts of the North Eastern Region (*The Economic Times*, 2022).

In a bid to streamline these processes, the central government issued the Indian Telegraph Right of Way Rules in 2016, providing for the expedited processing of applications for setting up underground and above-ground infrastructure (Government of India, 2016). The rules provide for mechanisms such as a single window clearance system (option of a single electronic application process for facilitating approvals by all authorities within a state), fixed timelines for processing,

and a cap on the fees for administrative expenses. However, the implementation of the rules still remains under the domain of the state governments, and the practices that have evolved have not been consistent or sufficiently effective (Dua, 2020; Standing Committee on Communications and Information Technology, 2021). Through the Draft Indian Telecommunication Bill, 2022, the central government is now proposing legislative changes to streamline the right-of-way permissions involving both public and private properties.[4]

In another attempt to address the right-of-way backlogs, the government recently launched the GatiShakti Sanchar Portal, which is described as a "collaborative institutional mechanism between all stakeholders" to act as a single interface for right-of-way approvals (Department of Telecommunications, n.d.). If successful, this initiative can play an important role in the delivery of the BharatNet project and in the forthcoming roll-out of 5G services, which require fiberisation levels far beyond the country's current levels.

5. Conclusion

India is characterised by great diversity in the living conditions of its population—a diversity that is also reflected in variations in digital access, digital skills, and digital outcomes. The importance of addressing these differences risks being obscured in digitisation initiatives founded upon aggregate national statistics that paint a picture of robust mobile penetration and internet access. This article has sought to highlight some of the state-level differences present in respect of digital access, skills, and infrastructure across India—as a basis to dispel assumptions about the homogeneity and growing universality of the country's digital transformation. However, the article's objective is not to make a case against the pursuit of digital transformation on these grounds. Rather, the aim is to highlight the realities of India's federal structure and the varying levels of digital readiness that need to be accounted for in the design and implementation of the country's growing number of digital initiatives.

References

Aulakh, G. (2021, December 16). DoT Secy asks field units to expedite RoW permissions for fast broadband rollout. *The Live Mint*. https://www.livemint.com/industry/telecom/dot-secy-asks-field-units-to-expedite-row-permissions-for-fast-broadband-rollout-11639651989591.html

Barboni, G., Field, E., Pande, R., Rigol, N., Schaner S., & Troyer Moore, C. (2018). *A tough call: Understanding barriers to and impacts of women's mobile phone adoption in India*. Harvard Kennedy School. https://epod.cid.harvard.edu/sites/default/files/2018-10/A_Tough_Call.pdf

Barua, S. (2011, December 13). Why Northeast is lagging behind the rest of India? *The Shillong Times*. https://theshillongtimes.com/2011/12/13/why-northeast-is-lagging-behind-the-rest-of-india

4 See https://dot.gov.in/relatedlinks/indian-telecommunication-bill-2022

Bhardwaj, A. (2021). HDI: How states fare in human development. Center for Economic Data and Analysis (CEDA), Ashoka University. https://ceda.ashoka.edu.in/hdi-how-states-fare-in-human-development

Department of Telecommunications. (n.d.). About us. https://sugamsanchar.gov.in/about-us

Dua, T. R. (2020). Right of way (RoW) for digital communications infrastructure. Presentation on behalf of Tower and Infrastructure Providers Association. BRI Workshop with East and Northeast Region States. https://icrier.org/pdf/2-March-2020/TAIPA.pdf

Gokhale, N. A. (2022, February 3). The neglected states of the nation. *Outlook*. https://www.outlookindia.com/website/story/the-neglected-states-of-the-nation/220206

Government of Himachal Pradesh. (2018). District domestic product of Himachal Pradesh 2011-12 to 2015-16. https://himachalservices.nic.in/economics/pdf/distt_dp_2015-16.pdf

Government of India. (2016). The Indian Telegraph Right of Way Rules, 2016. https://dot.gov.in/sites/default/files/Consolidated%20Indian%20Telegraph%20Right%20of%20Wa-1_1.pdf?download=1

Government of India. (n.d.a). Digital India. https://www.digitalindia.gov.in

Government of India. (n.d.b). Direct Benefit Transfer (DBT): About us. https://dbtbharat.gov.in/page/frontcontentview/?id=MTc

GSMA. (2020). Rights of way (RoW) in India: A submission from the GSMA relating to India mobile sector. https://dipa.co.in/reports/GSMA%20Paper%20on%20RoW%20Sep%272020.pdf

