

THE ROLE OF AI IN COMBATING POLITICAL DEEP FAKES IN AFRICAN DEMOCRACIES

(Group 1: Uzo Mkparu, Winnie Kungu, Chialuka Prisca-mary Onuoha & Kayode Adekoya)

A. Introduction

As Artificial intelligence (AI) becomes increasingly pervasive, its impact now permeates all facets of life and society including the tenets of democracy. In particular, the emergence of deepfakes has ushered in the capability to significantly impact the integrity of democratic processes, the way we engage with democratic institutions, and our political attitudes. Deepfakes are AI-powered tools that can mimic and manipulate video and audio to the degree that is indistinguishable from real content.¹ These are typically used to spread disinformation, misinformation, and propaganda which is used to manipulate public opinion, political discourse, and even create unrest. These distortions are further amplified when microtargeting techniques are used on citizens to mimic the susceptibilities of the groups being targeted.² Ultimately, misinformation through deepfakes could lead to the significant risk of undermining the legitimacy of democracies.³

Over the past few decades, some democratic institutions across Africa have come under threat due to a myriad of vulnerabilities such as being young and fragile democracies, the threat and occurrence of military *coup d'états*, and clear societal inequalities. In this environment, the emergence and proliferation of deepfakes adds another dimension of risk to these weakened democracies which when extrapolated presents itself as an AI-induced existential threat to African democracies. Recent examples of political deepfakes include a manipulated audio that surfaced during the Nigeria General Election in February 2023⁴ purported to be of a presidential candidate planning to rig the election and also

¹ Yisroel Mirsky and Wenke Lee. (2020) (The Creation and Detection of Deepfakes: A Survey” ACM Computer. Survey. < <https://arxiv.org/pdf/2004.11138>>

² Dobber, T., Metoui, N., Trilling, D., Helberger, N., & de Vreese, C. (2021). Do (Microtargeted) Deepfakes Have Real Effects on Political Attitudes? *The International Journal of Press/Politics*, 26(1), 69-91. <https://doi.org/10.1177/1940161220944364>

³ Bennett, W. L., & Livingston, S. (2018). The disinformation order: Disruptive communication and the decline of democratic institutions. *European Journal of Communication*, 33(2), 122-139. <https://doi.org/10.1177/0267323118760317>

⁴ Allie Fuck, Adrian Shahbaz and Kian Vesteinsson, The Repressive Power of Artificial Intelligence (2024) < [1](https://freedomhouse.org/report/freedom-net/2023/repressive-power-artificial-intelligence#:~:text=During%20the%20February%202023%20Nigerian,integrity%20of%20the%20electoral%20system.></p></div><div data-bbox=)

that of the widely reported pro-junta activists that made the rounds in Burkina Faso in January 2023.⁵

The relatively low adoption of AI in Africa compared to the Global North has not helped the threat level from political deepfakes as attacks on African democracies could potentially be orchestrated by foreign actors and entities outside of the region. This feeds into the imbalance of power between the Global South and Global North and further creates instability and catastrophic risk for the region. This low adoption also means that there is less awareness about the existence of political deepfakes and hence more susceptibility to its adverse effects. Interestingly, warnings about the dangers and proliferation of political deepfakes have been shown to have the unintended consequences of making citizens disbelieve even real political media thereby creating the potential for the weaponizing of deepfake information.⁶ Overall, the African countries with fledgling democracies and with low AI adoption could be disproportionately at risk from this threat.

Conversely, it would appear that AI is the technology best poised to tackle the menace of deepfakes through detection and other means. The following sections will explore in more detail the emergence and applications of deepfakes, opportunities, challenges, and limitations of the current AI landscape in Africa, and the potential role of AI in mitigating the societal risk of deepfakes to African democracies.

B. The Rise of the Deepfake Threat

In the digital age, the proliferation of deepfake technology poses a significant threat to the integrity of democratic processes.

