Jurnal Informatika Ekonomi Bisnis

http://www.infeb.org

2025 Vol. 7 Iss. 2 Hal: 226-231 e-ISSN: 2714-8491

The Impact of Artificial Intelligence Implementation on Business Ethics in Corporate Decision-Making

Rahmi Rachmawati Sari[1]✉, Nurhaeni Sikki[2], Fitri Chintiyani[3], Kinanti Vionanda Pusparani[4], Liska Zahara Putri[5], Yulia Puspita Sari[6], Leni Nuraeni HR[7], Jeanne Nurtami Depi[8], Nurul Fitria Nasution[9]

[1,2,3,4,5,6,7,8,9]Pascasarjana Universitas Sangga Buana

rahmirachmawati3@gmail.com

Abstract

The integration of artificial intelligence into corporate decision-making processes has introduced complex ethical challenges that demand urgent scholarly and practical attention. This study explores how AI implementation influences business ethics across five key dimensions: algorithmic bias, transparency, responsibility diffusion, privacy, and human autonomy. Employing a qualitative research design, data were collected through in-depth interviews with 25 executives and ethics officers from multinational corporations across diverse sectors. Thematic analysis revealed widespread concerns about the opacity of AI systems, ethical disengagement, and the absence of clear accountability structures. Although some organizations have begun adopting ethical AI governance frameworks, the effectiveness of these measures varies significantly, with many efforts remaining reactive and fragmented. The findings highlight a critical gap between ethical aspirations and organizational realities, emphasizing the need for systemic reforms that embed ethics into the design, deployment, and oversight of AI technologies. This study contributes to the growing body of literature on AI ethics by offering empirical insights and actionable implications for policy, leadership, and future research.

Keywords: Artificial Intelligence, Business Ethics, Corporate Decision-Making, Algorithmic Accountability, Ethical AI

1. Introduction

The integration of artificial intelligence (AI) into corporate environments has revolutionized organizational processes, including decision-making, risk assessment, and strategic planning [1]. As AI systems increasingly assume roles previously reserved for human judgment, questions about the ethical ramifications of such delegation have emerged. Scholars have argued that AI can introduce biases, erode accountability, and obscure transparency in decision-making processes [2]. This growing ethical concern stems from the fact that AI technologies often operate through opaque algorithms, making it difficult for stakeholders to discern the rationale behind critical corporate decisions [3]. As companies embrace data-driven tools to enhance efficiency and competitiveness, ethical frameworks guiding traditional business practices face significant challenges in adapting to this paradigm shift [4].

Business ethics, traditionally rooted in moral philosophy and stakeholder theory, is undergoing a transformation in response to the rise of autonomous systems. Numerous studies have explored how corporate ethics are influenced by technological interventions [5] [6] [7] revealing a shift in how organizations conceptualize responsibility and fairness. For instance, AI-powered decision systems may inadvertently reproduce or amplify existing societal inequities due to biased training data or flawed algorithmic design [8]. As a result, scholars and practitioners alike are advocating for the embedding of ethical principles into the design and deployment of AI in business [9] arguing that ethical lapses can erode stakeholder trust and ultimately affect long-term corporate sustainability [10].

Moreover, ethical concerns are not limited to data bias but extend to issues of surveillance, autonomy, and informed consent in digital environments. The literature emphasizes that ethical AI implementation must account for broader organizational values and cultural contexts to ensure responsible decision-making [11][12][13]. The need for explainability and accountability in AI models used in corporate governance has led to the emergence of frameworks such as ethical-by-design and responsible AI [14]. These frameworks seek to harmonize technological advancement with moral obligations, thereby supporting organizational legitimacy and social acceptance [15]. Nevertheless, empirical investigations into how these frameworks are operationalized within corporate decision-making remain limited, particularly across varying industry sectors and organizational sizes [16].