GSMA. (2022). *The mobile gender gap report 2022*. https://www.gsma.com/r/wp-content/uploads/2022/06/The-Mobile-Gender-Gap-Report-2022.pdf

Kaka, N., Madgavkar, A., Kshirsagar, A., Gupta R., Manyika J., Bahl K., & Gupta, S. (2019). *Digital India: Technology to transform a connected nation*. McKinsey Global Institute. https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/digital-india-technology-to-transform-a-connected-nation

Khatri, H. (2022, July 21). Analyzing the mobile network experience in India's telecom circles. *Opensignal.* https://www.opensignal.com/2022/07/21/analyzing-the-mobile-network-experience-in-indias-telecom-circles

Marwah, A. M. (2019, July). RoW Roadblocks: Exploring solutions to expedite fibre roll-out. *tele.net*. https://tele.net.in/row-roadblocks

Ministry of Electronics and Information Technology (MeitY). (n.d.). UMANG: About this app. https://play.google.com/store/apps/details?id=in.gov.umang.negd.g2c&hl=en_IN&gl=US

Ministry of Finance. (2021, August 6). Know all about e-RUPI, the new digital payment instrument. Press Information Bureau Mumbai. https://pib.gov.in/PressReleasePage.aspx?PRID=1743056

Mothkoor, V., & Mumtaz, F. (2021, March 23). The digital dream: Upskilling India for the future. *Ideas for India.* https://www.ideasforindia.in/topics/governance/the-digital-dream-upskilling-india-for-the-future.html

Pandey, K. (2020, July 30). COVID-19 lockdown highlights India's great digital divide. *Down to Earth*. https://www.downtoearth.org.in/news/governance/covid-19-lockdown-highlights-india-s-great-digital-divide-72514

Parsheera, S. (2019a, July 5). Go to Lahaul-Spiti for "internet ka vanvaas". Digital India & telecom companies skipped it. *The Print*. https://theprint.in/opinion/go-to-lahaul-spiti-for-internet-ka-vanvaas-digital-india-telecom-companies-skipped-it/258707

Parsheera, S. (2019b, October 17). India's on a digital sprint that is leaving millions behind. *BBC World*. https://www.bbc.com/news/world-asia-india-49085846

PM India. (n.d.). Leveraging the power of JAM: Jan Dhan, Aadhar and Mobile. https://www.pmindia.gov.in/en/government_tr_rec/leveraging-the-power-of-jam-jan-dhan-aadhar-and-mobile

Prasad, R. S. (2022, June 2). The Digital India transformation. *The Indian Express.* https://indianexpress.com/article/opinion/columns/digital-india-eight-years-on-narendra-modi-5g-aadhaar-7948202

Press Trust of India. (2022, February 22). India to have 1 billion smartphone users by 2026: Deloitte. *The Hindu*. https://www.thehindu.com/business/india-to-have-1-billion-smartphone-users-by-2026-deloitte/article65075151.ece

Republic of India. (1950). Constitution of India. https://legislative.gov.in/sites/default/files/COI.pdf

Santuka, Y. (2021, June 26). The Story of CoWIN: How a vaccination strategy for Digital India failed to bridge the digital divide. *The Bastion*. https://thebastion.co.in/covid-19/the-story-of-cowin-how-a-vaccination-strategy-for-digital-india-failed-to-bridge-the-digital-divide

Singh, N., Zhou, Y., Williams. K., Kendall, J., & Kaushik, P. D. (2013). Bridging the digital divide in rural India: Lessons from a survey in four states. *Review of Market Integration*, *5*(1), 1–42. https://doi.org/10.1177/0974929213496499

Standing Committee on Communications and Information Technology. (2021). *India's preparedness for 5G*. Twenty-first Report. Seventeenth Lok Sabha. https://eparlib.nic.in/bitstream/123456789/799780/1/17_Information_Technology_21.pdf

Srivastava, R. (2021, March 18). Over 25000 villages in India still lack internet connectivity, Lok Sabha told. *India Today*. https://www.indiatoday.in/technology/news/story/over-25000-villages-in-india-still-lack-internet-connectivity-lok-sabha-told-1780758-2021-03-18

Surya, S. (2021, March 1). Standing Committee report: India's preparedness for 5G summary. *PRS Legislative Service*. https://prsindia.org/policy/report-summaries/india-s-preparedness-for-5g

