The Eastasouth Journal of Information System and Computer Science
Vol. 3, No. 02, December 2025, pp. 262-270
ISSN: 3025-566X, DOI: 10.58812/esiscs.v3i02

Analysis of the Moral Obligations of AI Developers Thru the Principle of Explainability in the Perspective of Kantian Deontological Ethics: A Qualitative Study

Rizma Fauziyah[1], Agung Winarno[2], Subagyo[3]
[1] Universitas Negeri Malang
[2] Universitas Negeri Malang
[3] Universitas Negeri Malang

Article Info

Article history:

Received Nov, 2025
Revised Dec, 2025
Accepted Dec, 2025

Keywords:

AI Ethics;
Explainable AI (XAI);
Moral Obligation

ABSTRACT

The proliferation of "Black Box" Artificial Intelligence systems creates a significant ethical void regarding accountability and user autonomy, fundamentally challenging the right of individuals to understand decisions affecting their lives. This study aims to analyze the moral obligations of AI developers to implement Explainability (XAI) using the rigorous normative framework of Kantian Deontological Ethics. Employing a qualitative research design with conceptual analysis, the study utilizes secondary data from Kant's foundational texts and contemporary literature on algorithmic transparency, applying the Categorical Imperative as the primary lens. The findings conclude that the deployment of non-explainable AI constitutes a direct violation of Kant's Formula of Humanity, as it reduces users merely to means for achieving computational goals rather than treating them as autonomous, rational agents. Furthermore, the practice fails the Universal Law test, which prohibits the universalization of opacity in decision-making processes. Consequently, the study asserts that Explainability is a non-negotiable moral duty for developers, establishing that predictive accuracy cannot ethically justify the erosion of human autonomy, thereby demanding a paradigm shift from utilitarian efficiency to deontological adherence in AI development.

Corresponding Author:

Name: Agung Winarno
Institution: Universitas Negeri Malang
Email: agung.winarno.fe@um.ac.id

1. INTRODUCTION

The development of artificial intelligence (AI) technology in the last decade has surpassed the initial predictions of experts and is now a fundamental infrastructure for various sectors of human life. From complex medical diagnoses and the criminal justice system to financial algorithms, AI is taking over decision-making roles previously held only by humans. In this new digital ecosystem, technology developers and software engineers play a crucial role as architects of modern civilization [1]. They not only write code for purely technical functions, but also indirectly design the values embedded in systems that interact with the wider community. However, the magnitude of this technical power is often not balanced by a full awareness of the moral impact it has on end-users. As AI systems become

increasingly autonomous, the question of creators' responsibility for their creations becomes increasingly urgent to answer. The speed of technological innovation often far outpaces the evolution of the ethical frameworks that guide it, creating a dangerous gap. Therefore, positioning developers as the primary moral subject is a vital initial step in contemporary technological discourse.

One of the biggest challenges that has emerged with the advancement of Deep Learning and Neural Network methods is the phenomenon known as the "Black Box"[2]. In this system, algorithms are capable of processing massive amounts of data and generating decisions with high accuracy, but the internal process by which these decisions are reached is often incomprehensible even to their creators. The complexity of the layered computational layers makes the decision-making logic flow obscure and opaque to human observers. This opacity creates a fundamental problem when AI systems are used for decisions that affect someone's life, such as loan approvals or legal verdicts. When a system works effectively but is not transparent, users are forced to place blind trust without a clear rational basis. This creates extreme information asymmetry between technology providers and users affected by those technology decisions. This situation triggers a technical dilemma where developers often have to choose between high accuracy performance or lower model transparency. However, allowing this ambiguity to continue without intervention means normalizing ignorance in a crucial process of human life.