The implementation of AI also raises ethical dilemmas in board-level decision-making where trade-offs between profitability, employee welfare, and societal impacts must be carefully weighed. Studies have shown that while AI can enhance predictive accuracy and optimize performance, it can also displace human judgment and obscure responsibility lines, complicating ethical assessments [17] [18] [19]. Additionally, scholars argue that the power dynamics inherent in AI deployment can lead to asymmetries in

Accepted: 28-04-2025 | Revision: 25-05-2025 | Publication: 27-05-2025 | doi: 10.37034/infeb.v7i2.1127

knowledge and control, thereby weakening democratic participation in corporate governance [20]. These findings highlight a pressing need for ethical literacy among decision-makers who deploy AI systems [21] [22], particularly in contexts where decisions bear significant consequences for stakeholders, including consumers, employees, and the public.

Despite increasing academic attention, there remains a research gap in understanding how AI implementation specifically affects ethical decision-making practices within corporations. While theoretical frameworks have proposed normative principles for ethical AI [23], few studies have empirically examined the extent to which these principles are reflected in organizational routines and decision-making processes. Furthermore, existing literature often lacks a comparative perspective that considers sectoral or cultural differences in ethical AI deployment [24]. Addressing this gap is critical to formulating evidence-based guidelines that promote ethical AI adoption and prevent unintended harm. The present study aims to investigate the ethical implications of AI implementation in corporate decision-making, focusing on the extent to which ethical frameworks are applied, the challenges faced by organizations, and the consequences for corporate responsibility.

This research contributes to the expanding field of AI ethics by offering empirical insights into how ethical considerations are navigated in real-world corporate settings. It builds on previous theoretical contributions while bridging the gap between normative principles and applied practices. Ultimately, this study seeks to inform policymakers, corporate leaders, and AI developers about the ethical responsibilities associated with AI implementation in business decision-making. By critically engaging with cross-disciplinary literature and conducting rigorous empirical analysis, the research aspires to advance both academic discourse and practical guidance on aligning AI innovation with ethical corporate governance.

2. Research Method

This study employs a qualitative research design to explore the ethical implications of artificial intelligence implementation in corporate decision-making. Qualitative methodology is appropriate for capturing the complexity and context-dependent nature of ethical phenomena, particularly when organizational behaviors and values are involved [25]. Through an interpretivist paradigm, this research seeks to understand how corporate actors perceive and navigate ethical challenges posed by AI technologies within their decision-making processes. Data were collected through semi-structured in-depth interviews with key decision-makers, including executives, managers, and ethics officers from a diverse sample of multinational corporations operating in finance, technology, and manufacturing sectors. This purposive sampling approach ensures that participants have direct exposure to AI-driven systems and experience in ethical deliberation [25]. Interviews were guided by an open-ended protocol that explored themes such as ethical dilemmas, decision transparency, accountability, and value alignment in the context of AI adoption. All interviews were audio-recorded with informed consent and transcribed verbatim for analysis.

To ensure rigor and trustworthiness, thematic analysis was employed as the primary analytic technique, allowing for the identification, analysis, and interpretation of patterns within qualitative data [26]. The analysis was conducted iteratively, beginning with open coding, followed by axial coding to explore relationships among themes related to AI ethics in decision-making. NVivo software was used to assist in organizing and managing the coding process, enhancing reliability and traceability of analytic decisions. Throughout the research, credibility was maintained through triangulation with corporate policy documents and AI ethics guidelines, while member checking was conducted with selected participants to validate interpretations. Ethical approval for the study was obtained from an institutional ethics review board, and confidentiality of all respondents was strictly preserved. By situating the analysis within the broader context of ethical AI frameworks and organizational governance literature [27], the study provides a rich, nuanced understanding of how ethical concerns are perceived, negotiated, and operationalized within contemporary corporate environments.

3. Result and Discussion

Prevalent Ethical Concerns in AI-Driven Corporate Decision-Making (N = 25) on Table 1.

