https://rijournals.com/research-in-education/

RESEARCH INVENTION JOURNAL OF RESEARCH IN EDUCATION 5(2):65-71, 2025

©RIJRE Publications

ONLINE ISSN: 1115-9014

PRINT ISSN: 1597-8656

https://doi.org/10.59298/RIJRE/2025/526571

The Ethics of AI: Philosophical Perspectives

Kakembo Aisha Annet

Faculty of Education, Kampala International University, Uganda

ABSTRACT

As artificial intelligence (AI) systems become increasingly integrated into human affairs, the ethical frameworks guiding their development and deployment demand critical reassessment. This paper examines the intersection of AI and philosophy, challenging reductive "common-sense ethics" by engaging deeply with foundational ethical theories. Tracing the historical development of AI ethics, it examines how various philosophical approaches utilitarianism, deontology, virtue ethics, and social contract theory inform current debates on machine morality, algorithmic responsibility, and human-AI relations. The paper also scrutinizes issues of bias, surveillance, and fairness, arguing for ethically robust AI governance mechanisms that are context-sensitive and philosophically grounded. Drawing from both classic moral philosophy and contemporary applications, this paper seeks to move beyond superficial ethical codes toward a deeper understanding of the values that should guide AI development in increasingly autonomous and high-stakes environments.

Keywords: Artificial Intelligence, AI Ethics, Philosophy of Technology, Utilitarianism, Deontological Ethics, Virtue Ethics.

INTRODUCTION

AI systems present ethical challenges that are often oversimplified by promoting a common-sense ethical AI approach. This perspective neglects the deeper philosophical issues involved as AI technology becomes more prevalent. The ethics of AI is not merely about individual actions or decisions guided by ethical principles; it encompasses fundamental dilemmas tied to the technology's development and production, often hidden by reductive viewpoints. Key figures in philosophy, ethics, and AI recognize significant developments and misinterpretations related to AI. The potential risks of superintelligent, weakly autonomous systems highlight an existential threat if developed without proper safeguards. Meanwhile, contemporary weak artificial general intelligence (AGI) already poses immediate ethical concerns, influencing various sectors like journalism and finance. It is crucial to establish regulatory mechanisms, methods for auditability, and emergency "kill switches" for such AI systems. The "computational blind spot," where decisions occur beyond conscious oversight, is critical for social media platforms using machine learning algorithms [1, 2].

Historical Context of AI Ethics

Emerging from applications in robotics, control theory, and simulation of natural life systems, the scientific discipline of artificial intelligence (AI) arose in the mid-1950s. In symposiums where researchers typically introduced their work in the terms of their disciplines and applied it to problems within them, the Dartmouth workshop in 1956 was unique in that pioneering areas were outlined and a bold attempt was made to tackle a problem shared by many disciplines: the development of a "thinking machine." The workshop was historic; it was the birthplace of AI. The number of AI peer-reviewed research articles per year increased tenfold from 1981 to 1990. From 2000 to 2009, the average revenue per year of the AI firms reached a peak. The 2010s saw explosive developments in applied domains, such as speech recognition, language processing, and computer vision. A direct outcome of those developments is rising concerns over the ethical implications of those publications. On the one hand, while AI ethical implications can be traced back at least to 1985, it only recently became a hotly debated issue. On the

other hand, the influx of AI ethics papers does not mean that researchers agree on a set of hot topics to work on. Instead, there is a diversity of opinions on what ethical implications are important and should be addressed. There are a few notable reviews that catalogue AI ethics papers, primarily from the perspective of existing literature. However, a more elementary question, such as the historical divergence of AI ethics papers, is worthy of being answered; hence, understanding the trajectory of AI ethics development warrants research. To that end, a bibliometric analysis of literature that publishes AI ethics papers is conducted. The three-phased development suggests that AI ethics was in its incubation phase from 2000 to 2009. With the advent of the AI narrative, the focus switched to making AI human-like machines from 2010 to 2019. The current focus of research, making AI human-centric machines, is set for ongoing investigations [3, 4].