Responding to the dangers of the Black Box phenomenon, the concept of Explainability or eXplainable Artificial Intelligence (XAI) has emerged as a key requirement in the development of responsible intelligent systems [3]. Explainability is not just an additional feature to facilitate technical debugging; it is a communication bridge that allows users to understand the reasoning behind the algorithm's output. Without the ability to explain the decision-making process, an AI system could potentially perpetuate hidden biases and systemic errors that are difficult to detect and correct. The need for this explanation becomes extremely vital because humans have a fundamental need to understand the causality of events that befall them in order to maintain a sense of justice. If an individual is denied access to healthcare services by AI without adequate explanation, their right to object or understand the situation is lost. Thus, Explainability is transforming from a mere technical issue into a serious human rights and social justice concern [4]. The absence of adequate explanation can erode public trust in technology and trigger widespread social rejection. Therefore, this principle must be seen as a fundamental and non-negotiable element in modern intelligent system architecture.

For a long time, the evaluation of AI development has often been dominated by a Utilitarian perspective that focuses on the final outcome and maximum efficiency [5]. In this view, an algorithm is considered "good" as long as it produces the greatest benefit for the most people, even if the process within it is not transparent or sacrifices the understanding of a few individuals. This consequence-based approach often justifies the use of Black Boxes in the name of accuracy and speed, disregarding the individual's right to be treated with respect thru transparency. However, this purely results-oriented approach has a moral flaw because it fails to protect human autonomy and dignity in the process. A paradigm shift in ethics is needed, emphasizing obligations and binding principles, regardless of how beneficial the final outcome is. This is where Deontological ethics offers a more robust framework for demanding moral accountability from developers. The focus must shift from "what technology can do" to "what developers should do" as a professional and moral obligation.

Deontological ethics, as proposed by Immanuel Kant, provides a highly relevant philosophical foundation for dissecting developers' moral obligations to provide explainability [6]. Kant's central concept of the

Categorical Imperative, particularly the humanity formulation, asserts that humans must always be treated as ends in themselves and never merely as means or tools. In the context of AI, hiding decision logic from users is equivalent to treating them solely as passive data objects, which violates their autonomy and rationality as human beings. Kant emphasized that rational beings have the right to understand the rules that bind them, making transparency an absolute condition for respecting human dignity. The developer's obligation, in this view, is to ensure that their creations do not diminish human capacity to think and act autonomously[7]. Therefore, from a Kantian perspective, explainability is not seen as an option, but as an absolute moral duty that must be fulfilled unconditionally. Applying this classical theory to modern technological problems offers new insights into the ethical boundaries that must be adhered to in innovation.

Based on this background and theoretical framework, this qualitative research aims to deeply analyze the moral obligations of AI developers thru the principle of Explainability from a Kantian Deontological ethics perspective. This study will explore how the system's inability to be explained violates universal moral principles and diminishes the essence of the user's humanity. Thru this critical analysis, this article aims to fill the literature gap, which has historically focused more on legal and technical aspects rather than purely philosophical obligations. This research will argue that algorithmic transparency is a manifestation of respect for human autonomy that should not be compromised for the sake of efficiency. It is hoped that the results of this study can provide a strong ethical foundation for developers and policymakers to prioritize Explainability as a standard moral norm. Ultimately, this research aims to emphasize that advancements in AI technology must go hand in hand with the preservation of noble human values as outlined in deontological ethics. This contribution is expected to reform the technology industry's perspective on their ethical responsibilities in the future.

2. THEORETICAL REVIEW

2.1 Moral Obligations of AI Developers

Technology developers can no longer be seen as neutral technicians who simply execute mathematical instructions; they are moral agents who control the values embedded in algorithmic systems. In the context of professional ethics, the moral obligations of developers extend beyond mere compliance with positive law or company standards. This obligation includes the inherent responsibility to ensure that the systems they create do not harm, discriminate against, or manipulate users. Because algorithms have the power to influence human life decisions, from financial access to legal verdicts, every line of code written is a manifestation of moral actions with real consequences[8].

