Department of Digital Business
Journal of Artificial Intelligence and Digital Business (RIGGS)
Homepage: https://journal.ilmudata.co.id/index.php/RIGGS
Vol. 4 No. 3 (2025) pp: 3212-3221
P-ISSN: 2963-9298, e-ISSN: 2963-914X

Human-Centered AI: Designing Intelligent Systems that Empower, Not Replace

Ahmad Nur Ihsan Purwanto
Department of Computer Science. Universitas Ary Ginanjar. Indonesia
ahmadnur.ihsan@esqbs.ac.id

Abstract

The rapid advancement of artificial intelligence (AI) has generated both optimism and concern regarding its impact on human society. While AI holds significant potential to augment human capabilities, fears of widespread automation and job displacement remain prevalent. This study explores the concept of human-centered AI, emphasizing the design of intelligent systems that empower individuals rather than replace them. Drawing on a qualitative library research approach, the paper analyzes ethical principles, including fairness, accountability, transparency, and inclusivity, as well as practical design strategies such as human-in-the-loop frameworks, explainable AI, and participatory methods. The findings highlight a persistent gap between abstract ethical ideals and their practical translation into technical systems. To bridge this gap, the study proposes integrating normative values with engineering practices through actionable design principles and interdisciplinary collaboration. The research concludes that developing human-centered AI is essential to ensuring that technological progress fosters human flourishing, equity, and social justice in an era of accelerating digital transformation.

Keywords: *Human-Centered AI, Ethical AI, Human Empowerment, Explainable Artificial Intelligence, Participatory Design*

1. Introduction

In an ideal technological landscape, artificial intelligence (AI) should function as a collaborative partner that enhances human capabilities while respecting human values. Intelligent systems are envisioned not merely as tools of automation but as enablers of creativity, decision-making, and problem-solving [1]. The normative expectation is that AI development aligns with human-centered principles, ensuring fairness, transparency, and inclusivity, so that technological advancement contributes to social welfare and empowers individuals rather than displacing them. This paradigm, often referred to as Human-Centered AI, prioritizes human aspirations such as self-efficacy, social participation, and environmental protection in the design and implementation of AI systems [2], [3]. This approach integrates robust AI algorithms with human-centered design methodologies to create technologies that amplify, augment, and enhance human performance and decision-making capabilities [4]. This involves a concerted effort to move beyond merely explaining AI operations to enabling profound user agency, allowing individuals to adapt and even co-design the AI's internal mechanisms [5].

This ensures that AI serves as a true extension of human intent and capability, rather than an autonomous entity [6]. Such a philosophical grounding mandates that AI systems are developed with an acute awareness of their ethical implications and societal impact, fostering responsible innovation [7]. This necessitates a shift from viewing AI as an autonomous agent to understanding it as a sophisticated tool whose utility and ethicality are contingent upon human oversight and intentional application [6]. This fundamental reorientation towards human-centricity is crucial for cultivating trust and facilitating the widespread, beneficial adoption of AI technologies across diverse societal domains [8]. This paper explores the multifaceted dimensions of Human-Centered AI, delving into its theoretical underpinnings, practical applications, and the ethical considerations that guide its development [9]. It will examine how Human-Centered AI frameworks, which prioritize human values and cognitive processes, can mitigate the risks associated with AI, such as algorithmic bias, privacy infringements, and the erosion of human agency [10], [11].

However, the empirical reality reflects growing concerns that AI deployment often prioritizes efficiency and cost reduction over human empowerment. Automated systems increasingly replace human roles in manufacturing, services, and knowledge-based industries, triggering fears of job displacement and diminished human agency. Instances of algorithmic bias, opaque decision-making processes, and the concentration of technological power within a limited number of corporations further exacerbate the gap between the envisioned promise of AI and its

practical outcomes [12]. This divergence highlights the need for a paradigm shift in how AI systems are designed and implemented. This shift necessitates a renewed focus on integrating human cognitive biases, which inherently influence AI fairness, into the design process to address hidden pathways of human-to-AI biases [13]. This paper argues that a robust Human-Centered AI framework, prioritizing ethical alignment and participatory design, is imperative to bridge this gap, ensuring AI systems genuinely empower rather than undermine human capabilities [14]. Furthermore, this approach ensures that AI systems are not merely technically proficient but also socially beneficial, promoting human flourishing and societal well-being [6].

Previous studies have emphasized the ethical dimensions of AI, explored algorithmic accountability, and proposed frameworks for responsible innovation [6]. While these works contribute valuable insights, they often remain abstract or primarily normative, with limited focus on concrete design approaches that integrate human-centered values into technical development processes [15]. Moreover, existing research tends to concentrate on either ethical principles or technical efficiency, leaving a research gap in bridging these domains to develop actionable models of AI design that both safeguard human dignity and optimize system performance. This paper aims to bridge this gap by presenting a comprehensive Human-Centered AI framework that integrates ethical considerations directly into the design and deployment lifecycle, thereby moving beyond theoretical discussions to offer practical methodologies for fostering responsible and empowering AI systems [16]. This framework specifically addresses the inherent challenges of human fallibility and bias in AI development by embedding mechanisms for continuous ethical reflection and value alignment throughout the design process [17].

Specifically, this involves methods for identifying and mitigating epistemically violent biases that arise from underrepresentation in training data, as seen in the poor performance of facial recognition systems on African American women [18]. This framework will operationalize how to foster ethical AI usage and explore how AI can serve as a mirror, reflecting human biases and moral flaws to enable continuous ethical learning and improvement [19]. Ultimately, the goal is to cultivate AI systems that are not only technologically advanced but also ethically robust, promoting a symbiotic relationship between humans and intelligent machines. Such a relationship is predicated on embedding ethical principles from the outset of AI development, ensuring alignment with human values and societal expectations [20]. This necessitates a reassessment of existing moral standards in light of the profound risks and benefits AI presents, recognizing that ethical and responsible AI are crucial for fostering user and stakeholder trust [21], [22].

