Journal of Posthumanism
2025
Volume: 5, No: 5, pp. 204–216
ISSN: 2634-3576 (Print) | ISSN 2634-3584 (Online)
posthumanism.co.uk

DOI: https://doi.org/10.63332/joph.v5i5.1316

Ethically Aligned Artificial Intelligence: Investigating Algorithmic Bias, Human Identity, and Posthuman Ethics through a Data-Driven Philosophical Lens

Awad Alyousef[1], Asem Omari[2]

Abstract

Human society is being transformed by AI at an unprecedented speed-and in this context questions regarding fairness, agency, and identity become very tricky. This paper presents an interdisciplinary study into algorithmic bias and its ethical implicature in the relationship of data science and philosophy. By bringing into account concepts from posthumanism and moral theory, the research scrutinizes the redefinition of human identity within an age of autonomous systems and artificial consciousness. It proposes a data-oriented mechanism to the detection and mitigation of bias in AI systems, substantiated with case studies in health care and recruitment. One of the emphases of this work is on transparency, accountability, and inclusive governance, ultimately arguing for ethically aligned AI that promotes human dignity concurrent with evolving technological realities. Implications of this study point toward diversifying the development of a new ethical paradigm with the capability to regulate AI design, deployment, and policy.

Keywords: *media convergence era; broadcasters and presenters; dilemmas and opportunities; radio and television; digital technology*

Introduction

In the realm of algorithmic governance and intelligent machines, the ethical alignment of artificial intelligence with human values has become one of the top issues challenging researchers, policymakers, and societies today. It compels the increasing presence of artificial intelligence (AI) in human affairs with ethical considerations on the responsibility of use so that they do not infringe against fundamental human values (Ayinla, 2024). Incrementally, as AI permeates the human world in health, finance, and governance, the ethical dilemmas that may stem from this gradual innovation need to be proactively anticipated. Development and deployment of AI should, therefore, be in alignment with prevailing societal norms and ethical principles, thus engendering trust and sprouting unintended adverse consequences. Additionally, AI and robotics are like a ramp of promise to change everything in the way industries and individuals relate and yet present some very serious ethical challenges confronting responsible innovation (Grover, 2025). These include algorithmic bias, accountability, and societal implications of the integration of AI - all of which demand a structured approach to ethical AI governance. It is already being recognized that AI systems can reveal people's political orientations, thus bringing along enormous potential risks that the AI community should be aware of and actively work to mitigate to avoid misuse and discrimination (Peters, 2022). For

[1] Information Systems Department. College of Computer and Information Sciences. Prince Sultan University, Riyadh, Saudi Arabia, Email: aalyousef@psu.edu.sa, (Corresponding Author)
[2] Computer Information Science, Higher Colleges of Technology Al Ain, United Arab Emirates, Email: aomari@hct.ac.ae

this reason that an appropriate safeguard and ethical guidelines need to be set for protecting individual's privacy against the manipulation of their political opinions.

Research Problem and Objectives

Algorithmic bias in AI systems leads to unfair discrimination possibly based on social identity, and thus creating detection and mitigation methods ensures that these systems yield just results (Peters, 2022). It can take different forms: biases in gender, race, or politics, and it can essentially harm societies-general as well as individuals. It is therefore very important to find and fix sources of algorithmic bias before they form a permanent fixture in AI systems. The paper comprises the review of ethical dilemmas in the development of AI-based systems and ways to make algorithmic decision-making processes more transparent, fair, and accountable (Akinrinola, 2024). These ethical dilemmas should, therefore, be used to reveal the ethical factors that should steer the development and deployment of AI systems. The ultimate goal is to reconcile technological development with people by promoting responsible development and alignment with personal rights, so as to secure harmonious coexistence with AI (Ayinla, 2024). This is a multidisciplinary challenge that must tackle moral, social, and even legal implications of AI technologies.

Scope and Significance

Thus the study revolves around ethical inquiries on AI and robotics, which essentially brings more significant arguments into its agendum of complete policy advocacy, regulatory supervision, and multidisciplinary collaboration so that artificially intelligent techniques work for humanity conscie~ntiously and ethically in the future (Grover, 2025). Thus, one part of the research is examining those dilemmas to induce a suitable framework for governing ethical AI that will be aimed towards fairness, accountability, and transparency. Arguably, it concerns realizing algorithmic bias, which is crucial in creating a generalizable and fair technology for AI that impacts more downside population groups possibly most through biased algorithms (Mittermaier, 2023). Fairness of AI systems would also contribute to the non-existent replication of current inequalities and hence social justice. Thus-this research propels that necessary development of comprehensive ethics so that technological advancement touches the imagination and needs of society as a whole and that AI technologies are developed and used for the welfare of all members of society (Zuhri, 2024). The study will contribute to a framework for governing ethical AI to achieve enhancement of responsible innovation and trust in AI systems.

