Explainability Imperative of Generative Artificial Intelligence Navigating the Moral Dilemma of AI in Nigeria and Charting a Path for the Future

Emedo Chinyere Christian, PhD

Department of Philosophy, Faculty of Arts, Nnamdi Azikiwe University, Awka, Anambra State, Nigeria.

Abstract

This paper explores the explanability imperative in the context of Generative Artificial Intelligence (GAI) and its crucial role in addressing the concerns posed by AI technology in Nigeria. This underscores the ethical necessity for AI systems, especially generative ones to provide clear and understandable explanations for their decisions and actions. Although the advent of generative AI undoubtedly heralds the future and however, has also exposed Nigerian society to new vulnerabilities that seemingly are detrimental to our epistemic agency and peaceful political settings. Employing the phenomenological method of philosophical inquiry here, we discovered that this new technology has posed big threats to the future world, and that Nigeria falls amongst this new technology users. To navigate the moral dilemma caused by Generative Artificial Intelligence, this paper suggests many proactive approaches like the development of localized AI explainability standards, the regulatory frameworks, and educational initiatives to promote awareness and understanding of AI systems in Nigeria. By prioritizing the Explanability Imperative, Nigeria can chart a path towards a future whereby AI technologies aligned with societal values, upholds standard education, and as well contributes positively to the nation's development. This paper encapsulates the importance of AI explainability in Nigeria's AI landscape and its potential to shape a more ethically responsible and transparent AI future.

Keywords: *Generative AI, Epistemic Agency, Democracy, Opaque, and Explianbility AI.*

INTRODUCTION

The events in this contemporary time which have seen the rapid integration of artificial intelligence (AI) technologies as a pivotal aspect of human existence, have caused so much worry amongst the general public. The increasing use of Artificial Intelligence (AI) in making decisions for public affairs has provoked debate on the benefits and potential harmfulness of self-learning technologies ranging from the hopes of fully informed and objectively taken decisions for fear of the entrenching social injustice (Bostrom:2017, Pasquale:2015; O'Neil:2016;).

This issue has also elicited so many scholastic explorations in the field of philosophy, particularly in the field of ethics; as a step towards preventing negative outcomes and achieving responsible systems. Many of these scholars have argued on the needs to open the "black box" of AI decision-making and make it more transparent. This, they also contend, will enable the understandings of the functions of AI as well as receiving possible explanations from individuals' decisions (Floridi et al.:2018; Lepri et al.:2017; Wachter et al.:2017; Binns: 2018; De Laat: 2018).

However, little has been said from the lens of political epistemology about the effect of AI on people under the non-expert or new technology users' category like most African countries, specifically Nigeria. Thus, the question on the importance of explanations in the age of Artificial Intelligence dependency. To address this question, this paper brings political epistemology to bear; drawing from the concept of epistemic agency to support concerns about AI, knowledge and Democracy.

Therefore, this paper is divided into four sections. Firstly, we shall conceptually analyze the keywords of the study viz; Generative AI, Epistemic agency, Democracy; and Political epistemology. The second section will talk about why Generative AI negates epistemic agency which sabotaged Human Autonomy. The third section shall discuss why Artificial Intelligence is problematic to democracy, and why it is a threat to educational standards in Nigeria. The fourth section will focus on the explainability imperative and what needs to be done.

CONCEPTUAL CLARIFICATION

Generative Artificial Intelligence (GAI)

Artificial Intelligence is a complex concept that comprises a wide range of technologies and techniques. The discuss here, will focus on Artificial Intelligence in the sense of generative ones due to the context of this research paper. Thus, Generative AI refers to a category of artificial intelligence that involves creating or generating new content such as text, images, voices, or even music. Unlike traditional AI systems that rely on explicit programming, generative AI uses models

often based on neural networks, to learn patterns and relationships from data and generate novel outputs. This technology has applications in various fields, such as natural language generation, image synthesis, and creative tasks. Open AI's GPT-3, Bing, and Generative Adversarial Networks (GANs), are examples of GAI.

