Volume 3, No. 11 November 2025
p-ISSN 3032-3037 | e-ISSN 3031-5786

ADVANCES IN SOCIAL HUMANITIES RESEARCH

AI Ethics and Risk in Indonesia's Manufacturing Sector

Brilian Adi Saputra
Swiss German University, Indonesia
Email: brilianadi@gmail.com

Abstract
The manufacturing sector is a cornerstone of Indonesia's economy, contributing substantially to GDP and employment. As manufacturers increasingly adopt artificial intelligence (AI) and related Industry 4.0 technologies, they face new ethical challenges and operational risks. This paper reviews the ethical implications of AI deployment in Indonesian manufacturing, examines risk management strategies and governance practices, and surveys relevant policy frameworks. Drawing on recent literature and reports, we identify key ethical concerns (such as worker displacement, algorithmic bias, data privacy, and safety) and explore how manufacturers can incorporate risk management (e.g., ISO 31000 and FMEA-based methods) to build trustworthy AI systems. We also assess Indonesia's national AI strategy and regulatory context (including the 2020–2045 AI National Strategy and recent Kominfo guidelines), noting strengths and gaps in current governance. The discussion synthesizes best practices for responsible AI use in manufacturing, and the conclusion offers recommendations for policymakers and industry, such as updating standards, investing in AI ethics and talent, and fostering multi-stakeholder oversight.

Keywords: *manufacturing sector, artificial intelligence, ethical, risk*

Introduction

Indonesia's manufacturing sector is a major pillar of national development. In 2025, it remained the single largest contributor to GDP, growing 4.6% and underpinning overall economic growth (bps, 2025). With over 19 million workers employed and significant output in food, electronics, automotive, and other industries, manufacturing is seen as central to the country's vision (e.g., *"Making Indonesia 4.0"*) for industrial modernization. In this context, artificial intelligence (AI) and automation are viewed as key enablers of productivity and competitiveness. Studies estimate that AI could boost Indonesian manufacturing productivity by tens of percent over the next decade (IBM, 2025) (CSIS, 2020). However, the introduction of AI systems also raises ethical concerns and risks that can affect workers, consumers, and society (National Center for Biotechnology Information, 2025).

AI ethics broadly refers to ensuring that AI technologies are used in ways that respect values like human rights, fairness, accountability, and safety (Vyhmeister et al., 2023). In manufacturing, AI-driven robotics, predictive maintenance, and quality control algorithms promise efficiency but may also displace jobs, encode biases, infringe on

privacy, or create new safety hazards (Sivanathan et al., 2025). For example, automating a production line could reduce some repetitive jobs even as it creates technical roles, raising questions of worker retraining and social equity (Sivanathan et al., 2025). Autonomous equipment must remain safe and transparent to human operators, or else risks of accidents increase (Xu et al., 2025). Moreover, AI systems rely on data that may include personal or proprietary information, posing privacy and security challenges. Without clear guidelines and oversight, manufacturers may struggle to align innovation with ethical use (Mehrabi et al., 2023). As one recent study notes, "practitioners struggle with understanding and implementing [AI risk] framework. This lack of understanding exposes manufacturing to a multitude of risks, including the organisation, its workers, as well as suppliers and clients" (Alexandra et al., 2023).

Indonesia has recognized these challenges. The government launched a National Strategy on Artificial Intelligence (*Stranas KA*) for 2020–2045 to guide AI development with industry input. In late 2023, the Ministry of Communications and Information (*Kominfo*) issued a ministerial circular (No. 9/2023) on AI ethics, providing principles for system providers to mitigate AI risks (UNESCO, 2024) (Azhar, 2025). UNESCO and *Kominfo* are also collaborating on an *"AI Readiness Assessment"* to benchmark Indonesia's AI governance against global ethics standards. Despite these steps, comprehensive AI regulation is still in formation. Current laws (e.g., the 2008 Electronic Information and Transactions Law) cover AI only indirectly.