Theoretical Foundations

AI is deemed to have minimum human intervention capability in doing tasks, with algorithms to process data and make decisions, allowing automation and increased efficiency across innumerable fields. This definition shows the capability of AI systems to learn from data, for them to adapt in new situations, and for them to make decisions without explicit instructions from a human. In healthcare, AI is gaining more prominence as tools develop to predict surgical outcomes and assist surgeons in improving patient care and surgical precision (Mittermaier, 2023). To increase medical practices and improve patient outcomes, such applications suggest promising results for the world of technology. One group-based large coarse barrier in the text, artificial intelligence in studies: its aid in text analysis, thematic classification, and generation of digital responses, hence improving ease of access and substantiality in scholarly assessments

(Zuhri, 2024). This massive change is the potential to utterly redefine the way we understand and interpret complex texts, giving way to new insights and perspectives.

Posthumanism: Historical and Philosophical Context

Posthumanism acts by transmitting doubt into the anthropocentric stance on theory and practice: on the one hand, one must challenge older shared philosophical views of technology as some instrument that is entirely in the hands of humans to command and make use of (Estrada, 2024). First, the question would be pondered concerning how to live through the thought that man stands always at the core of nature as a fixed pole and that technology exists simply to help man. Posthuman ethics (with its) inclusion of distributed agency- wherein the effective building of an AI among humans co-create moral dynamics of their actions-that stands upon the tougher position (encan, 2024) forces in immediate questioning of classic anthropocentric moral frameworks that set human agency on top. This definition finally sees AI, not just as a tool/tool against humans, but an active agent in creating ethical outcomes. In conclusion, posthuman subjectivity blurs the traditional division between nature and design by embedding human and non-human logic towards a redefinition of humanity in the face of technology (McAvan, 2021). This reconfiguration questions traditional human self-perception, thereby pulling the strings from humans into machines.

Ethical Paradigms in Human–Machine Interactions

The idea is to imply frameworks on ethics for AI that take into account transparency, explainability, and the role of humans from the perspective of algorithmic decision-making. It is envisaged that AI systems will be held accountable and not regarded as "black boxes" (Nazeer, 2024). The policy frameworks are seen to be safeguards. They are to create trust in AI systems to avoid unintended ill effects. The integration of the ethical, social, and legal values with the technical development of AI is also important at the design stage, ensuring that the ethical considerations and the development of AI go hand in hand (Dignum, 2017). This is one sort of multidisciplinary approach involving ethicists, engineers, and policymakers. Among different things, this will ensure that AI systems function on societal values. It is imperative that the AI systems must align with the human values. This alignment should be supported by open and participatory processes involving multitudinal stakeholders in AI design and deployment (Huang, 2024). The value of this approach is that it upholds democratic AI ethics and serves as assurance of the alignment of AI systems with the values of the societies they serve.

Literature Review

Philosophical tools help one see the meaning of technology and the debate about AI as a tool for replacing thought; they offer criteria for judging AI's possible effects on human thought (Dutra, 2024). They help to question the claims made against AI and determine whether AI can indeed copy or surpass human intelligence to some degree. The motive for moral philosophy is to consider who or what deserves moral consideration-the moral status of intelligent and autonomous machines-that challenges anthropocentrism (Gunkel, 2012). Thus the matter of the rights and responsibilities of AI systems and that moral obligation toward humans presents itself. Another form of AI ethics may conceivably count as relational, radically empirical, and altruistic, one moving past regular deontological or utilitarian-type ones toward a concern with the relations between humans and AI (Gunkel, 2020). Focusing here would keep empathy, compassion, and care as paramount in human-AI interaction.

Prior Studies on AI Ethics and Moral Reasoning

AI ethics have emerged as an important research issue concerning ethical problems, privacy violations, and discrimination that has incited many guidelines and principles to alleviate these challenges (Huang, 2022). The research intends to make sure that AI systems are developed and utilized in ethical, responsible ways within the limits of societal values. Research concerning AI ethics usually looks through philosophical lenses, legal, or technical, almost neglecting the sociocultural one, hence emphasizing the need for a more holistic and interdisciplinary approach (Avnoon, 2023). This involves much more of the social, cultural, and political contestations around the paradigms within which AI systems are conceived, developed, and deployed. Contemporary advances in AI have led to what is being seen as a potentially far-reaching discourse on AI ethics, thus resulting in numerous ethics guidelines that seem to be growing awareness about AI's ethical implication (Hagendorff, 2020). Such guidelines show the frameworks for ethical AI development and deployment, but actual implementation and enforcement will define their power.

