

Gregory Gondwe*

CHATGPT and the Global South: how are journalists in sub-Saharan Africa engaging with generative AI?

<https://doi.org/10.1515/omgc-2023-0023>

Received April 8, 2023; accepted May 18, 2023

Abstract

Study purpose: This study explores the usage of generative AI tools by journalists in sub-Saharan Africa, with a focus on issues of misinformation, plagiarism, stereotypes, and the unrepresentative nature of online databases. The research places this inquiry within broader debates of whether the Global South can effectively and fairly use AI tools.

Design/methodology/approach: This study involved conducting interviews with journalists from five sub-Saharan African countries, namely Congo, DRC, Kenya, Tanzania, Uganda, and Zambia. The objective of the study was to ascertain how journalists in sub-Saharan Africa are utilizing ChatGPT. It is worth noting that this study is a component of an ongoing project on AI that commenced on September 19, 2022, shortly after receiving IRB approval. The ChatGPT project was initiated in January 2023 after discovering that our participants were already employing the Chatbot.

Findings: The study highlights that generative AI like ChatGPT operates on a limited and non-representative African corpus, making it selective on what is considered civil and uncivil language, thus limiting its effectiveness in the region. However, the study also suggests that in the absence of representative corpora, generative AI tools like ChatGPT present an opportunity for effective journalism practice in that journalists cannot completely rely on the tools.

Practical implications: The study emphasizes the need for human agencies to provide relevant information to the tool, thus contributing to a global database, and to consider diverse data sources when designing AI tools to minimize biases and stereotypes.

Article note: This article underwent double-blind peer review.

***Corresponding author: Gregory Gondwe**, Department of Communication, California State University San Bernardino, 5500 University Pkwy, San Bernardino, CA, 92407, USA, E-mail: gregory.gondwe@csusb.edu

Social implications: The social implications of the study suggest that AI tools have both positive and negative effects on journalism in developing countries, and there is a need to promote the responsible and ethical use of AI tools in journalism and beyond.

Originality/value: The original value of the study lies in shedding light on the challenges and opportunities associated with AI in journalism, promoting post-colonial thinking, and emphasizing the importance of diverse data sources and human agency in the development and use of AI tools.

Keywords: ChatGPT; generative AI; in-depth interviews; journalism; Nvivo; sub-Saharan Africa

1 Introduction

The increasing dependence on generative AI tools and the drive towards a connected and unified world has sparked debates on the ability of the Global South to effectively engage with new media technologies. Scholars such as Kothari and Cruikshank (2022) and Munoriyarwa et al. (2021) identify the lack of resources and technological skills among many users in the Global South as a significant barrier to the optimal utilization of AI tools. Nonetheless, the fact that the Global South, particularly Sub-Saharan Africa, can actively contribute to these ongoing debates underscores the relevance of the “Global Village” concept initially introduced by McLuhan in 1964. The heuristic framework envisions a global landscape in which technology facilitates the creation of an interconnected electronic network that transcends cultural, geographical, and potentially economic barriers. In this epoch, the notion of “Us” versus “Them” as a divisive marketing strategy is gradually becoming obsolete. Mbembe (2021) asserts that this interdependence is crucial for survival, emphasizing that technology should be centered on human agency to augment our capabilities.

Skeptics challenge McLuhan’s (1964) and Mbembe (2021) vision of the possibility of an interconnected world, highlighting the persistent inequalities perpetuated by technological advancements. Couldry and Mejias (2020) perceive the networked world as a manifestation of capitalism, where the Global South merely provides raw data that is exploited by the West. Munoriyarwa et al. (2021) contend that this approach is financially burdensome, limiting the participation of individuals from the Global South due to a lack of requisite technological skills (Kothari and Cruikshank 2022; Munoriyarwa et al. 2021). On a different note, Mhlambi (2020) argues that concerns surrounding technology are temporary and stem from five primary issues: exclusion of marginalized communities from decision-making processes in systems design, biases in data selection, failure to recognize societal interconnections, the commodification of digital identities, and data centralization. The implementation of

generative AI tools in Africa raises concerns about the creation of representative online databases, as previous studies have shown that global software categorizes certain information as either civil or uncivil, thereby limiting the inclusion of African content (Calabrese 2015; Gondwe 2021; Vargo and Hopp 2023).

The study set out to investigate the integration of generative AI tools, specifically ChatGPT, into the daily practices of journalists operating in five African countries, namely the Democratic Republic of Congo, Kenya, Tanzania, Uganda, and Zambia. By conducting interviews with journalists based in sub-Saharan Africa, the study seeks to explore both the challenges and potential benefits associated with the utilization of ChatGPT within contexts where databases are underrepresented. In this study, the term “database” refers to the collection of texts or materials from which the Chatbots extract data to generate information. The research will examine the extent of the Global South’s involvement in the use of generative AI, the representation of Global South corpus within these tools, and the potential concerns that journalists in the Global South may have regarding the utilization of generative AI tools in their work.

2 Literature and theoretical review

2.1 Generative AI and the internet in sub-Saharan Africa

Scholars contend that technological advancements are often developed without considering the context and needs of the Global South (Coudry and Mejias 2020; Mhlambi and Tiribelli 2023; Munoriyarwa et al. 2021; Nobel 2018). Consequently, metrics used to assess internet penetration and social media participation still rely on Western standards. For instance, data from the UN population division, Nielsen online, and the International Telecommunications Union project that Africa, a continent primarily composed of young people with a population of approximately 1.5 billion, has a 43.2 % internet penetration rate, with a 13.23 % growth. This puts the entire continent at 11.2 % of the internet world. While this data may have some validity, it often disregards the innovative and creative ways in which the Global South participates in technology.

Nemer’s (2022) book documents the ways in which people from Favela slums in Brazil engage with Western-designed technologies using creative and unconventional methods that most people in the West might disapprove of yet serve a purpose within their context. Similarly, studies have shown that most people in Africa adopt mundane ways of engaging with new technologies. For example, unlike Venmo or CashApp, mobile banking in Africa operates independently of traditional banking systems (Asongu and Nwachukwu 2016). Arguably, local people devise ways that are only applicable to their context.