Philosophical Foundations of Ethics

Ethics of Artificial Intelligence and Robotics is a substantial domain for the younger branch of applied ethics, whose focus is on the moral implications, effects, and rules for robots and artificial intelligence (AI). Among the most prominent concerns is whether AI ought to be socially, morally, or legally regarded as autonomous agents with rights or culpability. This question is intimately connected with ensuring that they act in accord with their end-users' ethical standards, either through (a) choosing to comply or by (b) reprogramming them. Related and equally consequential questions are those about the rights of AI in virtue of its capabilities or status as sentient beings. It is not clear that a social relation, rights, or responsibilities follow from which matters concerning humanity can be ignored. The (also) substantial effect of AI on in-person human relations is raising concerns about inner-communication norms and bystanders' rights. Mass data mining of social networks without consent raises issues with privacy, patenting, deceit, erosion of knowledge, or (de facto) whole-group manipulation. Coding the social network feeds or consequently choosing the sources regarding a region raises questions about censorship. Legitimate and illegitimate uses of AI's operational area might breed unintended security holes hampering habits with shrinking human input in decision loops or intelligent self-growth, let alone war machines programming autonomous methods of choosing victims or killing with increasing success rates, risking areas of deontological ethics. Finally, concerning revenue, AI issues access—marketed versus source code, and survival—project-increasing growth versus no-product death areas closely leading to AI's rights discussion [5, 6].

Utilitarianism and AI

There is a naïve assumption inherent in most discussions of artificial intelligence (AI): We want it to be beneficial. We want it to help, not hinder, humans. But what exactly does it mean for AI to be "beneficial," and how can one develop AI systems that are harmless, helpful, and honest? We cannot hope for a formal specification of "beneficial" that does not ultimately rely on human intuition, judgment, and insight. However, it is a question worth exploring, and there is a growing body of literature on the topic. Of course, the question must be fundamentally answered before any philosophically adequate system can be implemented. This question is not so much methodological but rather metaphysical: Why do we want our machines to be beneficial? What is the motivation behind it? Why not simply Maximal Cognition or Maximal Intelligence? Such questions have been addressed only at a surface level; i.e., in terms of guidelines and programming ethics, such as the off-switch problem or simple command not to harm humans. The beneficial AI movement may not create "beneficial AI" if the ethical assumptions motivating it are not made explicit. As mentioned earlier, reinforcement learning (RL) will likely be an important technique in the desired area of control over advanced AI. It is a form of learning by trial and error that is thought to yield capabilities and robustness that are virtually impossible to write down explicitly. A consequence of this learning technique is that the AI will ultimately learn to reason about itself and other agents, and that it will reason according to an entirely different morality than the one presumed by the designers. Also, it is not unreasonable to believe that a method that works in the RL setting will not work in the more general and unstructured setting. Surprisingly little has been said regarding this aspect of RL and how it bears on the ethics of AI and the philosophy of mind [7, 8].

Deontological Ethics and AI

Deontological ethics, or duty-based ethics, prioritize following an established standard of right behavior, emphasizing that moral obligations arise from greater moral principles, irrespective of consequences. In the context of action-making AIs, this entails two main components: a ruleset outlining duties for specific situations and a method to evaluate AI reasoning against these rules. Implementing this requires translating an ethical theory into logical axioms, often incorporating a first-order logical proof system in the AI. While the reasoning process concerning consequences is generally manageable, validating AI reasoning remains a significant, albeit surmountable, challenge. Established AI frameworks can assist in

normative reasoning and basic rules. Consequently, AI ethics can be managed within a similar timeframe as coherent action, but with distinct methods and thought processes. Beyond tractability, there are concerns regarding the accurate representation of duty within the systems. Each system's design influences its reasoning capabilities, potentially leading to outcomes deemed unacceptable by its creators. To mitigate this, personalizing normative reasoning could introduce varied moral perspectives, fostering skepticism and ethical variability. However, action-making AIs can develop a belief-desire-goal framework with tractable normative reasoning that aligns with broadly accepted values, suggesting an architecture that integrates a self-aware rule model alongside a societal ethical context, reflecting accepted norms [9, 10].