Telecom Regulatory Authority of India (TRAI). (2016, February). *Recommendations on implementation strategy for BharatNet*. https://www.trai.gov.in/sites/default/files/Recommendations%20on%20BharatNet%2001.02.2016%20FINAL.pdf

TRAI. (2022, May 4). *The Indian telecom services performance indicators: October–December, 2021*. https://www.trai.gov.in/sites/default/files/QPIR_05052022.pdf

The Economic Times. (2022, April 29). India's low fiberisation may lead to wastage of spectrum resource in 5G rollout. ET Government. https://government.economictimes.indiatimes.com/news/technology/indias-low-fiberisation-may-lead-to-wastage-of-spectrum-resource-in-5g-rollout-coai/91170918

Unique Identification Authority of India (UIDAI). (2022). A unique identity for the people. https://uidai.gov.in/images/Aadhaar_Brochure_July_22.pdf

World Bank. (n.d.). Rural population (% of total population) – India. https://data.worldbank.org/indicator/SP.RUR.TOTL.ZS

Brazil's over-centralised governance of digital transformation

Larissa Galdino de Magalhães Santos
Associate Researcher, CyberBRICS project, Fundação Getulio Vargas (FGV) Law School, Rio de Janeiro; and Associate Researcher, Operating Unit on Policy-Driven Electronic Governance, United Nations University (UNU-EGOV), Guimarães, Portugal
https://orcid.org/0000-0002-9116-892X

Abstract
This article analyses governance of Brazil's Digital Transformation Strategy (E-Digital) through a new public governance (NPG) lens. Based on the analysis, the author finds that governance of E-Digital is too centralised, with too much decision-making power resting with state actors at the federal level, led by the Presidency. This analysis of the Brazilian experience aims to contribute towards understanding the modalities necessary for democratic, sustainable governance of digital transformation in Global Southern contexts.

Keywords
digital transformation, Brazil, E-Digital, governance, new public governance (NPG)

Acknowledgement
This article draws on elements of a longer report prepared for the CyberBRICS[1] project, Fundação Getulio Vargas (FGV) Law School, Rio de Janeiro.

DOI: https://doi.org/10.23962/ajic.i30.14843

Recommended citation
De Magalhães Santos, L. G. (2022). Brazil's over-centralised governance of digital transformation. *The African Journal of Information and Communication (AJIC), 30*, 1-10. https://doi.org/10.23962/ajic.i30.14843

1 See https://cyberbrics.info. The BRICS countries are Brazil, Russia, India, China and South Africa.

1. Introduction

While Brazil scores well in terms of digitisation of government services, occupying seventh position among the 198 countries of the World Bank's GovTech Maturity Index 2020 (Dener et al. 2021), its progress towards digital transformation across all sectors of the economy and society is uneven. Internet access in the country is still unequal, with about 90% of the low-income population connected exclusively by cell phone (CETIC, 2021).

The Brazilian Digital Transformation Strategy (E-Digital) was launched in 2018.[2] Coordinated by the Ministry of Science, Technology, and Innovation (MCTI), E-Digital has the ambitious purpose of coordinating public policies capable of guaranteeing the adoption of technology for an economically developed, fair, and solidarity-based society. In all, about 100 initiatives have been established by E-Digital. Because it is the central public policy strategy in the context of Brazilian digital transformation, it is important to understand and evaluate E-Digital's governance model.

There is insufficient knowledge regarding the governance modalities that best support digital transformation (Chantillon, 2021), and most research in this area covers countries with high technological performance, such as Australia, the Republic of Korea, or Denmark (see Nielsen & Jordanoski, 2020). This article aims, among other things, to encourage more examination of digital transformation policy and implementation in Global Southern contexts.

In this article, I analyse Brazil's governance of digital transformation through a new public governance (NPG) lens. In their recent article on NPG, Pereira and Ckagnazaroff (2021) characterise this mode of governance as:

> a type of governmental institutional arrangement which, in articulating economic-financial, institutional-administrative and socio-political dimensions, establishes partnerships with civil society and the market through deliberative processes, and seeks innovative solutions to social problems. (p. 112)

As Pereira and Ckagnazaroff (2021) point out, NPG is associated with democratic inclusion that allows civil society and private sector actors to contribute to public policies in a partnership process in the search for solutions to social problems. The

2 See Decree No. 9.319 of 21 March 2018, at http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/decreto/D9319.htm. E-Digital was updated on 17 November 2022 by the Ministry of Science, Technology, and Innovation (MCTI), after the finalisation of this article (see https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/transformacaodigital).

ideal NPG approach is one in which multiple actors contribute and influence the public policy decision-making system, and consequently contribute to the delivery of public services.