Free Facebook Basics and zero-ratings, in combination with affordable Chinese phones, have also facilitated the Global South's involvement in groundbreaking technological innovations. Believing that human rights permeate the notion of the haves versus the have-nots, Internet.org, a not-for-profit organization with Facebook/Meta as its main founder, has sought to embark on providing access to the global world (Luchs 2016). In other words, Internet.org holds a strong belief that access to the Internet will inevitably foster discrimination in the future. With the increasing number of individuals gaining internet connectivity, the division will not only exist between those who possess and lack access but also between those who are connected and disconnected. In recognition of this potential issue, the organization was established several years ago, primarily concentrating on the global south. Its objective was to address the imminent influx of internet users from these regions, referred to as the "next billion consumers."

To extend the company's debated philanthropic mission, Mark Zuckerberg, the founder, and CEO of Facebook, made his inaugural visits to the global south, outlining plans to offer internet access through the utilization of solar-powered drones, such as the Aquila drone. Additionally, Facebook's Free Basics app now grants mobile phone users in most countries of the Global South the ability to access a text-based version of the platform without incurring any costs, thus called "Free-Facebook Basics" (Willems 2016). By November 2016, Facebook reported that its initiative, Internet.org, had successfully connected 40 million individuals to the Internet. Subsequently, in 2018, this number increased to nearly 100 million people (Constine 2018). According to the Internet.org webpage (accessed in May 2023), Free Basics was operational in 65 countries, including 30 countries within Africa.

This significant expansion was largely facilitated through a partnership with Airtel Africa, a subsidiary of the Indian telecommunications operator Airtel. This partnership expanded the notion of zero-rating service – which refers to one's ability to browse the Internet/Facebook apps, through their local telecommunication companies free of charge. And Free Facebook Basics is one example of a zero-rating service (Nothias 2020). Despite the ongoing contentious debates within India regarding the violation of net neutrality principles by Free Facebook Basics, which prohibits mobile operators from implementing differential tariffs for data services, Nothias (2020) and Willems (2016) have observed that the widespread availability of zero-rating initiatives has facilitated the expansion of the online public sphere. These authors note that zero-rating practices have played a crucial role in increasing internet usage and fostering active participation in the public sphere, particularly in the Global South. A notable example is the acknowledgment by Zambian President Hakainde Hichilema of the significant contribution made by the youth in utilizing Facebook and other social media platforms to promote free and fair elections during the 2021 tripartite elections.

However, the exponential growth observed in most countries of the Global South, especially sub-Saharan Africa, evinces a fundamental paradox. First, researchers have noted a correlation between increased digital footprints and data ownership. Couldry and Mejias (2020) argue that the Global South's participation in creating raw data is another form of colonization. Accordingly, the authors assert contend that the Global South only participates in creating raw, which corporations that control and own the platforms, collect, and sell at a higher profit. For example, a person spending numerous hours on social media might feel like entertainment, yet they are laboring for corporations by leaving behind digital footprints that are later sold to advertising companies or research institutes. The Cambridge Analytica scandal serves as an example of this argument. Further, researchers argue that there is a lack of a representative database in the sense that the West has criteria for what makes it into an online dataset. Data Scientist, Fu (2022), for example, argues that despite that most of the global population are non-English speakers, tech companies have few datasets for non-English languages. Fu explains that,

There is still a huge gap in technologies that can process and analyze non-English text, an issue for platforms like Facebook where nearly two-thirds of users use a language other than English. Communities outside of the Western sphere speak hundreds of different languages and that's a massive barrier to using algorithms for those people.

The problem is prevalent in Africa where most languages are not internationally recognized and thus compelled to use the colonizer's language. Because of the language barriers, most conversations in Africa are flagged as 'uncivil' (Calabrese 2015; Gondwe 2021; Vargo and Hopp 2023). In their studies, the authors argue that to moderate hate speech and political violence, online technologies tend to regulate civil versus uncivil conversations. Common examples include the 'topless protests' in South Africa that were introduced to raise awareness of how social media technologies censored cultures (Gillespie 2018; Gumede 2016). Essentially, the 'topless protests' refer to recent forms of activism in which bare-chested women use iconography from the past to confront homophobia, misogyny, and racism. The two arguments mirror Higgins et al. (2016) investigation of whether the production of social media content itself can 'resemiotize' pidgin languages. The authors posit their argument on the premise that most African languages, specifically pidgin, remain marginalized such that additional effort is necessary for some conversations to meet the database requirements. This is demonstrated in how researchers typically eliminate online data containing pidgin and switch-coded language when cleaning their data.

The challenge is that most official organizations, including the governments in sub-Saharan Africa, are still operating offline. Despite the capacity of their financial

resources, such institutions choose not to update their databases, thus allowing individual social media participants to proliferate the online corpora. As a result, most databases in sub-Saharan Africa are informed not by official content, but by digital footprints accidentally or unknowingly left behind by individual social media participants.

2.2 ChatGPT and the Global South

Harvard scholars Debbie Ginsburg and Jack Cushman recently performed an experiment on ChatGPT3, in which they asked their law students to write a hypothetical legal report/opening statement in Case X with ChatGPT3. When asked about their level of satisfaction with their output, most students indicated it was below average, except in the event where the Chatbot was asked specific and deeper questions. In other words, the relevance of generative AI was based on the depth and specificity of the questions in that the shallower the prompt, the more likely one would receive inconclusive or unfounded results (Gondwe 2023). What if the same experiment were conducted among student lawyers in sub-Saharan Africa? In their recent study, Jiao et al. (2023) identified three major aspects that inform the operation of generative AI tools: translation prompt, translation robustness, and multilingual translation.

The translation prompt refers to the idea that generative AI can only when a prompt is triggered. As a result, the quality of one's output depends on the quality of the questions. Translation robustness refers to the database size and the ability of the Chatbot to use such data for robustness. Generative AI tools are developed and trained on a large-scale dataset covering various domains. However, it remains unclear how robust ChatGPT3 can perform on domain-specific or noisy sentences. The third is the multilingual translation which refers to how well generative AI would perform on different language pairs. To further explain this phenomenon, Jiao et al. (2023) performed an analysis in which they compared ChatGPT to commercial software such as Google, DeepL, and Tencent. Their findings show ChatGPT performing competitively well with commercial translation products in translation prompts but lagged in translation robustness and multilingual translations. Essentially, the tool seemed to privilege Western languages and environments as opposed to other languages worldwide. In an event where generative AI tools are increasingly becoming popular among journalists in Global South, it is crucial to interrogate how the above status quo affects journalism performance.