In social media contexts, it can be used to search contents, make recommendations, recognize images or speech. Profile users target them with personalized advertisement, analyze sentiments in text, or create new content. The functions of GIA didn't stopped with the above, it considers for example, recommending and searching algorithms used in Facebook, Twitter, and Google, it also deepfakes videos and other media generated by machine learning algorithms in which a person in an existing image or video is (partly) replaced with someone else's likeness. For example, it is now possible to make a video of a politician and have him or her say or do things he never said. Also in the academic context, it is becoming a norm that people now depend on GAI like Open AI's GPT-3 to write research papers and other academic works. This leaves a begging question on the extent to which we should rely on these GAIs as a source of knowledge.

Epistemic Agency

Epistemic agency refers to an individual's capacity and ability to take active control over their beliefs, knowledge, and cognitive processes. It involves the exercise of intellectual autonomy, the ability to critically evaluate information, and the power to form and revise one's beliefs based on reasoned judgment. Epistemic agency pertains to the issue of exerting control over one's beliefs, as discussed by (Scholosser: 2019), and involves the examination of how these beliefs are shaped and modified. While individuals posses the repertoire to reflect on their beliefs, the extent of control they have over them remains unclear. Despite the peoples desire for control over their beliefs, the level of influence they can exert raises questions. Epistemology has long been engaged in discussion on the voluntary nature of belief formation and its connection to normative concepts like responsibility (Heil: 1983; Steup: 2001). In the realm of social epistemology, discussions also revolve around the impact of the social context on belief formation. Notably, beliefs are interconnected with a broader knowledge community, giving rise to the concept of "collective knowledge" (Goldman: 2001). In this context therefore, we shall shift our point to the development of political beliefs among citizens and the relevant epistemic agency within the sphere of technology.

Democracy

Democracy is a form of government where power lies in the hands of the people, either directly or indirectly by means of elected representatives. It is a political system where citizens have the right to participate in the decision-making processes that affect their lives. Theories and definitions of democracy exhibit a broad range, leading to significant variations in expectations placed on the citizens. Representatives perspectives can be categorized either as a"thin" content with citizens voting every 4 or 5 years, or a"thick" one requiring direct engagements in democratic processes, such as deliberation (Cohen: 1986; Boyte: 2017; Estlund: 2008) or agonistic struggle (Mouffe: 2016). The terms "thin" and "thick" have a historical context in political theory.

The "thin" conception of democracy which focuses on voting and representation is the aspect of democracy that we mean to discuss here, rather than the "thick" democracy which demands direct partispation of the citizens. This paper explores the knowledge foundation of democracy, with all variations assuming that citizens possess politically relevant knowledge. For instance, voters are expected to be informed about the political candidate's program. In a deliberative democracy, citizens suppose to know the issues and should exercise communicative rationality or engage in agonistic political struggles contesting different perspectives.

This inquiry however, does not only raises questions about the necessary knowledge in a democracy, but also explores the processes, know-how, and skills involved. Therefore, political epistemology which is interested in citizens' acquisition of politically relevant knowledge, and beliefs, addresses questions such as what constitutes knowledge and the relationship between truth and democracy. However, in this paper, the focus is on epistemic agency: whether citizens possess sufficient epistemic agency in the context of AI, rather than exploring the broader theme of truth and democracy.

Political Epistemology in the Context of GAI

Political democratic agency relies on epistemic agency, emphasizing the necessity for citizens in a democracy to have some degree of control over the development of their political knowledge. Reflections on one's beliefs, and the willingness to engage in public discourse becomes crucial in a deliberative and participative democracy. In this system, citizens are assumed to possess control over their beliefs. But, where they're subjected to brainwashing by an authoritarian regime, they will lose this control and, consequently the political agency in terms of voting or participating in democratic deliberations.

The manipulation of political beliefs poses a challenge to democratic processes like voting and deliberative democracy. These processes presume that citizens, whether as voters or participants in deliberation, maintain control over their political knowledge and beliefs, which they express through voting, argumentation, or engagement in agonistic struggle. In contrast, non-democratic, authoritarian, and totalitarian political orders, do not demand epistemic agency from citizens since their only expected role is to support the regime, discourage independent political beliefs formations, reflections, as well as discussions.