Deepfake technology utilizes AI algorithms to create realistic-looking images and videos that depict individuals saying or doing things they never did. This technology has advanced rapidly, making it increasingly difficult to distinguish between real and synthetic media. In the context of African democracies where information literacy varies widely, the potential for deepfakes to spread disinformation is particularly concerning.⁷

⁵ Africanews, Burkina: demonstration for "sovereignty" and in support of the junta in power < <https://www.africanews.com/2023/01/29/burkina-demonstration-for-sovereignty-and-in-support-of-the-junta-in-power/>>

⁶ Ternovski, J., Kalla, J., & Aronow, P. (2022). The Negative Consequences of Informing Voters about Deepfakes: Evidence from Two Survey Experiments. *Journal of Online Trust and Safety*, 1(2). <https://doi.org/10.54501/jots.v1i2.28>

⁷ The Africa Center for Strategic Studies' report, (2024), "Mapping a Surge of Disinformation in Africa" < Mapping a Surge of Disinformation in Africa – Africa Center for Strategic Studies>

Deepfakes, with their ability to fabricate seemingly authentic political content, exacerbate this vulnerability. Deepfakes erode trust in traditional media and political figures, making it difficult to discern truth from fiction.⁸ This can sow discord, manipulate public opinion, and ultimately undermine democratic processes.

Deepfake technology has a potential negative impact on political discourse in Africa. Some of these impacts include the following, First, fueling misinformation and political instability. Deepfakes can be used to create false narratives which can quickly spread through social media and other platforms thereby sparking hatred.

Second, the erosion of trust is a challenge occasioned by deepfakes. The inability to distinguish between real and fake content can lead to a general mistrust in media and official communications. This undermines the credibility of political institutions leading to increased cynicism and apathy among the electorate.

Third, deepfakes can be used for manipulation of elections. The AI tools could be used to discredit political candidates or influence public opinion during elections. For instance, the famous video of President Obama's deepfake.⁹ These tools are likely to portray a candidate in a negative light, potentially swaying voters' opinions and impacting the outcome of an election.

Fourth, deepfakes can amplify existing tensions. In regions where there is already social or political tension, deepfakes can exacerbate these issues by heightening existing risks.

Deepfake technology has the potential to cause havoc in the political arena, especially in the context of Africa's diverse and dynamic political landscape, which calls for vigilance and the development of tools to detect and combat deepfakes.

Examples of deepfakes in Africa include; in Burkina Faso in January 2023 unusual videos circulated on social media platforms featuring diverse individuals, seemingly endorsing the military junta. However, their movements were stiff, and their lips were out of sync with their words a classic characteristic of deepfakes.¹⁰

⁸ Nina Schick, (2022) Nina Schick on disinformation and the rise of synthetic media"< Nina Schick on disinformation and the rise of synthetic media - 80,000 Hours (80000hours.org)>

⁹ Wahington Post, (2024) "AI deepfakes threaten to upend global elections. No one can stop them," < <https://www.washingtonpost.com/technology/2024/04/23/ai-deepfake-election-2024-us-india/>>

¹⁰ADF Magazine, (2023) "Concern Grows as 'Deepfakes' Spread Misinformation" < <https://adf-magazine.com/2023/04/concern-grows-as-deepfakes-spread-misinformation/>>

In Gabon, President Bongo had been out of the country for medical treatment, leading to suspicions about his well-being. When a video of the President was eventually released, it raised questions due to its unusual nature.¹¹

In October 2023, cybercriminals exploited deepfake technology to impersonate African Union Commission chairperson Moussa Faki Mahamad creating realistic fake videos and audio recordings using deepfake technology.¹² They even engaged European leaders in video calls under his guise. The video was later debunked as a deepfake, but not before it had garnered significant traction and potentially swayed watchers.

These incidences underline the need for robust AI-powered solutions to detect and flag such content before it can cause widespread harm in Africa.

