Vol. 2 No. 2 2025 || E-ISSN 2997-7258

The Journal of Academic Science

journal homepage:
https://thejoas.com/index.php/

The Intersection of Ethics and Artificial Intelligence: A Philosophical Study

Jakoep Ezra Harianto
Sekolah Tinggi Teologi Lighthouse Equipping Theological School, Indonesia
Email: dr.jezra@gmail.com

KEY WORDS

Ethics, Artificial Intelligence, Philosophy, Moral Responsibility, Algorithmic Bias

ABSTRACT

The rapid development of artificial intelligence (AI) has had a significant impact on various aspects of human life, ranging from the economy, education, to health. However, these advances also raise complex ethical challenges, such as privacy concerns, algorithmic bias, moral responsibility, and potential misuse of technology. This research aims to explore the intersection between ethics and artificial intelligence through a philosophical approach. The method used in this study is qualitative with literature study (library research), examining various classical and contemporary ethical theories and their application in the context of AI development. The results of the study show that AI presents a new moral dilemma that cannot be fully answered by traditional ethical frameworks. For example, the concept of responsibility in AI becomes blurred when decisions are taken by autonomous systems without human intervention. Additionally, bias in AI training data indicates the need for strict ethical oversight in the design and implementation process of this technology. The study also highlights the need for a multidisciplinary approach in drafting ethical guidelines that are able to accommodate future AI developments. Thus, this research is expected to contribute to enriching the discourse on AI ethics and offering a deeper philosophical perspective in understanding the moral challenges faced.

1. Introduction

The development of artificial intelligence (AI) technology in the last decade has changed various aspects of human life, ranging from the economy, education, health, to the government system (Russell & Norvig, 2021). AI has the ability to process large amounts of data, make predictions, and even make decisions autonomously, which previously could only be done by humans (Goodfellow, Bengio, & Courville, 2016). Although this technology brings great benefits, there are also significant ethical challenges, such as algorithmic bias, privacy violations, and the potential loss of human autonomy in decision-making (Floridi et al., 2018).

The intersection between ethics and artificial intelligence (AI) is becoming increasingly relevant as the use of AI in various aspects of human life increases. AI is now not only limited to technical applications, but has also entered the realm of influencing moral decisions, such as in the judiciary, health, and public safety systems (Floridi et al., 2018). This is where the ethical challenge arises: how to ensure that the decisions made by machines remain in line with the moral values held by society? For example, algorithms used in labor recruitment or credit scoring systems can inadvertently reinforce biases present in the data, resulting in unconscious discrimination (O'Neil, 2016). This issue requires a clear ethical framework to govern how AI is developed and applied.

From the perspective of moral philosophy, artificial intelligence challenges many traditional ethical concepts, especially related to responsibility and autonomy. In conventional systems, responsibility for a decision can be clearly directed to a specific individual or institution. However, in the context of AI, when decisions are made by autonomous algorithms, the question arises: who is responsible for the errors or negative impacts caused by the system? (Bostrom & Yudkowsky, 2014). In addition, there is debate about the extent to which AI can be considered to have moral autonomy. Does a system capable of learning and adapting independently have the capacity to make ethical decisions, or is it just a tool that expands the human will? These questions are at the heart of philosophical discussions about AI ethics.

Furthermore, the relationship between ethics and AI also touches on broader global issues, such as unequal access to technology, the environmental impact of AI infrastructure, and the potential misuse of AI for unethical purposes, such as mass surveillance or information manipulation (Zuboff, 2019). Therefore, the discussion about AI ethics revolves not only on how this technology is used, but also on how it was designed from the beginning. A multidisciplinary approach involving philosophers, computer scientists, policymakers, and civil society is needed to create ethical guidelines that not only govern the use of AI, but also ensure that these technologies are developed with social, political, and environmental impacts in mind holistically.

Most research on AI ethics focuses on technical aspects, such as algorithmic bias mitigation or personal data protection (Crawford, 2021; O'Neil, 2016). However, in-depth philosophical studies of how classical ethical principles, such as deontology or utilitarianism, can be applied or adapted in the context of AI are still limited (Bostrom & Yudkowsky, 2014). This research aims to fill this gap by offering a comprehensive philosophical perspective on the moral dilemmas that arise from the use of AI.

The urgency of this research lies in the fact that AI is increasingly being used in decision-making that has a direct impact on human life, such as in the justice system, health, and labor recruitment (Eubanks, 2018). Without clear ethical guidelines, there is a risk that these technologies could reinforce injustices or even violate human rights (Zuboff, 2019). Therefore, a philosophical approach is needed to understand more deeply the moral principles that should guide the development and application of AI.

