

THE FUTURE OF KNOWLEDGE: ETHICAL IMPLICATIONS OF AI-DRIVEN KNOWLEDGE PRODUCTION

Cynthia Ganchok Ngwetar¹ & Jesse Iorfaga Agede²

¹Department of Philosophy, Federal University
Wukari, Taraba State, Nigeria
cganchok@gmail.com

²Department of Philosophy, Federal University
of Lafia, Nasarawa State, Nigeria
Jesse.agede@fulafia.edu.ng;
jessagede@yahoo.co.uk

Corresponding author ¹Department of
Philosophy, Federal University Wukari,
Taraba State, Nigeria
cganchok@gmail.com

Abstract

The integration of artificial intelligence (AI) into knowledge production has raised significant philosophical and ethical questions regarding the nature of understanding and the basis of epistemic authority. This paper employs a dialectical analysis, drawing upon contemporary epistemology and ethics, to examine how AI-generated outputs challenge traditional models of human cognition and the implications for intellectual autonomy. The analysis reveals that while AI systems are adept at processing vast datasets and identifying patterns, they inherently lack consciousness and intentionality, qualities central

to human comprehension. This absence raises concerns about the potential for AI to perpetuate biases present in training data, thereby reinforcing existing societal inequalities. Moreover, an overreliance on AI for critical thinking tasks may erode individual intellectual autonomy, as users become passive recipients of machine-generated conclusions. To address these challenges, a framework is proposed that emphasises ethical oversight, transparency in AI design, and the promotion of human-AI collaboration that respects and enhances human cognitive faculties. Ultimately, the paper argues that the future of knowledge should not be dictated by AI capabilities alone but should focus on integrating AI in ways that uphold human values and promote equitable access to information.

Keywords: Artificial Intelligence (AI), Knowledge Production, Epistemology, Ethics, Human Inquiry, Machine Outputs

Introduction

In recent times, the production, validation, and distribution of knowledge have undergone profound transformation, driven by the rise of artificial intelligence. AI systems now summarize scientific literature, generate plausible news articles, write code, answer complex philosophical questions, and even make predictions in medical and legal domains. While these capacities are remarkable, they also invite us to ask more fundamental and troubling questions: What counts as knowledge in a world

where machines generate it? Can machines know anything in the sense that humans do? And if not, what are the implications for relying on them as sources of truth?

At stake is not just the accuracy or performance of AI, but the very fabric of human knowledge; what it means, who controls it, and how it will evolve in the coming decades. The ethical implications of this shift are vast and deeply philosophical. They concern not just whether AI can be used responsibly, but whether the nature of knowing itself is being reshaped by systems that operate outside of human consciousness, experience, and social life. This paper explores these questions by critically examining the epistemological foundations of AI-driven knowledge, the growing authority of AI systems, the injustices encoded into their outputs, and the uncertain future of human inquiry in their wake.

Redefining Knowledge in the Age of AI

The advent of artificial intelligence (AI) has triggered a significant re-evaluation of the concept of knowledge, making it an epistemological issue. AI systems process, classify, and generate information, often with high accuracy and usefulness. But is that the same as knowledge? Traditionally, knowledge has been understood as justified true belief, an idea with roots in Plato's *Theaetetus* that has guided epistemological discourse for centuries. According to this view, for a belief to count as knowledge, it must

not only be true and believed but justified through some rational means. Machines, however, do not believe. They have no minds, no intentions, and no awareness of truth or falsehood. Their "outputs" are the result of pattern recognition, optimization, and probabilistic modelling, not reflective cognition (human thinking). However, as machines increasingly generate information and perform tasks that were once the exclusive domain of human cognition, it becomes necessary to scrutinise what constitutes knowledge in this new era.

What Counts as "Knowledge" When Machines Generate It?

The traditional epistemological model posits that knowledge comprises three components: belief, truth, and justification. This tripartite structure has been foundational in assessing human claims to knowledge. However, the emergence of AI challenges this paradigm, as machines lack beliefs and consciousness. They operate by processing data and executing algorithms, producing outputs based on patterns and statistical odds rather than understanding or awareness. This raises the question: can the information generated by AI be classified as knowledge?

