Managing Ethical Risks in Artificial Intelligence: A Project Management Framework for Responsible Innovation [1]

Rishiraj Kohli

Ph.D. Candidate, Information Technology (Artificial Intelligence)
University of the Cumberlands, USA

Abstract

While the widespread acceptance of Artificial Intelligence (AI) offers significant benefits, it also carries considerable ethical risk. Most governance approaches are currently grounded in principles, e.g., EU AI Act, OECD Principles, and IEEE standards, but there are limited practical tools to use after the ethics "principles" and potentially respond in the project management context. This paper proposes a project-management–based solution to address ethical risk in AI projects, it develops an ethical risk management framework comprising structured and operational steps to embed responsible innovation in the life of a project. The framework includes an Ethical Risk Score (ERS) system, process for involving stakeholder perspectives, and links well with the leading international guidelines emerging internationally. The framework was validated with expert input and demonstrated though a comparison to established governance frameworks with strength in applicability and operational clarity. Integration of ethical consideration has been shown through this work, that project structures can be a bridge between principle and practice in ethical governance strategies in AI projects. Future work will progress the framework proposed in this paper to the piloting stage to assess scaling and adaptability in a number of domains. This research helps advance the field toward implementation and metrics to enable accountable, trustworthy AI solutions. The study contributed valuable proposals to both research agendas in information systems and the emerging body or research in technology governance.

Keywords: *Artificial Intelligence; Ethical Risks; AI Governance; Trustworthy AI; Technology Ethics; Responsible Innovation Framework; Project Management; AI Risk Management; Ethical AI*

[1] How to cite this paper: Kohli, R. (2025). Managing Ethical Risks in Artificial Intelligence: A Project Management Framework for Responsible Innovation; *PM World Journal*, Vol. XIV, Issue X, October.

1. Introduction

One of the 21st century's most important technologies is rapidly becoming artificial intelligence (AI). It has already affected a wide range of sectors, including government, education, manufacturing, banking, and healthcare. AI is acknowledged for its capacity to boost productivity, make better decisions, and improve forecasting precision. However, there are growing worries about its moral ramifications.

The European Union's AI Act, the OECD AI Principles, and other ethics guidelines created by research and academic organizations are only a few of the international initiatives being taken to address these problems. However, stakeholders and organizations are left with little ethical advice for AI applications because these texts are still largely in the development stage.

Project management, which focuses on structured planning, execution, monitoring, and risk management, is a valuable but not very well-known way to add ethics to AI projects. PMI's PMBOK, PRINCE2, and Agile frameworks are examples of established methodologies that already have ways to find and reduce risks. The study focuses on three primary objectives: (i) to delineate the principal ethical risks linked to AI; (ii) to incorporate these risks into the phases of project management; and (iii) to suggest a framework for project managers to facilitate responsible innovation.

2. Materials and Methods

This research used a methodological process that combined a conceptual–analytical method with a case-informed validation method to create a framework to manage ethical risks in AI projects. The methodological process was divided into three stages: (i) literature review and synthesis; (ii) framework development; and (iii) validation via case analysis.

In the first stage, an extensive literature review was conducted using databases, which included Scopus, Web of Science, and IEEE Xplore, to identify existing AI ethical guidelines and risk management approaches and project management methods. The search terms employed were “AI ethics,” “responsible innovation,” “ethical risk management,” and “project management in AI.” Publications were prioritized from between 2015 to 2024 to capture the most recent developments in ethical project management.

In the second stage, synthesizing the information from the review led to the design of a project management framework that integrates ethical risk into the standard lifecycle of any project. The framework utilizes elements of established methods such as PMI’s

PMBOK, PRINCE2, and aspects of Agile while aligning to the ethical principles identified by OECD, UNESCO, and the European Union.

In the third stage, the framework underwent case-informed analysis to ensure validity and applicability, featuring three illustrative AI project scenarios in healthcare diagnostics, financial decision support, and autonomous mobility systems, all scenarios were selected as they represent a high level of ethical sensitivity and risk context diversity. Expert reviewer feedback gathered from professionals in AI governance and project management helped refine the framework further.