2.3 ChatGPT and journalism

Currently, research on the relationship between generative AI and journalism is still in its early stages, with little to no empirical studies conducted on the topic thus far. Pavlik's (2023) recent peer-reviewed article provides a starting point for reimagining the capabilities and limitations of Chatbots as they continue to develop into mature tools. However, various news articles and bloggers have touched on this topic, with many debates focusing on the negative aspects of Chatbots. Three key themes characterize these arguments.

2.3.1 The relationship between ChatGPT and misinformation

Most scholars perceive misinformation as a phenomenon that has always coexisted with media history and development (Tandoc et al. 2018; Wasserman 2020). However, the ease with which misinformation can spread through social media and other online platforms has increased (Helberger et al. 2022; Saldaña and Vu 2022). This has had a significant impact on public health, safety, and misinformation. Scholars have called for human intervention in both the design of technology and the creation of online content (Borchardt 2022; Lin and Lewis 2022; Tully et al. 2022).

In the context of generative AI, the level of human intervention continued to diminish. Because the model is designed with no inherent ability to distinguish accurate from inaccurate information, the probability that it could create false and biased content is high. Such repercussions could be detrimental, especially when generated texts are used in decision-making or to inform the public. This is because the chatbots rely on existing databases and language models, which scholars have criticized for their biases. Some scholars have argued that the chatbots mirror the emotions of its creators. For example, asking a chatbot to write a poem may reveal emotional balance or imbalances.

2.3.2 The relationship between ChatGPT and plagiarism in journalism

Most journalists in sub-Saharan Africa work under strenuous conditions that go beyond the normal journalistic routines (Mpagaze and White 2010). Apart from following a lead, digging up a story, interviewing sources, writing the piece, to editing and checking facts, most journalists are moonlighting (Mabweazara 2018). As a result, most journalists split their time between their practice and their quest for survival. This trend tends to compromise their dedication to the practice, thus leading to what some scholars have referred to as ineptness, lack of professionalism, and thus, corrupt practice (Mfumbusa 2008; Onyebadi and Alajmi 2023; Skjerdal

2018). Therefore, journalists perceive AI as a panacea to the problem since it lessens the amount of time they would spend on their stories. For example, with a single click, one would write a feature. Unfortunately, this euphoria is usually evanescent, and often accompanied by the fear of plagiarism in the field of journalism.

Studies have shown that AI language models can generate responses that mirror existing texts, thus raising questions about the originality of content. Others argue that AI language models help simplify the process of writing, making it easier for individuals to produce work that may not be their own. In journalism studies, this schism is demarcated between those who are for it and those who are against it. The gap widened with the introduction of new chatbots like ChatGPT3 and now 4, which can mimic a journalist in writing and intonation. Among several other factors that have made the Chatbot popular among journalists, is its ability to generate articles quickly and efficiently. As averred by one journalist working for a community media, generative AI has made it possible for them to generate articles and publish them quickly within a time frame of the electricity load shedding. However, there are concerns that its ability to generate human-like text may also facilitate plagiarism. Although there is a dearth of research on the relationship between generative AI and plagiarism among journalists, a few existing studies suggest the possibility that journalists may use this tool to produce articles without proper attribution (Blach-Ørsten et al. 2018). Like any language model, ChatGPT3 facilitates plagiarism among journalists by compromising ethical values and those with limited knowledge of the source of information.

2.3.3 The relationship between ChatGPT and stereotypes

Generative AI tools in media have shown measurable consequences for both users and the public sphere. While users can benefit from more relevant news and new ways of researching and writing stories, there are also concerns about the potential for selective exposure and selective access to information. This could result in detrimental effects on the public sphere, with ill-designed recommendation algorithms and a concentration of attention on a few platforms leading to a narrower and less diverse media market. Furthermore, the development of AI-driven technologies carries the risk of exploiting user vulnerabilities to manipulate, erode privacy, and institutionalize intellectual surveillance. There is also the potential for the unintentional or intentional creation of new digital inequalities. These risks cannot be ignored and must be carefully considered in the development and implementation of AI-driven tools in the media.

The relationship between generative AI and stereotypes in the media is an important and complex issue with potential consequences for both the media landscape and society. Existing studies suggest that AI language models, including

ChatGPT3 and 4, have the potential to perpetuate and challenge stereotypes (Carlson 2019; Noble 2018). Essentially, language models are built on a large language corpus through which machine-learning algorithms skim to create meaning.

In her book, Nobel (2018) contends that algorithms, at their best, perpetuate racial and gender profiling, misrepresentation, and even economic redlining. While some attempts have been made by chatbots to address the problems identified by Nobel (2018) with Google, a cursory search for the words “black girls” yielded positive results. However, this does not necessarily indicate a complete resolution of the issue. For instance, searches for African leaders on the chatbot returned a favorable analysis but with a caveat of corrupt practices attached to their names. This narrative is a relatively unknown terrain when searching for Western leaders, where most results are accompanied by positive attributes. Such narratives pose a challenging environment for African journalists who rely on existing datasets to report international news within and outside the African context. Prior research suggests that the use of AI among journalists may increase the likelihood of promoting misinformation, plagiarism, and stereotyping (Ali and Hassoun 2019; Shin 2022). Consequently, this study seeks to explore whether African journalists using generative AI in their daily practice are also susceptible to these vices or whether they employ AI as a safeguard against them. Additionally, it examines whether the lack of representative databases in the Global South contributes to a shield against engaging with these vices. Against this backdrop, the study poses the following research questions.

RQ1a: *What are the perceptions of sub-Saharan African journalists on the availability and quality of online databases?*

RQ1b: *How do these perceptions influence their sourcing and use of data in their journalistic practice?*

RQ2a: *How do journalists in sub-Saharan Africa navigate the challenges of unreliable internet connections and slow download speeds while utilizing new, generative AI tools?*

RQ2b: *What is the level of awareness among journalists regarding the relationship between initiatives such as Facebook Basics and zero-ratings with mobile phone companies, and how do they utilize these free internet services in their daily practice?*

RQ3a: *How does the use of generative AI tools among journalists in sub-Saharan Africa (with diminishing human intervention in its design and the presence of inherent biases in existing databases and language models) result in issues of misinformation, plagiarism, and stereotyping?*

RQ3b: *How do the issues of misinformation, plagiarism, and stereotype affect journalism practices in sub-Saharan Africa?*

3 Methods

This study involved conducting interviews with journalists from five sub-Saharan African countries, namely Congo, DRC, Kenya, Tanzania, Uganda, and Zambia. The objective of the study was to ascertain how journalists in sub-Saharan Africa are utilizing ChatGPT. It is worth noting that this study is a component of an ongoing project on AI that commenced on September 19, 2022, shortly after receiving IRB approval. The project was initiated in January 2023 after discovering that our participants were already using generative AI like ChatGPT.