C. AI as a Detector and a Weapon Against Disinformation

As we highlight the pivotal role of AI in addressing political deepfakes within African democracies, it is crucial to recognize the multifaceted nature of these challenges. It is also necessary to focus on how AI can combat the proliferation of deceptive content. This discourse explores the intricate mechanisms of detection and authentication techniques, scalability, efficiency, and continuous learning. By understanding these complexities, we aim to shed light on the innovative strategies and technologies that AI offers in safeguarding the integrity of democratic processes in Africa.

i. Detection and Authentication Techniques

Deepfakes, fueled by AI and deep learning techniques, have become increasingly sophisticated, making them challenging to detect. However, AI offers promising solutions for identifying manipulated content through both content-based and context-based methods.

Content-based techniques analyze facial features, voice patterns, and audio-visual artifacts to detect inconsistencies, while context-based methods assess the credibility of sources and information consistency. Platforms like Deepware and Sensity exemplify these techniques, employing advanced algorithms to flag deepfake content with high accuracy.¹³ Additionally, traditional approaches, such as the hack developed for identifying AI-generated pictures by counting the fingers or toes of humans in these artworks, contribute

¹¹ The Thinking Watermill Society, AI-Generated Deepfakes: A Growing Threat to Africa, < AI-Generated Deepfakes: A Growing Threat to Africa - The Thinking Watermill Society>

¹²Institute of Commonwealth, “Fake news, mistrust and global media” < <https://commonwealth.sas.ac.uk/blog/fake-news-mistrust-and-global-media>>

¹³Mukta, M.S.H.; Ahmad, J.; Raiaan, M.A.K.; Islam, S.; Azam, S.; Ali, M.E.; Jonkman, M.,(2023) "An Investigation of the Effectiveness of Deepfake Models and Tools", < <https://doi.org/10.3390/jsan12040061>>.

to the arsenal of detection techniques. These methods complement traditional algorithms and enhance the accuracy of deepfake detection.

ii. Scalability and Efficiency

Deploying AI for deepfake detection across diverse African contexts presents scalability challenges. However, advancements in AI algorithms and computational resources address these concerns. Platforms utilizing cloud-based AI, such as DeepTrace and Dessa, demonstrate scalable solutions capable of analyzing vast amounts of digital content efficiently.¹⁴ Additionally, edge computing technologies enable real-time detection, mitigating the spread of deepfake content across online platforms.

iii. Continuous Learning

The dynamic nature of the deepfake threat necessitates continuous learning and adaptation by AI systems. Machine learning algorithms, trained on evolving datasets, enhance detection accuracy over time. Truepic is a platform that leverages blockchain technology to create immutable records of authentic media, facilitating continuous learning and verification of content authenticity.¹⁵

Misinformation remains pervasive, exacerbated by the proliferation of deepfake technology. Despite this, efforts to increase public awareness and media literacy play a crucial role in mitigating its impact. Educational campaigns, such as those conducted by the African Union and UNESCO, empower citizens to discern authenticity from manipulated content.¹⁶ Moreover, interdisciplinary collaborations between AI researchers, psychologists, and educators offer holistic approaches to combating misinformation and safeguarding democratic discourse.

By fostering collaboration among stakeholders, including researchers, tech companies, policymakers, and educators, innovation flourishes, ensuring that AI systems remain resilient against emerging threats.

D. Addressing the Bottlenecks and the Future of AI in Democracy in Africa

AI technologies in their nature have massive benefits to humanity but also pose momentous risks.¹⁷ AI has exemplified research, public administration, and revolutionized trade.

¹⁴ Sally Adee (2020) "What Are Deepfakes and How Are They Created? Deepfake technologies: What they are, what they do, and how they're made" < <https://spectrum.ieee.org/what-is-deepfake> >

¹⁵ ScoreDetect, "Content Authenticity Verification Tools" < <https://www.scoredetect.com/blog/posts/content-authenticity-verification-tools> >

¹⁶ UNESCO, (2023) "Journalism, fake news & disinformation: handbook for journalism education and training".

¹⁷ AI governance in Africa: An overview of regulation and policy work on artificial intelligence in Africa by ALT Advisory (2022)

Equally, it has exposed deep-rooted risks in technologies such as bias, exploitation, and blatant manipulation of public opinion to undermine democratic processes. Does this mean that we cannot find a balance between good and bad in AI? Definitely not. Just like a fulcrum, the pivot of AI governance needs to be founded on the salient principles of human rights so that the levers can balance without crippling innovation. AI tools can be used to regulate the challenges that they present in order to ensure safety of the humans.