Consider the Gettier problem, introduced by philosopher Edmund Gettier, which presents scenarios where individuals have justified true beliefs that nevertheless fail to constitute knowledge due to the

element of luck or misleading appearances involved (Gettier, 1963). Similarly, AI systems may produce outputs that are true and appear justified based on their programming and data inputs, yet lack the foundation of genuine understanding. For instance, an AI language model might generate a grammatically correct and contextually relevant sentence without any comprehension of its meaning. This parallels Gettier cases, suggesting that the absence of understanding in AI outputs disqualifies them from being classified as knowledge in the human sense.

Philosophers like Hubert Dreyfus (1992) have long argued that AI lacks the embodied, contextual, and intuitive understanding that characterises human intelligence. Dreyfus, drawing on Heidegger and Merleau-Ponty, critiqued the assumption that cognitive processes could be reduced to rule-following or symbolic computation. Although today's AI systems, especially large language models, have surpassed early rule-based systems in power and flexibility, they remain disconnected from the experiential life-world that gives essence to human knowing. Their outputs may be useful, accurate, and even creative in appearance, but they do not emerge from conscious understanding.

Luciano Floridi (2011) proposes a more generous framing through his "philosophy of information." He sees knowledge not just in terms of belief, but as well-formed, meaningful, and truthful data processed

by an agent. From this perspective, AI systems may be seen as *contributors* to knowledge when their outputs are incorporated into human understanding. But the critical point is that AI does not know; it enables or mediates knowing for others. The origin of epistemic agency remains human, even if machines shape the terrain on which we think.

Can AI know?

The question of whether AI can possess knowledge necessitates an examination of the nature of knowing. If knowledge requires consciousness, intentionality, and understanding, then AI, as it currently stands, cannot be said to know in the human sense. AI systems lack subjective experiences and the capacities for self-reflection, which are integral to human knowledge. Even AI attests to this, because of this inquiry, ChatGPT was engaged and asked the question, "Can AI know?" and below was its response;

If by "know" you mean "have knowledge" in the same way humans do, involving awareness, understanding, and experience, then the answer is no, not really. AI like me doesn't have consciousness, subjective experience, or beliefs. I process patterns, retrieve data, and generate responses based on training — but I don't *know* things the way a person does.

But if you mean "know" in a more functional or pragmatic sense, like being able to access and apply information accurately, then **yes**, AI can "know" in that limited way. I can tell you the capital of France, summarize a novel, or analyze a trend, all based on the data I've been trained on or have access to. (ChatGPT)

Philosopher Immanuel Kant's ideas are applicable here. Kant argued that human cognition is not merely the mechanical processing of information but involves self-awareness and the ability to perceive the distinction between knowledge and lived experience (Kant, 1781/1998). Machines, devoid of such faculties, cannot attain genuine knowledge. They can simulate understanding and produce outputs that mimic human responses, but this is not tantamount to true knowledge. Furthermore, the tendency of AI systems to generate credible but incorrect information, known as "hallucinations," draws attention to their lack of understanding. These systems may produce erroneous outputs with high confidence, reflecting their inability to discern truth from falsehood (Marcus & Davis, 2019). This limitation highlights the absence of the critical evaluative processes that characterize human knowing.

Distinction between Human Understanding and Machine Outputs

Human understanding is characterised by intentionality (the capacity to be about something, to hold beliefs, desires, and to engage in conscious thought). In contrast, machines operate devoid of consciousness or intentionality; they manipulate symbols and data without any comprehension, understanding, or experience of the content. John Searle's Chinese Room argument illustrates this distinction effectively. In this thought experiment, a person inside a room follows syntactic rules to manipulate Chinese symbols without understanding their meaning, thereby suggesting comprehension without actual understanding (Searle, 1980). This analogy emphasises that, despite the appearance of understanding, machines merely process inputs to produce outputs without any grasp of meaning.

