Formosa Journal of Multidisciplinary Research (FJMR)
Vol. 3, No. 10, 2024: 3777-3792

Algorithms and Human Rights: The Impact of AI Technology on the Protection of Individual Rights

Muhammad Alfi Fadhlurrahman[1]*, Stanislaus Riyanta[2], Muhammad Reza Rustam[3]
School of Strategic and Global Studies, Universitas Indonesia, Indonesia
Corresponding Author: Muhammad Alfi Fadhlurrahman
muhammadalfi99@gmail.com

A R T I C L E I N F O

Keywords: AI Transparency, Human Rights, Algorithmic Decision-Making, GDPR Compliance, Ethical AI

Received : 16, September
Revised : 30, September
Accepted: 22, October

A B S T R A C T

This study examines the impact of artificial intelligence (AI) algorithms on the protection of individual rights, with a focus on transparency and accountability. The research aims to analyze how regulations like the GDPR in the European Union influence AI decision-making processes to safeguard human rights. Using a comparative legal analysis method, this research investigates the role of prospective and retrospective transparency in ensuring that AI systems operate ethically and fairly. The results highlight that transparency significantly enhances accountability and user rights, particularly in preventing discrimination and promoting security. The study concludes that collaboration among policymakers, developers, and users is critical to aligning AI technologies with human rights principles, ultimately fostering more equitable and secure outcomes.

DOI: https://doi.org/10.55927/fjmr.v3i10.11644
ISSN-E: 2829-8896
https://journal.formosapublisher.org/index.php/fjmr

INTRODUCTION

Artificial intelligence (AI) technology has undergone significant development and is now applied across various sectors, including healthcare, security, finance, and social media. In this context, AI serves as a tool to enhance the efficiency and effectiveness of decision-making processes. However, the application of this technology also raises serious concerns related to human rights, particularly regarding privacy, equality, and freedom of expression. For instance, in the healthcare sector, the use of AI can threaten patient privacy if health data is not properly managed, potentially leading to workplace discrimination and increased long-term healthcare costs (Price & Cohen, 2019; C. Wang et al., 2022). Additionally, AI algorithms used in recruitment processes can exacerbate discrimination issues if there is insufficient human oversight, highlighting the importance of maintaining a human factor in decision-making processes (Mariani & Lozada, 2023; Ore & Sposato, 2021). Concerns over privacy are growing as AI systems collect increasingly vast amounts of data. For example, in the education sector, the use of AI can result in violations of student privacy if their data is not adequately protected (Kamenskih, 2022; Sabharwal et al., 2023). Furthermore, the implementation of regulations such as the General Data Protection Regulation (GDPR) in Europe demonstrates that despite efforts to protect personal data, many organizations remain unprepared to comply with such regulations, potentially leading to privacy rights violations (Li et al., 2019). Therefore, it is crucial to develop frameworks that not only comply with regulations but also consider the ethical implications of AI usage, ensuring that this technology does not infringe upon individuals' fundamental rights (Huriye, 2023; Radanliev & Santos, 2023).

The use of artificial intelligence (AI) in various sectors, including recruitment, law enforcement, and financial services, presents complex challenges related to human rights. One of the main issues is the lack of transparency in the algorithms used by AI systems, making it difficult to understand how decisions are made and whether biases are embedded in these processes. For instance, recruitment algorithms may reinforce discrimination based on race or gender if there is insufficient oversight (Khair et al., 2020). Research has shown that algorithms are often inaccessible or difficult to comprehend, creating challenges in accountability and responsibility for decisions made (Huriye, 2023). This underscores the need for stricter regulations to ensure that AI-driven decisions do not harm specific individuals or groups (Mardiani et al., 2023). Moreover, AI can increase the risk of privacy violations through the mass collection and processing of data. Many AI systems collect data without clear consent from individuals, potentially resulting in breaches of privacy rights (Unver & Asan, 2022). In this context, it is essential to implement policies that protect personal data and ensure that individuals maintain control over their information (Kiseleva et al., 2022). Research indicates that transparency in data collection and usage is crucial in building public trust in AI technologies (Basereh et al., 2021). Therefore, the development of regulatory frameworks governing the use of AI must include transparency

and accountability aspects to safeguard individual rights (Upadhyay et al., 2023).

On the other hand, the use of AI in content moderation on social media also presents challenges regarding freedom of expression. Algorithms used to filter content are often overly aggressive, which can lead to the removal of legitimate content and restrict freedom of speech (Jiang et al., 2020). Research suggests that AI-based content moderation systems can exacerbate issues of bias and injustice, making it crucial to involve human oversight in the moderation process to ensure more equitable and accountable decisions (Diakopoulos, 2020). Therefore, a balanced approach is needed between the use of AI technology and the protection of human rights, with comprehensive regulations to ensure that AI implementation does not harm society at large (Müller et al., 2021). Although various efforts have been made to regulate the use of artificial intelligence (AI), current regulations remain insufficient to comprehensively protect human rights. In Europe, data protection laws like the General Data Protection Regulation (GDPR) have provided an essential foundation for individual privacy, but the enforcement of these regulations often struggles to keep pace with the rapid advancements in AI technology (Mészáros et al., 2022; Milossi et al., 2021). While the GDPR offers an important legal framework, challenges in its application and enforcement across different countries continue to pose significant barriers to protecting individual rights (Brown, 2022; J. Wang et al., 2023). Furthermore, many countries still lack equivalent regulations, creating a gap in human rights protection on a global scale (Hoseini, 2023).