AI technologies drive disinformation at two critical stages namely, content creation and content dissemination.¹⁸ At these stages, the information is likely to be manipulated to drive a specific agenda by perpetuating stereotypes or creating harmful content such as disinformation and misinformation. However, all is not lost, the third stage of content moderation is paramount in fighting disinformation. In the application of AI technologies, humans remain essential to the process, and thus, trust and safety should be at the forefront of any AI development.

Some countries in Africa have realized the benefits of AI tools to mitigate the challenges of misinformation. For instance, in Zambia, the European Union and the United Nations Development Programme (UNDP) set up a joint initiative to identify and fight harmful content during the general elections of August 2021. To respond to the threat posed by disinformation, the EC-UNDP Joint Task Force and the Democracy Strengthening (DSZ) project developed 'iVerify', a fact-checking and response mechanism.¹⁹

AI tools such as text analysis algorithms and image analysis or recognition have been used to fight disinformation by monitoring inappropriate information.²⁰ Behavioral analysis flags trolling and bots. However, these tools are not perfect and may fail to detect harm or malicious intentions.²¹ The United Nations has also urged for the implementation of actions against misinformation and cybersecurity.²²

The use of human moderators is critical in the detection of misinformation because they assist in flagging out harmful content. Even though AI tools may be efficient in detecting disinformation, human moderators continue to remain at the core of many moderation

¹⁸ Eduardo Albrecht, Eleonore Fournier-Tombs, and Rebecca Brubaker, *Disinformation and Peacebuilding in Sub-Saharan Africa: Security Implications of AI-Altered Information Environments* (New York: United Nations University and Interpeace, 2024)

¹⁹ European Commission, *Fighting Disinformation to Support Elections in Zambia* < https://international-partnerships.ec.europa.eu/news-and-events/stories/fighting-disinformation-support-elections-zambia_en >

²⁰ Pervaiz Akhtar et al *Detecting fake news and disinformation using artificial intelligence and machine learning to avoid supply chain disruptions* (2022) < <https://link.springer.com/article/10.1007/s10479-022-05015-5> >

²¹ Alla Katsnelson, "Identifying Misinformation's Intent," *Columbia Engineering*, 2023,

²² United Nations (2020). UN tackles 'infodemic' of misinformation and cybercrime in COVID-19 crisis. Retrieved from <https://www.un.org/en/un-coronavirus-communications-team/un-tackling-%E2%80%99infodemic%E2%80%99-misinformation-and-cybercrime-covid-19>

processes.²³ Additionally, digital content provenance has been praised for its strengths in establishing the origin of specific content, thereby enabling content moderators to flag out manipulated media.²⁴

Africa is a consumer of major digital technologies in the global market including AI tools. This comes at a cost such as the infrastructural gap that cannot effectively seize harmful AI tools and which results in imbalances in consumer protection. The domination of foreign AI technologies and solutions poses a gap in the developmental priorities of African countries.

Investing in infrastructures such as AI-powered tools for detection and countering disinformation will ensure the sanctity of the information and protection of the society from the influence of disinformation and misinformation driven by malicious AI.

It is imperative for Africa to explore multisectoral approaches such as public-private partnerships to facilitate the development of AI.²⁵ Africa has the chance to invest in AI research to curb misinformation and ensure the safety of human consumers in support of ongoing collaborative efforts by international organizations such as UNESCO.

Policy advocacy has steered the regulatory landscape in calling for human-centered approaches in the regulation of AI. African democracies should prioritize the development of regulatory frameworks for the responsible use of AI in politics to curb misinformation.