Table 1. Prevalent Ethical Concerns in AI-Driven Corporate Decision-Making (N = 25)

Ethical Theme	Number of Mentions (N=25)	Percentage (%)
Algorithmic Bias	18	72%
Lack of Transparency	21	84%
Responsibility Diffusion	15	60%
Privacy Concerns	19	76%
Autonomy Undermining	17	68%

A recurring concern raised by participants was the prevalence of algorithmic bias in AI-assisted decision-making. Of the 25 corporate executives and ethics officers interviewed, 72% explicitly mentioned bias as an ethical issue, as shown in Table 1. These participants expressed apprehension about AI systems perpetuating historical inequalities embedded in training data or learned through discriminatory patterns. Several respondents noted that hiring algorithms tended to favor candidates based on demographic traits that inadvertently reflected past discriminatory hiring trends. These findings align with prior research indicating that biased data inputs can result in prejudiced outputs, undermining fairness in corporate environments [28].

Despite increased awareness, most companies lacked formal mechanisms to audit or mitigate algorithmic bias. Respondents admitted that while technical teams occasionally flagged bias concerns, few organizations had standardized ethical review procedures embedded into AI development workflows. This gap in ethical

infrastructure reinforces earlier claims that organizational accountability often lags behind technical innovation [20]. Several participants also described tensions between data science departments and compliance units, indicating that ethical accountability was often fragmented and reactive rather than anticipatory. This aligns with Dignum's argument that organizational design must evolve to incorporate ethics-by-design principles if AI is to be deployed responsibly [29].

Moreover, interviewees highlighted how algorithmic bias can intersect with broader diversity, equity, and inclusion (DEI) initiatives, either supporting or subverting them. For instance, one HR director described how an AI recruitment tool was initially celebrated for efficiency but later revealed to disadvantage female candidates for leadership roles-a pattern traced back to biased language in historical performance evaluations. Such real-world implications emphasize that algorithmic fairness must be contextually evaluated, not merely optimized mathematically [23].

The most cited concern, raised by 84% of participants, was the lack of transparency in AI decision-making. Executives emphasized difficulties in understanding how AI systems arrived at particular decisions, especially when using complex machine learning models such as deep neural networks. These concerns reflect ongoing scholarly debates about the black-box nature of AI and the ethical importance of explainability [2][28]. Many respondents noted that opacity undermined trust, not only within the organization but also among clients, regulators, and the public.

In practice, a lack of explainability manifested in various ways. For instance, one financial services firm had deployed an AI credit scoring tool but was unable to provide clients with a clear rationale for loan rejections. This absence of transparency led to regulatory scrutiny and reputational risk. Similar cases have been documented in the literature, where unexplained outcomes from AI systems create accountability vacuums and legal ambiguities [25] [30].

While some companies experimented with explainable AI (XAI) tools, their adoption remained limited. Participants mentioned that technical explanations often failed to satisfy non-technical stakeholders, such as compliance officers or end-users. The disconnect between technical validity and ethical acceptability illustrates what Morley et al. refer to as the interpretability gap. Bridging this gap requires not only algorithmic solutions but also organizational capacity building, including training for ethics literacy and interdisciplinary collaboration [30].

Furthermore, several respondents emphasized the importance of aligning AI transparency with existing corporate governance frameworks. One general counsel from a multinational corporation stressed that legal compliance increasingly requires demonstrating that automated decisions are traceable and justifiable. This trend reflects the broader regulatory momentum behind AI governance, particularly in jurisdictions advancing AI-specific legal regimes, such as the EU's AI Act [7] In this context, transparency is not merely a technical feature but a normative requirement with implications for legal risk and ethical integrity.

Closely tied to transparency is the issue of responsibility diffusion, reported by 60% of participants. Several respondents observed that when AI systems influenced or even made decisions, it became unclear who was ultimately accountable. This phenomenon, often described as moral crumple zones [12], occurs when responsibility is distributed among technical developers, corporate executives, and AI outputs, leading to ethical ambiguity.

One risk officer recounted an internal investigation into a pricing algorithm that had dynamically adjusted premiums in ways that disproportionately affected elderly customers. Although no single party intentionally encoded discrimination, responsibility was diffused across multiple teams—data scientists, product designers, and marketing managers—each assuming someone else had verified the ethical implications. This finding echoes studies that critique current corporate structures for lacking clearly defined roles when AI is embedded into operations [11][16].