Recent societal and technological developments prompt questions about the sufficiency of the epistemic foundation of existing democratic forms, with concerns about citizens' education and vulnerability to misinformation. The use of artificial intelligence (AI) and big data by governments and tech companies, particularly through digital social media, adds a new dimension to these challenges, potentially impacting the knowledge basis of the democracy. The influence of AI on political knowledge and knowledge formation requires further investigation.

While the erosion of epistemic agency has been observed for some time with concerns present in the Western democracy due to sensationalists' media reporting, philosophers and political theorists contribute to the discussion with arguments and conceptual work. This paper also, explores the relationship between GAI knowledge and democracy focusing on the risks and dangers associated with GAI. Although AI may have positive effects on democracy (an under-studied area), but this paper does not delve into that aspect. Understanding the challenges is considered crucial for those advocating "AI" for democracy.

WHY GAI IS DETRIMENTAL TO EPISTEMIC AGENCY AND HUMAN AUTONOMY

Among the biggest concerns about the wide adaptation of GAI by the general public is the potential negation of the epistemic agency and autonomy. Epistemic agency as discussed in the first phase of this paper refers to individuals' active role in determining and acquiring knowledge, while autonomy involves the ability to make independent decisions. This underscores the moral agency of human beings which distinguishes them from other phenomena such as lower animals and AI robotics, this is because human beings alone are capable of performing moral actions. An important rudiment of human beings moral status is that they are responsible agents, and can be attributed the responsibility, because for now, only human beings possess the two philosophical conditions for the attribution of responsibility, the so-called 'control' condition and the 'epistemic' condition (Fischer and Ravizza 1998; Rudy Hiller 2018).

Therefore in the context of this paper, GAI negates our epistemic agency and autonomy in many ways which we shall buttress in the following paragraphs.

Autonomous Information Generation

Generative Artificial Intelligence can autonomously generate contents which leads to a surge in information creation without direct human involvements. This autonomous content generation can undermine individuals' traditional roles in researching, verifying, and interpreting information eroding their epistemic agency.

Misinformation and Manipulation

The inherent nature of Generative Artificial Intelligence provokes concerns about the potential for generating misinformation or biased content. When individuals rely heavily on AI-generated information, there is the risk of being exposed to distorted perspectives, impacting their ability to form accurate beliefs and make informed decisions.

Loss of Control Over Narrative

GAI can contribute to the creation of narratives that may not align with diverse human perspectives or values. This loss of control over the narrative landscape diminishes human autonomy, as individuals may find themselves been influenced by AI-generated contents that don't necessarily reflect their own beliefs or priorities.

Ethical Concerns and Unintended Consequences

Generative Artificial Intelligence systems may inadvertently perpetuate societal biases present in training data, exacerbating existing inequalities. Ethical concerns surrounding the deployment of GAI further complicate the relationship between individuals and the information environment, impacting both epistemic agency and autonomy

Therefore, while Generative AI holds tremendous potential, careful consideration and proactive measures are essential to mitigate its detrimental effects on epistemic agency and the human autonomy. Balancing the benefits of AI-driven information generation with the preservation of individuals agency and autonomy is a sine qua on for fostering a healthy and informed society.

THE PROBLEMATIC EFFECT OF GAI ON DEMOCRACY AND ACADEMIC STANDARDS

The occurrences of fake news and misinformation as widely recognized, often serving political agendas, are exemplified by the notorious Pizzagate incident in the year 2016. For example, members of online message board spread false information alleging that Bill and Hillary Clinton operated a pedophile sex ring in a pizza restaurant, leading to real-world consequences such as death threats. This debunked story resurfaced yet again in 2020 on Instagram and TikTok, endorsed by QAnon.