Virtue Ethics in AI Development

As AI development becomes mainstream, various thinkers discuss its ethical implications, emphasizing that those involved should ensure AI benefits humanity. This discussion includes principles like 'non-maleficence', 'fairness', and 'explainability'. To achieve ethical AI, it's essential to integrate second-order virtues like courage and vigilance alongside first-order principles. Recently, awareness of AI ethics has increased, highlighted by the European Commission's 2019 guidelines for AI in the EU. Numerous global initiatives also tackle AI ethics, seeking to summarize the pitfalls of AI applications. Crucial principles include non-maleficence, beneficence, fairness, privacy, transparency, accountability, and responsibility. However, current AI ethics guidelines face criticism for their limitations. Primarily, they often reflect existing practices since AI development is already assessed for legal compliance, evident in fairness regulations in finance under laws like the General Equal Treatment Act. Moreover, some ethics initiatives merely serve to justify the current state of technology, acting as PR tactics that mask the surveillance economy's reality, while promoting benign uses like improved healthcare through AI [11, 12].

Social Contract Theory and AI

Modern AI systems are trained via processes that can be seen as variations of reinforcement learning. A computer program has some internal state that it changes to come to conclusions, take actions, and predict outcomes to problems it faces in the future. A traditional engineering system is designed, and its operation is approved by a human engineer. AI systems, on the other hand, learn the meaning of the inputs and the actions they (and a human) can take at times through trial-and-error exploration. Some of the inputs to base these decisions on, as well as some of the values held, are still produced by humans. In high-stakes and/or uncontrolled environments, it is important to help humans understand how and why an AI-produced conclusion is arrived at. Trusted AI systems (as social agents) should help humans infer them using methods available to themselves. For implemented AI systems, it would be ideal that a third party can independently estimate their competency assurances, similar to how a human expert can be trusted if certified by abiding by high medical standards. Logically valid explanations can be mathematically simplified by showing that the linearly independent parents of a node of a directed acyclic graphical model are the same in the explanations as in the model. Technically, the use of proof assistants and high-assurance hardware can ensure implementation integrity [13, 14].

AI and Moral Responsibility

Recent advances in artificial intelligence (AI) present exciting possibilities alongside serious challenges. AI can now perform tasks previously deemed improbable, such as developing medications, driving cars, and creating intricate textual and audiovisual works. The broad implementation of AI raises urgent ethical considerations regarding responsible AI as systems operate without human input. While focus has historically been on beneficial applications, there is growing concern over AI's use in high-stakes environments, including defense, behavior analysis online, and military command. Such deployment significantly heightens the need to define what responsible AI entails, particularly concerning moral responsibility in human-AI collaborations. Determining how to attribute moral responsibility in these contexts, especially in defense and policing, is an important ethical issue, as these teams impact society in more consequential ways. Although there has been some progress in developing guidelines for responsible AI, much remains unaddressed, particularly about moral responsibility assignment in human-AI teams. Existing literature on social agents could provide insights, yet there has been little collaboration between these fields. This gap highlights a crucial area for further research, ensuring that as AI becomes more integrated into society, its ethical implications and responsibilities are thoroughly examined and addressed [15, 16].

Bias and Fairness in AI

This article addresses bias and fairness in AI by providing a comprehensive analysis of the concepts of fairness, equity, and discrimination in ethical and human rights contexts. It argues that the disconnect

between ethics, human rights standards, and legal anti-discrimination statutes means that legal questions about algorithmic discrimination are too narrow and leave unconsidered important and far-reaching ethical questions. AI Fairness should distinguish between these normative domains and the kinds of fairness concerns that arise in them. Case studies about predictive policing/data justice, fair hiring, and AI-driven job market tools are discussed. They address how ethical and regulatory frameworks differ and the implications both for practice and for further regulatory developments. The concept of fairness is taken to mean fair treatment of generative AI systems (both in terms of outward manifestations and treatment) as it relates to the fairness of socio-technical systems in which they are embedded, their biases, discrimination, and resource allocation effects. The concept of fairness will be distinguished into three basic categories: fairness in its normative, ethical sense in regards to AI systems; fairness in its human rights sense, concerned with algorithmic discrimination, non-harm, and equality; and fairness in its legal, statutory sense, governed by anti-discrimination laws. Fairness in these three domains diverges in terms of both scope and standards, and also responds to a multiplicity of contexts and considerations that are often different and arguably not reconcilable. Addressing the ethical issues around algorithmic bias without regard for the already existing human rights and legal frameworks will result in blindspots, and conversely, treating algorithmic discrimination as an issue of lower (or possibly zero) significance when compared against broader ethical questions will miss important and far-reaching implications for the use and further adoption of these technologies [17, 18].