To address this, some organizations introduced AI ethics committees or governance boards, yet these mechanisms were often ad hoc or lacked enforcement power. Participants expressed concern that without binding accountability structures, ethical responsibilities could be easily bypassed in the pursuit of efficiency or market advantage. This aligns with Jobin et al, who warn that voluntary ethical commitments in corporate AI strategies are insufficient without robust oversight mechanisms [8].

Notably, the diffusion of responsibility was more acute in firms where AI was integrated across multiple functions, such as marketing, logistics, and human resources. In such cases, ethical oversight required cross-functional collaboration, which was often hindered by siloed knowledge and conflicting priorities. Several interviewees acknowledged that ethical failures were rarely the result of isolated decisions but emerged from systemic interactions-a dynamic that demands holistic ethical governance models rather than fragmented controls [30].

Privacy concerns were mentioned by 76% of participants, particularly in sectors handling sensitive customer or employee data. Many respondents emphasized the tension between AI's appetite for large, granular datasets and emerging legal and ethical standards for data protection. These insights reflect the growing importance of privacy-by-design principles and compliance with regulations like the General Data Protection Regulation (GDPR) in shaping corporate AI strategies [9] [13].

Interviewees described AI applications that collected behavioral data for customer profiling, productivity monitoring, or predictive analytics. In several cases, employees were unaware of the extent to which their digital footprints were used to inform managerial decisions. Such practices raise serious ethical questions about consent, autonomy, and surveillance in the workplace [29]. One HR manager shared that after introducing a performance prediction tool, employee trust deteriorated due to perceptions of being constantly monitored, despite assurances of anonymized data use.

Additionally, the ethical salience of privacy varied across cultural and regional contexts. Respondents from firms operating in Asia and Latin America reported fewer employee concerns about digital surveillance, citing different cultural attitudes toward privacy. This supports earlier studies that emphasize the contextual and normative dimensions of data governance ethics [2]. However, most participants agreed that global firms must harmonize privacy practices to meet the highest ethical and legal standards across jurisdictions.

Some companies adopted privacy impact assessments (PIAs) and engaged external data ethics consultants to mitigate risks. Nevertheless, these measures were often implemented post-hoc or in response to external pressure rather than internal ethical initiative. The findings underscore the necessity for proactive, value-driven approaches to privacy in AI systems-approaches that go beyond legal compliance to address ethical legitimacy and stakeholder expectations [3].

The integration of AI into core business functions has raised profound concerns regarding the erosion of human autonomy in decision-making. Approximately 68% of respondents expressed apprehension that AI systems, particularly those with prescriptive or automated capabilities, were increasingly supplanting human judgment, thereby diminishing the role of ethical deliberation. One senior executive described a scenario in which managerial discretion was overridden by predictive maintenance algorithms that mandated operational changes without room for human override. These accounts support concerns raised in the literature regarding the risk of moral deskilling, where overreliance on AI tools leads to the weakening of human ethical sensitivity and agency [15].

Participants frequently noted that while AI offered significant efficiency gains, it also introduced a form of ethical disengagement. Employees, especially those in middle management, reported feeling disempowered when asked to implement recommendations made by algorithms without sufficient context or rationale. This detachment created a compliance-oriented mindset rather than one grounded in critical ethical reflection, echoing findings by Green and Viljoen who argue that AI-driven governance often reframes complex ethical issues into simplified procedural tasks [18]. The disconnection between AI outputs and human evaluative reasoning leads to a diffusion of moral responsibility, making it harder to hold any party accountable for ethically contentious decisions [2].