3.1 Sampling

Participants for this study were recruited through convenient sampling. According to Wimmer and Dominick (2000), convenience samples can be helpful in collecting exploratory information and may produce useful data. Since the current study was an exploratory one, a convenience sample was deemed appropriate. We contacted attendees of the 2020 and 2021 African Investigative Journalism Conferences with a request to participate in the study. The list was obtained from the Conference website containing names and contact information. In an event where the contact information was unclear, online information was sought. The focus was on individuals using generative AI, particularly in sub-Saharan Africa.

The initial project attracted a total of 43 journalists from five countries, but only 17 indicated that they had used the chatbots as of January 2023. As noted earlier, the project began as a general AI study but included questions on generative AI as unveiled to the world. Therefore, the analysis is based on 17 in-depth, semi-structured interviews with journalists working primarily for the local media. The journalists were either permanent or freelancers with relatively influential news outputs. All the interviewees said that they have worked for more than five years, with the majority being men (11 men, 6 women). The semi-structured interviews were conducted via zoom and automatically recorded and transcribed. The length of each interview was between 52 min and 103 min.

3.2 Data analysis

Data analysis was conducted using Nvivo software, which is a computer software package developed by QSR International. Nvivo is specifically designed for qualitative data analysis (QDA) and facilitates the organization, analysis, and extraction of insights from unstructured or qualitative data sources such as interviews, open-

ended survey responses, journal articles, social media content, and web content. It is particularly useful when conducting in-depth analysis on both small and large volumes of data. The utilization of Nvivo software enables researchers to effectively manage and analyze diverse forms of qualitative data, as outlined by McNiff (2016). Transcribed data was first organized in 17 PDF documents to match the 17 interviewees and imported into Nvivo software. Each transcript was given a pseudonym to conceal the identity of the respondents. Following our research questions and respondents' answers, we formulated five (5) themes, namely, Internet and Zero-rating services, Nature of online databases, ChatGPT and Misinformation, ChatGPT and Plagiarism, ChatGPT and Stereotypes. New sub-themes were developed as nodes supporting the main five themes.

4 Findings

The findings for this study are organized around our main research questions described above. We grouped our research questions, i.e., RQ1a and RQ2b into one, following the themes generated during our interview process and data analysis process. Nonetheless, each RQ was addressed within the framework. Topics, such as Internet and Zero-rating services, the Nature of online databases, ChatGPT and Misinformation, ChatGPT and Plagiarism, and ChatGPT and Stereotypes, that mirrored the participant's responses guided the presentation of our findings. Consequently, the results were structured to address the research questions and these identified themes. Overall, the findings indicate that the majority of journalists expressed enthusiasm regarding the potential capabilities of generative AI. However, they also expressed disappointment due to the limitations of generative AI tools within their specific local environment.

4.1 Journalists' perception of the nature of online databases

Research questions RQ2a and RQ2b aimed to comprehend the perceptions held by sub-Saharan journalists regarding the quality of online databases and their utilization within their journalistic practices. Responses on the nature of online databases were divided. While some respondents indicated their presence, others argued that they were almost non-existent. When asked for details, the group that argued for the availability of the database based their justification on the increased social media presence. "Where do you think we get data from when crowd coding or crowd-sourcing? Do you think all the digital footprints disappear? Forget about the 'right to be forgotten'. "You can never undelete what you have posted", they argued. On the

other hand, those who believe in the absence of online databases noted that what goes online is not what the majority would believe.

Accordingly, one of the respondents said, “I think we must move away from the belief that everything we post online makes it into a global database. We need to start thinking about useful information. For example, developed countries are very particular with what they call databases.” In other words, the second group believed that there should be a definition of what we call databases. One participant even indicated how it was hard to find news content from their government media but easy to find the information posted by a local individual who only posts without thinking of repercussions.

When asked about what kind of content they look for, the majority, including those from the first group indicated that they go for clear and relevant content. As one person had averred, “I am a professional journalist. I don’t go for petty gossip. When I am crowdsourcing, I look for reliable and credible people to give me information. I don’t waste time with illogical thinkers or attention seekers”. In general, there was widespread agreement among the respondents regarding the criteria for quality data/information in journalism. Content that utilized local languages or code-switching was typically disregarded and viewed as irrelevant and unprofessional. For instance, journalists from the Democratic Republic of Congo (DRC) and Zambia were particularly concerned about the quality of French and English comments. Comments written in local languages or a combination of local languages and French or English were seen as challenging to handle. A journalist from Zambia expressed difficulty in understanding individuals from the Copperbelt, a province in Zambia, due to their use of a language that resembled hooliganism. Consequently, only content written in proper colonial English or French grammar was deemed suitable for analysis and inclusion in a larger corpus.

It is worth noting that the Copperbelt region has developed a distinct language by blending various languages as a form of protest against the difficulties of life and occasionally to challenge oppressive regimes. Regrettably, there has been minimal recognition by scholars and the media of the significance of such linguistic efforts. This lack of recognition can be attributed to the preference of journalists and other scholars for what they perceive as civil or uncivil while prioritizing the colonial languages.

4.2 AI, internet connections, and zero-rating services in sub-Saharan Africa

RQ2a and RQ2b examined how journalists in sub-Saharan Africa navigate the challenges posed by unreliable internet connections and slow download speeds while

utilizing new generative AI tools. The purpose was to assess their level of awareness regarding the connections between initiatives such as Facebook Basics and zero-ratings with mobile phone companies. The questions also sought to explore how journalists incorporate and utilize these free Internet services in their daily journalistic practices. Regarding the intricate role of journalism in sub-Saharan Africa, our research findings reveal that journalists worldwide place great importance on the speed of changing technology and consistently strive to acquire new, emerging skills. Nearly all participants in our study emphasized the necessity for advanced skills to keep pace with rapidly evolving technologies, with a unanimous consensus on the challenges presented by unreliable internet connections. One journalist clarified that while access to the internet is generally available to media houses and freelancers, the issue lies with inconsistent connectivity and slow download speeds.