There is a consensus that misinformation and fake news pose a threat to democracy. They influence public opinions, create confusion about truth and reality, and erode the epistemic quality required for deliberative democracy. For instance, (Mckay and Tenove:2020) argued that online disinformation undermines a polity's ability to engage in fact-based communication. Also, there have been wide deliberations in recent times about the ideal approach to go about fake news and disinformation because of the sensitivity of the topics. For instance, some scholars have argued that the ideal way to react to misinformation like hate speech is counter-speech. This looks plausible, lepoutre (2023) argued that though, counter speech might be plausible, but measures should be taken to avoid hateful counter speech as it can generate serious harm which will dent the epistemic agency of the general public. The

political significance of truth, becomes evident in totalitarian regimes. Hannah Arendt (1951) in her book "The Origin of Totalitarianism", pointed out that it is a feature of these regimes to distort the truth through consistent lying, and this is used to establish and maintain a fictional world.. Artificial Intelligence however, introduces new challenges by enabling truth distortions using altered search engines, deepfake videos, and targeted misinformation spread via social media usage and big data analysis. Authoritarian regimes benefit from such manipulation, but liberal democracies are also at risk. Citizens' confidence in identifying misinformation, may diminish as AI-generated contents including deepfake videos, becomes harder to distinguish from reality. Thus the rapid and global dissemination of misinformation via social media amplifies these challenges.

A notable example involves a digitally altered video of Nancy Pelosi in 2019, shared by Donald Trump, illustrating the political relevance of deepfakes. Bringing it down to the Nigerian sphere, there were reports of cases where Generative Artificial Intelligence was deployed to spread fake news and disinformation, particularly during the pre and post-2023 general elections. For instance, the cases where Atiku Abuabakar's daughter was reported to threatening Nyesom Wike who apparently, was an opponent to his father during the PDP's primary elections with intimidation should his father win. There was also the case of an AI-generated voice call which allegedly was a leak from Mr Peter Obi, the Labour Party Presidential candidate declaring that the election was a Muslim-Christian religious war.

However, the problem extends beyond specific cases of lying, impacting trust in others and democratic society. Citizens may question their epistemic capacities when faced with increasingly believable AI-generated fakes, leading to a broader mistrust in their ability to distinguish truth from falsehood. This erosion of trust is compounded by the realization that AI puts others in control of political knowledge and beliefs formation undermining democratic agency. The future of Nigerian healthy political existence is posed with massive challenges in the age of new technologies like Generative Artificial Intelligence.

 On the other hand, it has a big implication for the academic standard of the Nigerian educational system. It is said that the education sector is among the critical sector that determines the level of development of a society, and determine the sustainability of every society's future. Thus, the rapid dependence on new technologies like GAI reduces the ability for critical thinking and engagement which improves the learning standard of educational systems.

Dependence on GAI for information leads to a decline in critical thinking skills, and makes individuals become passive consumers rather than actively engaging with and questioning the content presented to them. This reduced engagement threatens both epistemic agency and human autonomy by fostering a more passive relationship with knowledge. Also, GAI systems may inadvertently perpetuate societal biases present in training data, exacerbating existing inequalities, especially in the realm of education. Ethical concerns surrounding the deployment of GAI further complicate the relationship between individuals and the information environment, impacting both epistemic agency and autonomy, which in turn entrenches epistemic injustice.

There is a shared responsibility between the stakeholders, the policy makers, and the general public to confront these challenges, specifically in efforts to protect epistemic agencies. In this context therefore, we shall discourse the imperative of explanations as a lens through which Nigeria can confront the challenges posed by these technologies.

THE IMPERATIVE OF EXPLAINABILITY AI: TOWARDS SOLVING THE MORAL DILEMMA IN NIGERIA

The increasing integration of Artificial Intelligence (AI) into various aspects of Nigerian sectors raises ethical concerns, particularly regarding the moral dilemma posed to new tech users who dominate Nigerian demography. One crucial aspect of AI ethics is the imperative in the explainability, which refers to the ability of AI systems to provide understandable and transparent explanations for their decisions and actions. The issue of Explanation is an ever-reoccurring topic in philosophy because of its relevance in informing knowledge about things. For Salmon (1978), causal explanation is valuable because it explains the how and why's an effect occurs and consequently, provides information regarding when and where the relationship can be replicated. Also for Bufacchi (2018), we need to know what justifies a practice, we need reasons that are or could be publicly endorsed for accepting a practice since practice cannot be self-justifying.