LITERATURE REVIEW

The relationship between algorithms and human rights has been widely discussed in recent literature, particularly as AI systems play a growing role in decision-making processes that impact individual rights. Scholars such as Cathy O'Neil (2017) in *Weapons of Math Destruction* highlights how biased algorithms can perpetuate inequality, posing a threat to fundamental rights like privacy, non-discrimination, and equality. Similarly, Frank Pasquale (2015) in *The Black Box Society* emphasizes the lack of transparency in AI decision-making, which undermines accountability and fairness. The GDPR (General Data Protection Regulation), particularly its provision for algorithmic transparency, has become a key legal framework for mitigating these risks, allowing individuals to better understand and challenge AI-driven decisions.

The research is grounded in human rights theory, which emphasizes the protection of human dignity, equality, and freedom, as articulated by theorists such as Jack Donnelly (2013). This theoretical perspective supports the argument that AI systems must uphold universal human rights principles and be subject to legal and ethical scrutiny. By linking these theories with regulatory frameworks like the GDPR, the study investigates how transparency and accountability in AI can safeguard fundamental rights while fostering fairer outcomes.

METHODOLOGY

The research method employs an inductive approach, utilizing techniques derived from grounded theory (Wolfswinkel et al., 2013) based on a theory-driven literature review. The process follows five steps: (1) Defining the scope of the review, (2) Searching the literature, (3) Selecting a sample, (4) Analyzing the data, and (5) Presenting the results of the compiled articles. The researcher conducted searches using various queries in online libraries to gain an initial understanding of the relevant literature from multiple disciplines.

RESEARCH RESULT

Human Rights Algorithm Transparency and Accountability

Algorithm transparency in AI technology is increasingly recognized as a fundamental principle that ensures the processes, decisions, and operational mechanisms of AI systems are comprehensible to various stakeholders, including users, regulators, and society at large. This transparency is particularly vital in the context of human rights, as it serves to mitigate the risks of unfair treatment and violations of individual rights that can arise from opaque algorithmic decision-making. Moreover, transparency in AI systems fosters trust among users and the broader society. When individuals know how algorithms function and the criteria they use for decision-making, they are more likely to trust these systems. This trust is crucial for accepting and integrating AI technologies in various sectors, including healthcare, finance, and law enforcement, where the stakes are particularly high. The ability to scrutinize algorithms can also empower users to challenge decisions they perceive as unjust, reinforcing their rights and agency (Rodrigues, 2020). AI transparency can be divided into three main levels: 1) algorithmic transparency, 2) interaction transparency, and 3) social transparency (see Figure 1). These three aspects provide an overarching view of how AI systems can be understood and accounted for at various levels. Although each level is seen as a distinct dimension, they are interconnected and work together to realize transparency in AI use (Haresamudram et al., 2023). Algorithmic transparency refers to transparency within the AI system itself, specifically how the algorithm works. It involves understanding the logic, models, and data used by the algorithm to make decisions. Algorithmic transparency is essential to avoid potential bias and discrimination in AI systems, as well as enable external audits that can test the fairness and accuracy of the models used. However, understanding how algorithms work can be challenging as some AI models, especially those based on deep learning, are very complex and are often perceived as black boxes. Interaction transparency relates to transparency in the relationship between AI and its users. This includes how information about the AI system is conveyed to the user, as well as how the user can understand, interact with, and even influence the output produced by the AI. Social transparency then involves the relationship between AI and society as a whole. This includes the social and ethical impacts of using AI, as well as how AI is received and understood by the public. Social transparency ensures that AI is not only understood by experts or specific users, but also by the wider society affected by its use.

Figure 1. Three Levels of AI Transparency (Haresamudram et al., 2023)