In this article, governance structure is understood to be the set of mechanisms used to direct, evaluate, and monitor the management and conduct of public policies and services, taking into account the interests of society and stakeholders. Governance structures are organised by rules, roles, and forms of articulation between decision instances, forming a multifaceted system of inter-organisational and intersectoral relationships (Jessop, 1993).

Section 2 of this article provides more detail on the NPG framework, section 3 describes the governance structure of Brazil's E-Digital strategy, section 4 evaluates the patterns of governance of E-Digital, and section 5 provides conclusions.

2. New public governance (NPG)

The transition from the hierarchical and bureaucratic approach of public management[3] to the systemic approach of the NPG demands a new form of organisation from the state: an institutional arrangement based on relationships within and outside the government that configure structures according to national contexts (Pereira & Ckagnazaroff, 2021). According to Pereira and Ckagnazaroff (2021), public governance can be analysed in terms of the following dimensions: (1) main values related to governance, such as integrity, transparency, and efficiency; (2) the structure of the institutional system that is responsible for promoting interaction to obtain results, through networks, power, and coordination; and (3) processes of evaluating and monitoring the results of public policies.

In terms of the NPG framework, a country's governance of digital transformation can be evaluated in terms of the degree to which the institutional arrangement allows: the power of the state to shift to other actors; networks of actors to form and articulate towards a goal; and knowledge and information to be coordinated in the search for solutions (Pereira & Ckagnazaroff, 2021).

According to Nielsen and Jordanoski (2020, p. 288), governance models of cooperation and intergovernmental coordination tend to be neglected as a success factor for national digital transformation strategies. Strong governance models, with clear roles and responsibilities of all institutions and with formal intersectoral decision-making bodies, are able to foster intergovernmental coordination and cooperation. Likewise,

3 See, for example, Osborne (2010), for a discussion of evolution from new public management (NPM) to NPG.

high levels of inclusion of all sectors of society at all levels of government, including civil society and private actors, can increase the chances for success in implementation of digital transformation strategies.

In recent years, Brazil has been promoting notions of public governance in order to establish mechanisms for evaluating the degree to which government policy implementation and service delivery are generating public value. Article 2 of Brazil's Decree No. 9.203 of 2017 on Public Governance specifies that public value constitutes the products and results generated by public-interest state responses to societal demands.

However, the updating of government ecosystem strategies for digital transformation signals the unbalanced appropriation of governance models, in most cases disconnected from the guidelines of the Public Governance Policy of the Brazilian government, and from solid principles of solid public governance, such as openness and transparency, inclusion, participation, gender equality, and diversity (OECD, 2020b).

It can be argued that, in addition to government efforts to promote digital transformation strategies, a governance structure capable of contributing to the good performance of public policies must have the following attributes: (1) inclusion of multiple actors in positions with access to power; (2) forums and collegial settings that guarantee the discussion of the problems and challenges of under-represented groups; (3) networks between the local and global, private and public actors, under formal coordination; and (4) training of transformational leaders to coordinate networks for a solid digital transformation strategy.

3. Governance structure of Brazil's E-Digital

In terms of the March 2018 Decree that established Brazil's E-Digital framework, E-Digital is the core component of the country's National System for Digital Transformation (SinDigital), with SinDigital composed of E-Digital and "its thematic axes and its governance structure".[4] The focus of this article is on the "governance structure" set out in the Decree.

The governance of, and decision-making structure for, E-Digital are coordinated by the Civil House of the Presidency of the Republic. The Civil House acts as the coordinator of E-Digital, and holds the presidency of the Interministerial Committee for Digital Transformation (CITDigital), the powerful committee that monitors,

4 See Article 1, Decree No. 9.319 of 21 March 2018 at http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/decreto/D9319.htm

evaluates, and directs E-Digital. The Civil House of the Presidency also has the power to choose many of Interministerial Committee's members, who come from the following bodies:

- Civil House of the Presidency
- Ministry of Foreign Affairs
- Ministry of Economy
- Ministry of Education
- Ministry of Communications
- Ministry of Science, Technology, and Innovation (MCTI)
- General Secretariat of the Presidency
- Office of Institutional Security of the Presidency

The Executive Secretariat of the Interministerial Committee is led by E-Digital's core ministry, the MCTI, which means that the MCTI has a leading role in the implementation, monitoring, and updating of the strategy.[5]

The Interministerial Committee is responsible for establishing the priority actions of E-Digital; maintaining the coherence of the initiatives of different bodies; sharing information on the impact of sectoral initiatives; monitoring and evaluating the results of E-Digital; articulating and monitoring government, state, and municipal programmes; and issuing recommendations, updates, reviews, and deliberations on E-Digital.

