

TECHNOLOGY

The Algorithmic Compass: Navigating Ethical Decision-Making in the Age of AI-Driven Management

by Ch. Mahmood Anwar



Image Credit | Louis Deconinck

A guide to ethical AI in management

✓ **INSIGHT** | FRONTIER 02 Jun 2025

The digital revolution has brought us to a time where artificial intelligence (AI) is now a part of our everyday organizational life and managerial decision-making relies on it. AI is helping with things like speeding up recruitment processes and predicting future market trends. This highlights AI's crucial role in improving how we work by making tasks more effective and efficient. However, along with AI's power, there is a vital need to focus on ethics and fairness. As AI becomes more widespread, managers face challenges related to bias, transparency about how AI functions, and its impact on society. It is crucial for managers to be aware of these important issues while using AI in their work, especially in decision-making (Williamson & Prybutok, 2024).

RELATED CMR ARTICLES

Peter Cappelli, Prasanna Tambe, Valery Yakubovich, “**Artificial Intelligence in Human Resources Management: Challenges And a Path Forward**,” *California Management Review*, 61/4 (2019): 15–42.

Vegard Kolbjørnsrud, “**Designing the Intelligent Organization: Six Principles For Human-AI Collaboration**,” *California Management Review*, 66/2 (2023): 44–64.

One of the significant ethical issues we are addressing today is algorithmic bias. AI systems learn from historical data, which means they can continue and even worsen existing societal biases. This can result in unfair treatment of people. For example, recruitment algorithms based on past hiring practices might unintentionally exclude qualified candidates from minority groups (Dastin, 2018). To prevent this, it is crucial to take active steps to review and improve AI systems to ensure they are fair and ethical. According to Wachter *et al.* (2021), the problem of algorithmic fairness is not just about technology; it is also deeply connected to politics and society. Therefore, solving this issue needs a broad approach that includes different people and groups.

Many AI systems work like ‘black boxes,’ which means people cannot see what happens inside. This makes it hard to know how they make choices, raising concerns about clarity and responsibility. Explainable AI (XAI) is very important because it helps build trust, allowing managers to understand why AI makes certain decisions. According to Miller (2019), explaining things is a big part of how humans think and understand. It is important to show why and how AI does what it does to help humans. Mittelstadt (2019) points out that when we talk about the ethics of AI, it is not only about what AI can do technically, but also about what it should do ethically. This means it is important to make sure that how we develop AI is in line with what society values and believes is right. Managers should ask AI sellers for XAI features and make sure they learn how to understand and use these explanations correctly.

Below, I shall present important aspects that should be considered by managers to ensure ethical use of AI (see Figure 1):

Bias Mitigation

AI is increasingly being used in hiring, evaluating work performance, and understanding different customer groups. While this offers many advantages, it is important to first address any biases that might be present in these systems. To tackle these biases, we must carefully examine the data we are using. Mehrabi *et al.* (2021) point out that bias can enter at any stage of data handling, so regular checks are essential to ensure fair representation of different groups. In hiring, this means analyzing the training data to avoid historical biases against certain demographics. For job performance evaluations, it is crucial to gather feedback from diverse sources to prevent reliance on a limited, possibly biased dataset. For understanding customer segments, we need to be cautious about which features we use. For instance, using a zip code can inadvertently introduce socioeconomic biases. To make AI fairer, we can use ‘adversarial debiasing’ techniques, which train models to ignore sensitive attributes like race or gender, as suggested by Madras *et al* (2018). Additionally, being transparent about data sources and the reasons behind our decisions helps us improve continuously. Regular feedback from stakeholders involved is vital to spot any new biases and ensure AI systems reflect the ethical values of the organization.

Explainable AI (XAI)

Explainable AI (XAI) goes beyond just being clear; it is about building trust and helping managers make well-informed decisions. To effectively use Explainable AI (XAI), you need a strategic plan. First, you should figure out why you need explanations. Is it for meeting regulations, earning user trust, or fixing problems with the model? Understanding this helps you decide how detailed the explanations need to be. As Molnar (2020) mentions, “interpretability is not a single concept,” and choosing the right XAI method depends on your specific need. Next, focus on using model-agnostic methods like Local Interpretable Model-agnostic Explanations (LIME) or SHapley Additive exPlanations (SHAP) because they are versatile and work well with complex models that are hard to interpret. However, if you are working with models like decision trees, which are already easier to understand, use their natural clarity to your advantage. Also, it is important to design ‘explanation interfaces’ that suit the level of expertise of the user. According to Ribeiro *et al.* (2016), visualizing local explanations and highlighting important features can greatly help users grasp specific scenarios. Another important step is establishing a feedback loop. Regularly ask users if the explanations are clear and useful, and make improvements based on their feedback. Additionally, make sure to integrate XAI into every stage of the AI workflow, from data preparation to the deployment of models. This practice is crucial for regularly checking for biases and for taking necessary actions early on. Finally, it is essential to document the XAI processes in detail, including the methods used and their limits. This documentation supports accountability and makes it easier for others to replicate your approach. By following these steps, organizations can turn XAI from just an idea into a practical tool that helps build AI systems people can trust and understand.