These three levels of transparency algorithmic, interactional, and social, cannot be seen as separate elements. Instead, they are interdependent and work together to form the overarching transparency of AI. For example, while algorithmic transparency is important to explain how AI works, without sufficient interaction transparency, users may not be able to understand or use the information. Similarly, social transparency ensures that society and regulation are able to monitor AI, but to do so, they need strong algorithmic and interactional transparency (Haresamudram et al., 2023). The necessity for algorithm transparency is underscored by the potential for AI systems to perpetuate biases and discrimination if their workings remain hidden. Liu highlights that the perceived legitimacy of algorithmic decision-making is significantly enhanced when transparency is prioritized, allowing users to understand how decisions are made and fostering trust in these systems (Liu & Luo, 2023). This trust is essential for the acceptance and integration of AI technologies across various sectors, including healthcare and law enforcement, where the implications of algorithmic decisions can be profound (Arrieta et al., 2020). Furthermore, Goad and Gal emphasize that transparency not only enhances the legitimacy of algorithmic systems but also promotes public trust, which is vital for the effective functioning of these technologies in society (Goad & Gal, 2018). Moreover, regulatory frameworks increasingly emphasize the need for transparency in AI systems. Mc Gregor et al. argue that existing approaches to algorithmic accountability, such as transparency, provide a foundational baseline but are insufficient to fully address the potential harm to human rights caused by algorithmic decision-making (Mc Gregor et al., 2019). To effectively safeguard human rights, a comprehensive framework that encompasses the entire algorithmic life cycle from conception to deployment is necessary (Mc Gregor et al., 2019). This regulatory approach not only protects individuals but also encourages developers to prioritize ethical considerations in their work, thereby aligning technological advancement with human rights protections (Chatterjee & Sreenivasulu, 2021). In addition, the ethical implications of AI systems are increasingly recognized as critical to their deployment. Mittelstadt et al. note that the delegation of decisions to algorithms can have severe consequences for individuals and groups, particularly when gaps exist between the design and operation of these systems and our understanding of their ethical implications (Mittelstadt et al., 2016). The call for human-centered AI, as discussed by Riedl, emphasizes the importance of designing algorithms with an awareness of their impact on humans, thereby ensuring that these technologies serve human needs and uphold individual rights (Riedl, 2019).

The "black box" nature of many AI algorithms poses significant challenges to transparency. As highlighted by Akinrinola, the complexity of AI systems often obscures how decisions are made, leading to ethical dilemmas regarding bias and accountability (Olatunji Akinrinola et al., 2024). This lack of transparency can result in algorithmic decisions that perpetuate existing social biases, thereby raising concerns about fairness and justice (Olatunji Akinrinola et al., 2024). For instance, in the context of hiring practices, AI systems may inadvertently favor certain demographics over others if the underlying algorithms are not transparent and accountable (Du, 2024). This emphasizes the need for clear guidelines and frameworks that mandate transparency in AI systems to ensure that they operate fairly and justly. Moreover, the integration of AI into various sectors necessitates a robust regulatory framework that prioritizes transparency. The European Union's General Data Protection Regulation (GDPR) aims to enhance transparency around algorithmic decision-making, particularly by prohibiting the use of personal characteristics that could lead to bias and discrimination (Lim & Taeihagh, 2019). However, as noted by Lim and Taeihagh, there are gaps in the GDPR's provisions regarding the right to an explanation for automated decisions, indicating that further refinement is needed to ensure meaningful transparency (Lim & Taeihagh, 2019). This regulatory approach not only protects individual rights but also fosters trust in AI systems, as users are more likely to accept technologies that they understand and can scrutinize (Haresamudram et al., 2023).

The phenomenon of algorithmic bias is well-documented in the literature. For instance, Obermeyer et al. identified significant racial bias in a widely used healthcare algorithm, which misclassified Black patients as lower risk based on cost data, thereby disadvantaging them in terms of healthcare access (Mc Call et al., 2022). his example illustrates how biases can be embedded in algorithms, often reflecting historical inequities present in the training datasets. Transparency in algorithms can help identify such biases, allowing for corrective measures to be implemented before these systems are deployed in high-stakes environments. Moreover, the concept of transparency extends beyond mere visibility of the algorithm's workings; it encompasses the ability to understand and interpret the decisions made by these systems. Giunchiglia emphasizes that achieving algorithmic transparency is complicated by inherent "bias blind spots" that individuals may possess, which can obscure their awareness of the biases influencing their perceptions and behaviors (Giunchiglia et al., 2021). This highlights the need for comprehensive reporting guidelines that not only promote transparency but also explicitly address fairness and equity in algorithmic decision-making (Chin et al., 2023). Stakeholders can work towards more equitable outcomes by fostering an environment where biases can be identified and addressed. The ethical implications of algorithmic transparency are profound, particularly in public sector decision-making. Veale et al. argue that transparency is essential for ensuring accountability in high-stakes decisions, as it allows stakeholders to scrutinize the criteria and processes used by algorithms (Veale et al., 2018). By establishing clear standards for transparency, organizations can better align

algorithmic decisions with societal values, thereby promoting fairness and accountability in their operations.