Moreover, the risk of autonomy erosion was particularly evident in domains involving stakeholder interaction, such as customer service or employee evaluation. In some firms, AI tools were used to provide sentiment analysis of employee emails or to assess tone during video meetings for performance reviews. While marketed as tools to enhance objectivity, such systems were criticized by participants for discouraging open communication and fostering a climate of self-censorship. These concerns align with research on AI-mediated environments that disrupt traditional notions of trust, consent, and interpersonal responsibility [23]. The ethical consequences of diminished autonomy are not merely abstract; they impact employee morale, workplace culture, and the overall legitimacy of organizational decisions.

Several respondents acknowledged that human-in-the-loop systems were proposed as a solution to retain some degree of human control. However, in practice, these systems often functioned as formalities rather than substantive checks on AI authority. This observation is consistent with critiques that highlight the performative nature of human oversight in many algorithmic systems, where decisions are rarely overturned or critically assessed [25]. Effective preservation of autonomy, therefore, requires more than symbolic gestures-it necessitates reconfiguring organizational workflows to prioritize ethical reasoning alongside technical accuracy.

In response to the growing recognition of AI's ethical implications, organizations have adopted a range of adaptation strategies, though their scope and depth vary significantly. A majority of respondents indicated that their companies had begun to formalize ethical AI governance, either through internal policies, ethics committees, or adherence to external frameworks such as the OECD AI Principles or the EU AI Act proposals. These initiatives reflect a broader trend toward institutionalizing AI ethics as a function of corporate governance [27]. However, the degree to which these efforts translate into meaningful organizational change remains inconsistent.

Participants from companies in highly regulated industries, such as finance and healthcare, reported more advanced ethical protocols. These included mandatory algorithm audits, documentation of model logic, and cross-functional review boards involving ethicists, technologists, and legal experts. One notable example was a European insurance firm that had developed a responsibility matrix to assign ethical accountability at each stage of AI development, from data sourcing to deployment. Such practices are aligned with calls for operationalizing AI ethics through clear roles, accountability metrics, and value-sensitive design principles [8].

Conversely, respondents from less regulated sectors often described a more reactive or minimalist approach. Ethical considerations were typically addressed post-deployment or only when prompted by external scrutiny. A technology startup, for example, admitted to relying on public feedback and investor pressure as primary triggers for ethical reviews. This mirrors earlier findings that voluntary ethical commitments, while symbolically valuable, lack the enforcement mechanisms necessary for systemic change [12]. The inconsistency between proactive and reactive approaches across firms raises concerns about the scalability and durability of AI ethics in the absence of binding regulatory frameworks.

Notably, several organizations had invested in capacity-building initiatives aimed at fostering an ethical culture. These included workshops on AI ethics for managers, collaborative ethics sprints involving diverse stakeholders, and internal forums for voicing ethical concerns related to technological deployment. Such measures indicate a shift from compliance-driven ethics toward culture-driven engagement, a trend supported by scholars advocating for normative internalization of ethical values within organizational systems [25]. However, participants cautioned that cultural change requires sustained leadership commitment, resource allocation, and integration into performance metrics-elements often lacking in fast-growing or resource-constrained firms.

Interestingly, organizations that reported higher ethical maturity also emphasized the importance of iterative learning and feedback mechanisms. For example, one global logistics company implemented ethical impact assessments as part of their agile development cycle, allowing for real-time ethical evaluation and system redesign. This adaptive approach supports the emerging view of ethics as a dynamic process embedded in socio-technical systems rather than a fixed set of rules [1] [30]. By embedding reflexivity and stakeholder participation into AI development, such organizations demonstrate pathways for reconciling innovation with ethical resilience [31].

4. Conclusion

This study reveals that the implementation of artificial intelligence in corporate decision-making presents significant ethical challenges, particularly concerning algorithmic bias, transparency, responsibility diffusion, privacy, and the erosion of human autonomy. While some organizations have begun to adopt ethical frameworks and governance structures, many still struggle to translate principles into practice due to a lack of regulatory pressure, internal capacity, and cultural alignment. The findings underscore the need for proactive, context-sensitive, and interdisciplinary approaches to AI ethics that go beyond compliance, embedding ethical reasoning into organizational processes. Ultimately, fostering ethical AI adoption requires a systemic reconfiguration of corporate structures, responsibilities, and values to ensure that technological advancements align with societal expectations and moral accountability.