2. Research Methods

This study adopts a qualitative research design with a library research approach. The qualitative orientation is chosen because the research focuses on conceptual exploration, critical analysis, and interpretation of scholarly works rather than empirical measurement [23]. Library research serves as the primary strategy, enabling a systematic examination of academic literature, policy documents, and reports related to human-centered artificial intelligence, ethical AI frameworks, and design methodologies [24]. Through this approach, the study seeks to synthesize theoretical perspectives and identify normative and practical insights that contribute to the development of a human-centered AI framework [25].

The data collection method involves an extensive review of secondary sources. Relevant materials are gathered from peer-reviewed journals, academic books, conference proceedings, policy papers, and institutional reports published within the last five years to ensure the currency and relevance of information [26]. Databases such as Scopus, Web of Science, and IEEE Xplore are utilized to retrieve scholarly works, while policy and ethical guidelines are sourced from reputable organizations, including UNESCO, OECD, and the European Commission. Inclusion criteria are based on the relevance to themes of AI ethics, human-centered design, empowerment, and accountability in AI systems [27].

Data analysis is conducted using a qualitative content analysis method. The process involves three stages: reduction, categorization, and interpretation. First, selected sources are reviewed to extract key concepts and arguments related to human-centered AI. Second, these findings are categorized into thematic clusters such as ethical principles, technical design considerations, and socio-economic implications. Finally, interpretive analysis is applied to identify patterns, contradictions, and gaps within the literature. This analytical process allows for the construction of a conceptual framework that integrates ethical imperatives with practical design strategies for intelligent systems [28].

Through this methodological approach, the study not only maps existing discourse on human-centered AI but also provides a critical synthesis that addresses the research gap between normative principles and technical design practices. This comprehensive methodology, drawing on extensive literature review, ensures the framework is

DOI: https://doi.org/ 10.31004/riggs.v4i3.2450

grounded in current ethical discourse and technological advancements, facilitating the development of AI systems that truly align with human values and societal well-being [28], [29].

3. Results and Discussions

3.1. Ethical Foundations of Human-Centered AI

Human-centered AI is rooted in the ethical imperative to ensure that technological innovation aligns with fundamental human values. Ethical principles such as fairness, accountability, transparency, and inclusivity have been consistently emphasized in scholarly debates and policy frameworks [30]. Fairness entails the avoidance of discriminatory outcomes in algorithmic decision-making, while accountability ensures that system developers and operators remain responsible for the consequences of AI applications. Transparency refers to the clarity of processes and decisions generated by AI, which is essential to foster public trust [31]. Inclusivity underscores the importance of ensuring that AI benefits are distributed equitably across diverse social groups. By anchoring AI design to these ethical foundations, intelligent systems can serve as tools for human flourishing rather than as sources of social inequality or exclusion [32].

The alignment of AI with human values and interests is a primary consideration in its development, demanding a human-centered approach from conception through deployment [33]. This paradigm emphasizes the integration of human needs, capabilities, and values into the design, development, and deployment phases of AI systems, contrasting with purely technology-driven approaches that might overlook societal impacts [10]. This involves a deep consideration of ethical implications and a commitment to mitigating potential harms, ensuring that AI serves humanity's best interests [7]. The shift towards human-centered AI emphasizes not merely technological prowess but also the cultivation of systems that enhance human capabilities, promote well-being, and respect ethical boundaries [34]. This requires a deliberate effort to embed human dignity, privacy, and autonomy within AI's core architecture, moving beyond mere compliance to proactive ethical integration [33].

Such integration extends beyond technical specifications to encompass the socio-technical interactions, acknowledging that AI systems operate within complex human ecosystems [6]. This holistic perspective ensures that AI development considers not only individual user experience but also broader societal impacts, fostering applications that are both effective and ethically sound [35]. A critical aspect of ethical AI lies in ensuring fairness and mitigating biases to prevent discriminatory outcomes, especially within sensitive sectors such as finance, healthcare, and criminal justice, where AI decisions profoundly affect individuals [36]. Transparency further underpins public confidence by enabling users to comprehend the rationale behind AI-driven conclusions, thereby fostering trust and reducing skepticism [37]. Moreover, the ethical development of AI necessitates accountability mechanisms that clearly delineate responsibility for AI system behaviors, particularly when errors or harms occur, to ensure redress and uphold public trust [17].

This emphasis on responsible AI usage, aligned with Human-Centered AI principles, promotes human flourishing by positioning AI as a tool to augment, rather than diminish, human autonomy and capabilities [6]. This proactive approach necessitates that AI developers and users alike recognize their roles in shaping AI's impact, ensuring its deployment aligns with the overarching goal of societal benefit and individual empowerment. This paradigm shift underscores the necessity for continuous ethical reflection and adaptive governance frameworks to navigate the evolving landscape of AI technologies, ensuring their trajectory remains aligned with human welfare [38]. It is therefore crucial to foster environments where ethical principles guide research and development, promoting societal good and sustainability while actively preventing harm to all stakeholders [39]. This includes a focus on beneficence and non-maleficence, ensuring AI systems are designed to do good while minimizing potential harm or discrimination, which necessitates continuous evaluation of data quality and bias [40].

Furthermore, designing AI systems that are robust against adversarial attacks and are interpretable in their decision-making processes reinforces their ethical soundness and trustworthiness [41]. The integration of ethical considerations throughout the AI lifecycle, from conception to deployment and maintenance, is paramount for building public trust and ensuring responsible innovation [42]. This comprehensive approach also requires proactive measures to address issues such as algorithmic bias and lack of transparency, which are significant ethical concerns in AI development [43]. The development of responsible AI demands active engagement beyond mere compliance, requiring a commitment from all stakeholders to embed ethical frameworks into the entire system development lifecycle to ensure societal trust [44].

DOI: https://doi.org/ 10.31004/riggs.v4i3.2450

3.2. Human Empowerment versus Automation and Replacement

A central tension in the discourse on AI concerns its dual capacity to empower humans and to replace them. On one hand, AI systems can enhance human performance by augmenting cognitive and physical tasks, thereby enabling greater productivity and creativity [45]. On the other hand, the increasing deployment of AI in routine and knowledge-intensive tasks raises fears of job displacement and diminished human agency [46]. This paradox highlights the need to critically evaluate the societal consequences of automation. Human empowerment requires the deliberate design of AI systems that support skill development, decision-making, and problem-solving rather than displacing workers. By prioritizing augmentation over substitution, AI can function as a catalyst for expanding human potential while mitigating the risks of unemployment and inequality [47].