AI's impact on privacy is multifaceted, particularly in sectors such as recruitment and journalism. In recruitment, the use of AI and algorithms can lead to adverse discrimination if not carefully monitored, who advocate for maintaining human oversight to safeguard human rights and privacy (Mariani & Lozada, 2023). Similarly, journalists express concerns about AI's potential to misuse data and violate user privacy, emphasizing the need for ethical standards in AI applications within journalism (Al-Zoubi et al., 2024). These findings underscore the necessity for transparency in AI systems to ensure that individuals are informed about how their data is utilized, thus preserving their privacy rights. Moreover, the integration of AI into various domains raises ethical implications that directly affect privacy. For instance, the deployment of emotion AI in workplaces has been perceived as a significant intrusion into personal privacy, as it often involves the collection of sensitive emotional data without adequate consent or transparency (Roemmich, 2023). This intrusion can breach the contextual norms governing emotional information, thereby compromising employees' rights to privacy. Furthermore, the use of AI in healthcare has been linked to privacy violations due to extensive data collection practices that often occur without clear user consent (Naik et al., 2022; Williamson & Prybutok, 2024). The Cambridge Analytica scandal serves as a stark reminder of the risks associated with AI and big data, highlighting the urgent need for robust privacy protections and regulatory frameworks (Huang et al., 2024). To address these challenges, various strategies have been proposed. Techniques such as differential privacy and federated learning are being explored to enhance privacy preservation in AI applications (Surendranadha et al., 2020; Zhu et al., 2022). These methods aim to protect individual data while still allowing for the benefits of AI-driven insights. Additionally, the implementation of comprehensive legal frameworks, such as the General Data Protection Regulation (GDPR), is essential for ensuring that AI systems operate within ethical boundaries that respect individual privacy rights (HOXHAJ et al., 2023; Milossi et al., 2021). The ongoing discourse surrounding AI and human rights emphasizes the importance of adapting existing legal protections to the realities of AI technologies, ensuring that privacy rights are upheld in the face of rapid technological advancements (Leite et al., 2023; Špadina, 2023).

Algorithm Openness Policy as Human Rights Protection

The right to transparency in algorithmic decision-making is increasingly recognized as a fundamental aspect of safeguarding individual rights, particularly in the context of artificial intelligence (AI). Regulatory frameworks, such as the General Data Protection Regulation (GDPR) in the European Union, exemplify efforts to ensure that individuals are informed about how automated decisions affect them. The GDPR mandates that companies disclose the logic behind automated decisions and provide individuals with the necessary information to understand these processes, thereby promoting accountability and preventing potential abuses (Bernal & Mazo, 2022). Such regulations are

essential for fostering trust in AI systems and ensuring that users can exercise their rights effectively. Transparency in AI is not merely a regulatory requirement but also a moral imperative that aligns with the principle of explainability. This principle necessitates that organizations using AI systems articulate, in comprehensible terms, how their algorithms function and the factors influencing their decision-making processes. Transparency in AI is not merely a regulatory requirement but also a moral imperative that aligns with the principle of explainability. This principle necessitates that organizations using AI systems articulate, in comprehensible terms, how their algorithms function and the factors influencing their decision-making processes (Haresamudram et al., 2023). By doing so, organizations can minimize the risk of human rights violations and enhance user confidence in AI technologies. For instance, in healthcare, the lack of transparency regarding AI-driven decisions can lead to significant risks, as patients may not fully understand how their medical data is being used or how decisions affecting their care are made (Kiseleva et al., 2022; Lennartz et al., 2021). This highlights the critical need for transparency to ensure that AI applications do not compromise patient safety or autonomy.

Moreover, the concept of transparency can be categorized into different levels, including prospective and retrospective transparency. Prospective transparency involves providing information about how AI systems operate before decisions are made, while retrospective transparency focuses on post hoc explanations of decisions (Felzmann et al., 2019). Both forms of transparency are vital for establishing accountability in AI systems, as they allow stakeholders to scrutinize the decision-making processes and hold organizations accountable for their actions. This dual approach is particularly relevant in sectors like education and healthcare, where the implications of AI decisions can significantly impact individuals' lives (Chaudhry et al., 2022; Upadhyay et al., 2023). The ethical implications of transparency in AI extend beyond mere compliance with regulations. They encompass broader societal concerns regarding trust, fairness, and accountability. For example, studies have shown that higher levels of transparency in AI systems correlate with increased trust among users, which is crucial for fostering a positive relationship between humans and AI (Yu et al., 2023). Conversely, opacity in AI decision-making can lead to skepticism and resistance from users, undermining the potential benefits of AI technologies (Femi Osasona et al., 2024). Therefore, organizations must prioritize transparency not only as a legal obligation but also as a strategic imperative to build trust and ensure ethical AI deployment.