Noteworthy, the Global North is at the forefront in the shaping of various AI strategies. Notwithstanding, African voices are key in shaping the global dialogue in AI. African democracies should align their national AI policies with broader developmental agendas and continental policies such as the AU AI Strategy. African Countries such as Rwanda, Benin, Egypt, Morocco, Mauritius, Tunisia, Sierra Leone, and Senegal have taken a proactive role in developing national AI strategies.²⁶ African Countries have significantly invested in data protection and privacy frameworks with 36 out of 55 African countries (65%) having data protection law.²⁷

E. Conclusion and Summary of Recommendations

²³ Center for Data Ethics and Innovation, the Role of AI in Addressing Misinformation on Social Media Platforms, (2021)

²⁴ The Royal Society and BBC, “Generative AI, Content Provenance and a Public Service Internet,” Royal Society, < https://royalsociety.org/-/media/policy/projects/digital-content-provenance/digital-content-provenance_workshop-note_.pdf>

²⁵ SMARTAfrica, (2021) Blueprint on Artificial Intelligence for Africa, 1(e.d).

²⁶ Melodi Musoni, (2024) “Envisioning Africa’s AI governance landscape in 2024”.

²⁷ Data Protection Africa, Mapping the progress (and delays) for data protection in Africa, available at < Mapping the progress (and delays) for data protection in Africa>

In order to fully harness the role of AI in combating political deepfakes in African democracies, a multi-pronged approach is recommended as follows:

Firstly, there is the need for stakeholders to recognize the significant risk posed by deepfakes and their potential to, both undermine democracy and to be weaponized. In response, deliberate prioritization of regulation of AI technologies is required.

Secondly, the creation of a multi-sectoral group of key stakeholders spanning all sectors of society. These experts should be drawn from AI researchers, ethicists, the tech community, regulators, the media, citizens groups, civil society, global and regional multilateral organizations, policymakers, and the government to ensure a broad and inclusive range of perspectives and voices.

Thirdly, establishment and strengthening of the necessary infrastructure to create and sustain an environment that leverages and builds capacity in AI, and which also facilitates the ability to keep up with rapidly evolving advancements.

Fourth, there is the need for event-based awareness campaigns to run along events such as elections and other political processes in order to counteract campaigns by bad actors.

Finally, enhancing the effectiveness of AI by tailoring solutions to the unique circumstances and prevalent ethical considerations of African democracies is key. These techniques should be grounded in the need to preserve and protect human rights and safety.

There are still gaps in research on African democracies and the effect of deepfakes on these democracies, but some broad lessons can be learnt from insights gathered in the Global North and other Global South regions. Specific ethical considerations unique to Africa should also guide further research and subsequently lead to the development of hitherto unexplored perspectives and innovative solutions to this problem. African democracies should also proactively engage in global policy discussions shaping AI and not wait to be invited to the policy discussion tables. This essay should form the foundation for such further work on this subject.

(Word Count: 2905 inclusive of footnotes)

REFERENCES

1. Yisroel Mirsky and Wenke Lee. (2020) (The Creation and Detection of Deepfakes: A Survey” ACM Computer. Survey. < <https://arxiv.org/pdf/2004.11138>>.
2. Dobber, T., Metoui, N., Trilling, D., Helberger, N., & de Vreese, C. (2021). Do (Microtargeted) Deepfakes Have Real Effects on Political Attitudes? *The International Journal of Press/Politics*, 26(1), 69-91. <https://doi.org/10.1177/1940161220944364>.
3. Bennett, W. L., & Livingston, S. (2018). The disinformation order: Disruptive communication and the decline of democratic institutions. *European Journal of Communication*, 33(2), 122-139. <https://doi.org/10.1177/0267323118760317>
4. Allie Fuck, Adrian Shahbaz and Kian Vesteinsson, The Repressive Power of Artificial Intelligence (2024) < <https://freedomhouse.org/report/freedom-net/2023/repressive-power-artificial-intelligence#:~:text=During%20the%20February%202023%20Nigerian,integrity%20of%20the%20electoral%20system.>>
5. Africanews, Burkina: demonstration for "sovereignty" and in support of the junta in power < <https://www.africanews.com/2023/01/29/burkina-demonstration-for-sovereignty-and-in-support-of-the-junta-in-power/>>
6. Ternovski, J., Kalla, J., & Aronow, P. (2022). The Negative Consequences of Informing Voters about Deepfakes: Evidence from Two Survey Experiments. *Journal of Online Trust and Safety*, 1(2). <https://doi.org/10.54501/jots.v1i2.28>
7. The Africa Center for Strategic Studies' report, (2024), "Mapping a Surge of Disinformation in Africa" < Mapping a Surge of Disinformation in Africa – Africa Center for Strategic Studies>