Privacy and Surveillance Concerns

By altering data collection processes, repurposing data, aggregating it, exploiting noise in the data pipelines, adding noise to the data, or implementing predefined data usage and sharing protocols, privacy-intrusive behaviour may be avoided. It can be anticipated that observing and evaluating these behaviour modifications over time will lead not only to a better understanding of the motivations driving them but also to better means of discouraging them altogether. In a similar vein, the normative evaluation of AI practice en masse may indeed be beneficial. Philosophers are encouraged to embrace novel interpretive and methodological innovations to better understand the present cultural-historical manifestation of AI. Since privacy-intrusive behaviour and the justifications rooted in potential AI harms may vary in significance and narrative across time, space, and context, richer normativities and modes of governance could be explored. The AI industry landscape offers numerous compelling case studies for richly accumulating ethnographic material for interpretation. Action research should be prioritised, as it can yield vivid and multilayered empirical knowledge while also having a transformative impact on social reality. Ideal types of AI ethically conducive cultural formations, institutions, or processes within the AI landscape can be theorised and compared inductively with particular cases. Some ideals may consist of means of deliberation and consensus-building, such as certain configurations of interest mediation or procedures for ethical deliberation, while others might spotlight the visions or incentives that agents in the AI sector are expected to make a meaningful commitment to. Normative visions, interdisciplinary cooperation, macro-institutional structures, plurality of mediators and winning narratives for AI use, and configuration of audiences and power relations around AI are just some of the themes that may help expand the ideal types. Agencies within the AI industry should also enact applied philosophy, as AI could and should be made in ways that correspond to the ideals of a pluralistic, collaborative, and constructive public culture of use. Compliance with the recent legislation proposals made by the EU and the UK could serve as baseline steering principles [19, 20].

The Role of AI in Society

Since COVID-19, global digital connectivity has surged, with remote work, e-learning, online meetings, video conferencing, e-shopping, and e-banking/gaming thriving. Interactions now largely occur through major tech platforms. The pandemic spurred advancements in robotics, marking Japan's inaugural use of typhoon rescue robots and multiple countries' trials for drone delivery. Robotic deliveries were piloted in universities and hospitals, and humanoid robots were tested in customer service and elder care. AI emotion recognition systems have been trialed in hospitals and schools, and the development of self-driving cars has accelerated. AI systems also manage mask detection and body temperature monitoring through advanced surveillance technology. Although certain authorities are quickly formulating AI regulations, only a handful of countries have proposed new laws. AI poses threats to society in economic, political, and psychological domains, with some nations providing nearly hourly updates on calamities, disasters, and conflicts. This analysis often leads to the creation of memes and politically charged messages that amplify societal expectations. Social behaviors can be predicted through social media, potentially serving as early warning systems for unrest in public environments. In response to the Social 5.0 era, new leadership models and stringent regulations for monopolistic tech firms are essential. The

EU has established seven requirements for trustworthy AI, emphasizing that it must avoid harm, be unbiased, respect autonomy, protect privacy, remain transparent, and enhance human capabilities. Ultimately, the risks associated with AI must remain within acceptable limits [21, 22].