References

[1] Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., & Vayena, E. (2018). AI4People-An Ethical Framework for A Good AI Society: Opportunities, Risks, Principles, and Recommendations. *Minds and Machines, 28(4), 689–707*. DOI: https://doi.org/10.1007/s11023-018-9482-5 .

[2] Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The Ethics of Algorithms: Mapping the Debate. *Big Data & Society, 3(2).* DOI: https://doi.org/10.1177/2053951716679679 .

[3] Ajunwa, I. (2020). The Paradox of Automation as Anti-Bias Intervention. *Cardozo Law Review, 41(4), 1671–1742.* DOI: https://doi.org/10.2139/ssrn.2747236 .

[4] Ananny, M., & Crawford, K. (2018). Seeing Without Knowing: Limitations of The Transparency Ideal and Its Application to Algorithmic Accountability. *New Media & Society, 20(3), 973–989*. DOI: https://doi.org/10.1177/1461444816676645 .

[5] Dignum, V. (2019). Responsible Artificial Intelligence: How to Develop and Use AI In A Responsible Way. *Springer*. DOI: https://doi.org/10.1007/978-3-030-30371-6 .

[6] Cowls, J., & Floridi, L. (2018). Prolegomena to A White Paper On An Ethical Framework for A Good AI Society. *Minds and Machines, 28(4), 689–707*. DOI: https://doi.org/10.1007/s11023-018-9482-5 .

[7] Barocas, S., Hardt, M., & Narayanan, A. (2019). Fairness And Machine Learning: Limitations and Opportunities. DOI: https://doi.org/10.48550/arXiv.1908.09635 .

[8] Jobin, A., Ienca, M., & Vayena, E. (2019). The Global Landscape of AI Ethics Guidelines. *Nature Machine Intelligence, 1*(9), 389–399. DOI: https://doi.org/10.1038/s42256-019-0088-2 .

[9] Lepri, B., Oliver, N., Letouzé, E., Pentland, A., & Vinck, P. (2018). Fair, Transparent, and Accountable Algorithmic Decision-Making Processes. *Philosophy & Technology, 31, 611–627*. DOI: https://doi.org/10.1007/s13347-017-0279-x .

[10] Ryan, M., & Stahl, B. C. (2020). Artificial Intelligence Ethics Guidelines for Developers and Users: Clarifying Their Content and Normative Implications. *Journal of Information, Communication and Ethics in Society, 18(2), 261–284*. DOI: https://doi.org/10.1108/JICES-12-2019-0138 .

[11] Stahl, B. C., Timmermans, J., & Mittelstadt, B. D. (2021). Ethics of Technology: Theoretical Foundations and Practical Application. *Technology in Society, 65*, 101620. DOI: https://doi.org/10.1016/j.techsoc.2021.101620 .

[12] Smuha, N. A. (2019). The EU Approach to Ethics Guidelines for Trustworthy Artificial Intelligence. *Computer Law Review International, 20(4), 97–106.* DOI: https://doi.org/10.9785/cri-2019-200402 .

[13] Wachter, S., Mittelstadt, B., & Floridi, L. (2017). Why A Right To Explanation of Automated Decision-Making Does Not Exist In The General Data Protection Regulation. *International Data Privacy Law, 7*(2), 76–99. DOI: https://doi.org/10.1093/idpl/ipx005 .

[14] Winfield, A. F., & Jirotka, M. (2018). Ethical Governance Is Essential to Building Trust In Robotics and Artificial Intelligence Systems. *Philosophical Transactions of the Royal Society A, 376(2133), 20180085*. DOI: https://doi.org/10.1098/rsta.2018.0085 .

[15] Mökander, J., Axente, M., Casado, M., Creus, J., Dignum, V., & Dignum, F. (2021). Operationalizing AI Ethics: Barriers, Enablers and Next Steps. *AI and Ethics, 1*, 431–445. DOI: https://doi.org/10.1007/s43681-021-00047-5 .