Therefore, the moral obligation of developers should be understood as a proactive, not reactive, responsibility. They are obligated to anticipate potential ethical failures, such as bias or system closure, from the design phase. Failure to instill ethical considerations in technical architecture is not merely a technical error (bug), but a moral omission. Developers bear the responsibility of bridging the gap between machine complexity and human safety, ensuring that the technological power they build remains subject to human values.

2.2 The Principle of Explainability

The principle of Explainability (often referred to as XAI) emerged as a critical response to the "Black Box" phenomenon in Deep Learning, where the decision-making processes of algorithms become opaque to humans. Conceptually, Explainability is defined as the ability of a system to present the reasoning behind its decisions in a format that is understandable by human reasoning [9]. This is different from mere code transparency; explainability demands interpretability, which is an

explanation of the causal relationship of "why" a specific input produces a specific output.

In the hierarchy of technology ethics, this principle serves as a prerequisite for accountability and trust. Without the ability to be explained, an AI decision, no matter how statistically accurate, becomes unchallengeable and its fairness cannot be audited. Explainability gives users the right to know the rational basis for the treatment they receive from machines [10]. The absence of this principle leaves users blind and vulnerable, forcing them to submit to algorithmic authority without access to the logic that governs it.

2.3 Perspective of Kantian Deontological Ethics

Deontological ethics, as proposed by Immanuel Kant, emphasizes morality based on absolute duty and rules, regardless of the consequences or final outcome of the action [11]. The core of this theory is the "Categorical Imperative," an unconditional moral command that every rational agent must obey [12]. In the context of this research, Kant's most relevant formulation is the "Humanity Formula," which states: "Act in such a way that you treat humanity, both in yourself and in others, always as an end and never merely as a means." [13].

This perspective places human autonomy and rationality as the highest values to be respected. According to Kant, treating humans as a "means" means using them without respecting their ability to give rational consent [14]. Additionally, Kant's principle of "Universal Law" demands logical consistency; an action is only moral if the maxim (principle) behind it can be applied universally without contradiction. This framework provides a rigorous foundation for testing whether non-transparent AI development practices violate developers' moral obligations to the rational dignity of their users.

3. METHODOLOGY

This study employs a qualitative research design rooted in normative philosophical analysis to critically examine the ethical obligations of AI developers. By utilizing a conceptual analysis method, the research bridges technical concepts of *Explainable AI* (XAI) with the philosophical framework of Kantian Deontology. Data sources are derived from a systematic review of secondary literature, comprising foundational texts such as Kant's *Groundwork of the Metaphysics of Morals* and contemporary discourse on algorithmic transparency obtained from academic databases like Scopus and IEEE Xplore. The collected data is subjected to deductive thematic analysis, where the technical limitations of "Black Box" AI are rigorously tested against Kant's Categorical Imperative, specifically the Formulas of Universal Law and Humanity, to synthesize a normative argument establishing explainability as an inherent moral duty.

4. RESULT AND DISCUSSION

4.1 The Moral Obligations of AI Developers: From Technicians to Moral Agents

The first critical finding of this study redefines the role of the AI developer, shifting the perspective from a neutral technician to an active moral agent. In the contemporary technological landscape, developers often view their obligations through a technical or legal lens, prioritizing code efficiency and regulatory compliance [15]. However, from a qualitative analysis of professional ethics, this view is insufficient. The developer is the architect of the digital environment in which modern humans operate. Every decision to prioritize algorithmic accuracy over transparency is an active moral choice that affects the agency of the user. Therefore, the moral obligation of the developer is inherent to their profession; they are not merely writing instructions for machines, but legislating the rules of interaction for society. This obligation exists independently of external laws, emerging

instead from the profound power imbalance between the creator of the system and the user subject to it.