The Posthuman Condition: From Cyborgs to Conscious Code

Posthumanism, a field that portrays ambiguous boundaries between artificial and human intelligences, studies the egalitarian relationship of humans with machines, opposing the traditional anthropocentric view with regard to human intelligence (Ding, 2024). The posthuman portrayal raises questions regarding consciousness, identity, and agency in the age of AI. Within the posthuman epoch, re-agency is reconceived as networked, thereby also defying the anthropocentric frameworks of ethics prioritizing human agency and responsibility (encan, 2024). This really reframes the fact that AI is not just a passive tool but also an active agent whose actions will yield an ethical outcome. The monster concept destabilizes notions of fixity and metaphysics of presence to engage modern notions of sex, gender, and sexuality to reflect the posthuman quality of identity fluidity and hybridity (McAvan, 2021). In these challenges, one should raise questions as to how identity is constructed and the extent to which technology itself shapes it. However, very few studies present an integrated perspective that combines data-driven detection of bias with philosophical interpretation related to posthuman ethics and identity. This paper attempts to address this void.

Recent Studies on Ethical AI and Domain-Specific Applications

A growing body of research has recently focused on the ethical alignment of artificial intelligence (AI) in decision making, interpretability, and, more broadly, fairness across central domains. Indeed, Al-Omari et al. (2025) studied the responsible AI regulatory mechanisms in higher education, showcasing how ethical governance structures could be encoded proactively in the mitigation of algorithmic bias and other value alignments-the latter being among the concerns in the ethical realization of large language models (LLMs). In the legal domain, where both transparency and moral justification are required, Hassan et al. (2024) showed how a deep learning-based summarization model can improve interpretability of complex legal texts and, hence, ethical AI decision making in the highest stakes contexts. The import of preprocessing strategies such as stemming in improving the NLP model's accuracy for classification tasks in law is emphasized by Jabbar et al. (2024) for fairness and consistency in linguistic outputs.

According to Ammar et al. (2024), it is an ethical imperative to carry out fine-tuning at the level of the actual task in transformer models like BERT and GPT when predicting Arabic legal judgments. The authors stress cultural sympathy and moral coherence. Rehman et al. (2025)

suggested hybrid deep learning methods to improve facial emotion recognition (FER) systems, emphasizing the moral obligation to develop affective computing applications that are accurate in the human emotional state domain. Within the scope of cloud computing, Gaber and Alenezi (2024) proposed serverless architectures, in particular Function-as-a-Service (FaaS), which should provide a scalable environment for ethically aligned AI systems attacking challenges in security, accountability, and transparency of the systems.

Moreover, in 2024, Alyousef and Al-Omari analyzed AI applications for health care, emphasizing how ethical concerns—especially patient data privacy and algorithmic discrimination—demand coordinated global frameworks to maintain safety and moral integrity. In a similar vein, Semary et al. (2023) demonstrated that bias mitigation during sentiment analysis of user-generated content may be successfully implemented with transformer-based models such as RoBERTa thereby showing instead the flexibility of these architectures for ethically sensitive tasks.

Methodology

The approach will continue using philosophical inquiry to address questions on the alignment of AI into normative and technical dimensions, offering a framework for ethical development and deployment in AI (Gabriel, 2020). Critical inquiry into ethical, social, and legal implications posed by AI technologies is then undertaken. Philosophical analysis would help in understanding ethical dilemmas accompanying AI development, that is, transparency, fairness, and accountability, which would lead to understanding such considerations informing the design and use of AI systems (Akinrinola, 2024). This requires a thorough grasp on ethical principles as well as their applicability to technologies such as artificial intelligence. The methodology would therefore include content analysis of classical and contemporary Islamic literature on ethics and philosophy, presenting an entirely new angle of ethical implications of AI in terms of Islam (Zuhri, 2024). Justice, fairness, and the common good are stressed in the perspective of AI in development and deployment.

Analytical Framework

It has a systematic framework to detect and measure algorithmic bias with possible practical solutions to ensure an AI system that is fair and equitable (Sharma, 2025). The framework involves identifying sources of bias, measuring how they affect the system, and then applying strategies to either reduce or eliminate that bias. The analytic framework centers on viewing and representing algorithmic ethics within ideologies, discourses, and worldviews and yields a comprehensive understanding of the ethical dimension of artificial intelligence (Avnoon, 2023). This framework is expected to require multidisciplinarity and must design social, cultural, political contexts within which AI systems are constructed and implemented. Using thematic analysis of recent work, the study probes changes that AI might bring and ethical questions raised, specifying important themes and emerging issues in AI ethics. Such analysis could assist in developing ethical guidelines and public policy concerning AI.