Several previous studies have discussed ethics in AI development. Floridi et al. (2018) proposed ethical

principles for AI that focus on transparency, fairness, and non-discrimination. Bostrom & Yudkowsky (2014) discuss the potential existential risks of AI superintelligence, while Eubanks (2018) highlight how algorithms can reinforce social injustice. While this contribution is important, studies that explicitly link classical ethical theories to AI development practices are still rare.

The novelty of this study lies in the philosophical approach used to analyze ethical dilemmas in AI. This research not only discusses how ethics are applied in the development of AI, but also explores how AI challenges traditional ethical concepts, such as moral responsibility and human autonomy (Coeckelbergh, 2020). Thus, this research offers a new perspective that can enrich discussions about AI ethics.

This research aims to critically analyze the intersection between ethics and AI, focusing on how classical ethical principles can be applied or adapted in the context of modern technology. The benefits of this research are that it makes a theoretical contribution to the study of moral philosophy, while also providing practical guidance for technology developers and policymakers to ensure that AI is used ethically and responsibly.

2. Methodology

This study uses a qualitative approach with the type of literature study research to explore the intersection between ethics and artificial intelligence (AI) in a philosophical context. This method was chosen because the research aims to analyze and interpret various classical and contemporary ethical theories, as well as their applications in the development and application of AI (Moleong, 2021). Literature studies allow researchers to collect and review data from a variety of scientific sources to deeply understand how ethical principles interact with evolving AI technologies.

The data sources in this study consist of primary and secondary sources. Primary sources include classic ethical philosophy books such as Immanuel Kant's work on deontology and John Stuart Mill's on utilitarianism, as well as contemporary texts that address AI ethics, such as Bostrom & Yudkowsky's (2014) work on the existential risks of AI and Floridi et al. (2018) on the ethical framework for AI. Secondary sources include scientific journal articles, policy reports from international institutions such as the OECD and UNESCO, as well as publications from leading academic institutions that address ethical challenges in the application of AI (Crawford, 2021).

The data collection technique was carried out by browsing the literature from various academic databases, such as JSTOR, SpringerLink, and IEEE Xplore, to identify relevant researches. This process includes searching for articles with keywords such as AI ethics, algorithmic bias, autonomy in AI, and philosophical perspectives on AI. In addition, the researcher also utilized cross-references from the articles found to expand the scope of the literature analyzed (Bowen, 2009).

The data analysis method used is qualitative content analysis with a descriptive-analytical approach. The collected data is classified based on key themes, such as moral responsibility, algorithmic autonomy, and bias in AI systems. Furthermore, the data is critically analyzed to explore how classical ethical theories can be applied or even challenged by AI developments. Researchers also compare various philosophical perspectives to identify similarities, differences, and unique contributions to understanding AI ethics (Creswell & Poth, 2018). With this approach, the research is expected to be able to provide new insights into the moral challenges presented by artificial intelligence and how ethical principles can be adapted to face these challenges.

3. Result and Discussion

The following table presents the findings of 10 selected articles that have been selected based on their relevance and contribution to the study of ethics and artificial intelligence (AI) from a philosophical perspective. These articles are analyzed based on

research methods, study focuses, and key findings that make important contributions to AI ethics discussions.

No	Author & Year	Title	Findings
1	Bostrom, N., & Yudkowsky, E., 2004	*The Ethics of Artificial Intelligence*	AI presents a huge risk to humanity; a strict ethical framework is needed.
2	Floridi, L. et al., 2018	*AI4People—An Ethical Framework for a Good AI Society*	Proposing the principles of transparency, fairness, and non-discrimination in AI.
3	Eubanks, V., 2018	*Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*	The use of AI in public policy exacerbates social inequality.
4	O'Neil, C., 2016	*Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*	Algorithms reinforce social injustice if not ethically supervised.
5	Zuboff, S., 2019	*The Age of Surveillance Capitalism*	AI is used for the exploitation of personal data that violates the right to privacy.
6	Cordella, A., & Bonina, C. M., 2012	*A Public Value Perspective for ICT Enabled Public Sector Reforms*	There is a need for the integration of public values in the design and use of AI in the public sector.
7	Moor, J. H., 2006	*The Nature, Importance, and Difficulty of Machine Ethics*	Machines need to be programmed with ethical principles that can be adapted to the moral context.
8	Crawford, K., 2021	*The Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*	AI has a huge ecological and political impact; need broader ethical considerations.
9	Coeckelbergh, M., 2020	*AI Ethics*	The debate about whether AI can have moral autonomy or just a human tool.
10	Burton, E., Goldsmith, J., & Mattei, N., 2018	*Ethical Considerations in Artificial Intelligence Courses*	It is necessary to integrate ethics in AI education for responsible development.