Arrangements have been modified since 2018 by several amendment decrees.[6] Civil society, the scientific community, and the private sector are now represented by the Advisory Council for Digital Transformation, which has only an advisory role in relation to the Interministerial Committee. The overall impact of the amendment decrees has been to further centralise control of E-Digital in the Interministerial Committee.

Thus, E-Digital's governance at the strategic level is centralised in the Civil House, linked to the Presidency of the Republic. Despite the existence of decision-making bodies to facilitate cooperation and coordination, all decisions and documents must be considered by the Presidency. The head of the Interministerial Committee, the main body of power and information, is appointed by the Civil House.

5 See https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/transformacaodigital for the MCTI's E-Digital update on 17 November 2022, which occurred after the drafting of this article.

6 Decree No. 10.332 of 2020 instituted the Digital Government Strategy for 2020 to 2022. Decrees No. 9.804 of 2019 and No. 10.782 of 2021 changed the discipline of the governance structure and the implementation definitions of E-Digital, altered the recreation of the Ministry of Communications, and altered the powers of the MCTI.

Given the cross-cutting nature of E-Digital's activity axes, the Interministerial Committee can institute technical support activities through specific committees or technical chambers that will act in the monitoring and follow-up of E-Digital axes. According to the Committee's Monitoring Reports,[7] there are, for example, committees to discuss the regulatory environment for start-ups; e-commerce and exports; and the internet of things (IoT).

The Interministerial Committee is mandated to ensure dialogue among its Executive Secretariat, its other members, and the other bodies involved in digital transformation initiatives. The Interministerial Committee determines the composition of the non-state advisory council (comprising civil society, the scientific community, and the private sector). Therefore, even though information circulates between the Interministerial Committee and the non-state representatives comprising the advisory council, the power of implementation lies clearly with the federal government and related state bodies.

Among the state actors, while the responsibilities are shared between different actors in the implementation of E-Digital, there is a clear imbalance of power between those actors positioned close to the apex of power in the Presidency, and those farther from the apex who are at the forefront of execution. Since 2018, decision-making power and relevant information have been concentrated in the Interministerial Committee.

Altogether, seven meetings[8] of the Interministerial Committee were held between May 2018 and March 2020 to discuss and deliberate on the creation of committees and subcommittees; to update or reassess the composition of the Advisory Council; to update E-Digital; and to establish the Digital Government Strategy 2022–2026. Work Plans were prepared for 2018–2019 and 2021–2022, and there were three reports, in 2018, 2019, and 2020, on the implementation of E-Digital.

The question then arises: How is the E-Digital governance structure impacting the performance of Brazil's ecosystem for digital transformation?

4. Over-centralisation of digital transformation governance

The governance structure of Brazil's E-Digital strategy is, as we have seen above, highly centralised in a few actors at the federal level. Multiple actors can provide technical support and monitor the strategy, but for the most part their roles are

7 See, for example, https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/transformacaodigital/arquivosestrategiadigital/citdigital_relatorio-de-acompanhamento-2020-rev-12-2020.pdf

8 The reports from all Interministerial Committee meetings between May 2018 and March 2022 can be found in the E-Digital repository: https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/transformacaodigital/estrategia-digital-repositorio. The repository also makes available all government degrees and resolutions by advisory structures linked to E-Digital.

only advisory. There are, for example, no formal incentives for cross-cutting or decentralised policy initiatives, such as for smart cities, technology in agriculture, or IoT, to be adopted at the non-federal level by municipalities or states.