Data Sources (Texts, Theories, Case Studies)

The study examines peer-reviewed papers presented in leading conferences to determine recent trends in the area of AI ethics, highlighting the current research frontiers and emerging issues within the field (Birhane, 2022). This will inform future research directions and promote collaboration among researchers in AI ethics. Examples refer to literature, industry reports, and real-life instances through which the relationship between technological development and

societal values can be understood in a wider and better view of the ethical implications of AI (Nazeer, 2024). These sources give insightful lessons on the kinds of issues and potentials of AI, and the ethical considerations that are necessary for informing technology development and delivery. The study includes actual case examples to derive real-life applications of different moral frameworks in techno-ethical dilemmas, presenting practical guidance for ethical decision-making in AI (Panchal, 2024). These case studies also serve to illustrate the complexity surrounding AI ethics and the need to adopt multiple perspectives.

AI and Consciousness: Ontological Reflections

Opened up by the possibility of conscious AI, the discussion proposes an agnostic approach to possible ethics and legal frameworks with regards to the emerging consciousness of machines (Kikis, 2023). Such an approach presupposes focus on developing ethical and legal frameworks able to accommodate both conscious and non-conscious AI systems and avoids assumptions about the nature of consciousness. This study also delves into defining consciousness as not only a concept but also ethical issues raised when replicating or simulating consciousness in machines, acknowledging the intricate nature of consciousness-making it difficult to copy it artificially (Tiwari, 2025). It also calls for humility and caution in developing AI systems. Autonomy, personhood, and rights might be quickly indicated in an AI context but require a shift in human perspectives about the nature of possible AI autonomy and their relationship with humans to ensure that they coexist sustainably (Kikis, 2023). And that raises the fundamental issue on the definition of personhood and what rights are to be accorded to the AI systems.

Functionalism vs. Phenomenology

The research studies philosophical theories of consciousness such as functionalism and phenomenology to explore the extent to which and the limits of AI can imitate human-like consciousness. Functionalism treats consciousness as a result of functional organization, then phenomenology stresses individual experience of consciousness. These two philosophies engage in a deep dive of concepts of consciousness as propelled by a world that is increasingly technology driven and offer different directions towards the nature of consciousness and how it relates to technology (Tiwari, 2025). Importantly, they are useful in knowing how to build AI systems that approximate more human values and ethical principles. Such an investigation questions whether there would emerge a world reconstructed by philosophy based on technological parameters or humanist tradition, probably questioning the impact of artificial intelligence over what peoples understand as well as their place in it (Dutra, 2024). Such questioning also shows that critical attention is important with the values and assumptions guiding AI development.

Personhood in Synthetic Agents

The inquiry focuses on the moral standings of intelligent and autonomous machines questioning the traditional idea of technology only as a tool and also studying the possibility of AI as a moral agent (Gunkel, 2012). This consideration also leads to questioning the obligations of these systems and the ethical responsibilities humans have towards them. The present study engages in the possibility of AI as a responsible subject under serious discussions, that delve into the ethics that might be breached in AI authorship, as well as the difficulties with assigning accountability for AI-generated content (Yldz, 2023). The outcome of this exploration is a need for new ethical frameworks that could be addressing the inherent challenges posed by AI. The investigation also questions whether AI ought to be construed as a tool or as a value in itself,

which influences its acceptability as an author and raises issues such as the nature of creativity and originality in AI (Yldz, 2023). This view goes against the conventional ideas of authorship along with the extent of human creativity in the artistic expression.

Ethical Dimensions of Artificial Intelligence

Autonomy and Accountability

Ethical queries over accountability and liability arise with autonomous decision-making AI systems as they cause harm or make biased decisions (Nazeer, 2024). Thus, the requirements are clear demarcation of responsibilities and a system of redress when AI systems go wrong. Algorithmic bias, accountability in the event of failures, and potential misuse or misapplication of systems are among the most pressing ethical, social, and regulatory challenges that must be faced before AI can be developed and deployed in an ethically responsible manner. Multidisciplinary approaches are needed, involving ethicists, engineers, and policymakers, as well as the public. A structured framework for ethical AI governance is proposed to guide transparency in policy and regulation for justified and responsible service of humanity by AI, as well as to cultivate trust in such systems (Grover, 2025). It explains the need for fairness, accountability, and transparency in developing and deploying AI systems.

Bias, Justice, and Algorithmic Opacity

It violates everyone by their social identity through algorithmic bias and puts forth the fact that it should be recognized and examined by the AI community so that inequalities are not carried forward (Peters, 2022). The bias could be of different types, e.g., gender, race, or political, and could negatively affect them and other groups. Bias infusion in the AI system may lead to derived dataset and assumption-associated biased constructs requiring steps to minimize causes and mitigate bias so that AI systems could offer fairness and equity (Sinwar, 2023). Careful attention must give to this process while collecting data, designing models, and evaluating them. Thus, explainable AI (XAI) can improve algorithmic error as well as bias transparency and help in identifying them on their existence in algorithmic models besides advocating user trust on AI systems (Andrade, 2024). With the developments in XAI, such techniques make the understanding of AI systems' decision-making processes easier for users and thus the identification of correction of biases. Table 1 points that below presents the common types of algorithmic bias seen in the observed reality of AI applications together with their ethical ramifications.