This perspective necessitates a shift from viewing AI solely as a tool for efficiency gains to recognizing its potential for fostering human growth and societal resilience. This involves developing intelligent automation governance systems and fostering public AI literacy to ensure that the adoption of AI aligns with broader societal well-being and does not lead to excessive technologization that undermines human autonomy [48], [49]. This approach requires a multidimensional strategy combining ethics, regulation, innovation, and education to shape AI development in alignment with human and social values [35]. This strategy would foster a future where AI serves as a collaborative partner, enhancing human capabilities and enabling individuals to pursue higher-order tasks requiring creativity, critical thinking, and empathy. Such a paradigm necessitates continuous evaluation of AI's societal impact, ensuring that its evolution remains congruent with principles of equity, sustainability, and human flourishing. The challenge is further compounded by the ethical paradoxes AI presents, where the promise of scientific advancement and efficiency coexists with concerns about human dependence and obsolescence [50].

This underscores the urgency for robust policy frameworks and educational initiatives that prepare the workforce for evolving job markets and promote human-AI collaboration [51], [52]. Such frameworks should prioritize the development of adaptive skills and lifelong learning opportunities to enable individuals to thrive in an AI-integrated economy, ensuring that humans retain agency and oversight [6], [53]. A key aspect of this is designing AI systems with a clear human-centric purpose, aiming to maximize AI's utility for humanity by placing humans at the core of system design, thereby shifting away from perceiving AI as autonomous teammates [6]. This proactive design philosophy ensures that AI functions as a supportive tool, enhancing human capabilities and enabling individuals to focus on complex, high-value tasks that demand uniquely human attributes such as creativity and empathy [54]. This perspective diverges from previous notions that cast AI as an independent entity, instead emphasizing its role as an extension of human intellect and capability [55].

This Human-Centered AI approach aligns with the principle of designing systems to enhance individual and societal capacities [6]. The overarching aim is to leverage innovative algorithms and methodologies to achieve this, thereby maximizing AI's utility for humanity. This design philosophy ensures that AI serves as a tool for empowerment rather than replacement, fostering human-AI collaboration [56]. This framework promotes the idea that AI, when developed with human needs and values at its core, can augment human intelligence and creativity, leading to unprecedented levels of innovation and problem-solving [57]. This focus on human flourishing through AI integration also necessitates a re-evaluation of educational paradigms, emphasizing skills that complement AI capabilities, such as critical thinking, creativity, and social-emotional intelligence [6].

This approach underpins the development of a "living" AI policy, ensuring ongoing adaptability to technological advancements and societal shifts while prioritizing human agency and critical thinking [6], [58]. This policy aims to foster responsible AI usage, aligning with the principles of Human-Centered AI by emphasizing that humans maintain autonomy and accountability over AI's goals and outcomes [6]. This approach posits that human flourishing is the ultimate objective, viewing AI as a means to achieve this rather than an end in itself. This requires a clear definition of Human-Centered AI, which encompasses the design of AI systems that prioritize human capabilities, needs, and values, while also addressing ethical principles such as fairness, accountability, interpretability, and transparency in their governance [8]. This definition broadens the scope beyond algorithms in isolation to consider the intricate processes of human-algorithm collaboration, fostering systems of human-machine hybrid intelligence that integrate the strengths of both while mitigating their weaknesses [59].

3.3. Design Principles for Human-Centered Intelligent Systems

The translation of ethical ideals into practice requires the articulation of concrete design principles. Central among these is the concept of human-in-the-loop, which ensures that critical decision-making processes remain under human oversight. Explainable AI represents another crucial principle, aimed at making algorithmic outputs interpretable and understandable to users [60]. Additionally, participatory design or co-design frameworks

DOI: https://doi.org/ 10.31004/riggs.v4i3.2450

encourage the involvement of stakeholders—including end users, domain experts, and policymakers—in the development process. This interdisciplinary integration fosters systems that are not only technically efficient but also socially responsive [61]. By embedding these principles into design methodologies, AI systems can reflect both computational rigor and human-centered values, thereby ensuring their relevance and acceptance in diverse contexts.

Such a holistic approach to design standards is vital for supporting the industrialization of AI technology while adhering to a Human-Centered AI framework [62]. This ensures auditability and accountability throughout the design, development, and deployment phases, aligning with domain-specific regulations and standards [63]. This includes the integration of mechanisms for continuous monitoring and evaluation of AI system performance and societal impact. Furthermore, the iterative refinement of AI models based on real-world feedback loops is crucial for achieving adaptability and resilience in dynamic environments, reinforcing the notion of AI as a continuously evolving entity. The success of such a framework hinges on the collective commitment of designers, developers, and users to prioritize human autonomy and well-being, fostering a symbiotic relationship where AI augments human potential rather than diminishing it [6].

This paradigm shift necessitates a rethinking of traditional human-computer interaction models, moving beyond mere interaction to a deeper conceptualization of human involvement that includes considering sociotechnical implications and potential failures [64]. It mandates the creation of robust ethical guidelines and frameworks to navigate the complexities of AI integration into human society [65]. One such framework advocates for a participatory approach in AI development and governance, allowing those impacted by AI systems to actively shape their design and deployment, thus promoting more responsible and human-centric outcomes [66]. This approach acknowledges the inherent challenges AI poses to traditional participatory design methods, particularly regarding the comprehensive understanding of AI's capabilities and limitations by all stakeholders [67]. This participatory design ethos for Human-Centered AI also emphasizes the importance of ongoing societal dialogue and public education to bridge the knowledge gap and foster informed engagement [6].

The active participation of diverse stakeholders, including end-users, is increasingly recognized as crucial for ensuring that AI systems align with societal values and address potential negative impacts, such as fairness and equity breakdowns [68]. This requires a comprehensive "living" AI policy that is subject to ongoing discussions, considering both technical utility and normative and social perceptions [6]. This participatory turn in AI design aims to integrate diverse perspectives, moving beyond mere consultation to genuinely empower affected communities in the development, evaluation, and deployment of AI systems [69], [70]. This collaborative approach can mitigate the risks of AI misuse by fostering a shared understanding of its limitations and ethical considerations. This collaborative approach transforms the relationship from a designer-user dynamic to one of co-creators, fostering self-determination and community empowerment in the AI lifecycle [71].