Global Perspectives on AI Ethics

Recent efforts to tackle ethical concerns about AI and ML have generated a vast array of documents proposing ethical guidelines and policy frameworks across various levels. The production of these documents has surged, reflecting the urgency in addressing ethical governance of AI and ML systems and applications. This proliferation has led to a complex landscape of emerging AI policy frameworks alongside numerous ethical guidelines, characterized by diverse styles and themes. Annotating this extensive collection presents a significant challenge, similar to the "London bus" problem, where diversity can complicate understanding. During the COVID-19 pandemic, many governments quickly formed expert advisory bodies on the ethical use of AI and ML in combating the virus, highlighting the sudden rise in interest in applied AI ethics, evident in the numerous publications that followed. It is crucial to recognize pre-existing case studies on algorithmic justice and AI ethics from before the pandemic, albeit in a more protective context. While AI ethics put forward by the global north often claim universality, they frequently neglect local conditions where these ethics are applied. For AI to be ethical, it must be rooted in local contexts that reflect social, political, and cultural differences not found in the global north. Navigating ethical debates across countries is challenging, and there is growing concern that a market-driven approach to AI ethics and fairness fails to resonate in non-Western nations. Efforts to create datasets for auditing fairness in non-Western languages need to address how fairness is understood in context, considering trust, relationships, and power dynamics. Before addressing potential biases or unfair results, it is essential to define and measure fairness for non-Western languages. Establishing locally grounded AI ethics is a necessary first step for developing a globally relevant social technology aimed at justice [23, 24].

Future Challenges in AI Ethics

This section addresses challenges in AI ethics, focusing on future collaboration within the AI ethics community. It explores whether ethical AI can be achieved, suggesting that moving beyond principles-based AI ethics is necessary to tackle new AI issues. Emphasizing the need for principled solutions to ethical challenges posed by AI, there's an urgent demand for ethical oversight due to the sociotechnical problems arising from current AI systems. However, ethical deliberation often occurs privately, creating a disconnect with external audiences. The AI ethics community produces varying recommendations, with high-level principles addressing new issues, yet these principles alone may not guarantee ethical AI systems. The diversity and rapid evolution of AI technologies present significant challenges, raising concerns over data manipulation, privacy breaches, discrimination, and censorship. Understanding this complexity is vital for responsibly managing AI technologies. Ethical frameworks provide general priorities, but deeper case study approaches are essential to navigate the intricacies of existing AI systems. Fresh case studies and simple ethical discussions are crucial, as current initial studies typically emanate from close collaboration between AI developers and ethicists [25, 26].

Case Studies in AI Ethics

AI has powerful capabilities to capture, analyze, and use vast amounts of information about humans and human behavior. It also offers incredible opportunities for businesses to gather and apply information about individuals. Although AI systems do not have goals, desires, awareness, or agency, they can help humans pursue their goals. Nevertheless, there is a wide diversity of human goals, not all of which are benign; some are evil, and some may even be inadvertently consequentialist. For example, AI systems that analyze trends in internet search and browsing activity are being used to diagnose depression and related disorders on a large scale. On the one hand, this is benevolent and may save thousands of lives. On the other hand, it may also be used by governments to detect dissent and opposition to their rule. Such goals were not imagined by the designers of the analysis systems and were unintended by advertisers, but they follow from the use of the systems. Intelligence, tools that pursue instrumental goals, can be adapted to diverse ends, not all of which are benign. This is especially pronounced in the case of the decision-making AI (AI DM) systems capable of distributing resources, censoring information, and determining access and privilege. The ability to build such AI DM systems, or even more narrowly specialized instances, is now being developed by universities, research institutions, and even companies outside the world's tech centers. AI ethics is to reflect on the ethical implications of AI-powered applications, such as the use of a Google Street View car outside one's property, and to discuss how to mitigate potential hazards, especially when the consequences are unintended. The goal of AI ethics is to build ethically acceptable job application screening systems, face recognition systems for video surveillance, and systems

for deployment on fighter jets. A different kind of ethics looks at the very development of AI systems capable of learning and acting in the world, addressable by algorithmic means and software tools [27, 28].