Type of Bias	Example	Ethical Implication
Gender Bias	Job recommendation systems suggesting nursing roles predominantly to women	Reinforces gender stereotypes and limits equal job opportunities
Racial Bias	Facial recognition systems misidentifying darker-skinned individuals	Violates rights, leads to wrongful arrests, undermines trust
Political Bias	Algorithms classifying users' political views based on search behavior	Risks opinion manipulation and infringes on freedom of expression

Type of Bias	Example	Ethical Implication
Data Imbalance	Healthcare AI trained mostly on male patient data	Poor performance for underrepresented groups, unequal care outcomes
Confirmation Bias	News feed algorithms reinforcing pre-existing user beliefs	Encourages polarization and undermines critical thinking

Table 1: Mapping Bias in Machine Decision-Making: Ethical Dimensions of Algorithmic Discrimination

Moral Agency in Posthuman Entities

The research explores the issue of whether intelligent and autonomous machines may in fact have proper moral responsibilities, thereby contesting mainstream anthropocentric views that limit moral agency to human beings (Gunkel, 2012). This raises the matter of the criteria for moral agency and how AI systems might meet these criteria. Posthuman ethics takes into consideration the moral implications of AI and the changing landscape of agency, envisioning agency as a networked phenomenon involving ,occas ins human actors and nonhuman ones (encan, 2024). This redefinition implies that AI systems are no longer purely passive instruments, rather they are active agents that can influence ethical outcomes. Since a posthuman world will be populated by human and AI entities co-crafting ethical dynamics, ethical considerations must reconfigure major ideas around individual responsibility and assign greater salience to collaboration in responsibility (encan, 2024). Such collaboration demands a common understanding of ethical principles and an agreement on responsible AI development.

Human Identity in the Age of AI

Redefining Humanity in Technological Societies

Research discusses redefinitions of humanity within our technological societies, the examination of hybrid beings, the dissolution or expansion of the self, rebuffing classical conceptions of human identity, and the possibility of coexistence between humans and AI alike (Ding, 2024). This line of inquiry implicates questions for humanity's future and finds its answers based on technology's influence on our understanding of selfhood. The research also considers ethical implications of AI, particularly with respect to the changing relationship between AI and human identity, calling for critical engagement with the implications of AI on our sense of self and our relationships with others (Grover, 2025). These engagements must be multidisciplinary, incorporating considerations of the ethical, social, and psychological ramifications of AI. The analysis examines AI's influence on world affairs while further investigating philosophical debates on such concepts as consciousness and human uniqueness, thereby offering some international flavor as to the ethical and societal implications of AI (Al-Omari et al. , 2025). In a sense, this argument underlines the need for international collaborative efforts in addressing the challenges spawned by AI.

Hybrid Beings: Embodiment and Virtuality

This posthuman era literally falls out from human-to-nature and culture, interweaving human and non-human elements of the contemporary life, thus blurring the boundaries between the real and the virtual and raising questions concerning the embodiment of such (McAvan, 2021). This blurring challenges the traditional cognizance of human identity and the relationship between

body and self. Clocking new definitions of humanity in which sex, gender, and sexuality are becoming entangled with biological support and informational technology, these new definitions challenge traditional feelings of identity and make inquiries on the role technology plays in our understandings of self. In other words, the entanglement then calls for an ethical inquiry of the definitions bent by technology on human identity. Modern subjectivity has come to be indexed by hybridity, the mingling of human and non-human elements that poses problems for selfhood definitions, challenging traditional notions of human identity, and interrogating consciousness and agency in AI (McAvan, 2021). This hybridity asks for redefinitions of what it means to be human in a technological world.

The Dissolution or Expansion of Self

Posthumanism challenges normative ideas concerning what it means to be human, therefore becoming the very ground for reflection concerning the catalogue of evolving dynamics with which an advanced technological society has to come to grips to question the traditional anthropocentric view of humanity (Ding 2024). With such a challenge comes a critical examination of those values and assumptions informing our comprehension of self. The research takes a stand regarding existential questions raised by an AI consciousness level regarding humans and machines in an egalitarian relationship with such questions about what consciousness is and whether AI can even reach humanlike intelligence. Such exploration demands nuanced understanding of philosophical theories of consciousness and the ethics of making consciousness machines. The mere advent of AI brings forth the acknowledgment; here there are no longer robots about humans: they must therefore share their life with AI for an even more just and durable world between humankind and AI (Ding 2024). This change in human perspectives requires commitment to responsible AI development as well as a shift in human perspectives. Such philosophical-analytical methodologies afford the possibility of deepening understanding in the area of how algorithmic systems reflect and shape evolving ethical norms and human identities.