This holistic engagement ensures that AI development is not only technologically advanced but also ethically robust and socially beneficial, fostering trust and widespread adoption [34]. This includes a nuanced consideration of primary and secondary stakeholders, ensuring that participatory ideals like informedness, consent, and agency are realized across the broader AI ecosystem, extending beyond just immediate end-users [72]. This broader scope of stakeholder engagement is critical for addressing the complex socio-technical challenges presented by advanced AI systems, particularly in highly regulated domains where public trust and accountability are paramount [73]. Furthermore, given the opacity and complexity often inherent in advanced AI models, fostering deliberative governance mechanisms becomes essential to bridge knowledge boundaries between experts and the public, thereby building trust and ensuring that AI development is democratically accountable and responsive to societal needs [74].

3.4. Bridging Ethical Ideals with Practical Implementation

Although ethical frameworks for AI are abundant, their practical translation into technical systems remains a challenge. Many guidelines remain at the level of abstract principles, lacking actionable strategies for developers and engineers [75]. Bridging this gap requires mechanisms that integrate ethical ideals with operational practices, such as embedding fairness metrics into algorithmic evaluation, designing transparent system architectures, and instituting accountability mechanisms at both organizational and technical levels [76]. Case studies of AI deployment in areas such as healthcare, finance, and governance reveal both successes and shortcomings, illustrating the complexities of applying human-centered principles in real-world contexts. Developing a conceptual framework that unites normative ideals with technical design practices is therefore essential to advance AI systems that genuinely empower rather than replace human actors [77]. This transition from principles to practice demands a multi-faceted approach, encompassing not only the integration of ethical considerations

DOI: https://doi.org/ 10.31004/riggs.v4i3.2450

throughout the AI lifecycle but also a fundamental shift in organizational cultures and technical methodologies [78].

This involves cultivating habits of critical reflection and deliberation among researchers, technologists, and innovators from the earliest stages of design through to deployment, moving beyond mere compliance with off-the-shelf tools and documentation-centered governance [79]. This includes developing robust mechanisms for continuous evaluation and feedback, ensuring that AI systems remain aligned with human values as they evolve and adapt to new contexts [80]. This continuous alignment necessitates a dynamic and iterative design process that accounts for emergent properties of AI systems and their long-term societal impacts [81]. Furthermore, establishing clear accountability mechanisms and audit trails for AI-driven decisions is crucial to maintaining trust and addressing potential harms [82]. Such measures are particularly vital in sensitive applications, where AI decisions can have profound implications for individuals and society [43].

Moreover, given the inherent complexity and often black-box nature of advanced AI models, robust explainability frameworks are paramount to ensure that decisions made by AI systems are comprehensible and justifiable to human stakeholders [83]. This emphasis on explainability is critical for fostering user trust and enabling effective human oversight, particularly in high-stakes domains where autonomous decision-making could lead to significant consequences [84]. This aligns with the principle of explicability, which encompasses both epistemological intelligibility and ethical accountability, crucial for maintaining public confidence in AI technologies [85]. Ultimately, human-centered AI is about augmenting human capabilities and promoting human flourishing rather than automating human intelligence, requiring intentional design and responsible usage from all stakeholders.

This necessitates a profound shift from a purely technological lens to one that deeply embeds societal values, ethical considerations, and human well-being into the core of AI system development and deployment, prioritizing human agency and beneficial outcomes [86]. This perspective contends that the true measure of AI's success lies not in its autonomous capabilities but in its capacity to enhance human potential and decision-making processes. Consequently, this paradigm shift necessitates a rigorous framework for assessing AI systems, moving beyond mere performance metrics to include comprehensive evaluations of their societal impact, ethical implications, and alignment with human values [87]. Such a framework would move beyond conventional technical validation to incorporate multidisciplinary insights, thereby ensuring that AI serves as a force for societal good and human empowerment rather than merely an efficient tool [11].

This includes a proactive approach to anticipating and mitigating unintended consequences, emphasizing the importance of foresight in AI governance [88]. This forward-looking perspective requires ongoing research into potential future harms and benefits, ensuring that ethical considerations are integrated throughout the entire lifecycle of AI development and deployment, from conceptualization to post-deployment monitoring [14]. Furthermore, the continuous evolution of AI necessitates adaptive regulatory frameworks capable of responding to emerging challenges and opportunities, ensuring that human oversight remains viable even as AI complexity increases [89]. This involves a dynamic interplay between technological innovation and ethical deliberation, fostering a collaborative ecosystem where developers, policymakers, and end-users collectively shape the trajectory of AI to maximize societal benefit.

4. Conclusion

The pursuit of human-centered AI represents a crucial paradigm shift in the development of intelligent systems, emphasizing empowerment rather than replacement. This study underscores that ethical foundations such as fairness, accountability, transparency, and inclusivity must guide the trajectory of AI innovation to safeguard human dignity and social justice. At the same time, the tension between human empowerment and automation illustrates the urgent need to reorient AI from a paradigm of substitution toward one of augmentation, where technology enhances human capacities rather than displacing them. Principles such as human-in-the-loop, explainable AI, and participatory design emerge as actionable strategies that bridge ethical ideals with technical implementation. However, the persistence of a gap between abstract ethical guidelines and practical deployment calls for a conceptual framework that integrates normative values with engineering practices. By addressing this gap, AI can evolve into a transformative instrument that not only advances technological progress but also fosters equitable social outcomes. Ultimately, the design of human-centered AI systems is essential to ensuring that artificial intelligence contributes to human flourishing in an era of rapid digital transformation. Future research and practice in human-centered AI should prioritize the operationalization of ethical principles into measurable technical standards that can be consistently applied in system design and deployment. Collaboration between computer scientists, ethicists, policymakers, and end-users is essential to ensure that AI technologies remain socially responsive and contextually relevant. Policymakers should establish clear regulatory frameworks that

DOI: https://doi.org/ 10.31004/riggs.v4i3.2450

encourage transparency and accountability while preventing harmful practices of automation that risk displacing human labor. Developers are encouraged to adopt participatory design methodologies, ensuring that diverse perspectives inform the creation of intelligent systems. By fostering interdisciplinary collaboration and embedding ethical guidelines into technical practice, AI can evolve into a tool that strengthens human agency, enhances decision-making, and promotes social equity.