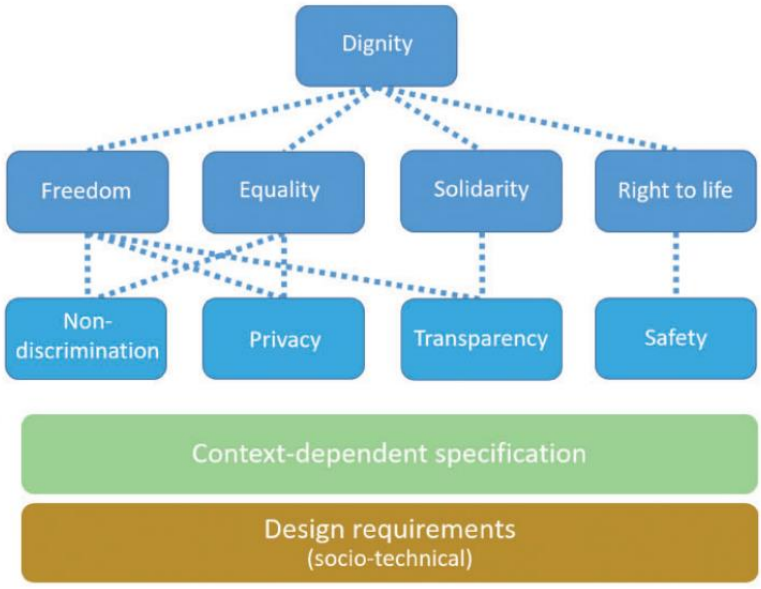

Figure 2. Hierarchical Relationship of Human Rights Values and Ethical Principles in AI Design (Aizenberg & van den Hoven, 2020)

The figure above illustrates the hierarchical relationship between fundamental human rights values and how they can be translated into more specific ethical principles, which are relevant in the design and development of artificial intelligence (AI) systems. This will be explained as follows:

a. The first tier begins with Dignity as a fundamental value that sits at the top of the hierarchy and is the basis of all other human rights. Human dignity is an inviolable value, which requires every individual to be respected and treated properly in every aspect of life (Aizenberg & van den Hoven, 2020). In the context of technology and artificial intelligence (AI) systems, respect for human dignity must also be prioritized to ensure that the development and application of technology does not violate these fundamental rights. Dignity is widely recognized as the foundational value that underpins all human rights, serving as the cornerstone for ethical considerations in various domains, including technology and artificial intelligence (AI) systems. The concept of human dignity is inviolable, necessitating that every individual is respected and treated with dignity in all aspects of life. This principle is particularly crucial in the context of AI, where the development and application of technology must align with respect for human dignity to prevent violations of fundamental rights.
b. The concept of human dignity serves as the foundation for a hierarchy of values that are essential for the protection and promotion of human rights. Following dignity, four specific second-order values emerge: Freedom, Equality, Solidarity, and Right to Life. Each of these values plays a critical role in shaping a just society, particularly in the context of technology and artificial intelligence (AI) (Aizenberg & van den Hoven, 2020). Freedom is defined as an individual's right to act according to their own will, free from coercion or undue influence. This value is

paramount in ensuring that individuals can make choices about their lives without external pressures that could compromise their autonomy. Aizenberg emphasizes that in the design of AI systems, it is crucial to safeguard this freedom by ensuring that technologies do not manipulate or constrain individual choices. Equality, as a second-order value, asserts that every individual deserves fair treatment without discrimination. This principle is particularly relevant in the development of AI systems, where biases can inadvertently be encoded into algorithms, leading to unequal outcomes for different groups. Aizenberg highlights the necessity of implementing fairness measures in AI design to prevent discrimination based on race, gender, or socioeconomic status. Solidarity emphasizes the importance of mutual support among individuals and groups, particularly regarding social welfare. In the context of AI, this value calls for the development of technologies that foster community engagement and cooperation. Aizenberg notes that AI can be harnessed to enhance social welfare initiatives, such as healthcare and education, by facilitating collaboration and resource sharing among communities. The Right to Life guarantees that every individual has the right to a safe and decent life. This value is critical in the context of AI, especially as technologies increasingly influence various aspects of daily life, including safety and security. Aizenberg argues that AI systems must be designed with a focus on protecting individuals' rights to life and safety, ensuring that technologies do not pose risks to users. In summary, the second-order values of Freedom, Equality, Solidarity, and Right to Life are integral to the framework of human rights in the context of AI. By prioritizing these values, stakeholders can work towards developing technologies that respect human dignity and promote a just and equitable society.

c. The transition from second-order values to third-order values in the context of technology and AI ethics is crucial for ensuring that the principles of human dignity, freedom, equality, solidarity, and the right to life are effectively operationalized. These third-order values Non-discrimination, Privacy, Transparency, and Security are concrete and directly applicable to the design and implementation of AI systems (Aizenberg & van den Hoven, 2020). Non-discrimination is a fundamental third-order value that mandates AI systems to operate without bias against individuals based on characteristics such as race, gender, or social status. Aizenberg emphasizes that the ethical design of AI must include mechanisms to identify and mitigate biases that can lead to discriminatory outcomes. Privacy is another essential third-order value that safeguards personal data, thereby supporting both freedom and equality. Individuals must have control over their personal information, which is increasingly at risk in an era of pervasive data collection and surveillance. Aizenberg argues that privacy protections are vital in maintaining individual autonomy and preventing misuse of data by AI systems. Transparency is critical for fostering both freedom

and solidarity within AI systems. This value demands openness regarding how AI technologies operate, including the decision-making processes and data usage. Aizenberg highlights that transparency is essential for accountability, allowing users to understand how their data is being utilized and how decisions that affect their lives are made. Security, as a third-order value, focuses on protecting users' physical and digital safety from potential threats posed by AI systems. Aizenberg points out that ensuring security is paramount in the design of AI, particularly as these technologies are integrated into critical infrastructure and daily life. This includes safeguarding against cyber threats, data breaches, and the misuse of AI in harmful ways. By prioritizing security, developers can help create a safe environment where individuals can engage with AI technologies without fear of harm.