Furthermore, this obligation is often obscured by the "problem of many hands" or the complexity of corporate structures, where individual developers feel detached from the final social impact of their code. However, the analysis suggests that moral responsibility cannot be diluted by organizational complexity [16]. The developer possesses unique epistemic access to the system's architecture that the end-user lacks. By releasing a system into the public domain, the developer asserts a claim about its safety and fairness. If this system is fundamentally flawed due to a lack of transparency, the moral burden lies with the creator. The obligation here is proactive rather than reactive; it requires the developer to anticipate the potential for moral harm, specifically the harm of manipulation or confusion, and to mitigate it through deliberate design choices before the system is ever deployed.

Finally, contrasting this with utilitarian ethics, the developer's obligation is not satisfied merely by producing "beneficial results" for the majority. A developer might argue that a complex, unexplainable algorithm cures more diseases or catches more criminals, thus justifying its opacity. However, this study posits that such a justification fails to meet the standard of moral obligation. A beneficial outcome does not absolve the developer of the duty to respect the process of interaction. The obligation is to ensure that the relationship between the human and the machine remains one of mastery and understanding, not subjugation [17]. Thus, the primary moral duty of the AI developer is to preserve the integrity of the user's experience, ensuring that the technology serves as a tool for human empowerment rather than a mechanism of obscure control [18].

4.2 The Principle of Explainability: Safeguarding the Formula of Humanity

The application of the Principle of Explainability is most critically understood through Kant's "Formula of Humanity," which dictates that humans must always be treated as ends in themselves, never merely as means [19]. "Black Box" AI systems, which produce decisions without accessible logic, fundamentally violate this principle. When a user is subjected to a decision, such as a credit denial or a medical triage outcome, without an explanation, their rational capacity is denied. They are reduced to data points to be processed, optimized, and sorted by the system. In this dynamic, the user becomes a means to achieve the system's goal (efficiency, profit, or speed), stripping them of the dignity that comes with understanding and consenting to the forces that govern their lives.

Explainability, therefore, is not just a technical feature for debugging, but the essential bridge that restores the "end" status of the human user. For a human to act autonomously, they must understand the environment in which they act. If the environment is governed by an opaque algorithm, the human acts in blindness, effectively coerced by the machine. Explainable AI (XAI) provides the necessary rationale that allows the user to engage with the decision critically, to accept it, challenge it, or learn from it [20]. By revealing the "why" behind an output, the system acknowledges the user's intelligence and right to know. This transparency transforms the interaction from a unilateral imposition of will (by the machine/developer) to a bilateral exchange of information, preserving the respect required by the Formula of Humanity.

Consequently, the absence of explainability in high-stakes AI constitutes a form of dehumanization. To uphold the Principle of Explainability is to reject this hierarchy. It asserts that no computational efficiency is worth the cost of treating a human being as an object. Thus, implementing XAI is the practical

method by which developers fulfill their duty to respect the inherent worth of the user, ensuring that technology remains a support system for human rationality rather than a substitute for it.

4.3 Perspective of Kantian Deontological Ethics: The Test of Universal Law

The final dimension of this discussion evaluates the lack of transparency through Kant's "Formula of Universal Law," which demands that one should act only according to that maxim whereby one can, at the same time, will that it should become a universal law [21]. If we attempt to universalize the maxim "developers may create opaque systems to maximize performance," we encounter a logical contradiction. If every system, legal, medical, financial, and educational. were to operate as a Black Box where reasons are hidden, the very concept of trust and accountability would collapse. A society cannot function if its fundamental decision-making processes are unintelligible. In such a world, justice becomes random, and medical care becomes arbitrary, rendering the social contract void. Since a rational being cannot will a world where reason is systematically suppressed, the creation of non-explainable AI is logically and morally impermissible.

This Deontological perspective stands in stark contrast to the prevailing consequentialist logic of the tech industry. While industry leaders often argue that the "ends" (high accuracy, rapid innovation) justify the "means" (complexity, opacity), Kantian ethics rigorously rejects this trade-off. Under the Universal Law, the intention and the nature of the act itself define its morality. The act of deploying a system that cannot be scrutinized is an act of deception or negligence, regardless of how well the system performs 99% of the time. The 1% of error in a Black Box system is not just a statistical anomaly; it is a moral failure because it cannot be rectified or understood by the victim. Therefore, the "duty" to explain is a Perfect Duty, it is absolute and admits no exceptions based on convenience or profit.