Alexandra et al. (2023) highlighted that manufacturing practitioners often lack a clear understanding of AI risk frameworks, resulting in inconsistent implementation of ethical standards and exposing industries to safety, privacy, and bias-related risks. However, their study focused on multinational corporations in Western economies and did not explore the regulatory or cultural dimensions influencing AI ethics in Southeast Asia. Similarly, Vyhmeister et al. (2023) emphasized the global challenge of aligning AI innovation with ethical governance, noting that most AI ethics guidelines remain normative and lack practical application models in industrial contexts (Mehrabi et al., 2023). Despite identifying accountability and transparency as critical issues, their research did not provide specific insights into how national strategies—such as Indonesia's *Stranas KA* or *Kominfo*'s Circular No. 9/2023—translate into ethical implementation within the manufacturing sector.

This academic paper reviews the literature on AI ethics and risk in manufacturing, with special attention to the Indonesian context. We survey scholarly research on AI ethics frameworks in industry, examine governance and policy developments in Indonesia, and discuss how manufacturing firms can manage AI risks. The goal is to provide a systematic analysis suitable for policymakers, industry leaders, and researchers interested in promoting responsible AI in Indonesia's manufacturing sector, offering both theoretical contributions to AI ethics literature and policy-relevant insights for Indonesian regulators and manufacturers aiming to balance innovation with social responsibility.

Methodology

This paper employed a qualitative, conceptual methodology based on a comprehensive literature review and policy analysis (SecurePrivacy, 2025). We surveyed peer-reviewed journals, conference proceedings, and institutional reports from roughly the past decade (2014–2025), focusing on the intersection of AI ethics, risk management, and the manufacturing sector. Databases such as IEEE Xplore, SpringerLink, and Google Scholar were searched with keywords including "AI ethics manufacturing", "industry 4.0 Indonesia", "AI governance policy", and "AI risk management". Credible sources include academic articles (e.g. those from the AI and Ethics journal), think tank and government reports (e.g. IBM Center studies, UNESCO analyses), and official Indonesian documents. From this corpus, we extracted information on four topical areas: (1) ethical implications of AI in manufacturing (worker and societal impact, algorithmic issues, data considerations); (2) risk management strategies and frameworks relevant to AI in industry (including references like ISO 31000, FMEA, and the NIST AI RMF); (3) governance practices in both corporate and regulatory contexts (transparency, accountability mechanisms, multi-stakeholder models); and (4) policy and legal frameworks specific to Indonesia (national AI strategy, data laws, AI guidelines). Where possible, findings from general literature were interpreted in the Indonesian manufacturing context. For example, generic risks like "loss of productivity due to AI error" were considered alongside local factors such as Indonesia's workforce structure and technology adoption rates. As a conceptual study, no new empirical data were collected. Instead, the analysis synthesizes insights from existing sources to paint a coherent picture of the current state and gaps (Exabeam, 2025). We organized the findings thematically, as presented in the Discussion section, to highlight connections between ethical concepts, risk management, and governance. Throughout, we paid attention to the ASEAN/Indonesian context by incorporating regional reports, statements by Indonesian officials, and local case examples where available.

Discussion

Ethical Implications for Manufacturing. Deploying AI in manufacturing involves multiple ethical dimensions. A first set of concerns is human-cantered ethics. Automation and AI-driven optimization can improve efficiency but may do so at the expense of human workers if not managed responsibly. Ethical AI deployment should prioritize worker welfare and human oversight. For example, predictive maintenance AI might notify a worker of impending equipment failure, aiding rather than replacing the technician. By contrast, a fully autonomous system that removes workers from the loop risks deskilling and potential unemployment. As UNESCO's policy advisors observe, Indonesia should ensure AI serves community well-being with a human-centric approach.

A second cluster of issues is fairness and bias (National Center for Biotechnology Information, 2025). In manufacturing, biases can enter when AI systems make decisions about resource allocation or quality standards. An illustrative scenario is recruitment: if an AI system filters job applications for factory work using historical hiring data, it may perpetuate past biases (e.g. gender or ethnicity imbalances). Similarly, algorithms that

predict which parts fail most often might disadvantage smaller suppliers whose limited data do not reflect typical patterns. Mitigating such bias requires transparent algorithms and diverse data (World Journal of Advanced Research and Reviews, 2025). The AI ethics literature stresses the need for data governance to address bias. In Indonesia, where Bahasa Indonesia and many local languages are underrepresented in data, ensuring linguistic and cultural fairness in AI is also a concern. Manufacturers should audit their AI for unequal impacts, for example by analysing error rates across different product lines or worker groups.