Another journalist shared their experience of difficulties uploading photos or videos but noted that alternatives such as live coverage through Facebook or the use of WhatsApp have proven to be beneficial. Overall, most participants expressed satisfaction with the level of internet penetration in the region, despite acknowledging its imperfect nature. When asked about Facebook Basics and zero-ratings, most journalists reported a lack of awareness regarding the relationship between these initiatives and mobile phone companies, although many had utilized free internet services. To clarify, while many journalists acknowledged the utilization of zero-rating services, they were unaware that these services were provided by Facebook with the intention of addressing disparities in internet access worldwide. In essence, the journalists lacked knowledge regarding the specific role played by Facebook in offering zero-rating services as a means to mitigate global internet connectivity inequalities.

4.3 ChatGPT and misinformation

RQ3a and RQ3b set out to investigate the use of generative AI tools among journalists in sub-Saharan Africa, considering the decreasing human intervention in their design and the presence of inherent biases in existing databases and language models. The objective was to examine the implications of such usage on the prevalence of misinformation, plagiarism, and stereotyping. Additionally, we sought to understand how these issues impact journalism practices in sub-Saharan Africa. The relationship between Misinformation and ChatGPT was assessed by asking participants about the degree of trust in the output and whether participants were aware of its inability to distinguish accurate from inaccurate information. The responses were all clustered on individual experience and the amount of time they have used the chatbot. Those who were interviewed in January showed a lack of awareness about

its inaccurate ability. However, many of them had only used the chatbot in a way they would use Google search engines or Wikipedia. Therefore, the fact that it could even organize thoughts was impressive to them.

However, those interviewed subsequently seemed to have gathered enough information to understand that they could not completely trust the chatbot. As one respondent had indicated, “I have learned that you cannot completely trust ChatGPT. I introduced it to one of my friends studying accounting and I was surprised when he told me that the chatbot gave him wrong information”. Accordingly, the friend asked ChatGPT the following question, “What is $12.4\% * \$142,800$?” ChatGPT3 responded with the following answer: 12.4% of $\$142,800$ is equal to $\$17,744.32$. To calculate this, you simply multiply the percentage by the number: $12.4\% * \$142,800 = 12.4/100 * \$142,800 = \$17,744.32$. A quick verification of the calculations indicated that ChatGPT3 was wrong as $12.4\% * \$142,800 = \$17,707.20$. “If ChatGPT3 can make these careless errors that google and even a simple calculator can do, what of things we do not know about?”, the respondent asked. In short, as more information came to light about ChatGPT3, respondents began to question its accuracy, and thus reliability.

4.4 ChatGPT and plagiarism

To understand the possibility of plagiarism using ChatGPT among journalists, we asked questions about the time they dedicate to the practice, whether they are moonlighting, and most of all, how they used the chatbot. Like questions on misinformation, we on issues reflecting the trust journalists place in ChatGPT output, and whether they find it necessary to add their own content to a story written by ChatGPT. Most respondents indicated that they were not satisfied with the output, suggesting that the chatbot could not write a complete and relevant story. One respondent commented, “I think ChatGPT is for big media like CNN, BBC, and Aljazeera. It is yet to recognize our environments. I have asked it to write a story about the events making news in my community and it knows nothing”.

Another respondent indicated that they had asked ChatGPT to write a feature story regarding the President of Zambia and his abolition of the death penalty. “I was so disappointed that the story generated by ChatGPT was shallow. It felt like reading something from Wikipedia – and as a journalist, we don’t publish anything from such sites”, said the respondent. Another respondent argued that ChatGPT was only useful for organizing your thoughts and not a tool for providing you with information. “Google is the best! If I want information, I go to Google and then ask ChatGPT to organize that information for me – period!”, they indicated. In short, most participants did not believe you could plagiarize with ChatGPT – at least not for journalism practices in Africa.

4.5 ChatGPT and stereotypes

To inquire about how generative AI, particularly, ChatGPT perpetuated stereotypes, we examined the specific biases within the African context. Among them were poverty, corruption, gender, family values, and journalists' perception of the West. For example, we asked journalists about how such issues tend to emerge in their work. One respondent opened the conversation by pointing to how Africa was presented. "Every time I ask ChatGPT to help write a feature story about my country, Congo, it ends with the mention of war, corruption, and poverty. I wonder why this is necessary", said the respondent. Another respondent shared a similar experience suggesting that although ChatGPT would say positive things about Africa, it always ended with the "however" which pointed to the levels of poverty. "It is worse when you ask the chatbot about an African leader – except Mandela, everyone has corruption attached to them. This is not the case for Western leaders." Imagine, even King Leopold II is fairly described than some existing African leaders – I think there are a lot of biases.

On the other hand, most journalists supported the fight for gender rights and family values – indicating that they were extremely cautious when dealing with such topics. But one journalist indicated that ChatGPT would never help write a feature story that is against Western narratives. "You know, despite the protests, Ugandans are still very conservative about LGBTQ issues. We have heard our president speak against it, and I asked ChatGPT to write a feature supporting the President's statements. I was surprised when ChatGPT began to lecture me on the rights of LGBTQ. I thought I just need to command, and ChatGPT needed to listen. But I was wrong. ChatGPT tells you what ChatGPT thinks is right". In short, the respondents acknowledged the perpetual stereotypes embedded within AI.

5 Discussion

In this study, we aimed to explore and examine the ways in which journalists in sub-Saharan Africa are engaging with generative AI in mundane ways. Specifically, we investigated the relationship between journalists' use of ChatGPT and the prevalence of misinformation, plagiarism, and stereotypes. Our inquiry was situated within the broader context of debates questioning whether the Global South can effectively and fairly engage with AI tools. We aimed to interrogate how journalists in contexts with limited online databases informed by selective algorithm biases interact with the chatbot. To establish our case, we conducted two quasi-experiments with students and community members from the US and compared the findings with in-depth

interviews with journalists from five sub-Saharan countries, including Congo DRC, Kenya, Tanzania, Uganda, and Zambia.