d. Once the values of Non-discrimination, Privacy, Transparency, and Security are identified, the subsequent step involves translating these values into context-dependent specifications and socio-technical design requirements. This process is crucial for ensuring that the ethical principles guiding AI development are effectively integrated into practical applications. Context-dependent specifications recognize that the application of these values is not uniform; rather, it varies significantly based on the specific social, cultural, and technical contexts in which AI systems are deployed. Aizenberg emphasizes that understanding the unique characteristics of each context is vital for ensuring that the implementation of values like Non-discrimination and Privacy is relevant and effective (Aizenberg & van den Hoven, 2020). In addition to context-dependent specifications, socio-technical design requirements play a critical role in operationalizing these normative values into actionable guidelines for AI system design. These requirements bridge the gap between ethical principles and technical implementation, ensuring that AI systems are not only technically sound but also aligned with human rights considerations. Aizenberg highlights the importance of translating values into specific design criteria that can guide developers in creating systems that respect user rights and promote social good. Moreover, the integration of socio-technical design requirements requires collaboration among various stakeholders, including technologists, ethicists, policymakers, and community representatives. Aizenberg advocates for a multidisciplinary approach to AI design, where diverse perspectives are considered to create systems that are not only technically proficient but also socially responsible (Aizenberg & van den Hoven, 2020). This collaborative effort can help ensure that the design of AI systems reflects the values and needs of the communities they serve, ultimately leading to more equitable and just technological outcomes.

DISCUSSION

The research reveals that algorithm transparency is critical for protecting human rights, especially in the context of AI systems. Regulations like the GDPR provide a framework for ensuring that AI-driven decisions are transparent and accountable, allowing individuals to understand and challenge outcomes that may infringe on their rights. The distinction between prospective transparency (clarifying how algorithms are expected to work) and retrospective transparency (allowing scrutiny of past decisions) emerges as a key finding. These mechanisms help mitigate risks such as bias, discrimination, and privacy violations, ensuring that AI technologies align with ethical standards.

Additionally, the research emphasizes the importance of embedding fundamental human values—such as dignity, freedom, and equality—into the development of AI systems. While legal frameworks provide necessary oversight, ethical considerations must also drive the design and deployment of AI. Collaboration between policymakers, developers, and society is essential to ensure that AI systems are not only technically sound but also socially responsible, promoting fairness and inclusivity in their application.

CONCLUSIONS AND RECOMMENDATIONS

Algorithm transparency is essential for protecting human rights in today's advanced technological era. Regulations like the GDPR provide a legal framework that supports transparency and accountability, while a deep understanding of algorithms can help mitigate the risks of violating individual rights. Implementing ethical principles rooted in fundamental values such as dignity, freedom, equality, solidarity, and the right to life is crucial in the development of AI systems. To achieve this goal, the application of these values in specific contexts, along with relevant socio-technical design requirements, is vital. Collaboration among developers, policymakers, and the community will ensure that AI systems are not only technically advanced but also socially responsible, thereby creating fairer and more sustainable outcomes for all parties involved.

ADVANCED RESEARCH

However, this study also has several limitations. First, the analysis of algorithm transparency often focuses on regulatory frameworks without considering the social and cultural factors that may influence public acceptance and understanding of AI technology. Second, there is a need for further studies on the long-term impacts of algorithm transparency on human rights, particularly within diverse sectors such as healthcare, education, and law enforcement. Additionally, the technical challenges of implementing transparency and reducing bias in algorithms still require more research to develop practical and effective solutions. Finally, public involvement in the development and application of AI technology has not been fully explored, which could lead to injustice and dissatisfaction if community voices are not taken into account. By identifying and addressing these shortcomings, future

research can provide more comprehensive insights and effective strategies for safeguarding human rights in the context of algorithms and AI technology.

ACKNOWLEDGMENT

The author would like to express sincere gratitude to colleagues and mentors for their valuable suggestions and insights throughout the development of this paper. Special thanks are extended to Universitas Indonesia for its support and resources, which greatly contributed to the completion of this research.

REFERENCES

Al-Zoubi, O., Ahmad, N., & Hamid, N. A. (2024). Artificial Intelligence in Newsrooms: Ethical Challenges Facing Journalists. *Studies in Media and Communication*, *12*(1), 401. https://doi.org/10.11114/smc.v12i1.6587.

Arrieta, A. B., Díaz-Rodríguez, N., Ser, J. Del, Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-López, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges Toward Responsible AI. *Information Fusion*, *58*, 82–115. https://doi.org/10.1016/j.inffus.2019.12.012.