AI systems are inherently complex, it is hard to understand how they work. Even developers may find it challenging to explain how specific inputs are transformed into outputs, leading to questions about the intelligibility and transparency of machine-generated information (Pasquale, 2015). This lack of transparency is very much unlike human decision-making, which can be explained and understood by individuals. Humans, though imperfect, can articulate the reasoning behind a decision, engage in self-reflection, and even revise their judgments upon

further reflection. In contrast, AI systems offer no such self-explanation; they are bound by the parameters set during training and may generate outputs that appear reasonable without any underlying rational basis. (Burrell, 2016).

What further separates human cognition from machine processing is the fact that humans have internal motivations, emotions, and values that drive their thoughts and actions, whereas machines operate solely on programming and data, lacking personal motivations or feelings. Human understanding is often purpose-driven and permeated with intentionality that reflects personal experiences, cultural backgrounds, and ethical considerations. This motivational drive leads us to ask “why” and “how,” thereby engaging in meta-cognitive reflection. AI, however, has no motivations, desires, or goals of its own. It merely simulates output production based on patterns observed in data. When asked to generate an explanation or to solve a problem, an AI system does not “want” to understand; it is merely executing its coded instructions. Without desire or self-directed curiosity, the rich dimension of human understanding remains out of reach for machine processes.

Another dimension where the distinction becomes crucial is in the realm of error recognition and learning. Humans are not infallible, but the process of recognising errors and adapting based on feedback is an important aspect of our

understanding. When humans make a mistake, they often reflect on the error, learn from it, and change their internal mindsets. In contrast, AI systems require new data or algorithmic adjustments from human engineers to correct errors. Although some machine learning models are designed to adapt through processes like back propagation, this adaptation is not equivalent to human learning. It is purely a matter of optimizing parameters rather than engaging in reflective understanding. This difference states that while AI may adjust its outputs based on training, it does so without any conscious involvement or self-correction that would constitute genuine understanding.

Moreover, the epistemic gap between human understanding and machine outputs is intensified by the manner in which context is processed. For instance, consider that human understanding often involves metaphors, analogies, and contextual cues that enrich our interpretation of information. A poet, for example, can weave layers of meaning into a text that resonate deeply with readers due to shared cultural and emotional contexts. An AI system, however, processes text on a purely syntactic level. While modern natural language processing models have become impressively proficient at generating texts that appear coherent and contextually appropriate, they do so by statistically predicting words based on vast datasets. They do not, however,

engage in the kind of interpretation and creative insight that humans derive from lived experience.

The philosophical implications of this gap have far-reaching consequences for how we conceive of knowledge in the digital age. If knowledge is more than a correct answer or a statistically reliable output, and if it involves a degree of intentionality, embodiment, and social situatedness, then the role of AI must be understood as supportive rather than substitutive. That is, while AI systems can aid in the collection and processing of data, they cannot replace the uniquely human capacity for meaning-making. This understanding aligns with the views of scholars who advocate for a “philosophy of information” that situates AI as a tool that complements human reasoning, rather than as an independent epistemic agent (Floridi, 2011).

Who Decides What Knowledge Is Legitimate?

Customarily, legitimacy in knowledge has been established through rigorous processes that involve peer review, debate among experts, and adherence to established epistemological standards. Human experts, by virtue of their training, experience, and engagement with lived realities, have historically been the arbiters of credible information (Collins & Evans, 2007). However, the integration of AI systems in producing and organizing knowledge has disrupted these traditional processes. The decision as

to what counts as legitimate knowledge is gradually being decentralised from academic and institutional gatekeepers to decentralised, algorithm-driven systems.

The philosophical basis of knowledge legitimacy, as earlier mentioned, can be traced back to classical epistemology, where the idea of justified true belief has been the benchmark for knowledge (Gettier, 1963). In the digital age, however, the criteria for justification have shifted. AI systems can generate information and make predictions based solely on patterns present in big data. This raises the critical question of whether machine-generated outputs meet the criteria for true and justified belief or whether they fall short due to the lack of human contextualisation and understanding.