Explainability is fundamental for building trust between AI systems and users. Users are likely to embrace and trust AI technologies when they understand how these systems arrive at their decisions. Lack of transparency can lead to skepticism and reluctance among users, especially in critical domains like healthcare, finance, and political choice.

Explanability therefore, is essential for holding AI systems accountable for their actions. Users, policymakers, and regulatory bodies need insights into the decision-making processes to ensure ethical behavior. In cases where Artificial Intelligence systems are involved in consequential decision-making, such as in autonomous vehicles or hiring processes, transparent explanations are crucial for addressing potential biases and preventing discrimination.

Users may face moral dilemma, specifically when interacting with Artificial Intelligence system which make decisions affecting them without providing clear explanations. This lack of transparency challenges the principles of informed consent. Understanding how AI system reach decisions is vital for users to make informed choices about their engagements with technology.

New technology users may encounter moral dilemma when AI system produces unintended and undesirable consequences. Without explainability, users may feel powerless and question the ethical implications of AI decisions. Providing explanations allows users to comprehend the rationale behind AI actions and empowers them to address or mitigate unintended outcomes.

Therefore addressing the moral dilemma requires collective responsibility among the policymakers, the developers, and the general public.

First of all, governments and regulatory bodies play important roles in mandating transparency and explainability standards for AI systems. Implementing and enforcing clear regulations can guide the ethical development and deployment of AI technologies.

Also, the developers must prioritize ethical considerations in Artificial Intelligence design, ensuring that systems are inherently interpretable. Designing algorithms with transparency in mind contributes to ethical usage and diminishes the moral dilemma faced by users.

There is the need for democratization of the users' education to confront this concern. Educating users about Artificial Intelligence technologies and their decision-making processes is crucial too. Transparent communication about how AI impacts their lives and, the provision of tools for understanding complex algorithms can empower users to navigate moral challenges effectively.

In conclusion, the imperative of explainability in AI is crucial for addressing the moral dilemma posed to new technology users like Nigerians. Transparent AI systems don't only build trust and confidence, but also uphold ethical standards and empower users to make an informed decisions. Again A concerted effort from developers, regulators, and educators is necessary to ensure that AI technologies align with ethical principles and contribute positively to society.

CONCLUSION

In this paper, we have provided an overview of various ways in which the foundation of democracy is at the risk of erosion, due to the implementation of Generative Artificial Intelligence. The risks encompass issues like spread of fake news, misinformation, as well as phenomena such as epistemic bubbles and the misrepresentation of statistical information. While these phenomena pose challenges in diverse ways, the focus have been on their impacts on epistemic agency, which subsequently poses a threat to political agency in a democratic system, specifically, that of Nigeria.

Acknowledging the complexity of empirical reality and recognizing that attacks on epistemic agency existed before the advent of AI, however, we have identified certain philosophical arguments that support concerns about these phenomena in the context of democracy. If these phenomena are indeed a reality, and if these arguments hold, it is important for those who value democracy in Nigeria to consider taking measures to prevent, limit, or mitigate these problems. This involves safeguarding and potentially enhancing both the epistemic and political agency in the context of AI and related technological and social phenomena.

For instance, we emphasized that educational initiatives in Nigeria should incorporate lessons on statistics and its limitations. Citizens should be encouraged to develop their reasoning repertoire, embrace epistemic doubt, think critically, and understand diverse perspectives. Similarly, efforts should be made to nurture individuals as responsible epistemic agents, fostering virtues that counteract increased epistemic arrogance and the dismissal of others' credibility, as proposed by Gunn and Lynch.

Furthermore, developers of AI and the organizations they work for can play a role in addressing these issues. Modifications to AI algorithms on social media platforms could disrupt epistemic bubbles, as suggested by Bozdag and van den Hoven (2021). They note limitations in existing tools, emphasizing the need for a broader range of democracy's models. Therefore, bridging the gap between democracy theory and technical development is essential, and policymakers in Nigeria should provide a framework that supports interdisciplinary research. Additionally, policy initiatives should encourage and mandate AI developers and companies to create AI that is resilient to threats against democracy. Also, we suggested that the general public should also understand that they have the duty of exercising their inherent moral agency which involves making independent and critical thinking a priority, because it will help in protecting their epistemic agencies against the misinformation. The future of Nigeria depends on the choices we make today, we must preserve quality knowledge for the next generation by making sure we critically harness the good side of these new technologies as a means or tool, and not depend on them, because they risk sabotaging our epistemic agencies.