The Role of Stakeholders in AI Ethics

The responsibility for the ethical use of AI is politically and philosophically complex. Contractors implementing AI might argue they are simply using a provided tool without knowledge of its coding or data parameters. Conversely, software companies might claim that responsibility lies solely with top management, who direct company policies. The design and deployment of machine learning systems present accountability issues on both technical and ethical fronts. While non-technical ethical concerns like rights violations are debatable, technical challenges related to agency, reliability, and unintended harm from AI use are significant. Stakeholders in AI encompass creators, users, and affected individuals. Corporations that develop AI systems often evade accountability for their applications, citing the lack of strategic diversity considerations and the misuse of AI by malicious actors as reasons for this detachment. The influence of AI architecture on behavior complicates traditional notions of moral responsibility, a subject not sufficiently explored by philosophers. Many scholars pause at attributing moral agency to algorithms, making it essential to avoid oversimplification. Yet, the pressing issues surrounding accountability necessitate rigorous examination of legal frameworks and ethical norms. Dismissing these concerns in favor of organizational accountability may overlook the nuanced responsibilities of those who embed and deploy these technologies [29, 30].

CONCLUSION

The ethical challenges posed by artificial intelligence cannot be sufficiently addressed through superficial or narrowly legalistic frameworks. As AI technologies evolve, a deeper philosophical engagement is necessary to anticipate and shape the implications of their integration into society. Utilitarian, deontological, and virtue-based approaches offer distinct lenses through which to evaluate AI behavior and its impact, while social contract theory highlights the need for trust and accountability between AI developers and the societies they serve. The risk of bias, loss of privacy, and moral ambiguity in autonomous systems underscores the urgency of developing interdisciplinary ethical standards that are both practically enforceable and philosophically coherent. By fostering dialogue across ethics, technology, and policy, we can guide AI toward a future aligned with human dignity, fairness, and collective well-being.

REFERENCES

1. Müller VC. Ethics of artificial intelligence 1. InThe Routledge social science handbook of AI 2021 Jul 12 (pp. 122-137). Routledge.
2. Mittelstadt B. Principles alone cannot guarantee ethical AI. Nature machine intelligence. 2019 Nov;1(11):501-7.
3. van Assen M, Muscogiuri E, Tessarin G, De Cecco CN. Artificial intelligence: A century-old story. InArtificial Intelligence in Cardiothoracic Imaging 2022 Apr 22 (pp. 3-13). Cham: Springer International Publishing. [HTML]
4. Krauss P. What is Artificial Intelligence?. InArtificial Intelligence and Brain Research: Neural Networks, Deep Learning and the Future of Cognition 2024 Jul 13 (pp. 107-112). Berlin, Heidelberg: Springer Berlin Heidelberg. [HTML]
5. Candrian C, Scherer A. Rise of the machines: Delegating decisions to autonomous AI. Computers in Human Behavior. 2022 Sep 1;134:107308.
6. Formosa P. Robot autonomy vs. human autonomy: social robots, artificial intelligence (AI), and the nature of autonomy. Minds and Machines. 2021 Dec;31(4):595-616.
7. Stahl BC, Andreou A, Brey P, Hatzakis T, Kirichenko A, Macnish K, Shaelou SL, Patel A, Ryan M, Wright D. Artificial intelligence for human flourishing–Beyond principles for machine learning. Journal of Business Research. 2021 Jan 1;124:374-88.
8. Kühl N, Schemmer M, Goutier M, Satzger G. Artificial intelligence and machine learning. Electronic Markets. 2022 Dec;32(4):2235-44.
9. Dilbar K. The Principles of Deontological Ethics. Web of Teachers: Inderscience Research. 2024 May 8;2(5):56-9.
10. Van Slyke JA, Peterson G, Brown WS, Reimer KS, Spezio M, editors. Theology and the science of moral action: Virtue ethics, exemplarity, and cognitive neuroscience. Routledge; 2012 Nov 12.
11. Pasham SD. A Review of the Literature on the Subject of Ethical and Risk Considerations in the Context of Fast AI Development. International Journal of Modern Computing. 2022 Dec 16;5(1):24-43.