One key factor in legitimising knowledge is accountability. When knowledge is generated by human experts, accountability mechanisms such as academic debates, public scrutiny, and ethical standards allow society to assess and trust the information provided. In contrast, AI systems may produce outputs with high statistical accuracy without offering clear insights into the reasoning process behind those outputs. As such, the traditional social contract of accountability is challenged when legitimacy is increasingly determined by algorithmic processes (O'Neill, 2002). Deciding legitimacy now involves not only assessing the content of knowledge

but also inspecting the process by which it is generated. Transparent and explainable AI is essential to bridging this gap, as it would allow stakeholders to understand the underlying logic behind AI-generated outputs. Until such transparency is achieved, the question of who decides what knowledge is legitimate remains fraught with uncertainties and power imbalances.

Shifting Trust from Experts to Algorithms

One of the most profound shifts in contemporary society is the reallocation of trust, from human experts, whose expertise is built over years of experience and peer recognition, to algorithms that operate with a veneer of objectivity and infallibility. This shift is both a reflection of technological progress and a response to the limitations of human judgment, which can be biased, inconsistent, or influenced by external factors. Nevertheless, the transition is not without risk.

Trust in algorithms is often predicated on the assumption that they are neutral judges of truth, operating solely on data-derived probabilities. However, the reality is more complex. Algorithms are only as objective as the data and assumptions that underlie them. As Noble (2018) points out, AI systems can sustain and even worsen existing inequalities because they often reflect historical biases and cultural prejudices present in the training data.

Despite this, the trust shift remains, as many people perceive algorithms as free from the subjective errors that can plague human judgment.

The charm of AI lies in its promise of efficiency and consistency. In fields such as medicine, for example, automated diagnostic tools can process large datasets to identify patterns that might not need human experts. Yet this very reliance on efficiency often comes at the expense of accountability. When errors occur, determining responsibility becomes complicated; is it the algorithm, the developer, or the institution that deployed the system? Trust, therefore, becomes distributed among multiple, often doubtful, stakeholders rather than resting solely on the shoulders of trusted experts (Bostrom, 2014).

In addition to reliability, the phenomenon of shifting trust is also linked to cognitive biases in human decision-making. Research in behavioral economics and psychology has revealed that people are prone to over-trusting technology, especially when its outputs seem intuitive or when the decision-making process appears rational (O'Neill, 2002). This technological optimism can lead to a dangerous over-reliance on algorithmic decisions, reducing opportunities for critical thinking and individual judgment. The shift of trust from human expertise to algorithms, therefore, not only transforms the epistemic landscape but also reconfigures power relations within society.

Ultimately, the call for transparency in AI is not merely a technical requirement; it is a fundamental ethical imperative. Floridi (2011) argues that transparency and explainability are essential to ensuring that AI systems contribute meaningfully to collective knowledge without obscuring the reasoning behind their decisions. If the logic behind an algorithmic output cannot be explained, then it becomes challenging to evaluate its validity or to challenge its conclusions. This lack of intelligibility can undermine the legitimacy of AI as a tool for knowledge production and decision-making.

Efforts to enhance transparency often focus on developing “explainable AI” (XAI) methodologies. XAI seeks to demystify the processes within AI systems so that stakeholders, be they developers, policymakers, or end users, can understand, interpret, and even contest the outcomes produced (Doshi-Velez & Kim, 2017). Despite progress in this area, there remains a significant gap between the ideal of complete transparency and the practical realities of complex AI systems. The challenge is to strike a balance between the sophistication of algorithms and the need for clarity in their operation, ensuring that transparency does not compromise performance or operational efficiency. Without robust transparency, there is a risk that AI systems may be used to legitimize decisions that are fundamentally unjust or that perpetuate systemic biases.

Human Dependence on AI for Critical Thinking and Ethical Responsibility

As AI systems become more prevalent, human dependence on these tools for critical thinking and decision-making continues to grow. While AI can support and enhance human reasoning by providing data-driven ideas and answers and reducing cognitive load, this dependence also presents significant challenges. Relying heavily on AI may dull the human capacity for independent thought, leading to a scenario where people become passive consumers of information rather than active participants in knowledge creation (Stiegler, 2010).