Analysis and Discussion

The current section analyzes both conceptual and empirical findings of the study, finally positioning the findings in a broader discourse within AI Ethics, Posthuman Philosophy, and Algorithmic Accountability. The discussion merges findings from philosophical coexisting themes, previous works on AI ethics, and the posthuman condition to give an extensive account of the ethical implications posed by Artificial Intelligence and how these implications impact human identity (Gunkel, 2020), (Hagendorff, 2020), (McAvan, 2021). The merger, therefore, allows for a holistic understanding of the challenges and opportunities posed by AI. An integrative analysis of ethical frameworks and technical solutions has been introduced to confront algorithmic bias in the AI systems giving real-world relevance for further bias mitigation and fairness (Nazeer, 2024), (Sharma, 2025). This analysis must be multidisciplinary, requiring ethical considerations in tandem with technical aspects. The authors advocate greater transparency, fairness, and accountability in the development of AI, emphasizing the need for ethical guidelines and regulatory compliance to ensure the responsible and ethical use of AI (Akinrinola, 2024), (Grover, 2025). The guidelines and regulations should be developed with the help of a wide range of representatives and should regularly be updated to reflect the shifting landscape of AI.

Implications for Posthuman Ethics and Policy

In a posthuman ethics context, the research debates the moral responsibility of life in a world mediated by technology and raises questions on the traditional notions of human agency and responsibility (Encana, 2024). This new understanding will allow the ethical implication of AI and technology in general to be evaluated in so far as they shape our moral landscape. The study aids policy by addressing the ethical and legal challenges in contemporary debates of AI through the lens of humanity and its implications, working toward a framework for ethical AI governance (Dutra, 2024). The framework should be built on the principles of fairness, accountability, and transparency and developed in consultation with an array of stakeholders. The implications call for a major reassessment of current practices in AI development, policy-making, and ethical guidelines, engendering a future where technology embodies human values and supports an equitable and sustainable society (Panchal, 2024). This review must involve a commitment to Responsible Innovation and an alternative outlook towards what current assumptions and practices deem acceptable.

Limitations of the Inquiry

The limitations of the inquiry include a tendency to focus on certain cultural contexts while neglecting an empirically based, socio-cultural perspective, indicating the need for research addressing the various cultural and social contexts in which AI is developed and deployed (Bahir, 2021). Such research should involve different stakeholders, engaging them with sensitivity toward the unique needs and values of different communities. The study states that this remains quite difficult given that AI is difficult to comprehend, while another challenge posed is the failure to adequately discern the qualities of AI systems, acknowledging the context of complexity that makes human-like intelligence very elusive to ascertain (Kikis, 2023). Such a realization further highlights the need to nurture caution and humility in developing AI systems. The limitations concentrate their efforts on ethical experiences and applications that are sensitive asymmetries, thus highlighting the opportunity to examine the ways in which AI could solidify existing inequalities or engender new discrimination (Birhane, 2022). The above indeed requires working hard for equity and justice in the development and deployment of AI.

Conclusion

This study examines the nexus of artificial intelligence, ethical alignment, and posthuman identity through the integration of data science methodologies with philosophical analysis. Focusing on the issue of algorithmic bias, the study makes the point that, given their innate autonomy, such systems could work excessively to reproduce inbuilt societal underpinnings. This bias demands technical interventions—namely, explainable AI models and governance structures meant to reduce algorithmic discriminatory tendencies—as well as a reevaluation of those viewpoints concerning agency, responsibility, and fairness in an age where human action is increasingly mediated by algorithmic systems. In light of posthumanist positions, the study champions the diverse body conceptualizations that hybrid human-machine encounters could mean for what it is to be human. Philosophical frameworks (especially Jeffersonian functionalism and existential phenomenology) could then be summoned to provide some remarkable critiques explaining the plausibility and implications of artificial consciousness. These critiques demand an ethical paradisal framework that more capitalistically contemplates on what is valuable in the world not only in terms of "use value" or efficiency but also in its dignity, cultural sensitivity, and sustainability in a morality logic. This study, which pleads for the ethical evolution of AI politics, was required to engage in transparent, transparent, and local-

relatable AI frameworks. By sequestering any harm and ensuring an equitable solution, the situation dictates that the placement of AI systems in law, healthcare, and affective computing calls for ethico-legal governance.