Acknowledgment

The researcher conveys sincere gratitude to Universitas Ary Ginanjar for its substantial support and assistance throughout this study. The financial contributions, provision of academic resources, and the creation of a conducive scholarly environment were instrumental in the successful completion of this research. Such support not only enabled the effective execution of the study but also demonstrated the University's steadfast dedication to advancing knowledge, fostering innovation, and nurturing the academic development of its scholars. The recognition of this research's significance, particularly in contributing to ongoing discussions on artificial intelligence and communication strategies, is deeply appreciated. The researcher acknowledges that the successful realization of the objectives and outcomes of this work would not have been possible without the invaluable encouragement and institutional commitment of Universitas Ary Ginanjar.

Reference

[1] A. N. I. Purwanto and A. A. Hanif, "Strategic Synergy: Integrating Business Management with Computer Science for Competitive Advantage," *TechComp Innov. J. Comput. Sci. Technol.*, vol. 1, no. 1, pp. 10–18, 2024, doi: https://doi.org/10.70063/techcompinnovations.v1i1.23.

[2] B. Shneiderman, "Human-Centered Artificial Intelligence: Three Fresh Ideas," *AIS Trans. Human-Computer Interact.*, pp. 109–124, 2020, doi: 10.17705/1thci.00131.

[3] C. Isensee, K. Griese, and F. Teuteberg, "Sustainable artificial intelligence: A corporate culture perspective," *Nachhalt. | Sustain. Manag. Forum*, vol. 29, pp. 217–230, 2021, doi: 10.1007/s00550-021-00524-6.

[4] B. Shneiderman, "Tutorial: Human-Centered AI: Reliable, Safe and Trustworthy," pp. 7–8, 2021, doi: 10.1145/3397482.3453994.

[5] M. Raees, I. Meijerink, I. Lykourentzou, V.-J. Khan, and K. Papagelis, "From explainable to interactive AI: A literature review on current trends in human-AI interaction," 2024, *Elsevier BV*. doi: 10.1016/j.ijhcs.2024.103301.

[6] P. A. H. Aure and O. Cuenca, "Fostering social-emotional learning through human-centered use of generative AI in business research education: an insider case study," *J. Res. Innov. Teach. Learn.*, vol. 17, no. 2, pp. 168–181, 2024, doi: 10.1108/jrit-03-2024-0076.

[7] J. Auernhammer, "Human-centered AI: The role of Human-centered Design Research in the development of AI," *Proc. DRS*, vol. 1, 2020, doi: 10.21606/drs.2020.282.

[8] A. Sigfrids, J. Leikas, H. Salo-Pöntinen, and E. Koskimies, "Human-centricity in AI governance: A systemic approach," *Front. Artif. Intell.*, vol. 6, 2023, doi: 10.3389/frai.2023.976887.

[9] N. K. Corrêa *et al.*, "Worldwide AI ethics: A review of 200 guidelines and recommendations for AI governance," 2023, *Elsevier BV*. doi: 10.1016/j.patter.2023.100857.

[10] M. Tahaei, M. Constantinides, and D. Quercia, "A Systematic Literature Review of Human-Centered, Ethical, and Responsible AI," *arXiv (Cornell Univ.*, 2023, doi: 10.48550/arxiv.2302.05284.

[11] M. Pflanzer, Z. Traylor, J. B. Lyons, V. Dubljević, and C. S. Nam, "Ethics in human–AI teaming: principles and perspectives," *AI Ethics*, vol. 3, no. 3, pp. 917–935, 2022, doi: 10.1007/s43681-022-00214-z.

[12] M. A. Mondol, M. A. Uddaula, M. S. Hossain, and M. A. Siddika, "AI-Powered Frameworks for the Detection and Prevention of Cyberbullying Across Social Media Ecosystems," *TechComp Innov. J. Comput. Sci. Technol.*, vol. 2, no. 1, pp. 1–15, 2025, doi: https://doi.org/10.70063/techcompinnovations.v2i1.59.

[13] A. Vakali and N. Tantalaki, "Rolling in the deep of cognitive and AI biases," *arXiv (Cornell Univ.*, 2024, doi: 10.48550/arxiv.2407.21202.

[14] H. Shen, T. Knearem, R. Ghosh, Y.-J. Yang, T. Mitra, and Y. Huang, "ValueCompass: A Framework of Fundamental Values for Human-AI Alignment," *arXiv (Cornell Univ.*, 2024, doi: 10.48550/arxiv.2409.09586.

[15] D. J. Larkin, L. J. Jaspersen, and K. Pandža, "Balancing anticipatory and deliberative governance in public-private partnerships for responsible innovation: The role of corporate innovation capabilities," *Organ. Stud.*, 2025, doi: 10.1177/01708406251336023.

[16] J. Ayling and A. Chapman, "Putting AI ethics to work: are the tools fit for purpose?," *AI Ethics*, vol. 2, no. 3, pp. 405–429, 2021, doi: 10.1007/s43681-021-00084-x.

[17] J. Borenstein and A. Howard, "Emerging challenges in AI and the need for AI ethics education," 2020, *Springer Nature*. doi: 10.1007/s43681-020-00002-7.

[18] B. Mbalaka, "Epistemically violent biases in artificial intelligence design: the case of DALLE-E 2 and Starry AI," *Digit. Transform. Soc.*, vol. 2, no. 4, pp. 376–402, 2023, doi: 10.1108/dts-01-2023-0003.

DOI: https://doi.org/ 10.31004/riggs.v4i3.2450

[19] D. De Cremer and D. Narayanan, "How AI tools can—and cannot—help organizations become more ethical," 2023, *Frontiers Media*. doi: 10.3389/frai.2023.1093712.