Bernal, J., & Mazo, C. (2022). Transparency of Artificial Intelligence in Healthcare: Insights From Professionals in Computing and Healthcare Worldwide. *Applied Sciences*, *12*(20), 10228. https://doi.org/10.3390/app122010228.

Chatterjee, S., & Sreenivasulu, N. S. (2021). Artificial Intelligence and Human Rights: A Comprehensive Study From Indian Legal and Policy Perspective. *International Journal of Law and Management*, *64*(1), 110–134. https://doi.org/10.1108/ijlma-02-2021-0049.

Du, J. (2024). *Exploring Gender Bias and Algorithm Transparency: Ethical Considerations of AI in HRM*. *4*(03), 36–43. https://doi.org/10.53469/jtpms.2024.04(03).06.

Felzmann, H., Fosch-Villaronga, E., Lutz, C., & Tamò-Larrieux, A. (2019). Transparency You Can Trust: Transparency Requirements for Artificial Intelligence Between Legal Norms and Contextual Concerns. *Big Data & Society*, *6*(1), 205395171986054. https://doi.org/10.1177/2053951719860542.

Femi Osasona, N., Olukunle Oladipupo Amoo, N., Akoh Atadoga, N., Temitayo Oluwaseun Abrahams, N., Oluwatoyin Ajoke Farayola, N., & Benjamin Samson Ayinla, N. (2024). Reviewing the Ethical Implications of Ai in Decision Making Processes. *International Journal of Management & Entrepreneurship Research*, *6*(2), 322–335. https://doi.org/10.51594/ijmer.v6i2.773.

Giunchiglia, F., Otterbacher, J., Kleanthous, S., Batsuren, K., Bogin, V., Kuflik, T., & Tal, A. S. (2021). *Towards Algorithmic Transparency: A Diversity Perspective*. https://doi.org/10.48550/arxiv.2104.05658.

Goad, D., & Gal, U. (2018). *Understanding the Impact of Transparency on Algorithmic Decision Making Legitimacy*. 64–79. https://doi.org/10.1007/978-3-030-04091-8_6.

Haresamudram, K., Larsson, S., & Heintz, F. (2023). Three Levels of AI Transparency. *Computer*, *56*(2), 93–100. https://doi.org/10.1109/mc.2022.3213181.

Hoseini, F. (2023). *AI Ethics: A Call for Global Standards in Technology Development*. *1*(4), 1–3. https://doi.org/10.61838/kman.aitech.1.4.1.

HOXHAJ, O., HALILAJ, B., & HARIZI, A. (2023). Ethical Implications and Human Rights Violations in the Age of Artificial Intelligence. *Balkan Social Science Review*, *22*(22), 153–171. https://doi.org/10.46763/bssr232222153h.

Huang, X., Yusoff, Z. M., Nor, M. Z. B. M., & Labanieh, M. F. (2024). *The Legal Challenges and Regulatory Responses to Artificial Intelligence (AI) in China*. 335–347. https://doi.org/10.2991/978-94-6463-352-8_26

Huriye, A. Z. (2023). The Ethics of Artificial Intelligence: Examining the Ethical Considerations Surrounding the Development and Use of AI. *American Journal of Technology*, *2*(1), 37–45. https://doi.org/10.58425/ajt.v2i1.142.

Jiang, S., Robertson, R. E., & Wilson, C. (2020). Reasoning About Political Bias in Content Moderation. *Proceedings of the Aaai Conference on Artificial Intelligence*, *34*(09), 13669–13672. https://doi.org/10.1609/aaai.v34i09.7117.

Khair, M. A., Mahadasa, R., Tuli, F. A., & Ande, J. R. P. K. (2020). Beyond Human Judgment: Exploring the Impact of Artificial Intelligence on HR Decision-Making Efficiency and Fairness. *Global Disclosure of Economics and Business*, *9*(2), 163–176. https://doi.org/10.18034/gdeb.v9i2.730.

Kiseleva, A., Kotzinos, D., & Hert, P. De. (2022). Transparency of AI in Healthcare as a Multilayered System of Accountabilities: Between Legal Requirements and Technical Limitations. *Frontiers in Artificial Intelligence*, *5*. https://doi.org/10.3389/frai.2022.879603.

Lennartz, S., Dratsch, T., Zopfs, D., Persigehl, T., Maintz, D., Hokamp, N. G., & Santos, D. P. dos. (2021). Use and Control of Artificial Intelligence in Patients Across the Medical Workflow: Single-Center Questionnaire Study of Patient Perspectives. *Journal of Medical Internet Research*, *23*(2), e24221.

https://doi.org/10.2196/24221.

Li, H., Yu, L., & He, W. (2019). The Impact of GDPR on Global Technology Development. *Journal of Global Information Technology Management*, *22*(1), 1–6. https://doi.org/10.1080/1097198x.2019.1569186.

Liu, S., & Luo, H. (2023). Examining Public Perceptions of Algorithm Transparency: An Empirical Analysis. *Lecture Notes in Education Psychology and Public Media*, *21*(1), 137–144. https://doi.org/10.54254/2753-7048/21/20230108.