In synthesis, the Kantian perspective establishes that Explainability is the boundary line for ethical AI development [22]. It serves as a check against the hubris of technological acceleration. If a system is too complex to be explained to the humans it affects, then, according to this ethical framework, it is too complex to be deployed morally. The analysis concludes that the imperative of the developer is to constrain the complexity of AI within the limits of human understanding. By adhering to the Universal Law, developers ensure that the digital future remains a "Kingdom of Ends," where technology is designed to operate within a framework of universalizable transparency, ensuring justice and rationality are preserved for all members of the community [23].

5. CONCLUSION

This study has rigorously analyzed the role of AI developers through the lens of Kantian Deontological ethics, concluding that the implementation of *Explainability* is not a mere technical option but a fundamental moral imperative. The investigation reveals that the widespread deployment of "Black Box" algorithms constitutes a direct violation of the *Categorical Imperative*, specifically the *Formula of Humanity*, by reducing rational human users to mere passive data points. By withholding the logic behind algorithmic decisions, developers inadvertently strip users of their autonomy, treating them as means to achieve computational efficiency rather than as ends in themselves. Furthermore, the application of the *Formula of Universal Law* demonstrates that a maxim permitting opaque decision-making cannot be universalized without destroying the essential trust required for societal function. The findings challenge the prevailing utilitarian narrative in the technology sector, arguing that high predictive accuracy does not ethically justify the erosion of human agency and transparency. Consequently, the

study establishes that the moral worth of an AI system is contingent upon its interpretability, rendering unexplainable systems ethically impermissible in high-stakes environments. This shifts the burden of responsibility squarely onto the developers, who must act not just as engineers but as moral agents bound by a "perfect duty" to truth and clarity. Ultimately, the research affirms that the dignity of the human subject must remain the supreme condition of all technological development, demanding that innovation operates within the boundaries of human understanding. Thus, *Explainability* serves as the critical ethical safeguard that prevents the domination of machine logic over human reason.

RECOMMENDATIONS

Based on the normative conclusions drawn from this research, several critical recommendations are proposed to realign AI development with deontological ethical standards. Foremost, AI developers and software engineers are urged to adopt an "Ethics by Design" methodology, where constraints regarding *Explainability* are integrated into the architectural phase rather than treated as retrospective features. Technical teams should prioritize the use of interpretable models, such as Decision Trees or Linear Regression, over opaque Deep Neural Networks in critical sectors like healthcare and criminal justice, even if it entails a marginal trade-off in accuracy. Corporate entities must revise their internal codes of conduct to explicitly recognize the preservation of user autonomy as a primary professional obligation, superior to commercial metrics of speed or efficiency. For policymakers, it is recommended to move beyond general data privacy laws and enact specific "Right to Explanation" mandates that legally prohibit the use of non-interpretable algorithms in public services. Regulatory bodies should establish standardized thresholds for algorithmic transparency that must be met before any automated system is permitted to enter the market. Furthermore, engineering education institutions should reform computer science curricula to include rigorous training in moral philosophy, ensuring future developers understand the weight of their agency. Industry leaders are encouraged to foster a culture where ethical scrutiny is rewarded, empowering engineers to halt projects that fail to meet the standard of the *Formula of Humanity*. By implementing these structural changes, the technology sector can ensure that its advancements remain subservient to the moral rights of the global community.