[20] D. Rossini, D. Croce, S. Mancini, M. Pellegrino, and R. Basili, "Actionable Ethics through Neural Learning," in *Proceedings of the AAAI Conference on Artificial Intelligence*, Association for the Advancement of Artificial Intelligence, 2020, pp. 5537–5544. doi: 10.1609/aaai.v34i04.6005.

[21] X. Ferrer, T. van Nuenen, J. M. Such, M. Coté, and N. Criado, "Bias and Discrimination in AI: A Cross-Disciplinary Perspective," *IEEE Technol. Soc. Mag.*, vol. 40, no. 2, pp. 72–80, 2021, doi: 10.1109/mts.2021.3056293.

[22] A. Batool, D. Zowghi, and M. Bano, "AI governance: a systematic literature review," *AI Ethics*, 2025, doi: 10.1007/s43681-024-00653-w.

[23] S. J. Gentles, C. Charles, D. Nicholas, J. Ploeg, and K. A. McKibbon, "Reviewing the research methods literature: principles and strategies illustrated by a systematic overview of sampling in qualitative research," 2016, *BioMed Central*. doi: 10.1186/s13643-016-0343-0.

[24] T. Capel and M. Brereton, "What is Human-Centered about Human-Centered AI? A Map of the Research Landscape," pp. 1–23, 2023, doi: 10.1145/3544548.3580959.

[25] Y. Nakao, L. Strappelli, S. Stumpf, A. Naseer, D. Regoli, and G. Del Gamba, "Towards Responsible AI: A Design Space Exploration of Human-Centered Artificial Intelligence User Interfaces to Investigate Fairness," *Int. J. Hum. Comput. Interact.*, vol. 39, no. 9, pp. 1762–1788, 2022, doi: 10.1080/10447318.2022.2067936.

[26] A. Sulemana, E. A. Donkor, E. K. Forkuo, and S. Oduro–Kwarteng, "Optimal Routing of Solid Waste Collection Trucks: A Review of Methods," 2018, *Hindawi Publishing Corporation*. doi: 10.1155/2018/4586376.

[27] L. Zhui, F. Li, X. Wang, Q. Fu, and R. Wei, "Ethical Considerations and Fundamental Principles of Large Language Models in Medical Education: Viewpoint," *J. Med. Internet Res.*, vol. 26, 2024, doi: 10.2196/60083.

[28] E. Prem, "From ethical AI frameworks to tools: a review of approaches," 2023, *Springer Nature*. doi: 10.1007/s43681-023-00258-9.

[29] D. K. Gao, A. Haverly, S. Mittal, J. Wu, and J. Chen, "AI Ethics: A Bibliometric Analysis, Critical Issues, and Key Gaps," *arXiv*, 2024, doi: 10.48550/ARXIV.2403.14681.

[30] P. S. Varsha, "How can we manage biases in artificial intelligence systems – A systematic literature review," *Int. J. Inf. Manag. Data Insights*, vol. 3, no. 1, p. 100165, 2023, doi: 10.1016/j.jjimei.2023.100165.

[31] A. Dhopte and H. Bagde, "Smart Smile: Revolutionizing Dentistry With Artificial Intelligence," 2023, *Cureus, Inc.* doi: 10.7759/cureus.41227.

[32] S. Kondra, S. Medapati, M. Koripalli, S. R. S. C. Nandula, and J. Zink, "AI and Diversity, Equity, and Inclusion (DEI): Examining the Potential for AI to Mitigate Bias and Promote Inclusive Communication," *J. Artif. Intell. Mach. Learn.*, vol. 3, no. 1, 2025, doi: 10.55124/jaim.v3i1.249.

[33] Y. Yue and J. Z. Shyu, "An overview of research on human-centered design in the development of artificial general intelligence," *arXiv (Cornell Univ.*, 2023, doi: 10.48550/arxiv.2309.12352.

[34] S. Schmager, I. O. Pappas, and P. Vassilakopoulou, "Understanding Human-Centred AI: a review of its defining elements and a research agenda," 2025, *Taylor & Francis*. doi: 10.1080/0144929x.2024.2448719.

[35] E. Hernández, "Towards an Ethical and Inclusive Implementation of Artificial Intelligence in Organizations: A Multidimensional Framework," *arXiv (Cornell Univ.*, 2024, doi: 10.48550/arxiv.2405.01697.

[36] A. Nastoska, B. Jancheska, M. Rizinski, and D. Trajanov, "Evaluating Trustworthiness in AI: Risks, Metrics, and Applications Across Industries," *Electronics*, vol. 14, no. 13, p. 2717, 2025, doi: 10.3390/electronics14132717.

[37] J. Morley, A. Elhalal, F. Garcia, L. Kinsey, J. Mökander, and L. Floridi, "Ethics as a Service: A Pragmatic Operationalisation of AI Ethics," *Minds Mach.*, vol. 31, no. 2, pp. 239–256, 2021, doi: 10.1007/s11023-021-09563-w.

[38] P. H. Cheong and L. Liu, "Faithful Innovation: Negotiating Institutional Logics for AI Value Alignment Among Christian Churches in America," *Religions*, vol. 16, no. 3, p. 302, 2025, doi: 10.3390/rel16030302.

[39] E. Hermann, "Artificial intelligence and mass personalization of communication content—An ethical and literacy perspective," *New Media Soc.*, vol. 24, no. 5, pp. 1258–1277, 2021, doi: 10.1177/14614448211022702.

[40] Y. Kajiwara and K. Kawabata, "AI literacy for ethical use of chatbot: Will students accept AI ethics?," *Comput. Educ. Artif. Intell.*, vol. 6, p. 100251, 2024, doi: 10.1016/j.caeai.2024.100251.

[41] P. Radanliev, "AI Ethics: Integrating Transparency, Fairness, and Privacy in AI Development," *Appl. Artif. Intell.*, vol. 39, no. 1, 2025, doi: 10.1080/08839514.2025.2463722.

[42] L. Zhu, X. Xu, Q. Lu, G. Governatori, and J. Whittle, "AI and Ethics -- Operationalising Responsible AI," *arXiv (Cornell Univ.*, 2021, doi: 10.48550/arxiv.2105.08867.