Critical thinking is not solely about processing information accurately; it involves questioning assumptions, evaluating evidence, and considering ethical implications. Human expertise is shaped by experiences, cultural narratives, and ethical reflections, qualities that an algorithm cannot replicate. When people delegate critical thinking to AI, they risk losing the ability to check and validate the information presented and to challenge the outputs based on reason and moral judgment. This erosion of critical thinking may have dangerous implications for democracy and public discourse, especially in an era of rapidly evolving technology.

Furthermore, ethical responsibility in the context of AI use is not clearly defined. When an error occurs in

an AI-generated decision, it is often unclear who bears responsibility: the developers who created the model, the institutions that deployed it, or the AI system itself. As responsibility becomes unclear, ethical accountability is weakened, potentially paving the way for unchallenged misuse of AI in ways that damage public trust (Bostrom, 2014). These raise pressing questions about the nature of responsibility in a digital age where decisions are increasingly mediated by machines.

To address these issues, it is crucial to foster a culture of critical engagement with AI. Educators, policymakers, and technologists must work collaboratively to ensure that human judgment remains central in contexts where AI is deployed. This requires not only technical solutions, such as improved transparency and explainability, but also the cultivation of a critical mindset that emphasizes continuous questioning and ethical reflection. It is incumbent upon society to establish regulatory structures and ethical guidelines that hold both human and machine contributors accountable for the decisions made (O'Neil, 2016).

Ethical responsibility in AI systems also demands a reconfiguration of trust. Instead of uncritically transferring trust from experts to algorithms, society must develop hybrid models that leverage the strengths of both human and machine intelligence. This entails recognising that while AI can enhance

cognitive processes by offering rapid data analysis and predictive insights, the value of human creativity, empathy, and ethical deliberation remains irreplaceable. Again, Floridi (2011) advocates for a conception of AI that positions it as a complement to human judgment rather than a substitute. In this context, it is explained as a tool that must operate within clearly defined ethical boundaries.

Fairness and Epistemic Justice in AI Knowledge Production

The production of knowledge using artificial intelligence systems presents significant challenges regarding fairness and epistemic justice. One critical concern is the embedded biases in AI training data. Since these systems learn from large datasets that reflect historical and cultural biases, the knowledge they generate can inadvertently reproduce or amplify these inequities (Noble, 2018). For example, if data predominantly represents one demographic over another, AI outputs may favor dominant narratives while sidelining minority perspectives.

This imbalance raises questions about representation and ownership of AI-generated knowledge. The question of who gets represented in such outputs and who ultimately owns them remains a contentious issue. The data used to train AI systems is often sourced from public or commercial databases that are biased toward the perspectives and experiences of

certain groups (Pasquale, 2015). Consequently, the resulting knowledge may not reflect the diversity of human experience and knowledge, thereby limiting the scope of what is recognized as valid or important in public discourse. This monopolisation of representation can contribute to a cycle where dominant narratives are continually reinforced, while marginalised voices remain excluded.

Moreover, this dynamic carries significant risks. When AI-generated knowledge reinforces dominant perspectives, it not only sustains existing power structures but also further marginalizes those already underrepresented. This form of epistemic injustice, where individuals or groups are unfairly deprived of the opportunity to contribute to or be recognized within the broader body of knowledge, downplays democratic ideals and ethical standards in society (Fricker, 2007). The reinforcement of dominant narratives limits the potential for critical debate and hinders the evolution of a more inclusive and varied understanding of complex issues.

Addressing these challenges requires a layered approach. Researchers and developers must strive for greater diversity in training datasets and incorporate ethical frameworks into AI design and implementation. Ensuring transparency in how data is selected and how models are built is crucial for mitigating bias and promoting fairness. Ultimately, the goal should be to create AI systems that

contribute to a more representative and just epistemic landscape. (Bouhouita-Guermech et al., 2023)

The Future of Human Inquiry: Navigating AI's Role in Intellectual Pursuits

In the dawning era of artificial intelligence, the basis of human inquiry undergoes a transformative shift. Traditional modes of questioning, learning, and discovery are increasingly interwoven with machine-generated data, algorithms, and automated systems that promise to boost our capabilities. Yet, as we integrate these technological advances into our intellectual lives, it is essential to ask not only what AI can do for us, but also how it changes the very nature of inquiry. In seeking to understand and navigate AI's role, we are compelled to balance optimism with vigilance, ensuring that our engagement with technology remains both ethically sound and deeply human.