Acknowledgments

The authors would like to acknowledge the support of Prince Sultan University for paying the Article Processing Charges (APC) of this publication. The authors would like to acknowledge the support of Prince Sultan University for their support making this publication successful.

References

Ayinla, B. S., Amoo, O. O., Atadoga, A., Abrahams, T. O., Osasona, F., & Farayola, O. A. (2024). Ethical ai in practice: balancing technological advancements with human values. None. https://doi.org/10.30574/ijsra.2024.11.1.0218

Grover, D. R. (2025). Ethical challenges in ai and robotics: balancing innovation and responsibility. None. https://doi.org/10.47392/irjaem.2025.0064

Peters, U. (2022). Algorithmic political bias in artificial intelligence systems. Philosophy & Technology. https://doi.org/10.1007/s13347-022-00512-8

Akinrinola, O., Okoye, C. C., Ofodile, O. C., & Ugochukwu, C. E. (2024). Navigating and reviewing ethical dilemmas in ai development: strategies for transparency, fairness, and accountability. None. https://doi.org/10.30574/gscarr.2024.18.3.0088

Mittermaier, M., Raza, M. M., & Kvedar, J. (2023). Bias in ai-based models for medical applications: challenges and mitigation strategies. None. https://doi.org/10.1038/s41746-023-00858-z

Zuhri, M. T., Sahlani, L., & Munawaroh, N. (2024). The ethics of artificial intelligence (ai) utilization in qur"anic studies: an islamic philosophical perspective. Asyahid Journal of Islamic and Quranic Studies (AJIQS). https://doi.org/10.62213/b6hexr21

G, S. G. N. (2024). The rise of ai and autonomous systems: transforming industries and navigating ethical challenges. International Journal for Research in Applied Science and Engineering Technology. https://doi.org/10.22214/ijraset.2024.65784

Estrada, D. (NaN). Human supremacy as posthuman risk. None. https://doi.org/10.25884/6Q27-6T77

encan, S. (2024). Posthuman ethics and distributed agency in kazuo ishiguros klara and the sun. None. https://doi.org/10.29110/soylemdergi.1541383

McAvan, E. (2021). Frankenstein redux. M/C Journal. https://doi.org/10.5204/mcj.2843

Nazeer, M. Y. (NaN). Algorithmic conscience: an in-depth inquiry into ethical dilemmas in artificial intelligence. International journal of research and innovation in social science. https://doi.org/10.47772/ijriss.2024.805052

Dignum, V. (2017). Responsible autonomy. None. https://doi.org/10.24963/ijcai.2017/655

Huang, L. T., Papyshev, G., & Wong, J. (2024). Democratizing value alignment: from authoritarian to democratic ai ethics. Springer Nature. https://doi.org/10.1007/s43681-024-00624-1

Dutra, D. J. V. & Brennand, E. G. D. G. (2024). Intelligence and philosophy. None. https://doi.org/10.4013/fsu.2024.251.08

Gunkel, D. J. (2012). The machine question: critical perspectives on ai, robots, and ethics. None. https://doi.org/10.7551/mitpress/8975.001.0001

Gunkel, D. J. (2020). Perspectives on ethics of ai. None. https://doi.org/10.1093/oxfordhb/9780190067397.013.35

Huang, C., Zhang, Z., Mao, B., & Yao, X. (2022). An overview of artificial intelligence ethics. Institute of Electrical and Electronics Engineers. https://doi.org/10.1109/tai.2022.3194503

Avnoon, N., Kotliar, D. M., & Rivnai-Bahir, S. (2023). Contextualizing the ethics of algorithms: a socio-