Based on the table above, it can be seen that artificial intelligence not only raises technical challenges, but also deep moral dilemmas. Themes such as moral responsibility in autonomous systems (Bostrom & Yudkowsky, 2014), algorithmic bias (O'Neil, 2016), and privacy exploitation (Zuboff, 2019) are the main focus of these studies. In addition, there is a strong push to integrate traditional ethical principles into AI development to prevent potential misuse of technology that could harm society at large.

Based on the results of the analysis of the 10 selected articles, it can be seen that artificial intelligence (AI) presents complex ethical dilemmas, especially in terms of moral responsibility and existential risks. Bostrom & Yudkowsky (2014) emphasized that the development of increasingly autonomous AI presents challenges in determining who is responsible for the decisions generated by AI systems. When AI is able to make decisions without direct human intervention, the line between human agency and machine agency becomes blurred. This has sparked the need for ethical frameworks capable of accommodating new dynamics in the relationship between humans and machines.

In addition to the issue of responsibility, algorithmic bias is one of the main findings raised by O'Neil (2016) and Eubanks (2018). Both studies show how the algorithms used in AI often reinforce existing social injustices, especially in labor recruitment systems, credit scoring, and law enforcement. The data used to train AI often contains historical biases, which if left unaddressed, can result in systematically discriminatory decisions. This demonstrates the need for close scrutiny in the design and implementation of algorithms to ensure that AI does not reinforce social inequalities.

Privacy and surveillance issues are also a major concern in AI ethics studies. Zuboff (2019) in The Age of Surveillance Capitalism describes how AI is used to exploit individuals' personal data for economic and political gain. In this context, AI is not only a tool of technological efficiency, but also a tool of mass surveillance that can threaten the rights of privacy and individual freedoms. Floridi et al. (2018) also underscore the importance of transparency in the use of AI to ensure that users understand how their data is collected and used.

Within a philosophical framework, AI challenges classic ethical concepts such as autonomy, responsibility, and justice. Coeckelbergh (2020) discusses whether AI can be considered to have moral autonomy or simply a tool that extends the human will. On the other hand, Moor (2006) introduced the concept of machine ethics, which emphasizes that machines need to be programmed with ethical principles in order to operate in an appropriate moral context. However, the main challenge lies in how to ensure that these principles are actually applied effectively in complex AI systems.

In addition to the technical and moral aspects, the ecological and political impacts of AI are also an important focus in several studies. Crawford (2021) in the Atlas of AI highlights how AI infrastructure, such as large data centers, has a significant ecological impact. The massive use of energy to process data not only poses environmental problems but also exacerbates global inequality, where developed countries control technological resources, while developing countries face the impact. This suggests that discussions about AI ethics should include social, political, and environmental dimensions.

Overall, the findings from this literature study suggest that the development and application of AI requires a multidisciplinary approach that integrates technical, ethical, and social perspectives. The development of a robust ethical framework must take into account the potential risks and negative impacts of AI on global societies, the environment, and power structures. By understanding the intersection between ethics and AI through the lens of philosophy, we can formulate policies and practices that not only encourage technological innovation, but also ensure that they are used for the common good.

Discussion and Analysis

Based on the results of the analysis of the 10 selected articles, it can be seen that artificial intelligence (AI) presents complex ethical dilemmas, especially in terms of moral responsibility and existential risks. Bostrom & Yudkowsky (2014) emphasized that the development of increasingly autonomous AI presents challenges in determining who is responsible for the decisions generated by AI systems. When AI is able to make decisions without direct human intervention, the line between human agency and machine agency becomes blurred. This has sparked the need for ethical frameworks capable of accommodating new dynamics in the relationship between humans and machines.

In addition to the issue of responsibility, algorithmic bias is one of the main findings raised by O'Neil (2016) and Eubanks (2018). Both studies show how the algorithms used in AI often reinforce existing social injustices, especially in labor recruitment systems, credit scoring, and law enforcement. The data used to train AI often contains historical biases, which if left unaddressed, can result in systematically discriminatory decisions. This demonstrates the need for close scrutiny in the design and implementation of

algorithms to ensure that AI does not reinforce social inequalities.

Privacy and surveillance issues are also a major concern in AI ethics studies. Zuboff (2019) in The Age of Surveillance Capitalism describes how AI is used to exploit individuals' personal data for economic and political gain. In this context, AI is not only a tool of technological efficiency, but also a tool of mass surveillance that can threaten the rights of privacy and individual freedoms. Floridi et al. (2018) also underscore the importance of transparency in the use of AI to ensure that users understand how their data is collected and used.

Within a philosophical framework, AI challenges classic ethical concepts such as autonomy, responsibility, and justice. Coeckelbergh (2020) discusses whether AI can be considered to have moral autonomy or simply a tool that extends the human will. On the other hand, Moor (2006) introduced the concept of machine ethics, which emphasizes that machines need to be programmed with ethical principles in order to operate in an appropriate moral context. However, the main challenge lies in how to ensure that these principles are actually applied effectively in complex AI systems.