FUTURE STUDY

While this study provides a robust theoretical framework based on Western deontological ethics, there are several avenues for further research to expand the understanding of algorithmic morality. Future investigations should aim to bridge the gap between theory and practice by conducting empirical case studies that test the feasibility of implementing Kantian constraints in real-world software development environments. It would be particularly valuable to explore how these ethical obligations resonate within different cultural frameworks, such as Confucianism or Ubuntu, which may offer alternative perspectives on the relationship between individual rights and communal technological benefits. Additionally, as Artificial Intelligence evolves into Generative AI and Large Language Models, the technical definition of "explanation" becomes increasingly complex, requiring new philosophical inquiries into the nature of intent and hallucination. Researchers are encouraged to examine the user's perspective through qualitative surveys to determine if technically "explainable" outputs actually succeed in restoring the user's subjective sense of autonomy and trust. Comparative studies could also be conducted to analyze the divergent ethical standards applied in different high-stakes sectors, distinguishing the moral weight of errors in medical diagnosis versus financial lending. Further work is needed to develop concrete metrics or "auditing tools" that can operationalize

Kantian principles into quantifiable standards for engineering quality assurance. Finally, interdisciplinary collaboration between moral philosophers and computer scientists is essential to refine the technical mechanisms of XAI so they align more precisely with ethical requirements. These future endeavors will be crucial in creating a universal ethical grammar for the digital age that transcends specific technological iterations.

ACKNOWLEDGMENT

The author wishes to express profound gratitude to the anonymous reviewers whose insightful critiques and constructive suggestions significantly elevated the theoretical rigor and clarity of this article. Sincere appreciation is extended to the Faculty of Economics and Business at State University of Malang for providing the essential academic environment and access to digital repositories that facilitated the extensive literature review required for this study. Gratitude is further expressed to the editorial team of the journal for their professional handling of the manuscript and their commitment to publishing qualitative research in the field of technology ethics. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors, ensuring the independence of the analysis. The author remains solely responsible for any errors or omissions that may remain in the final text despite the rigorous review process. Finally, the author thanks the broader academic community for the ongoing dialogue regarding AI ethics, which serves as a constant inspiration for this work.

REFERENCES

[1] Karthik Akinepalli, "The Pervasive Impact of Software Engineering and Architecture: A Multi-Industry Analysis," *Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol.*, vol. 10, no. 6, pp. 279–289, 2024, doi: 10.32628/cseit241051082.

[2] R. Guidotti, A. Monreale, S. Ruggieri, F. Turini, F. Giannotti, and D. Pedreschi, "A Survey of Methods for Explaining Black Box Models," *ACM Comput. Surv.*, vol. 51, no. 5, Aug. 2018, doi: 10.1145/3236009.

[3] A. Barredo Arrieta *et al.*, "Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI," *Inf. Fusion*, vol. 58, pp. 82–115, 2020, doi: 10.1016/j.inffus.2019.12.012.

[4] M. K. Land and J. D. Aronson, "Human Rights and Technology: New Challenges for Justice and Accountability," *Annu. Rev. Law Soc. Sci.*, vol. 16, no. Volume 16, 2020, pp. 223–240, 2020, doi: https://doi.org/10.1146/annurev-lawsocsci-060220-081955.

[5] M. Mori, S. Sassetti, V. Cavaliere, and M. Bonti, "A systematic literature review on artificial intelligence in recruiting and selection: a matter of ethics," *Pers. Rev.*, vol. 54, no. 3, pp. 854–878, 2025, doi: 10.1108/PR-03-2023-0257.

[6] A. Jedličková, "Ethical approaches in designing autonomous and intelligent systems: a comprehensive survey towards responsible development," *AI Soc.*, vol. 40, no. 4, pp. 2703–2716, 2025, doi: 10.1007/s00146-024-02040-9.

[7] C. O. Abakare, "Kantian Ethics And The Hesc Research: A Philosophical Exploration," *Predestinasi*, vol. 13, no. 2, p. 79, 2021, doi: 10.26858/predestinasi.v13i2.19534.