12. Rakowski R, Polak P, Kowalikova P. Ethical aspects of the impact of AI: the status of humans in the era of artificial intelligence. Society. 2021 Jun;58(3):196-203.
13. Pryzant R, Iter D, Li J, Lee YT, Zhu C, Zeng M. Automatic prompt optimization with" gradient descent" and beam search. arXiv preprint arXiv:2305.03495. 2023 May 4.
14. He H, Shuang E, Ai L, Wang X, Yao J, He C, Cheng B. Exploiting machine learning for controlled synthesis of carbon dots-based corrosion inhibitors. Journal of Cleaner Production. 2023 Sep 20;419:138210. [HTML]
15. Gianni R, Lehtinen S, Nieminen M. Governance of responsible AI: From ethical guidelines to cooperative policies. Frontiers in Computer Science. 2022 May 24;4:873437.
16. Trocin C, Mikalef P, Papamitsiou Z, Conboy K. Responsible AI for digital health: a synthesis and a research agenda. Information Systems Frontiers. 2023 Dec;25(6):2139-57. springer.com
17. Wang Y, Ma W, Zhang M, Liu Y, Ma S. A survey on the fairness of recommender systems. ACM Transactions on Information Systems. 2023 Feb 7;41(3):1-43. acm.org
18. Teodorescu MH, Morse L, Awwad Y, Kane GC. Failures of Fairness in Automation Require a Deeper Understanding of Human-ML Augmentation. MIS quarterly. 2021 Sep 1;45(3).
19. Schilling-Vacaflor A, Lenschow A. Hardening foreign corporate accountability through mandatory due diligence in the European Union? New trends and persisting challenges. Regulation & Governance. 2023 Jul;17(3):677-93.
20. Almeida D, Shmarko K, Lomas E. The ethics of facial recognition technologies, surveillance, and accountability in an age of artificial intelligence: a comparative analysis of US, EU, and UK regulatory frameworks. AI and Ethics. 2022 Aug;2(3):377-87.
21. Zicari RV, Brodersen J, Brusseau J, Düdder B, Eichhorn T, Ivanov T, Kararigas G, Kringen P, McCullough M, Möslein F, Mushtaq N. Z-Inspection®: a process to assess trustworthy AI. IEEE Transactions on Technology and Society. 2021 Mar 17;2(2):83-97. ieee.org
22. Li B, Qi P, Liu B, Di S, Liu J, Pei J, Yi J, Zhou B. Trustworthy AI: From principles to practices. ACM Computing Surveys. 2023 Jan 13;55(9):1-46. acm.org
23. Van Norren DE. The ethics of artificial intelligence, UNESCO and the African Ubuntu perspective. Journal of Information, Communication and Ethics in Society. 2023 Jan 31;21(1):112-28. [HTML]
24. Gwagwa A, Kazim E, Hilliard A. The role of the African value of Ubuntu in global AI inclusion discourse: A normative ethics perspective. Patterns. 2022 Apr 8;3(4).
25. Scheufele DA, Krause NM, Freiling I, Brossard D. What we know about effective public engagement on CRISPR and beyond. Proceedings of the National Academy of Sciences. 2021 Jun 1;118(22):e2004835117. pnas.org
26. Wolrath Söderberg M, Wormbs N. Internal deliberation defending climate-harmful behavior. Argumentation. 2022 Jun;36(2):203-28.
27. Gupta A, Raj A, Puri M, Gangrade J. Ethical Considerations in the Deployment of AI. Tuijin Jishu/Journal of Propulsion Technology. 2024;45(2):1001-4055. researchgate.net
28. Trimmel KE, Kriechbaum M, Lazou R, Brudermann T. Between distributive and procedural justice claims: Reframing patterns of discursive resistance against climate action. Energy Research & Social Science. 2024 Mar 1;109:103424.
29. Radanliev P, Santos O, Brandon-Jones A, Joinson A. Ethics and responsible AI deployment. Frontiers in Artificial Intelligence. 2024 Mar 27;7:1377011. frontiersin.org
30. Tilala MH, Chenchala PK, Choppadandi A, Kaur J, Naguri S, Saoji R, Devaguptapu B, Tilala M. Ethical considerations in the use of artificial intelligence and machine learning in health care: a comprehensive review. Cureus. 2024 Jun 15;16(6). cureus.com

CITE AS: Kakembo Aisha Annet (2025). The Ethics of AI: Philosophical Perspectives. RESEARCH INVENTION JOURNAL OF RESEARCH IN EDUCATION 5(2):65-71. https://doi.org/10.59298/RIJRE/2025/526571