[43] A. Ferhataj, F. Memaj, R. Sahatcija, A. Ora, and E. Koka, "Ethical concerns in AI development: analyzing students' perspectives on robotics and society," *J. Inf. Commun. Ethics Soc.*, 2025, doi: 10.1108/jices-08-2024-0111.

[44] Y. Fu and Z. Weng, "Navigating the Ethical Terrain of AI in Education: A Systematic Review on Framing Responsible Human-Centered AI Practices," 2024, *Elsevier BV*. doi: 10.1016/j.caeai.2024.100306.

DOI: https://doi.org/ 10.31004/riggs.v4i3.2450

[45] Elbadiansyah, Z. H. Sain, U. S. Lawal, C. C. Thelma, and A. L. Aziz, "Exploring the Role of Artificial Intelligence in Enhancing Student Motivation and Cognitive Development in Higher Education," *TechComp Innov. J. Comput. Sci. Technol.*, vol. 1, no. 2, pp. 59–67, 2024, doi: https://doi.org/10.70063/techcompinnovations.v1i2.47.

[46] D. Patil, "Impact of Artificial Intelligence on Employment and Workforce Development: Risks, Opportunities, and Socioeconomic Implications," 2024, doi: 10.2139/ssrn.5010441.

[47] M. A. Chen and X. Wang, "Displacement or Augmentation? The Effects of AI Innovation on Workforce Dynamics and Firm Value," 2025, doi: 10.2139/ssrn.5246388.

[48] B. Lainjo, "The Role of Artificial Intelligence in Achieving the United Nations Sustainable Development Goals," *J. Sustain. Dev.*, vol. 17, no. 5, p. 30, 2024, doi: 10.5539/jsd.v17n5p30.

[49] R. Gupta, K. Nair, M. Mishra, B. Ibrahim, and S. Bhardwaj, "Adoption and impacts of generative artificial intelligence: Theoretical underpinnings and research agenda," *Int. J. Inf. Manag. Data Insights*, vol. 4, no. 1, p. 100232, 2024, doi: 10.1016/j.jjimei.2024.100232.

[50] S. Du and C. Xie, "Paradoxes of artificial intelligence in consumer markets: Ethical challenges and opportunities," *J. Bus. Res.*, vol. 129, pp. 961–974, 2020, doi: 10.1016/j.jbusres.2020.08.024.

[51] S. Hazra, B. P. Majumder, and T. Chakrabarty, "AI Safety Should Prioritize the Future of Work," 2025, doi: 10.48550/ARXIV.2504.13959.

[52] M. Özer, Y. Köse, G. Kucukkaya, A. Mukasheva, and K. Ciris, "Adapting to the AI Disruption: Reshaping the IT Landscape and Educational Paradigms," in *Communications in computer and information science*, Springer Science+Business Media, 2025, pp. 366–374. doi: 10.1007/978-3-031-85930-4_33.

[53] S. Dégallier-Rochat, M. Kurpicz-Briki, N. Endrissat, and O. Yatsenko, "Human augmentation, not replacement: A research agenda for AI and robotics in the industry," *Front. Robot. AI*, vol. 9, 2022, doi: 10.3389/frobt.2022.997386.

[54] M. Constantinides and D. Quercia, "The Impact of AI on Jobs: HCI is Missing," *arXiv (Cornell Univ.*, 2025, doi: 10.48550/arxiv.2501.18948.

[55] N. S. Azzaky, A. Salimah, and C. R. Saputri, "Revolutionizing Business: The Role of AI in Driving Industry 4.0," *TechComp Innov. J. Comput. Sci. Technol.*, vol. 1, no. 1, pp. 28–37, 2024, doi: https://doi.org/10.70063/techcompinnovations.v1i1.24.

[56] J. Shabbir and T. Anwer, "Artificial Intelligence and its Role in Near Future," *arXiv (Cornell Univ.*, 2018, doi: 10.48550/arxiv.1804.01396.

[57] J. Linares-Pellicer, J. Izquierdo-Domenech, I. Ferri-Molla, and C. Aliaga-Torro, "We Are All Creators: Generative AI, Collective Knowledge, and the Path Towards Human-AI Synergy," 2025, doi: 10.48550/ARXIV.2504.07936.

[58] J. J. Nay and J. W. Daily, "Aligning Artificial Intelligence with Humans through Public Policy," *SSRN Electron. J.*, 2022, doi: 10.2139/ssrn.4122518.

[59] J. Guszcza *et al.*, "Hybrid Intelligence: A Paradigm for More Responsible Practice," *SSRN Electron. J.*, 2022, doi: 10.2139/ssrn.4301478.

[60] J. Visave, "Transparency in AI for emergency management: building trust and accountability," *AI Ethics*, vol. 5, no. 4, pp. 3967–3980, 2025, doi: 10.1007/s43681-025-00692-x.

[61] A. A. T. J. Cassidy, A. Fuad, and M. U. A. A. Shofy, "Emerging Trends and Challenges in Digital Crime: A Study of Cyber Criminal Tactics and Countermeasures," *TechComp Innov. J. Comput. Sci. Technol.*, vol. 1, no. 1, pp. 38–45, 2024, doi: https://doi.org/10.70063/techcompinnovations.v1i1.25.

[62] C. Zhao and W. Xu, "Human-AI Interaction Design Standards," 2025, doi: 10.48550/ARXIV.2503.16472.

[63] A. Herrera-Poyatos, J. Del Ser, M. L. de Prado, F.-Y. Wang, E. Herrera-Viedma, and F. Herrera, "Responsible Artificial Intelligence Systems: A Roadmap to Society's Trust through Trustworthy AI, Auditability, Accountability, and Governance," 2025, doi: 10.48550/ARXIV.2503.04739.

[64] S. Chancellor, "Toward Practices for Human-Centered Machine Learning," *Commun. ACM*, vol. 66, no. 3, pp. 78–85, 2023, doi: 10.1145/3530987.

[65] V. Vakkuri, K. Kemell, and P. Abrahamsson, "AI Ethics in Industry: A Research Framework," *arXiv (Cornell Univ.*, 2019, doi: 10.48550/arxiv.1910.12695.