Ethical Frameworks

Traditional ethical frameworks are important foundation for business ethics, but they need to be adapted for managing AI. Utilitarianism, which aims to make the most people happy, struggles with AI because AI can have unexpected effects, and it is difficult to measure happiness in complex systems. Researchers like Mittelstadt *et al.* (2016) point out that AI often spreads responsibility among many agents, making it difficult to determine

who is accountable for outcomes. Deontology, which focuses on following rules and duties, also faces challenges in setting universal rules for AI. This is because technology changes quickly, and cultural contexts differ widely. Therefore, instead of sticking strictly to one ethical approach, a combination of ideas is needed. For example, a utilitarian approach could help in designing an AI system initially, while deontological principles could ensure compliance with data privacy regulations. Moreover, virtue ethics, which emphasizes good character and moral behavior, becomes increasingly important in creating an ethical culture around AI in organizations. Vallor (2016) emphasizes that cultivating ‘technomoral virtues’ such as transparency, accountability, and fairness is essential for responsible AI management. This involves integrating ethical considerations into every stage of AI development and deployment, promoting a thoughtful and proactive approach to handling the unique moral dilemmas that AI presents.

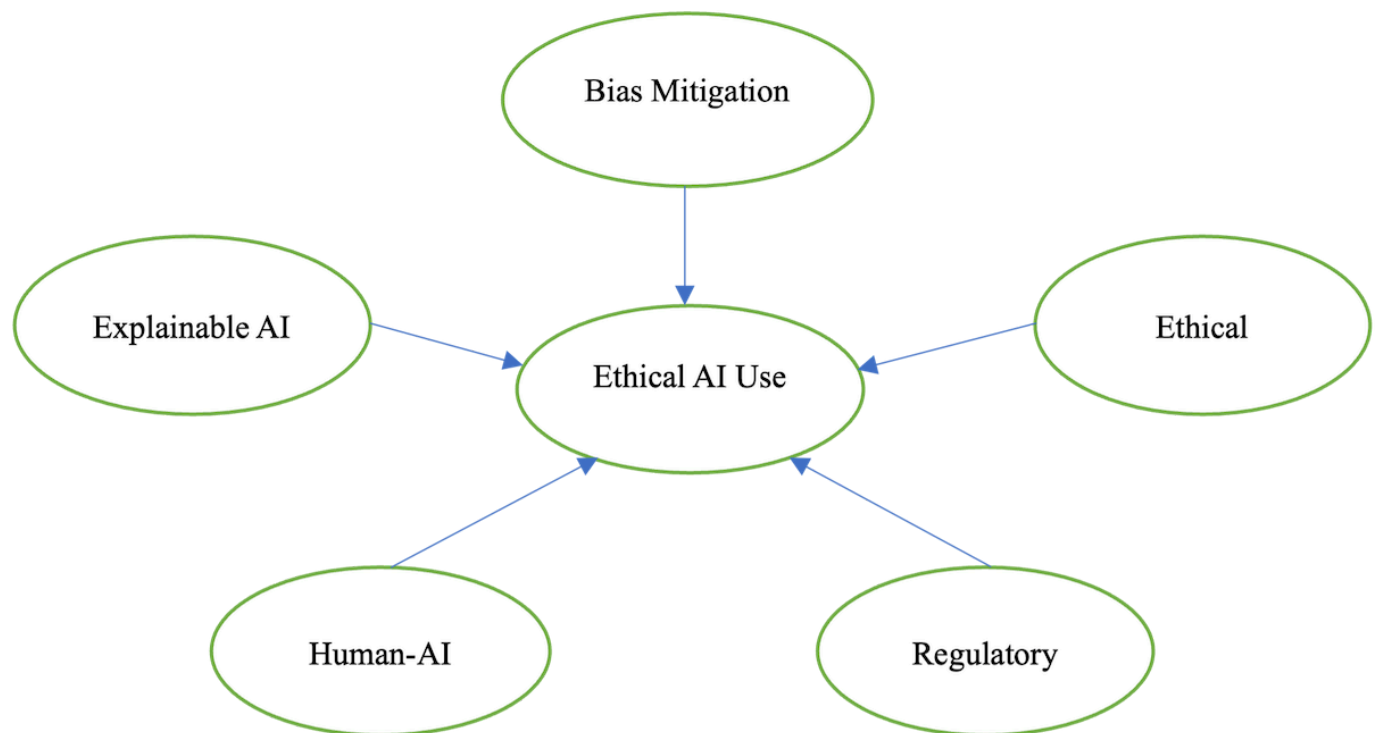


Figure 1: Important Aspects of Ethical AI Use

Human-AI Collaboration

To work well with AI, managers need to change their role from just giving orders to being partners with AI. It is important to use AI for tasks like analyzing data, finding patterns, and making predictions. Managers should focus on human skills such as empathy, creativity, and making ethical choices. As Rahwan (2018) states, AI should be a partner, not a rival, allowing managers to concentrate on important organizational decisions. Building a strong understanding between humans and AI is a key. This means training managers to know what AI is good at, what it is not, how to ethically interpret AI results, and how to provide feedback to improve AI. Natarajan *et al.* (2023) emphasize designing systems where both humans and AI can start and manage tasks together in a flexible manner. Making AI processes easy to understand helps building trust and smooth collaboration. Managers must be able to question AI's advice to maintain critical oversights. Setting up clear communication methods and feedback systems ensures easy information flow and continuous learning. By creating a respectful and understanding working culture, organizations can truly benefit from the collaboration between humans and AI.

Regulatory Compliance