Overall, E-Digital's governance can be said to be yielding only an adequate level of performance. In 2021, it was found that of the 100 actions stipulated by E-Digital, 23 actions had been completed, 60 actions were in progress, and 17 had not yet started (CGEE, 2021, pp. 97–110). Many profound digital inclusion challenges persist—challenges which, generally speaking, are tied to the country's socio-digital inequalities. On E-Digital's Infrastructure and Quality of Access Networks axis, there is still a pronounced imbalance between the regions of the country, especially between the North and Northeast, with regard to fibre optic coverage, access networks in schools, and extension of the mobile network. In E-Digital's Research, Development, and Innovation axis, there are persistent challenges in terms of internet access in public schools, the effectiveness of digital literacy projects, and technical training to enable transition to the digital economy (CGEE, 2021).

Some analyses have found that although E-Digital's structure has the necessary strategic, operational, and implementation scope, its decision-making power and circulation of information are concentrated too narrowly in the hands of government representatives. This diagnosis is reflected in recent assessments (see CGEE, 2021; OECD, 2018, 2020a) of the governance of Brazilian digital transformation, including E-Digital and the Digital Government Strategy.[9]

The governance framework should help the government to achieve its digital transformation goals. According to Chantillon (2021), public administration can create public value when carrying out its governmental activities, as long as governance of the activities includes all stakeholders. The public sector is one of many initiators of value creation processes, but it is not the only one. From this perspective, there is a relative flaw in the traditional view of creating public value in the Brazilian governance structure with regard to digital transformation.

Although governments are touted as the main entrepreneurs of innovation and explorers in the pursuit of public value for society, the public sector must take cognisance of citizens and customers' desire for organisations to adjust their business models. The reconfiguration of the public sector business model requires the incorporation of knowledge and experience of all potential partners—citizens, companies, and other non-state entities (Wirtz et al., 2021).

9 See Decree No. 10.332 of 2020, which instituted the Digital Government Strategy for 2020 to 2022.

In other words, true digital transformation is a cumulative process, using emerging technologies and dynamic stakeholder interactions, which results in the systematic reconfiguration of organisations and the most flexible and interactive business model, unlike the centralised model currently driving the Brazilian E-Digital strategy. The inevitable tensions generated by centralised governance and the inability to include stakeholders in decision-making have the potential to negatively influence the implementation of strategic actions. But radical decentralisation is also not the answer to the paradoxes of governance. Digital transformation efforts require both centralised management capabilities, grounded in a blend of hierarchy and unity, and decentralised capabilities grounded in diversity.

Although centralisation and decentralisation may seem to represent opposing approaches to the structure of public organisations, the two can (and must) exist simultaneously—in a necessary coexistence that will sometimes be comfortable and at other times be characterised by paradox and tension. Reform of a public institution may, depending the on stage the institution is at in its evolution, require increasing its centralising forces (via hierarchy, unification, and centralised governance), or it may require increased decentralising forces (via devolution, diversification, and sub-national governance) (Witesman, 2020). Centralised and decentralised structures can coexist based on the pursuit of public values.

5. Conclusions

To advance knowledge about the governance factors that can influence the performance of national digital transformation strategies, this article has analysed governance of the E-Digital framework in Brazil. Among other things, this article seeks to contribute to the literature on public management and governance in the Global South (see, for example, Hoque & Zakaria, 2014). Based on my analysis of E-Digital's governance modalities to date, I have made the argument above that E-Digital's governance is over-centralised in the hands of state actors, particularly federal state actors, and even more particularly in the Presidency.

What, then, is the model for a governance structure that supports sustainable digital transformation? The answer to this question depends on the status of the actors who engage in the strategising, operational organisation, and the implementation of public policies for digital transformation. What we do know is that a strong model includes the participation of multiple actors in decision-making, and the circulation of information and power to influence the ecosystem of actors involved with the strategies. Having a governance structure with well-defined rules and decision-making structures does not guarantee effective digital transformation. It is necessary to go further and consider the roles of non-state actors such as private sector representatives, citizens, scientists and academics, who will often have different perspectives and priorities from those of the government—and to allow these actors to serve as resources for optimal decision-making and implementation by public

managers. As Pereira and Ckagnazaroff (2021) point out in their discussion of the new public governance (NPG) model, the institutional arrangements of governance must foster networks of interdependencies between public and non-state actors grounded in the need to solve complex problems.