At its core, human inquiry has long been defined by our capacity for curiosity, reflection, and critical judgment. Whether in the pursuit of scientific discovery or artistic innovation, the process of inquiry is driven by the distinctive abilities of human consciousness: intentionality, context, and the ability to derive meaning. When we say that AI is taking on a role in intellectual pursuits, we should first recognise that, as we have earlier, these systems

operate very differently from human minds. And while these systems can simulate certain aspects of reasoning and generate impressive outputs, they do so without the consciousness or understanding that characterises human thought. Nonetheless, the power of AI lies not in its claim to genuine understanding but in its capacity to handle complex tasks that would overwhelm the human brain. (Schlagwein & Willcocks 2023) In research, for example, AI systems can analyse large amounts of data, identify patterns invisible to the naked eye, and even suggest hypotheses that prompt further inquiry by human experts. In the educational sphere, AI offers personalised learning experiences, adapting to individual student needs and learning styles. AI-driven platforms can analyse student performance data to identify strengths and weaknesses, enabling tailored instructional strategies. In scientific research, AI accelerates data analysis, identifies patterns, and predicts outcomes, thereby expediting the discovery process. (Hassan2024) For instance, AI has been instrumental in predicting protein structures, a task that traditionally required extensive time and resources. This capability significantly speeds up drug discovery and the development of new treatments. This collaboration between human and machine can lead to remarkable breakthroughs, accelerating the pace of discovery and expanding our intellectual horizons. By freeing ourselves from the toil of data processing, we are afforded the opportunity to focus on higher-order thinking and

creative pursuits. In this light, AI is not a replacement for human inquiry but a tool that complements and extends our capabilities. (Berber, 2024)

Furthermore, this collaboration raises the question: How do we maintain the human element in inquiry when AI takes on a larger role? The answer may lie in developing what can be termed “AI literacy,” an understanding of both the capabilities and the limitations of these systems. By becoming conversant with the principles of machine learning, the structure of algorithms, and the potential for bias, scholars and practitioners can critically assess AI outputs and incorporate them into a broader, human-centred structure of inquiry. In doing so, we reclaim the critical, evaluative process that defines human knowledge, ensuring that AI remains a tool rather than an inscrutable authority.

Ultimately, the integration of AI into human intellectual pursuits invites us to reflect on the nature and purpose of inquiry itself. Throughout history, human inquiry has been a dynamic process of questioning and reimagining the world. The advent of AI does not mark the end of this tradition; rather, it provides new tools and new challenges that can enrich our ongoing quest for knowledge. The task before us, therefore, is to ensure that as we embrace these tools, we remain committed to the values that have long defined meaningful inquiry: curiosity, critical reflection, and ethical responsibility.

In sum, AI has the potential to be an extraordinary partner in intellectual pursuits, augmenting our capabilities and expanding the horizons of what we can achieve. However, this potential comes with challenges that demand careful consideration. The lack of transparency of machine learning systems, the need for strong ethical frameworks, and the enduring importance of human consciousness all highlight that true knowledge remains a deeply human endeavour. To navigate this evolving landscape, we must cultivate AI literacy, encourage interdisciplinary collaboration, and steadfastly commit to the reflective process that defines our pursuit of understanding. By doing so, we not only preserve the essence of human inquiry but also harness the strengths of AI in a manner that is both productive and profoundly humane.

Conclusion

AI's integration into human intellectual pursuits presents both opportunities and challenges. Despite AI's capabilities, human engagement remains crucial in knowledge-making. Humans provide ethical guidance, contextual understanding, and emotional intelligence, qualities that AI lacks. Incorporating human judgment ensures that AI-generated knowledge aligns with societal values and ethical standards. Moreover, human involvement in interpreting AI outputs prevents the perpetuation of biases and ensures that diverse perspectives are

considered. This collaborative approach upholds the ethical value of human engagement in the knowledge-making process.