To effectively collaborate with AI, managers must transition from a traditional command-and-control approach to one that emphasizes partnership and cooperation. The rules surrounding AI ethics are changing quickly, so organizations must stay proactive in following them. There is a noticeable shift from simply having voluntary ethical guidelines to establishing strict laws, especially regarding data privacy and the fairness of algorithms. As mentioned by Yeung (2018), 'regulatory innovation' is necessary to match the fast-paced development of AI. One important step is to create an internal AI ethics board made up of legal, technical, and ethical specialists. This board's role is to stay updated on regulatory changes and check for any potential risks. Regular AI audits should be conducted by organizations, focusing on the sources of data, the transparency of algorithms, and the detection of biases. This is in line with recommendations from the EU's proposed AI Act (European Commission, 2021). Implementing strong data governance policies is essential, highlighting the use of minimal data and having clear

reasons for its use, which is important for following GDPR and similar regulations. Additionally, putting ‘human-in-the-loop’ systems in place is crucial, as it ensures that human oversight and responsibility are maintained in important AI-related decisions. Engaging with regulators and taking part in industry conversations about AI ethics can offer valuable insights into new best practices. Finally, it is important to document all processes related to AI development and implementation, including risk assessments and ways to handle potential issues, as this is key for proving compliance and building trust with stakeholders.

All the above aspects have been intelligently tabulated in Table 1.

Feature	AI Bias Mitigation	Explainable AI (XAI)	Ethical Frameworks	Human-AI Collaboration	Regulatory Compliance
Definition	Techniques and processes for reducing and eliminating unfair biases in AI systems.	Methods to make AI decision-making processes transparent and understandable to humans.	Guiding principles and values for the ethical development and deployment of AI.	Strategies for effectively combining human and AI capabilities to enhance decision-making and outcomes.	Adherence to laws, regulations, and standards governing the development and use of AI.

Feature	AI Bias Mitigation	Explainable AI (XAI)	Ethical Frameworks	Human-AI Collaboration	Regulatory Compliance
Key Goals	Ensure fairness and equity in AI outcomes.	Increase transparency and trust in AI.	Establish ethical guidelines for AI development.	Optimize the synergy between human and AI intelligence.	Meet legal and regulatory requirements.
	Prevent discriminatory or harmful impacts.	Enable accountability for AI decisions.	Promote responsible AI practices.	Enable effective human oversight and control.	Minimize legal and financial risks.
	Promote inclusivity.	Facilitate debugging and improvement of AI models.	Address societal impacts of AI.	Improve decision quality and efficiency.	Ensure compliance with industry standards.
	Data preprocessing and augmentation.	Model interpretation techniques (LIME, SHAP).	Developing ethical codes of conduct.	Human in loop systems.	Staying informed about relevant regulations (e.g., EU AI Act).
	Algorithmic bias detection and correction.	Visualization of AI decision processes.	Conducting ethical impact assessments.	Collaborative AI design and development.	Implementing compliance monitoring systems.
Key Practices	Fairness metrics and evaluation.	Providing rationale and explanations for AI outputs.	Implementing ethical governance structures.	Tools for human-AI interaction.	Documenting AI systems and processes.