8. Nina Schick, (2022) Nina Schick on disinformation and the rise of synthetic media" < Nina Schick on disinformation and the rise of synthetic media - 80,000 Hours (80000hours.org) >
9. Wahington Post, (2024) "AI deepfakes threaten to upend global elections. No one can stop them," < <https://www.washingtonpost.com/technology/2024/04/23/ai-deepfake-election-2024-us-india/> >
10. ADF Magazine, (2023) "Concern Grows as 'Deepfakes' Spread Misinformation" < <https://adf-magazine.com/2023/04/concern-grows-as-deepfakes-spread-misinformation/> >
11. The Thinking Watermill Society, AI-Generated Deepfakes: A Growing Threat to Africa, < AI-Generated Deepfakes: A Growing Threat to Africa - The Thinking Watermill Society >
12. Institute of Commonwealth, "Fake news, mistrust and global media" < <https://commonwealth.sas.ac.uk/blog/fake-news-mistrust-and-global-media> >
13. Mukta, M.S.H.; Ahmad, J.; Raiaan, M.A.K.; Islam, S.; Azam, S.; Ali, M.E.; Jonkman, M.,(2023) "An Investigation of the Effectiveness of Deepfake Models and Tools", < <https://doi.org/10.3390/jsan12040061> >.
14. Sally Adey (2020) "What Are Deepfakes and How Are They Created? Deepfake technologies: What they are, what they do, and how they're made" < <https://spectrum.ieee.org/what-is-deepfake> >
15. ScoreDetect, "Content Authenticity Verification Tools" < <https://www.scoredetect.com/blog/posts/content-authenticity-verification-tools> >
16. UNESCO, (2023) "Journalism, fake news & disinformation: handbook for journalism education and training".
17. ALT Advisory (2022), AI governance in Africa: An overview of regulation and policy work on artificial intelligence in Africa.
18. Eduardo Albrecht, Eleonore Fournier-Tombs, and Rebecca Brubaker, Disinformation and Peacebuilding in Sub-Saharan Africa: Security Implications of AI-Altered Information Environments (New York: United Nations University and Interpeace, 2024)

19. European Commission, Fighting Disinformation to Support Elections in Zambia < https://international-partnerships.ec.europa.eu/news-and-events/stories/fighting-disinformation-support-elections-zambia_en>
20. Pervaiz Akhtar et al Detecting fake news and disinformation using artificial intelligence and machine learning to avoid supply chain disruptions (2022) < <https://link.springer.com/article/10.1007/s10479-022-05015-5>>
21. Alla Katsnelson, “Identifying Misinformation’s Intent,” Columbia Engineering, 2023,
22. United Nations (2020). UN tackles ‘infodemic’ of misinformation and cybercrime in COVID-19 crisis. Retrieved from <https://www.un.org/en/un-coronavirus-communications-team/un-tackling-%E2%80%98infodemic%E2%80%99-misinformation-and-cybercrime-covid-19>
23. Center for Data Ethics and Innovation, the Role of AI in Addressing Misinformation on Social Media Platforms, (2021)
24. The Royal Society and BBC, “Generative AI, Content Provenance and a Public Service Internet,” Royal Society, < https://royalsociety.org/-/media/policy/projects/digital-content-provenance/digital-content-provenance_workshop-note_.pdf>
25. SMARTAfrica, (2021) Blueprint on Artificial Intelligence for Africa, 1(e.d).
26. Melodi Musoni, (2024) “Envisioning Africa’s AI governance landscape in 2024”.
27. Data Protection Africa, Mapping the progress (and delays) for data protection in Africa, available at < Mapping the progress (and delays) for data protection in Africa>