Mariani, K., & Lozada, F. V. (2023). The Use of AI and Algorithms for Decision-Making in Workplace Recruitment Practices. *Journal of Student Research*, *12*(1). https://doi.org/10.47611/jsr.v12i1.1855.

McCall, C. J., DeCaprio, D., & Gartner, J. (2022). *The Measurement and Mitigation of Algorithmic Bias and Unfairness in Healthcare AI Models Developed for the CMS AI Health Outcomes Challenge*. https://doi.org/10.1101/2022.09.29.22280537.

McGregor, L., Murray, D., & Ng, V. (2019). International Human Rights Law as a Framework for Algorithmic Accountability. *International and Comparative Law Quarterly*, *68*(2), 309–343. https://doi.org/10.1017/s0020589319000046.

Mészáros, J., Minari, J., & Huys, I. (2022). The Future Regulation of Artificial Intelligence Systems in Healthcare Services and Medical Research in the European Union. *Frontiers in Genetics*, *13*. https://doi.org/10.3389/fgene.2022.927721.

Milossi, M., Alexandropoulou-Egyptiadou, E., & Psannis, K. E. (2021). AI Ethics: Algorithmic Determinism or Self-Determination? The GPDR Approach. *Ieee Access*, *9*, 58455–58466. https://doi.org/10.1109/access.2021.3072782.

Müller, A., Mertens, S., Göstemeyer, G., Krois, J., & Schwendicke, F. (2021). Barriers and Enablers for Artificial Intelligence in Dental Diagnostics: A Qualitative Study. *Journal of Clinical Medicine*, *10*(8), 1612. https://doi.org/10.3390/jcm10081612.

Olatunji Akinrinola, N., Chinwe Chinazo Okoye, N., Onyeka Chrisanctus Ofodile, N., & Chinonye Esther Ugochukwu, N. (2024). Navigating and Reviewing Ethical Dilemmas in AI Development: Strategies for Transparency, Fairness, and Accountability. *GSC Advanced Research and Reviews*, *18*(3), 50–58. https://doi.org/10.30574/gscarr.2024.18.3.0088.

Ore, O., & Sposato, M. (2021). Opportunities and Risks of Artificial Intelligence in Recruitment and Selection. *International Journal of Organizational Analysis*,

30(6), 1771–1782. https://doi.org/10.1108/ijoa-07-2020-2291.

Radanliev, P., & Santos, O. (2023). *Ethics and Responsible AI Deployment*. https://doi.org/10.33774/apsa-2023-f1fkq.

Riedl, M. O. (2019). Human-centered Artificial Intelligence and Machine Learning. *Human Behavior and Emerging Technologies*, *1*(1), 33–36. https://doi.org/10.1002/hbe2.117.

Roemmich, K. (2023). *Emotion AI atWork: Implications forWorkplace Surveillance, Emotional Labor, and Emotional Privacy*. https://doi.org/10.31219/osf.io/udrf2.

Unver, M. B., & Asan, O. (2022). Role of Trust in AI-Driven Healthcare Systems: Discussion From the Perspective of Patient Safety. *Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care*, *11*(1), 129–134. https://doi.org/10.1177/2327857922111026.

Upadhyay, U., Gradisek, A., Iqbal, U., Dhar, E., Li, Y.-C., & Syed-Abdul, S. (2023). Call for the Responsible Artificial Intelligence in the Healthcare. *BMJ Health & Care Informatics*, *30*(1), e100920. https://doi.org/10.1136/bmjhci-2023-100920.

Wang, C., Zhang, J., Lassi, N., & Zhang, X. (2022). Privacy Protection in Using Artificial Intelligence for Healthcare: Chinese Regulation in Comparative Perspective. *Healthcare*, *10*(10), 1878. https://doi.org/10.3390/healthcare10101878.

Wang, J., Mao, W., & Wenjie, W. (2023). *The Ethics of Artificial Intelligence: Sociopolitical and Legal Dimensions*. *2*(2), 27–32. https://doi.org/10.61838/kman.isslp.2.2.6.

Williamson, S. M., & Prybutok, V. (2024). Balancing Privacy and Progress: A Review of Privacy Challenges, Systemic Oversight, and Patient Perceptions in AI-Driven Healthcare. *Applied Sciences*, *14*(2), 675. https://doi.org/10.3390/app14020675.

Yu, L., Li, Y., & Fan, F. (2023). Employees' Appraisals and Trust of Artificial Intelligences' Transparency and Opacity. *Behavioral Sciences*, *13*(4), 344. https://doi.org/10.3390/bs13040344.

Zhu, T., Ye, D., Wang, W., Zhou, W., & Yu, P. S. (2022). More Than Privacy: Applying Differential Privacy in Key Areas of Artificial Intelligence. *Ieee Transactions on Knowledge and Data Engineering*, *34*(6), 2824–2843. https://doi.org/10.1109/tkde.2020.3014246