Data privacy and security are another ethical domain. Manufacturing AI often uses large datasets, some of which may involve personal data (e.g. employee biometric or performance data, customer order information) (GDPR Local, 2025). Ethical deployment means complying with privacy principles. Indonesia's new Personal Data Protection Law (PDPL, enacted 2022) will require manufacturers to obtain consent and secure data when AI applications involve personal data. For instance, an AI-driven quality control camera system that incidentally captures workers' faces must handle that data under the PDPL. Ethical practice demands data minimization (only collecting what is needed) and anonymization where possible. Companies should also secure data against breaches, since IP theft and data leaks are major risks in industry 4.0 ecosystems. Recent reports emphasize that Indonesia's growing data infrastructure needs enhanced cybersecurity measures to maintain public trust

Risk Management Strategies. Effective mitigation of ethical and operational risks requires a structured approach (Vyhmeister & Castane, 2023). One recommended strategy is to adapt established risk management frameworks to the AI context. For example, the TAI-PRM framework (Trustworthy AI – Project Risk Management) explicitly combines ISO 31000 (risk management standard) and FMEA (Failure Mode and Effects Analysis) to identify "ethical hazards" in AI projects. Manufacturers can similarly treat potential issues like "algorithmic error", "data breach", or "lack of human oversight" as failure modes to be analysed. The TAI-PRM case study showed that this method allowed identification and tracking of multiple ethical failure modes in an EU industrial AI project, demonstrating how risk management can operationalize ethics Indonesian firms might adopt analogous processes: cross-functional teams (including engineers, ethicists, and line operators) could systematically list AI-specific risks in each stage of system development, and develop control measures (e.g. additional training data, supervision protocols, rollback procedures).

Another strategy is to use global AI risk frameworks like NIST's AI Risk Management Framework (AI RMF, 2023). The AI RMF, while not industry-specific, outlines risk categories (accuracy, robustness, privacy, security, etc.) that manufacturers can apply. For instance, under "robustness", companies would ensure AI continues to function safely under unexpected conditions. Under "privacy", they would assess how data is collected and stored. Although NIST is a US initiative, its risk taxonomy is widely referenced. ISO is also developing standards for AI (such as the forthcoming ISO/IEC 42001 for AI management systems) which Indonesian companies can align with. The key

point is that risk management should be proactive and continuous: as Brintrup et al. (2023) note, manufacturing organizations often lack clear implementation of AI-risk frameworks, leading to risks for workers and supply chains

Practically, risk management in Indonesian manufacturing may involve specific measures:

1. Governance structures: Companies could establish an AI oversight committee or designate a Chief AI Officer responsible for compliance. Regular audits of AI systems, similar to financial audits, could check for ethical adherence (bias testing, traceability of decisions).
2. Training and culture: Building an "AI risk-aware" culture is important. Workers and managers need training on both the benefits and potential harms of AI tools. For example, workers should know how an AI quality inspection system works and have channels to report malfunctions or unexpected outputs
3. Safety integration: In factories, AI safety must integrate with existing worker safety systems. New hazards introduced by AI (e.g. robot motion errors) should be treated similarly to mechanical hazards, with guards, emergency stops, and fail-safes designed into the AI system (National Center for Biotechnology Information, 2025).
4. Cyber-physical security: Since manufacturing systems are increasingly interconnected (IoT, cloud controls), risk management must encompass cyber-physical threats. This means encrypting communications, isolating critical systems, and having rapid incident-response plans for cyber-attacks.

Overall, the goal is to "harness AI's transformative potential while safeguarding against risks" by embedding ethics into every stage of design and operation.

Governance Practices and Policies. Governance refers to the rules, institutions, and processes that oversee AI use at a higher level (Sustainability, 2025). In the Indonesian manufacturing context, governance involves both corporate policies and government regulation. We have noted that the government has started to set principles (e.g. Kominfo's circular, data protection law). Companies must complement these with internal governance. For instance, many global firms institute AI ethics guidelines or boards (often involving legal, HR, and technical experts) to review new AI initiatives for compliance with values like fairness and privacy (Ghobakhloo et al., 2025). Indonesian manufacturers could follow suit, perhaps drawing on sectoral industry associations to develop best-practice guidelines tailored to local conditions (Vyhmeister & Castane, 2023).