[16] Gasser, U., & Almeida, V. A. F. (2017). A Layered Model for AI Governance. *IEEE Internet Computing*, *21*(6), 58–62. DOI: DOI: https://doi.org/10.1109/MIC.2017.4180835 .

[17] Voigt, P., & Von dem Bussche, A. (2017). The EU General Data Protection Regulation (GDPR). *Springer*. DOI: https://doi.org/10.1007/978-3-319-57959-7 .

[18] Green, B., & Viljoen, S. (2020). Algorithmic Realism: Expanding The Boundaries of Algorithmic Thought. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 19–31. DOI: https://doi.org/10.1145/3351095.3372840 .

[19] Fjeld, J., Achten, N., Hilligoss, H., Nagy, A., & Srikumar, M. (2020). Principled Artificial Intelligence: Mapping Consensus In Ethical and Rights-Based Approaches to Principles for AI. *Berkman Klein Center Research Publication*, (2020-1). DOI: https://doi.org/10.2139/ssrn.3518482 .

[20] Binns, R. (2018). Fairness In Machine Learning: Lessons From Political Philosophy. *Proceedings of the 2018 Conference on Fairness, Accountability, and Transparency*, 149–159. DOI: https://doi.org/10.1145/3287560.3287598 .

[21] Risse, M. (2019). Human Rights and Artificial Intelligence: An Urgently Needed Agenda. *Human Rights Quarterly*, *41*(1), 1–16. DOI: https://doi.org/10.1353/hrq.2019.0000 .

[22] Braun, V., & Clarke, V. (2019). Reflecting On Reflexive Thematic Analysis. *Qualitative Research in Sport, Exercise and Health*, *11*(4), 589–597. DOI: https://doi.org/10.1080/2159676X.2019.1628806 .

[23] Binns, R., Veale, M., Van Kleek, M., & Shadbolt, N. (2018). It's Reducing A Human Being to A Percentage: Perceptions of Justice In Algorithmic Decisions. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1–14. DOI: https://doi.org/10.1145/3173574.3173951 .

[24] Coeckelbergh, M. (2020). AI ethics. *MIT Press*. DOI: https://doi.org/10.7551/mitpress/12231.001.0001 .

[25] Doshi-Velez, F., & Kim, B. (2017). Towards A Rigorous Science of Interpretable Machine Learning. *arXiv preprint*. DOI: https://doi.org/10.48550/arXiv.1702.08608 .

[26] Elish, M. C. (2019). Moral Crumple Zones: Cautionary Tales In Human-Robot Interaction. *Engaging Science, Technology, and Society*, *5*, *40–60*. DOI: https://doi.org/10.17351/ests2019.261 .

[27] Holstein, K., Wortman Vaughan, J., Daumé III, H., Dudik, M., & Wallach, H. (2019). Improving Fairness In Machine Learning Systems: What do Industry Practitioners Need?. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–16. DOI: https://doi.org/10.1145/3290605.3300830

[28] Lipton, Z. C. (2018). The Mythos of Model Interpretability. *Communications of the ACM, 61(10), 36–43*. https://doi.org/10.1145/3233231 .

[29] Mittelstadt, B. D. (2019). Principles Alone Cannot Guarantee Ethical AI. *Nature Machine Intelligence, 1(11), 501–507*. DOI: https://doi.org/10.1038/s42256-019-0114-4 .

[30] Morley, J., Floridi, L., Kinsey, L., & Elhalal, A. (2021). From What to How: An Initial Review of Publicly Available AI Ethics Tools, Methods and Research to Translate Principles Into Practices. *Science and Engineering Ethics, 27(1), 1–31*. DOI: https://doi.org/10.1007/s11948-020-00239-7 .

[31] Taddeo, M., & Floridi, L. (2018). How AI can be a force for good. *Science, 361(6404), 751–752*. DOI: https://doi.org/10.1126/science.aat5991 .