[66] A. Parthasarathy, A. Phalnikar, A. Jauhar, D. Somayajula, G. Krishnan, and B. Ravindran, "Participatory Approaches in AI Development and Governance: A Principled Approach," *arXiv (Cornell Univ.*, 2024, doi: 10.48550/arxiv.2407.13100.

[67] C. T. Okolo, "AI in the 'Real World': Examining the Impact of AI Deployment in Low-Resource Contexts," *arXiv (Cornell Univ.*, 2020, doi: 10.48550/arxiv.2012.01165.

[68] F. Delgado, S. J. H. Yang, M. Madaio, and Q. Yang, "Stakeholder Participation in AI: Beyond 'Add Diverse Stakeholders and Stir,'" *arXiv (Cornell Univ.*, 2021, doi: 10.48550/arxiv.2111.01122.

[69] M. Young *et al.*, "Participation versus scale: Tensions in the practical demands on participatory AI," *First Monday*, 2024, doi: 10.5210/fm.v29i4.13642.

[70] F. Delgado, S. Yang, M. Madaio, and Q. Yang, "The Participatory Turn in AI Design: Theoretical Foundations and the Current State of Practice," pp. 1–23, 2023, doi: 10.1145/3617694.3623261.

DOI: https://doi.org/ 10.31004/riggs.v4i3.2450

[71] A. Birhane *et al.*, "Power to the People? Opportunities and Challenges for Participatory AI," pp. 1–8, 2022, doi: 10.1145/3551624.3555290.

[72] L. Ajmani, N. A. Abdelkadir, and S. Chancellor, "Secondary Stakeholders in AI: Fighting for, Brokering, and Navigating Agency," pp. 1095–1107, 2025, doi: 10.1145/3715275.3732071.

[73] R. Sieber, A. Brandusescu, S. Sangiambut, and A. Adu-Daako, "What is civic participation in artificial intelligence?," *Environ. Plan. B Urban Anal. City Sci.*, 2024, doi: 10.1177/23998083241296200.

[74] A. Buhmann and C. Fieseler, "Deep Learning Meets Deep Democracy: Deliberative Governance and Responsible Innovation in Artificial Intelligence," *Bus. Ethics Q.*, vol. 33, no. 1, pp. 146–179, 2022, doi: 10.1017/beq.2021.42.

[75] A. N. I. Purwanto, M. Fauzan, T. Widya, and N. S. Azzaky, "Ethical Implications and Challenges of AI Implementation in Business Operations," *TechComp Innov. J. Comput. Sci. Technol.*, vol. 1, no. 2, pp. 68–82, 2024, doi: https://doi.org/10.70063/techcompinnovations.v1i2.52.

[76] D. Peters, K. Vold, D. Robinson, and R. A. Calvo, "Responsible AI—Two Frameworks for Ethical Design Practice," *IEEE Trans. Technol. Soc.*, vol. 1, no. 1, pp. 34–47, 2020, doi: 10.1109/tts.2020.2974991.

[77] A. V. Abdussalam and G. A. Auladi, "Pushing Boundaries: AI and Computer Science in the Era of Technological Revolution," *TechComp Innov. J. Comput. Sci. Technol.*, vol. 1, no. 1, pp. 01–09, 2024, doi: https://doi.org/10.70063/techcompinnovations.v1i1.22.

[78] C. Burr and D. Leslie, "Ethical assurance: a practical approach to the responsible design, development, and deployment of data-driven technologies," *AI Ethics*, vol. 3, no. 1, pp. 73–98, 2022, doi: 10.1007/s43681-022-00178-0.

[79] G. A. Auladi and F. Muwahid, "Cybersecurity and Moral Responsibility: A Philosophical-Islamic Approach to Digital Trust," *TechComp Innov. J. Comput. Sci. Technol.*, vol. 2, no. 1, pp. 53–65, 2025, doi: https://doi.org/10.70063/techcompinnovations.v2i1.93.

[80] D. Schiff, B. Rakova, A. Ayesh, A. Fanti, and M. Lennon, "Principles to Practices for Responsible AI: Closing the Gap," *arXiv (Cornell Univ.*, 2020, doi: 10.48550/arxiv.2006.04707.

[81] T. Züger and H. Asghari, "AI for the public. How public interest theory shifts the discourse on AI," *AI Soc.*, vol. 38, no. 2, pp. 815–828, 2022, doi: 10.1007/s00146-022-01480-5.

[82] A. Toader, "Auditability of AI systems – brake or acceleration to innovation?," *SSRN Electron. J.*, 2019, doi: 10.2139/ssrn.3526222.

[83] M. A. Qasimi, "Solving the Travelling Salesman Problem Using the PSO Optimization," *TechComp Innov. J. Comput. Sci. Technol.*, vol. 1, no. 1, pp. 19–27, 2024, doi: https://doi.org/10.70063/techcompinnovations.v1i1.20.

[84] J. J. Cañas, "AI and Ethics When Human Beings Collaborate With AI Agents," 2022, *Frontiers Media*. doi: 10.3389/fpsyg.2022.836650.

[85] L. Floridi and J. Cowls, "A Unified Framework of Five Principles for AI in Society," *Harvard data Sci. Rev.*, 2019, doi: 10.1162/99608f92.8cd550d1.

[86] V. Dignum, "Responsible Autonomy," *arXiv (Cornell Univ.*, 2017, doi: 10.48550/arxiv.1706.02513.

[87] M. Khamassi, M. Nahon, and R. Chatila, "Strong and weak alignment of large language models with human values," *Sci. Rep.*, vol. 14, no. 1, 2024, doi: 10.1038/s41598-024-70031-3.

[88] J. R. Prasmanto, A. Ma'ruf, M. F. Assidiq, and M. F. Diyaulhaq, "Managing Digital Economic Corridors: International Cooperation and Regional Integration in the Digital Age," *TechComp Innov. J. Comput. Sci. Technol.*, vol. 2, no. 1, pp. 39–52, 2025, doi: https://doi.org/10.70063/techcompinnovations.v2i1.63.

[89] A. Holzinger, K. Zatloukal, and H. Müller, "Is Human Oversight to AI Systems still possible?," *N. Biotechnol.*, 2024, doi: 10.1016/j.nbt.2024.12.003.

DOI: https://doi.org/ 10.31004/riggs.v4i3.2450