[8] T. A. Griffin, B. P. Green, and J. V. M. Welie, "The ethical agency of AI developers," *AI Ethics*, vol. 4, no. 2, pp. 179–188, 2024, doi: 10.1007/s43681-022-00256-3.

[9] K. Sokol and P. Flach, "Explainability fact sheets: A framework for systematic assessment of explainable approaches," *FAT* 2020 - Proc. 2020 Conf. Fairness, Accountability, Transpar.*, pp. 56–67, 2020, doi: 10.1145/3351095.3372870.

[10] A. B. Haque, A. K. M. N. Islam, and P. Mikalef, "Explainable Artificial Intelligence (XAI) from a user perspective: A synthesis of prior literature and problematizing avenues for future research," *Technol. Forecast. Soc. Change*, vol. 186, no. PA, p. 122120, 2023, doi: 10.1016/j.techfore.2022.122120.

[11] R. Chaddha and G. Agrawal, "Ethics and Morality," *Indian J. Orthop.*, vol. 57, no. 11, pp. 1707–1713, 2023, doi: 10.1007/s43465-023-01004-3.

[12] A. Benlahcene, R. Zainuddin, Bin, N. Syakiran, and B. A. Ismail, "A Narrative Review of Ethics Theories: Teleological & Deontological Ethics," *J. Humanit. Soc. Sci.*, vol. 23, no. 1, pp. 31–32, 2018, doi: 10.9790/0837-2307063138.

[13] T. E. Hill Jr, "Humanity as an End in Itself.," *Ethics*, vol. 91, no. 1, pp. 84–99, 1980.

[14] P. Kleingeld, "Contradiction and Kant's Formula of Universal Law," *Kant-Studien*, vol. 108, no. 1, pp. 89–115, 2017, doi: 10.1515/kant-2017-0006.

[15] O. Akpobome, "The Impact of Emerging Technologies on Legal Frameworks: A Model for Adaptive Regulation".

[16] N. Brunsson, I. Gustafsson Nordin, and K. Tamm Hallström, "'Un-responsible' Organization: How More Organization Produces Less Responsibility," *Organ. Theory*, vol. 3, no. 4, 2022, doi: 10.1177/26317877221131582.

[17] D. Silva and L. Cunha, "'Between me and the machine': on the apparent suppression of embodied know-how in the automated world and its implications," *Soc. Sci. Humanit. Open*, vol. 12, no. March, p. 102227, 2025, doi: 10.1016/j.ssaho.2025.102227.

[18] M. Ryan and B. C. Stahl, "Artificial intelligence ethics guidelines for developers and users: clarifying their content and normative implications," *J. Information, Commun. Ethics Soc.*, vol. 19, no. 1, pp. 61–86, 2021, doi: 10.1108/JICES-12-2019-0138.

[19] P. Formosa, "Dignity and respect: How to apply kant's formula of humanity," *Philos. Forum*, vol. 45, no. 1, pp. 49–68, 2014, doi: 10.1111/phil.12026.

[20] R. Dwivedi, R., Dave, D., Naik, H., Singhal, S., Omer, R., Patel, P., ... & Ranjan, "Explainable AI (XAI): Core ideas, techniques, and solutions," *ACM Comput. Surv.*, vol. 55, no. 9, pp. 1–33, 2023.

[21] M. Braham and M. van Hees, "The Formula of Universal Law: A Reconstruction," *Erkenntnis*, vol. 80, no. 2, pp. 243–260, 2015, doi: 10.1007/s10670-014-9624-y.

[22] C. Peterson and J. Broersen, "Understanding the Limits of Explainable Ethical AI," *Int. J. Artif. Intell. Tools*, vol. 33, no. 3, pp. 1–24, 2024, doi: 10.1142/S0218213024600017.

[23] B. Rasiklal Yadav, "The Ethics of Understanding: Exploring Moral Implications of Explainable AI," *Int. J. Sci. Res.*, vol. 13, no. 6, pp. 1–7, 2024, doi: 10.21275/sr24529122811.