In addition to the technical and moral aspects, the ecological and political impacts of AI are also an important focus in several studies. Crawford (2021) in the Atlas of AI highlights how AI infrastructure, such as large data centers, has a significant ecological impact. The massive use of energy to process data not only poses environmental problems but also exacerbates global inequality, where developed countries control technological resources, while developing countries face the impact. This suggests that discussions about AI ethics should include social, political, and environmental dimensions.

Overall, the findings from this literature study suggest that the development and application of AI requires a multidisciplinary approach that integrates technical, ethical, and social perspectives. The development of a robust ethical framework must take into account the potential risks and negative impacts of AI on global societies, the environment, and power structures. By understanding the intersection between ethics and AI through the lens of philosophy, we can formulate policies and practices that not only encourage technological innovation, but also ensure that they are used for the common good.

4. Conclusion

The development of artificial intelligence (AI) has brought significant changes in various aspects of human life, from the economy to the justice system. However, these advances also give rise to complex ethical dilemmas, especially related to moral responsibility, algorithmic bias, and privacy. AI that is able to make decisions autonomously presents challenges in determining who is responsible for the impact of those decisions. Additionally, bias in AI training data suggests that these technologies, if not closely monitored, could reinforce existing social injustices.

From a moral philosophy perspective, AI challenges classic concepts such as autonomy, justice, and human rights. The debate over whether AI can have moral autonomy or simply be a tool for the human will is still ongoing. Additionally, the use of AI in mass surveillance and exploitation of personal data poses a serious threat to individual privacy rights. These findings underscore the need for a robust ethical framework, which not only focuses on the technical aspects, but also takes into account the social, political, and environmental impacts of AI applications.

To ensure that AI is developed and used ethically, a multidisciplinary approach involving philosophers, computer scientists, policymakers, and civil society is needed. This research emphasizes that AI ethics should be an integral part of the technology design and implementation process, not just an add-on after the technology has been developed. As such, AI can be a tool that not only drives innovation, but also strengthens social justice and protects individual rights.

Further research is suggested to explore more specific ethical frameworks in various contexts of AI applications, such as in the judicial, health, or education systems. In addition, empirical studies examining how AI ethics policies are implemented in different countries will provide more comprehensive insights into the effectiveness of existing regulations.

Further research also needs to focus on the ecological impact of AI infrastructure, given the large energy consumption in data processing and how this affects global inequality. In addition, a comparative study of cross-cultural perceptions of ethics in the use of AI could provide a richer understanding of how different moral values affect the development of this technology.

Finally, it is recommended to develop AI ethics education guidelines that can be integrated in computer science curricula. In this way, future technology developers will have a better understanding of their ethical responsibilities, thus being able to ensure that AI is used for the common good.

References

Bowen, G. A. (2009). Document analysis as a qualitative research method. Qualitative Research Journal, 9(2), 27-40.

Fung, A. (2017). Open Government and Transparency Reform: A Critical Review. Cambridge University Press.

Gupta, S. (2020). Digitalization and Fiscal Transparency: The Future of Government Financial Management. Journal of Public Administration Research and Theory, 30(2), 205-219.

Hood, C., & Heald, D. (2006). Transparency: The Key to Better Governance? Oxford University Press.

Kim, J. (2019). E-Government and Corruption: Empirical Evidence from Developing Countries. Journal of Policy Analysis, 41(3), 150-172.

Mardiasmo. (2018). Accountability and Transparency in State Financial Management. Journal of Accounting and Public Finance, 22(1), 55-75.

Marzuki, P. M. (2017). Legal Research. Jakarta: Kencana.

Moleong, L. J. (2021). Qualitative Research Methodology. Bandung: PT Remaja Rosdakarya.

Rahman, A. (2021). The Role of Digital Platforms in State Financial Management: Transparency vs. Data Security Challenges. International Journal of Public Administration, 12(4), 322-339.

Setyowati, R. (2022). Legal Challenges in Digitalization of State Financial Management in Indonesia. Indonesian Journal of Legal Studies, 18(2), 130-147.

Soekanto, S., & Mamudji, S. (2021). Normative Law Research: A Brief Review. Jakarta: Rajawali Press.

Susanto, H. (2021). E-Government Implementation in Indonesia: Between Policy and Practice. Jurnal Ilmu Administrasi Negara, 14(3), 220-238.

Tanzi, V. (2018). Government Financial Transparency: Opportunities and Risks in the Digital Era. IMF Working Papers, WP/18/25.

World Bank. (2020). Digital Transformation and Public Financial Management. World Bank Publications.