While examining the challenges associated with artificial intelligence (AI) in sub-Saharan Africa, this study also emphasizes the importance of effectively harnessing emerging media technology. In contrast to perspectives that perceive AI as purely automated, our research demonstrates that AI output relies on human agency. The findings are consistent with Jiao et al. (2023) findings about the ineffectiveness of chatbots in providing reliable translations. Against the backdrop, we conclude that since AI is developed by humans and subsequently fed with human input, its effectiveness will also rely on human agency. The implication is that the quality of output will also be contingent on the quality of the data. But as observed in the findings, it could be argued that the quality of data in sub-Saharan Africa is still questionable, therefore, questioning the quality, validity, and reliability of the perceived output. As most scholars have argued, Africa is only a reach source for raw data (both digital and physical as witnessed through the mushrooming of extractive industries, Couldry and Mejias 2020). At most, the continent contributes its digital footprints which are in turn produced into finished goods and resold back to us and the rest of the world. Essentially, this has always been a hallmark for understanding how Africa economically engages with the West.

In contrast, this study reveals opportunities for human agency to play a pivotal role, aligning with Mbembe's (2021) concept of the "epoch of indivisibility." Unlike in colonial times, there appears to be a growing recognition of the interconnectedness of our planet. As Mbembe argues, the dichotomy of "Us" versus "Them" no longer holds the same persuasive power, given that oppressors are now inevitably affected by the suffering of the oppressed. The key lies in creating an enabling environment where even the oppressed can survive. The COVID-19 pandemic has taught us valuable lessons about the nature of coexistence, emphasizing that developing a vaccine for a specific group is crucial to preventing further mutations. Similarly, technology, including AI, must also consider diverse environments beyond the West. Consequently, the effectiveness of AI tools, such as ChatGPT, will depend not only on the quality of data provided by the West but also on contributions from the rest of the world. Therefore, to employ AI effectively in journalism, it is imperative for humans to assume control and supply the tool with pertinent and beneficial information, thus contributing to a comprehensive global database.

Further, the findings also suggest the presence of confirmation biases within the AI tools utilized by journalists. Although not explicitly stated, participants implied that the chatbot consistently produced expected outputs. This issue is not unprecedented, as many AI tools have been accused of recording and analyzing conversations and online searches to tailor suggestions according to individual preferences (Vermeer and Trilling 2020; Weeks et al. 2022). The overt confirmation of biases

becomes apparent when the chatbot provides different responses to similar questions based on one's affiliations. Notably, ChatGPT tends to portray African countries and their leaders in a negative spotlight, with questions about them often concluding with anecdotes that connote poverty, disease, or corruption. These findings corroborate Nobel's (2018) critique of AI technologies.

While these observations may prove valuable for researchers exploring the interplay of AI stereotypes in the Global South, the study also raises a concern. The depiction of corruption, particularly among African leaders, relies on anecdotes that could perpetuate stereotypes. However, this observation presents an opportunity for postcolonial thinking, which considers the Global South as "a site of epistemic disobedience to the established hegemonic journalistic norms and mores in theory, pedagogy, and practice that are in service of global coloniality and the modern empire" (Moyo 2022, p. 1568). The critical insights derived from this study enable journalists to recognize the importance of decolonizing their content, even in the presence of biased AI.

The study's primary implications underscore the necessity of human agency at every stage of chatbot utilization. This implies that users must provide clear and pertinent information to the chatbot to enhance its outputs. The findings can also assist developers and marketers of generative AI tools in designing effective strategies to enhance customer satisfaction. Identifying stereotypical outputs is a step towards combating them, as misinformation and plagiarism are often intertwined with stereotypes. Therefore, technological designers and journalists should collaborate to mitigate these issues (Carlson 2019; Noble 2018). This study contributes to a nascent body of literature (Kothari and Cruikshank 2022; Munoriyarwa et al. 2021; Nothias 2020; Willems 2016) on how the Global South is and should engage with AI in a distinct manner.

6 Conclusions

The utilization of generative AI in the Global South presents a combination of opportunities and challenges, but its current status remains uncertain. The pursuit of responsible AI is not exclusive to the Global South, as scholars worldwide grapple with the establishment of binding values, irrevocable standards, and the symmetrical and diametrical of relational personhood (Mhlambi 2020, p. 20) in the development and application of AI products like ChatGPT. The absence of a fundamental consensus regarding the nature of the data to be incorporated into the global database poses a significant concern. It is important to acknowledge the technological contributions of China to the Global South, as their affordable mobile phones have facilitated internet access and reshaped the digital landscape (Gagliardone 2019). Additionally, Facebook Meta has offered free and zero-rated internet services,

despite facing a ban in India due to its failure to deliver on philanthropic promises (Mukherjee 2016; Nothias 2020; Willems 2016).

The significant presence of online data originating from the Global South warrants the active engagement of scholars investigating the region in ongoing discussions concerning generative AI. Essential considerations should encompass various issues, including the political economies associated with new technologies, ethical dimensions in AI design, and the extent of meaningful participation from the Global South in the design process. It is worth noting that although many software designers from the Global South are operating in the West, questions arise as to why they fail to incorporate their cultural values and the advantages offered by their home regions. Exploring this matter could serve as a starting point for discussions on designing an authentic and representative dataset. Achieving this vision requires the development of datasets that portray the Global South in a positive light.

Given the internet's rapidity and efficacy, this endeavor becomes even more crucial. Governments and other organizations in the Global South need to prioritize this undertaking by ensuring the availability of accurate and accessible information. Moreover, these governments must invest in technological education and literacy with an emphasis on curricula that support coding, computer science, and other disciplines related to new media technologies. By establishing computer labs in every school, the advantages outweigh the costs. Equipping the younger generation with advanced technologies will aid in dismantling stereotypes and fostering representation, thereby ensuring that AI systems operate on a precise and inclusive database crafted by the Global South.

However, it is important to acknowledge the limitations of this study. Treating the five countries as homogeneous entity neglects the variations within the cross-cultural context. We have not accounted for the differences in journalistic cultures across diverse contexts, as highlighted by Hanitzsch et al. (2011). In a cross-cultural setting, distinctions prove to be less straightforward than conventional wisdom and prior evidence may suggest. Future research should delve deeper into other variables, such as the presence of a free press, surveillance practices, and privacy concerns, to provide a more comprehensive understanding.