At the governmental level, multi-stakeholder involvement is emphasized by experts. The UNESCO analysis recommends involving civil society and academia in policy-making. For manufacturing AI, this might mean including labor unions, consumer groups, and industry bodies in discussions about AI standards. For example, when setting acceptable levels of machine automation, regulators should consider worker perspectives on job quality. It is also suggested to adopt a "regulatory sandbox" approach (as used in fintech), allowing companies to pilot AI innovations under relaxed rules but with oversight. This can speed adoption while monitoring unforeseen issues. Such balanced

regulatory methods are advised for Indonesia to foster innovation and address risks (Wellbrock et al., 2025).

Currently, Indonesia's main AI guideline (the Kominfo circular) is a starting point but limited in scope. It requires system providers to be "responsible and accountable" but does not specify enforcement mechanisms. Going forward, specialists recommend formalizing more detailed regulations or even an AI law. GovInsider reports that a draft regulation is in the works. This law could mandate risk assessments for high-impact AI (similar to the EU AI Act's risk categories) or require transparency measures (e.g. disclosing AI use in products). Meanwhile, existing regulations touch on related areas: for example, Indonesia's upcoming data protection regulation will impose obligations on data controllers (including those using AI), and industrial safety laws will apply to AI-driven machinery as they would to any equipment.

Indonesia also participates in international AI governance discussions, which influences domestic policy. Officials have stated a commitment to global cooperation and ethical AI for the public good. The ASEAN region has its own AI principles (in line with OECD principles) that call for inclusive growth and safeguarding human rights. Indonesian manufacturers should align with these by, for example, implementing international standard safeguards in their AI supply chains (e.g. ensuring foreign AI suppliers comply with Indonesian law and values).

Key Ethical and Risk Themes. To summarize, the discussion highlights several recurrent themes:

a. Transparency and Explainability: Black-box AI is ethically problematic. Manufacturing firms should strive to use explainable AI models where possible, especially for decisions affecting people or safety.
b. Fairness and Non-discrimination: AI outputs should be tested for bias regularly, for instance by checking if defect detection rates are similar across different product batches.
c. Human Oversight: Automated systems should be overseen by humans, with clear processes for intervention. Critical decisions (like shutting down a production line) should not be left entirely to AI without human review.
d. Accountability: Responsibility for AI decisions must be traceable. Companies need to document their AI development processes and decision logic so that, if a harmful outcome occurs, they can assign and manage liability.
e. Privacy and Security: As noted, data ethics is crucial. Data collected for AI should respect consent and data minimization. Robust cyber defenses are part of this theme.

These themes are supported by both academic sources and policy recommendations. For instance, IBM's analysis of the Indonesian context stresses that "governance is the key to build trust" in AI, involving all stakeholders. This underscores the need for clear rules, oversight, and engagement – otherwise trust erodes and risks increase.

Conclusion

Indonesia's manufacturing sector stands at a critical juncture where AI and automation offer significant productivity benefits but also introduce complex ethical and governance challenges unique to the country's regulatory, social, and industrial context. While frameworks like the National AI Strategy (*Stranas KA*) and *Kominfo*'s Circular No. 9/2023 provide important groundwork, they need to evolve into enforceable standards embedded in manufacturing practices. Future research should focus on developing context-specific, actionable models for AI ethics implementation, including the effectiveness of regulatory expansions, institutional coordination, and ethical risk management systems like ISO 31000. Additionally, studies could assess workforce training initiatives, the role of multi-stakeholder collaboration, and cybersecurity measures tailored for small and medium enterprises to create a holistic governance approach that balances innovation, competitiveness, and social inclusion in Indonesia's manufacturing sector.