References

Centre for Management and Strategic Studies (CGEE). (2021). *Diagnóstico da E-Digital: Atualização da Estratégia Brasileira para a Transformação Digital* [*E-Digital diagnostic*s. *Update of the Brazilian Strategy for Digital Transformation*]. https://www.cgee.org.br/documents/10195/734063/Diagnostico+E-digital.pdf/8c028cc0-4dae-4ff2-abf7-220acf40779a?version=1.2

Centro Regional de Estudos para o Desenvolvimento da Sociedade da Informação (CETIC). (2021). *Executive summary – Survey on the use of information and communication technologies in Brazilian households –ICT households 2020*. https://www.cetic.br/pt/publicacao/executive-summary-survey-on-the-use-of-information-and-communication-technologies-in-brazilian-households-ict-households-2020/

Chantillon, M. (2021). *A governance framework facilitating a digital transformation of the public administration*. PhD thesis. Instituut voor de Overheid, KU Leuven. https://lirias.kuleuven.be/retrieve/637247

Dener, C., Nii-Amponsah, H., Ghunney, L. E., & Johns, K. D. (2021). *GovTech maturity index: The state of public sector digital transformation.* World Bank. https://doi.org/10.1596/978-1-4648-1765-6

Hoque, M. J., & Zakaria, A. F. M. (2014). Challenges of public management in the Global South. *Southern Conjuncture*, *5*(21–22), 23–38. https://doi.org/10.22456/2178-8839.45448

Jessop, B. (1993). Towards a Schumpeterian workfare state? Preliminary remarks on post-Fordist political economy. *Studies in Political Economy*, *40*, 7–40. https://doi.org/10.1080/19187033.1993.11675409

Klievink, B., & Janssen, M. (2009). Realizing joined-up government — Dynamic capabilities and stage models for transformation. *Government Information Quarterly*, *26*(2), 275–284. https://doi.org/10.1016/j.giq.2008.12.007

Liva, G., Codagnone, C., Misuraca, G., Gineikyte, V., & Barcevičius, E. (2020). Exploring digital government transformation: A literature review. In *Proceedings of the 13th International Conference on Theory and Practice of Electronic Governance (ICEGOV 2020)* (pp. 502–509). Association for Computing Machinery. https://doi.org/10.1145/3428502.3428578

Mergel, I., Edelmann, N., & Haug, N. (2019). Defining digital transformation: Results from expert interviews. *Government Information Quarterly*, *36*(4), 101385. https://doi.org/10.1016/j.giq.2019.06.002

Nielsen, M., & Jordanoski, Z. (2020). Digital transformation, governance and coordination models: A comparative study of Australia, Denmark and the Republic of Korea. In *Proceedings of the 21st Annual International Conference on Digital Government Research (dg.o '20)* (pp. 285–293). Association for Computing Machinery. https://doi.org/10.1145/3396956.3396987

Organisation for Economic Co-operation and Development (OECD). (2018). *Digital government review of Brazil: Towards the digital transformation of the public sector*. https://doi.org/10.1787/9789264307636-en

OECD. (2020a). *Going digital in Brazil*. https://doi.org/10.1787/e9bf7f8a-en

OECD. (2020b). *Policy framework on sound public governance: Baseline features of governments that work well.* https://doi.org/10.1787/c03e01b3-en

Osborne, S. P. (2010). Introduction: The (new) public governance: A suitable case for treatment? In S. P. Osborne (Ed.), *The new public governance? Emerging perspectives on the theory and practice of public governance*. Routledge. https://doi.org/10.4324/9780203861684

Pereira, B. A. D., & Ckagnazaroff, I. B. (2021). Contributions to consolidate the new public governance. *Cadernos EBAPE.BR*, 19(1). https://doi.org/10.1590/1679-395120200104x

Tangi, L., Janssen, M., Benedetti, M., & Noci, G. (2021). Digital government transformation: A structural equation modeling analysis of driving and impeding factors. *International Journal of Information Management*, *60*, 102356. https://doi.org/10.1016/j.ijinfomgt.2021.102356

Wirtz, B., Langer, P., & Schmidt, F. (2021). Digital government: Business model development for public value creation – A dynamic capabilities based framework. *Public Administration Quarterly*, *45*(3), 232–255. https://doi.org/10.37808/paq.45.3.2

Witesman, E. M. (2020). Centralization and decentralization: Compatible governance concepts and practices. In *Oxford research encyclopedia of politics*. https://doi.org/10.1093/acrefore/9780190228637.013.1390

The African Journal of Information and Communication (AJIC)

Published by the LINK Centre
University of the Witwatersrand (Wits)
Johannesburg, South Africa
https://www.wits.ac.za/linkcentre

ISSN 2077-7213 (online version)
ISSN 2077-7205 (print version)