Appendix

Interview Guide Questionnaire

Thank you for participating in this study on the use of generative AI tools like ChatGPT by journalists in the Global South. The purpose of this study is to understand how African journalists use ChatGPT and the challenges they face while using this

tool. We will ask you a series of questions related to your experience with ChatGPT, and your insights will be valuable in helping us understand the benefits and challenges of using generative AI tools.

1. Can you tell us about your experience using ChatGPT for journalistic purposes, particularly in contexts where databases are underrepresented?
2. Based on your experience, how can the integration of generative AI tools like ChatGPT benefit journalists in the Global South, particularly in contexts where databases are underrepresented?
3. In your opinion, what are the main challenges that journalists in sub-Saharan Africa face when using generative AI tools like ChatGPT, and how can these challenges be overcome?
4. How do you think ChatGPT and other generative AI tools can help in addressing the issue of underrepresentation of the Global South in the corpus used by these tools?
5. To what extent is the Global South represented in the corpus used by generative AI tools like ChatGPT, and what are the potential concerns surrounding this representation for journalists in the Global South?
6. Have you noticed any instances of misinformation, plagiarism, or stereotyping while using ChatGPT for journalistic purposes? If yes, could you please provide some examples?
7. In your experience, how do you ensure that the content generated by ChatGPT is accurate and unbiased?
8. What measures would you suggest to ensure that generative AI tools like ChatGPT are used responsibly in the field of journalism?
9. Can you describe your experience using the internet (and zero-rating services) in your work as a journalist?
10. In your opinion, how accessible are online databases that you use in your work as a journalist?
11. What challenges have you encountered when using generative AI tools like ChatGPT, particularly regarding misinformation?
12. Have you ever encountered instances of plagiarism when using ChatGPT or any other generative AI tool? If so, can you describe the experience?
13. What are your thoughts on the potential for generative AI tools like ChatGPT to perpetuate stereotypes, particularly in the context of African journalism?
14. In your opinion, how can the integration of generative AI tools like ChatGPT benefit journalists particularly in contexts where databases are underrepresented?
15. What challenges have you encountered when using generative AI tools like ChatGPT, and how can these challenges be overcome?

Conclusion: Thank you for your valuable insights. Your responses will be used to contribute to the ongoing research on the use of generative AI tools in journalism and to help develop strategies for the responsible use of AI tools in the field.

References

- Ali, Waleed & Mohamed Hassoun. 2019. Artificial intelligence and automated journalism: Contemporary challenges and new opportunities. *International Journal of Media, Journalism, and Mass Communications* 5(1). 40–49.
- Asongu, Simplice A. & Jacinta C. Nwachukwu. 2016. The mobile phone in the diffusion of knowledge for institutional quality in sub-Saharan Africa. *World Development* 86. 133–147.
- Blach-Ørsten, Mette, John Møller Hartley & Marcus B. Wittchen. 2018. A matter of trust: plagiarism, fake sources, and paradigm repair in the Danish News Media. *Journalism Studies* 19(13). 1889–1898.
- Borchardt, Alexandra. 2022. Go, Robots, Go! the value and challenges of artificial intelligence for local journalism. *Digital Journalism* 10(10). 1919–1924.
- Carlson, Matt. 2019. News algorithms, photojournalism, and the assumption of mechanical objectivity in journalism. *Digital Journalism* 7(8). 1117–1133.
- Calabrese, Andrew. 2015. Liberalism's disease: civility above justice. *European Journal of Communication* 30(5). 539–553.
- Constone, Josh. 2018. Facebook's Internet.org has connected almost 100M to the Internet, TechCrunch. <https://techcrunch.com/2018/04/25/internet-org-100-million/> (accessed 12 May 2023).
- Couldry, Nick & Ulises A. Mejias. 2020. *The costs of connection: How data is colonizing human life and appropriating it for capitalism*. Redwood City: Stanford University Press.
- Fu, Gratiana. 2022. Toward ethical artificial intelligence in international development. Center for digital acceleration (CDA) insight. <https://dai-global-digital.com/cda-insights-2022-toward-ethical-artificial-intelligence-in-international-development.html> (accessed 5 May 2023).
- Gagliardone, Iginio. 2019. *China, Africa, and the future of the internet*. London: Bloomsbury Publishing.
- Gillespie, Tarleton. 2018. *Custodians of the internet: Platforms, content moderation, and the hidden decisions that shape social media*. New Haven: Yale University Press.
- Gondwe, Gregory. 2021. Online incivility, hate speech and political violence in Zambia: Examining the role of online political campaign messages. *Journal of African Media Studies* 13(1). 35–51.
- Gondwe, Gregory. 2023. ChatGPT3 in sub-Saharan Africa: A threat or an opportunity? Available at: SSRN <https://doi.org/10.2139/ssrn.4392918>.
- Gumede, William. 2016. Rise in censorship of the internet and social media in Africa. *Journal of African Media Studies* 8(3). 413–421.
- Hanitzsch, Thomas, Folker Hanusch, Claudia Mellado, Maria Anikina, Rosa Berganza, Ipek Cangoz. 2011. Mapping journalism cultures across nations: A comparative study of 18 countries. *Journalism Studies* 12(3). 273–293.
- Helberger, Natali, Marcel Van Drunen, Judith Moeller, Sander Vrijenhoek & Steffen Eskens. 2022. Towards a normative perspective on journalistic AI: Embracing the messy reality of normative ideals. *Digital Journalism* 10(10). 1605–1626.
- Higgins, Christina, Goki Furukawa & Haejoo Lee. 2016. Resemiotizing the metapragmatics of Konglish and pidgin on YouTube. In Philip Seargeant, Caroline Tagg, & Amy Brown (eds.), *Social media discourse, (dis) identifications and diversities*, 320–344. Routledge.