professional approach. New Media & Society. https://doi.org/10.1177/14614448221145728
Hagendorff, T. (2020). The ethics of ai ethics: an evaluation of guidelines. Springer Science+Business Media. https://doi.org/10.1007/s11023-020-09517-8
Ding, Z. (2024). Between man and machine: a posthumanist analysis of i, robot. Journal of Education, Humanities and Social Sciences. https://doi.org/10.54097/b1p5ar74
Gabriel, I. (2020). Artificial intelligence, values, and alignment. Springer Science+Business Media. https://doi.org/10.1007/s11023-020-09539-2
Sharma, R. (2025). Addressing algorithmic bias in statistical models: integrating technical solutions with ethical governance for fair ai systems. International Journal of Scientific Research in Computer Science Engineering and Information Technology. https://doi.org/10.32628/cseit251112316
Panchal, J. (2024). Ethics and artificial intelligence (ai) in the age of automation: reexamining moral frameworks in techno-ethical dilemmas. ShodhKosh Journal of Visual and Performing Arts. https://doi.org/10.29121/shodhkosh.v5.i1.2024.4851
Birhane, A., Ruane, E., Laurent, T., Brown, M. S., Flowers, J., Ventresque, A., & Dancy, C. L. (2022). The forgotten margins of ai ethics. None. https://doi.org/10.1145/3531146.3533157
Kikis, M. (2023). Legal framework for the coexistence of humans and conscious ai. Frontiers Media. https://doi.org/10.3389/frai.2023.1205465
Tiwari, S. K. & Tiwari, V. K. (2025). Philosophical analysis of consciousness as an intersection point of philosophy, culture and artificial intelligence. None. https://doi.org/10.53032/tvcr/2025.v7n1.07
Yldz, A. (2023). Ai as a co-author? we should also ask philosophical (and ethical) questions. European Journal of Therapeutics. https://doi.org/10.58600/eurjther1723
Sinwar, D., Sinha, A., Sapra, D., Singh, V., & Raghuwanshi, G. (2023). Assessing and mitigating bias in artificial intelligence: a review. Recent Advances in Computer Science and Communications. https://doi.org/10.2174/2666255816666230523114425
Andrade, O. M. D. & Alves, M. A. S. (2024). From "black box" to "glass box": using explainable artificial intelligence (xai) to reduce opacity and address bias in algorithmic models. Revista Thesis Juris. https://doi.org/10.5585/13.2024.26510
(2025). Artificial intelligence and foreign affairs. None. https://doi.org/10.3726/b22429
Bahir, S. R., Kotliar, D. M., & Avnoon, N. (2021). Contextualizing ai ethics in time and space. None. https://doi.org/10.5210/spir.v2021i0.12235
Kostick, K. M. & Gerke, S. (2022). Ai in the hands of imperfect users. Nature Portfolio. https://doi.org/10.1038/s41746-022-00737-z
Krushnasamy, V. S., Al-Omari, O., Sundaram, A., & others. (2025). LiDAR-based climate change imaging in geoscience using spatio extreme fuzzy gradient model. Remote Sensing in Earth Systems Sciences. https://doi.org/10.1007/s41976-025-00197-5
Nimma, D., Al-Omari, O., Pradhan, R., Ulmas, Z., Krishna, R. V. V., El-Ebiary, T. Y. A. B., & Rao, V. S. (2025). Object detection in real-time video surveillance using attention based transformer-YOLOv8 model. Alexandria Engineering Journal, 118, 482–495. https://doi.org/10.1016/j.aej.2025.01.032
Al-Omari, O., & Al-Omari, T. (2025). Artificial Intelligence and Posthumanism: A Philosophical Inquiry into Consciousness, Ethics, and Human Identity. Journal of Posthumanism, 5(2), 458–469. https://doi.org/10.63332/joph.v5i2.432
Rehman, A., Mujahid, M., Elyassih, A., AlGhofaily, B., & Bahaj, S. A. O. (2025). Comprehensive review and analysis on facial emotion recognition: Performance insights into deep and traditional learning with current updates and challenges. Tech Science Press. DOI: 10.32604/cmc.2024.058036
Al-Omari, O., Alyousef, A., Fati, S., Shannaq, F., & Omari, A. (2025). Governance and ethical frameworks for AI integration in higher education: Enhancing personalized learning and legal compliance. Journal

of Ecohumanism, 4(2), 80–86. DOI: 10.62754/joe.v4i2.5781

Alyousef, A., & Al-Omari, O. (2024). Artificial intelligence in healthcare: Bridging innovation and regulation. Journal of Ecohumanism, 3(8), 10582–10589. DOI: 10.62754/joe.v3i8.5673

Hassan, A. Q. A., Al-onazi, B. B., Maashi, M., Darem, A. A., Abunadi, I., Mahmud, A. (2024). Enhancing extractive text summarization using natural language processing with an optimal deep learning model. AIMS Press, 2024. DOI: 10.3934/math.2024616

Jabbar, A., Iqbal, S., Tamimy, M. I., Rehman, A., Bahaj, S. A., Saba, T. (2024). An analytical analysis of text stemming methodologies in information transformers. arXiv.

Ammar, A., Koubaa, A., Benjdira, B., Nacar, O., Sibaee, S. (2024). Prediction of Arabic legal rulings using large language models. Faculty of Electrical Engineering Banja Luka. DOI: 10.3390/electronics13040764

Gaber, S., & Alenezi, M. (2024). Transforming application development with serverless computing. International Journal of Cloud Applications and Computing, 14(1), 1-12. DOI: 10.4018/IJCAC.365288

Semary, N. A., Ahmed, W., Amin, K., Pławiak, P., & Hammad, M. (2023). Improving sentiment classification using a RoBERTa-based hybrid model. Frontiers Media S.A., December. DOI: 10.3389/fnhum.2023.1292010