Feature	AI Bias Mitigation	Explainable AI (XAI)	Ethical Frameworks	Human-AI Collaboration	Regulatory Compliance
Importance	Crucial for building trust and ensuring AI systems are used responsibly and ethically.	Essential for accountability, transparency, and building user trust in AI.	Fundamental for guiding the ethical development and deployment of AI in society.	Vital for maximizing the benefits of AI while maintaining human oversight and control.	Necessary for mitigating legal risks and ensuring responsible AI implementation.
					Compliance with data privacy regulations (GDPR).
Examples of application	Fairness in loan approval algorithms.	Explaining credit risk scores.	AI ethics guidelines in healthcare.	AI-assisted medical diagnosis with human review.	Adherence to industry-specific AI standards.
	Bias detection in facial recognition systems.	Providing reasons for medical diagnoses.	Responsible AI in autonomous vehicles.	Collaborative robots in manufacturing.	
	Equal representation in AI training data.	Showing how AI makes hiring decisions.	Ethical data usage in marketing.	Human oversight in AI-driven financial trading.	Reporting AI system information to regulatory bodies.

Table 1: Key Aspects of Ethical AI Use

Conclusion

To develop artificial intelligence (AI) responsibly, we need a simple and clear plan that focuses on fairness, openness, ethics, collaboration, and following rules. These parts are not separate; they work together to ensure AI is used the right way. It is important to tackle

bias, making AI systems easy to understand, base AI development on ethical principles, build strong partnerships between humans and AI, and keep up with new regulations. Doing these things helps us deal with AI's complexities. This well-planned approach is not just about safety; it is absolutely necessary for building trust, getting the most benefits from AI, and ensuring it has a positive and fair impact on society. The journey to ethical AI does not stop; it requires us to always pay attention and adapt as needed. The aim is to make sure technology helps people in a fair and sustainable way, now and in the future.

References

1. Dastin, J. **“Amazon scraps secret AI recruiting tool that showed bias against women,”** Reuters, (2018).
2. **“Proposal for a regulation of the European parliament and of the council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain union legislative acts,”** European Commission, (2021).
3. Madras, D., Pitassi, T., & Zemel, R., “Fairness through causal awareness: Learning causal relationships for fair decision-making,” Proceedings of the 35th International Conference on Machine Learning, PMLR, (2018): 3383–3392.
4. Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A., “A survey on bias and fairness in machine learning. ACM Computing Surveys,” 54/6 (2021): 1-35.
5. Miller, T., “Explanation in artificial intelligence: Insights from the social sciences,” Artificial Intelligence, 267 (2019): 1–38.
6. Mittelstadt, B. D., “Principles alone cannot guarantee ethical AI,” Nature Machine Intelligence, 1/11 (2019): 501–507.
7. Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L., “The ethics of algorithms: Mapping the debate. Big Data & Society,” 3/2 (2016).
8. Molnar, C., “Interpretable machine learning: A Guide for Making Black Box Models Explainable,” Leanpub, (2020).
9. Natarajan, M., Seraj, E., Altundas, B. et al., “Human-Robot Teaming: Grand Challenges,” Current Robotics Reports, 4 (2023): 81–100.
10. Rahwan, I., Society-in-the-loop: “Programming the algorithmic social contract,” Ethics and Information Technology, 20/1 (2018): 5–14.

11. Ribeiro, M. T., Singh, S., & Guestrin, C., “Why should I trust you?: Explaining the predictions of any classifier,” Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining ACM, (2016): 1135–1144.
 12. Vallor, S., “Technology and the virtues: A philosophical guide to a future worth wanting,” Oxford University Press, (2016).
 13. Wachter, S., Mittelstadt, B., & Russell, C., “Why fairness cannot be automated: Bridging the gap between data-driven AI and the law,” Computer Law and Security Review, 41 (2021): 105567.
 14. Williamson, S. M., & Prybutok, V., “The era of artificial intelligence deception: Unraveling the complexities of false realities and emerging threats of misinformation,” Information, 15/6 (2024): 299.
 15. Yeung, K., “Algorithmic regulation: A critical interrogation,” Regulation and Governance, 12 (2018): 505–523.
-



Ch. Mahmood Anwar [Follow](#)

Prof. Ch. Mahmood Anwar is a research consultant, editor, author, entrepreneur, and HR and project manager. His research interests include critical analysis of published business research, research methods, new constructs development and validation, theory development, business statistics, business ethics, and technology for business.