- Jiao, Wentao, Wei Huang Wang, Xinyan Wang, and Zhuowei Tu. 2023. Is ChatGPT a Good Translator? A Preliminary Study. *arXiv preprint arXiv*. 2301.08745.
- Kothari, Ammina, Sally Ann Cruikshank. 2022. Artificial intelligence and journalism: An Agenda for journalism research in Africa. *African Journalism Studies* 43(1). 17–33.
- Lin, Bowen & Seth C. Lewis. 2022. The one thing journalistic AI just might do for democracy. *Digital Journalism* 10(10). 1–23.
- Luchs, Inga. 2016. Free basics by Facebook. An interview with Nishant shah. *Spheres: Journal for Digital Cultures* 3(2016). 1–8.
- Mabweazara, Hayes M. 2018. When your ‘take-home’ can hardly take you home: Moonlighting and the quest for economic survival in the Zimbabwean press. *News making cultures in Africa: Normative trends in the dynamics of socio-political & economic struggles*, 99–117. London: Palgrave Macmillan.
- Mbembe, Achille. 2021. Out of the dark. *Essays on decolonization*, 7–41. New York: Columbia University Press.
- McLuhan, Marshall. 1964. *Understanding media. The extensions of man*. Canada: McGraw-Hill.
- McNiff, Jean. 2016. *You and your action research project*, 4th edn. New York: Routledge.
- Mfumbusa, Bernardine. 2008. Newsroom ethics in Africa: Quest for a normative framework. *African Communication Research* 1(2). 139–158.
- Mhlambi, Sabelo & Simona Tiribelli. 2023. Decolonizing AI ethics: Relational autonomy as a means to counter AI Harms. *Topoi* 42(1). 1–14.
- Mhlambi, S.abelo. 2020. From Rationality to Relationality: Ubuntu as an Ethical and Human Rights Framework for Artificial Intelligence Governance. Carr Center for Human Rights Policy Discussion Paper Series. 9. <https://carrcenter.hks.harvard.edu/publications/rationality-relationality-ubuntu-ethical-and-human-rights-framework-artificial> (accessed 19 November 2022).
- Mpagaze, Denis & Robert White. 2010. Tanzanian journalists’ ambivalent perception of their ethics: A ‘Jekyll and Hyde’ occupation. *African Communication Research* 3(3). 543–574.
- Moyo, Last. 2022. Introduction: Journalism studies and the Global South-theory, practice and pedagogy. *Journalism Studies* 23(13). 1567–1577.
- Mukerjee, Subhayan. 2016. Net neutrality, Facebook, and India’s battle to# SaveTheInternet. *Communication and The Public* 1(3). 356–361.
- Munoriyarwa, Allen, Sarah Chiumbu & Gilbert Motsaathebe. 2021. Artificial intelligence practices in everyday news production: The case of South Africa’s mainstream newsrooms. *Journalism Practice* 1–19. <https://doi.org/10.1080/17512786.2021.1984976>.
- Nemer, David. 2022. *The technology of the oppressed: Inequity and the digital Mundane in favelas of Brazil*. Cambridge: MIT Press.
- Noble, Safiya Umoja. 2018. Algorithms of the oppressed. In *Algorithms of oppression*. New York: New York University Press.
- Nothias, Toussaint. 2020. Access granted: Facebook’s free basics in Africa. *Media, Culture & Society* 42(3). 329–348.
- Onyebadi, Uche & Fawaz Alajmi. 2023. Unethical but not illegal: Revisiting brown envelope journalism practice in Kuwait. *Journalism* 24(3). 616–632.
- Pavlik, John. 2023. Collaborating with ChatGPT: Considering the implications of generative artificial intelligence for journalism and media education. *Journalism and Mass Communication Educator* 78(1). 84–93.
- Saldaña, Magdalena & Tien Hong Vu. 2022. You are fake news! Factors impacting journalists’ debunking behaviors on social media. *Digital Journalism* 10(5). 823–842.
- Shin, Donghee. 2022. How do people judge the credibility of algorithmic sources? *AI & Society* 37(2). 1–16.

- Skjerdal, Terje. 2018. Brown envelope journalism: The contradiction between ethical mindset and unethical practice. In Hayes Mawindi Mabweazara (ed.), *Newsmaking cultures in Africa: Normative trends in the dynamics of socio-political & economic struggles*, 163–183. London: Palgrave Macmillan.
- Tandoc Jr, Edson C., Zheng Wei Lim & Richard Ling. 2018. Defining “fake news”: A typology of scholarly definitions. *Digital journalism* 6(2). 137–153.
- Tully, Melissa, Dani Madrid-Morales, Herman Wasserman, Gregory Gondwe & Kioko Ireri. 2022. Who is responsible for stopping the spread of misinformation? Examining audience perceptions of responsibilities and responses in six sub-Saharan African Countries. *Digital Journalism* 10(5). 679–697.
- Vargo, Chris & Tobias. Hopp. 2023. Incivility on popular Politics and news Subreddits: An Analysis of in-groups, community Guidelines and Relationships with social media engagement. <https://scholarspace.manoa.hawaii.edu/items/7b31b719-4137-4437-9309-7074462e4f82/full> (accessed 6 March 2023).
- Vermeer, Susan & Damian Trilling. 2020. Toward a better understanding of news user journeys: A Markov chain approach. *Journalism Studies* 21(7). 879–894.
- Wasserman, Herman. 2020. Fake news from Africa: Panics, politics, and paradigms. *Journalism* 21(1). 3–16.
- Weeks, Brian E., Daniel S. Lane & Lauren B. Hahn. 2022. Online incidental exposure to news can minimize interest-based political knowledge gaps: evidence from two us elections. *The International Journal of Press/Politics* 27(1). 243–262.
- Willems, Wendy. 2016. Beyond Free Basics: Facebook, data bundles, and Zambia’s social media internet. Africa at LSE. Available at: <http://blogs.lse.ac.uk/africaatlse/>.
- Wimmer, Roger & Joseph Dominick. 2000. *Mass media research: An introduction*, 6th edn. Belmont, CA: Wadsworth.

Bionote

Gregory Gondwe

Department of Communication, California State University San Bernardino, San Bernardino, CA, 92407, USA

gregory.gondwe@csusb.edu

Dr. Gregory Gondwe, Assistant Professor of Journalism at California State University - San Bernardino and a Visiting Scholar for the 2022-2023 academic year with the Institute for Rebooting Social Media (RSM) at Harvard, is a dedicated researcher of emerging media trends in Africa. He has conducted in-depth studies on the implications of China’s involvement with African newsrooms and the prevalence of misinformation and fake news in Africa. Additionally, he has explored the impact of new media technologies on African journalism. Dr. Gondwe’s research focuses on social media trends in sub-Saharan Africa and their relationship with journalism, digital governance, community orchestration, censorship, and surveillance.