References

- Bærøe, K., Sandvik, H., & Solberg, E. (2020). Human intervention and monitoring in AI: Safeguards, data curation, and ethical instruction. *AI & Society*, 35(3), 567-579.
- Berber, A. (2024). *Shannon Vallor: THE AI Mirror: How to Reclaim Our Humanity in an Age of Machine Thinking*. Oxford: Oxford University Press, 2024. Paperback (ISBN 978-0-19-775906-6),
<https://doi.org/10.1007/s10677-024-10474-9>
- Bostrom, N. (2014). *Superintelligence: Paths, dangers, strategies*. Oxford University Press.
- Bouhouita-Guermech, S., Gogognon, P., & Bélisle-Pipon, J.-C. (2023). Specific challenges posed by artificial intelligence in research ethics. *Frontiers in Artificial Intelligence*, 6.
- Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society*, 3(1), 1–12.
<https://doi.org/10.1177/2053951715622512>
- Campbell, L., & Waddington, R. (2024). Reassessing academic integrity in the age of AI: A systematic review. *Higher Education Research & Development*, 43(1), 1-15.

- Collins, H., & Evans, R. (2007). *Rethinking expertise*. University of Chicago Press.
- Crawford, K. (2021). *Atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. Yale University Press.
- Dreyfus, H. L. (1992). *What computers still can't do: A critique of artificial reason*. MIT Press.
- Fricker, M. (2007). *Epistemic injustice: Power and the ethics of knowing*. Oxford University Press.
- Floridi, L. (2011). *The philosophy of information*. Oxford University Press.
- Foucault, M. (1980). *Power/Knowledge: Selected interviews and other writings 1972–1977* (C. Gordon, Ed.). Pantheon Books.
- Fügener, A., Müller, O., & Bärøe, K. (2021). Instructing AI systems to discern beneficial from harmful information: Ethical considerations. *AI & Ethics*, 1(2), 123–135.
- Gettier, E. L. (1963). Is justified true belief knowledge? *Analysis*, 23(6), 121–123.
- Hassan, S. A., Omar, A. I., & Ahmed, N. R. (2024). Exploring the Ethical Implications of AI in Public Health Research: A Comprehensive Analysis. *South Eastern European Journal of Public Health*, 108–115.
- Kant, I. (1998). *Critique of Pure Reason* (P. Guyer & A. W. Wood, Eds. & Trans.). Cambridge University Press. (Original work published 1781)
- Kanu, A. I., Addidi, T. & Kanu, C. C., “Artificial intelligence and cybercrime in Nigeria:

- Towards an Ethical framework". *Dialogue and Universalism*. 32(1) 143-158, 2024.
- Kanu, I. A. (2020). "Igwebuiké philosophy of science and technology". *Nnadiébué Journal of Education*. Vol. 5. No. 4, pp. 64-74.
- Marcus, G., & Davis, E. (2019). *Rebooting AI: Building Artificial Intelligence We Can Trust*. Pantheon Books.
- Marrone, M., et al. (2022). How does generative artificial intelligence impact student creativity? *Computers in Human Behavior Reports*, 6, 100-110.
- Muller, O., Bärøe, K., & Sandvik, H. (2019). Curating data for AI: Ethical implications. *Data & Society*, 6(4), 345-360.
- Mateja, M., & Heinzl, C. (2021). Extending human creativity with AI: Bridging the conceptual divide. *AI Open*, 2, 1-10.
- Nagel, T. (1974). What is it like to be a bat? *The Philosophical Review*, 83(4), 435-450.
- Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. New York University Press.
- O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown Publishing Group.
- Pasquale, F. (2015). *The black box society: The secret algorithms that control money and information*. Harvard University Press.
- Schlagwein, D., & Willcocks, L. (2023). 'ChatGPT et al.': The ethics of using (generative) artificial

- intelligence in research and science. *Journal of Information Technology*.
- Schmidt, E. (2022). *We're Not Using AI to Its Fullest Human Potential*.
- Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(3), 417–457.
- Stiegler, B. (2010). *Taking care of youth and the generations*. Stanford University Press.
- Vinchon, M., et al. (2023). AI creativity and the human-AI co-creation model. In *Advances in Intelligent Systems and Computing* (Vol. 1234, pp. 123-134).