References

Azhar, M. (2025). Indonesia calls for inclusive and equitable AI governance. *GovInsider Asia*. https://govinsider.asia/intl-en/article/indonesia-calls-for-inclusive-and-equitable-ai-governance

Brintrup, A., Pak, J., Ratiney, D., Pearce, T., Wichmann, P., Woodall, P., & McFarlane, D. (2023). Trustworthy, responsible, ethical AI in manufacturing and supply chains: Synthesis and emerging research questions. *AI and Ethics, 3*(1), 159–174. https://doi.org/10.48550/arXiv.2305.11581

Centre for Strategic and International Studies. (2020). *Emerging technology in Indonesia's manufacturing sector*. https://www.csis.or.id/publication/emerging-technology-in-indonesias-manufacturing-sector/

Exabeam. (2025). *Anonymization vs. pseudonymization: Key differences for data privacy*. https://www.exabeam.com/

GDPR Local. (2025). *How AI GDPR will shape privacy trends in 2025*. https://gdprlocal.com/

Ghobakhloo, M., Iranmanesh, M., Foroughi, B., Karuppiah, N., & Preece, C. (2025). Industry 5.0 and human-centric manufacturing: A systematic review and integrative framework. *Future Internet, 17*(9), Article 430. https://doi.org/10.3390/fi17090430

IBM. (2024). *Study finds Indonesian FSI, manufacturing companies leaning into AI, but skills gap and data governance hinder progress*. https://asean.newsroom.ibm.com

Indonesia, BPS-Statistics. (2025). *Indonesia's economy sustains growth amid global uncertainty - News and press release*. https://www.bps.go.id/en/news/2025/05/05/703/ekonomi-indonesia-tetap-tumbuh-di-tengah-ketidakpastian-global.html

Kazim, E., & Koshiyama, A. S. (2021). A high-level overview of AI ethics. *Patterns, 2*(9), Article 100314. https://doi.org/10.1016/j.patter.2021.100314

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2023). Fairness and bias in artificial intelligence: A brief survey of sources, impacts, and mitigation strategies. *Sci, 6*(1), Article 3. https://doi.org/10.3390/sci6010003

National Center for Biotechnology Information. (2025). *AI trustworthiness in manufacturing: Challenges, toolkits, and the path to Industry 5.0* [PMC Articles, PMC12298069]. U.S. National Library of Medicine. https://www.ncbi.nlm.nih.gov/pmc/

SecurePrivacy. (2025). *Compliance challenges at the intersection between AI & GDPR in 2025*. https://secureprivacy.ai/

Sivanathan, A., Foo, D., Lim, H. J., Gharakheili, H. H., & Sivaraman, V. (2025). AI trustworthiness in manufacturing: Challenges, toolkits, and the path to Industry 5.0. *Sensors, 25*(14), Article 4357. https://doi.org/10.3390/s25144357

Sustainability. (2025). Research on the impact of artificial intelligence on the resilience of the manufacturing industry chain. *Sustainability, 17*(21), Article 9775. https://doi.org/10.3390/su17219775

UNESCO. (2024). *Leading AI ethics: UNESCO and KOMINFO launch AI readiness assessment methodology in Indonesia*. UNESCO.

Vyhmeister, E., & Castane, G. G. (2023). Risk as a driver for AI framework development on manufacturing. *AI and Ethics, 3*(1), 155–174. https://doi.org/10.1007/s43681-022-00159-3

Vyhmeister, E., & Castane, G. G. (2025). TAI-PRM: Trustworthy AI-project risk management framework towards Industry 5.0. *AI and Ethics, 5*(2), 819–839.

Wellbrock, W., Malinovska, M., & Ludin, D. (2025). Ethical implications and potential opportunities and risks of artificial intelligence in supply chain management. *Discover Sustainability, 6*, Article 886. https://doi.org/10.1007/s43621-025-01808-3

World Journal of Advanced Research and Reviews. (2025). Algorithmic bias, data ethics, and governance: Ensuring fairness in AI systems. *WJARR, 25*(02), 1746–1763.

Xu, J., Sun, Q., Han, Q.-L., & Tang, Y. (2025). When embodied AI meets Industry 5.0: Human-centered smart manufacturing. *IEEE/CAA Journal of Automatica Sinica, 12*(3), 485–501. https://doi.org/10.1109/JAS.2025.125327