REFERENCES

1. Arendt, H. 1951/2017. The Origins of Totalitarianism. London: Penguin

2. Binns R (2018) Algorithmic accountability and public reason. Philos Technol 31(4):543–556

3. Bostrom N (2017) Superintelligence: paths, dangers, strategies. Oxford University Press.

4. Boyte, H.: John Dewey and citizen politics: how democracy can survive artificial intelligence and the credo of efficiency. Educ. Cult. 33(2), 13–47 (2017).

5. Bozdag, E., van den Hoven, J.: Breaking the filter bubble: democracy and design. Eth. Inf. Technol. 17, 249–265 (2015)

6. Bufacchi, Vittorio (2008). 'The Truth About Rights', Journal of Human Rights, Volume, issue 4, pp 311-326.

7. Cohen, J.: An epistemic conception of democracy. Ethics 97(1), 26–38 (1986).

8. De Laat PB (2018) Algorithmic decision-making based on machine learning from Big Data: Can transparency restore accountability? Philos Technol 31(4):525–541

9. Estlund, D.: Democratic authority. Princeton University Press, Princeton (2008).

10. Fischer, J.M., & Ravizza, M. (1998). Responsibility and control: A theory of moral responsibility. Cambridge University Press.

11. Floridi L, Cowls J, Beltrametti M, Chatila R, Chazerand P, Dignum V, Schafer B (2018) AI4People—An ethical framework for a good AI society: opportunities, risks, principles, and recommendations. Mind Mach 28(4):689–707

12. Goldman, A., Whitcomb, D.: Social epistemology: essential readings. Oxford University Press, New York (2011)

13. Heil, J.: Doxastic agency. Philos. Stud. 43(3), 355–364 (1983)

14. Lepoutre, M. Hateful Counterspeech. Ethic Theory Moral Prac 26, 533-554 (2023). https://doi.org/10.1007/s10677-022-10323-7

15. Lepri B, Oliver N, Letouzé E, Pentland A, Vinck P (2017) Fair, transparent, and accountable algorithmic decision-making processes. Philos Technol 2017:1–17

16. McKay, S., Tenove, C.: Disinformation as a threat to deliberative democracy. Polit.Res. Q. (2020). https://doi.org/10.1177/10659 12920938143.

17. Mouffe, C. Democratic politics and conflict: an agonistic approach. Politica comun 9. https://www.quod.lib.umich.edu/p/pc/12322227.0009.011?view=text;rgn=main (2016). Accessed 18 Dec 2023.

18. O'Neil C (2016) Weapons of math destruction: how big data increases inequality and threatens democracy. Broadway Books, New York.

19. Pasquale F (2015) The Black Box Society: the secret algorithms that control money and information. Harvard University Press, Cambridge

20. Rudy-Hiller, F. 2018. The epistemic condition for moral responsibility. Stanford Encyclopedia of Philosophy. Retrieved 14, Jan. 2024. https://plato.stanford.edu/entries/moral-responsibility-epistemic/ .

21. Salmon, W.C. (1978). Why ask, "Why?" An inquiry concerning Scientific Explanation Proceedings and Address of the American Philosophy Association, 51(6), 683 https://doi.org/10.203/3129654

22. Schlosser, Markus. 2019. Agency. The Stanford encyclopedia of philosophy. Edward N. Zalta (Ed.). https://www.plato.stanford. edu/archives/win2019/entries/agency/. Accessed 8 Jan 2024

23. Steup, M.: Knowledge, truth, and duty: essays on epistemic justification, responsibility, and virtue. Oxford University Press, Oxford (2001).

24. Wachter S, Mittelstadt B, Floridi L (2017) Why a right to an explanation of automated decision-making does not exist in the general data protection regulation. Int Data Privacy Law 7(2):76–99