Figure 1. Methodological Process Flow

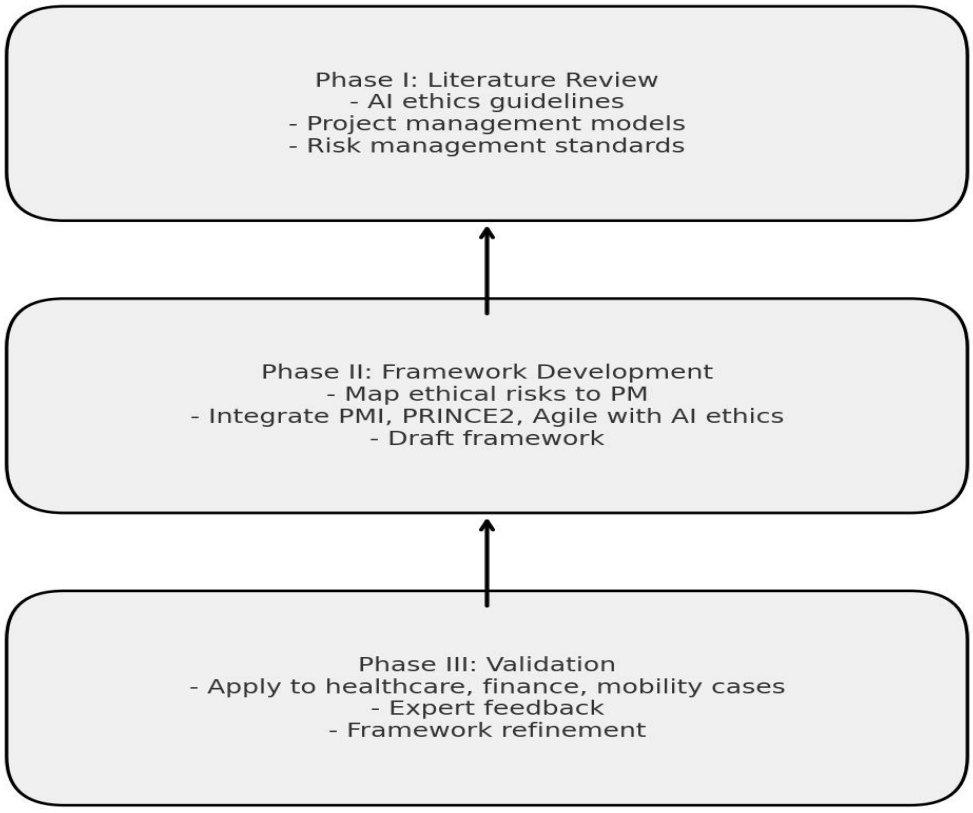

Figure 1. Methodological framework for developing and validating the AI ethical risk management model.

The methodological steps are summarized in Table 1.

Table 1. Overview of methodological phases in framework development

Phase	Description	Key Activities	Outputs
Phase I: Literature Review	Identification of existing approaches in AI ethics and project management	Database search (Scopus, WoS, IEEE); thematic coding	Synthesis of ethical risk themes
Phase II: Framework Development	Integration of ethical principles into project management workflows	Mapping PMI, PRINCE2, Agile with AI ethics guidelines	Draft project management framework
Phase III: Case-Informed Validation	Testing applicability of framework in ethically sensitive contexts	Application to healthcare, finance, and mobility cases; expert review	Refined and validated framework

3. Results

This research presents a Project Management Framework for Responsible AI Innovation that embeds ethical risk management into the project lifecycle. The framework is built on five project management stages: Initiation, Planning, Execution, Monitoring & Control, and Closure, each of which contains specific ethical risk management activities.

3.1 Framework Overview

The framework operationalizes ethical principles by incorporating them into project management tasks. Instead of treating ethics as a peripheral or follow-up consideration, ethical risks will be examined, evaluated, and alleviated at every stage of the AI project.

3.2 Key Elements of the Framework

Integrating Ethics into the PM Lifetime: Stages typical of the project management life cycle now include ethical checkpoints.

- **Risk-Based Ethics**: Ethical dilemmas can be seen as a category of risk that can be assessed, ranked, and treated in the same process as the use of a standard risk management tool.

- **Stakeholder Engagement**: Continuous engagement with various stakeholders ensures fairness and diversity.

- **Case-Informed Design**: The project has been developed through iterative cycles with case studies from the domains of healthcare, finance, and mobility AI.

3.3 Visual Representation

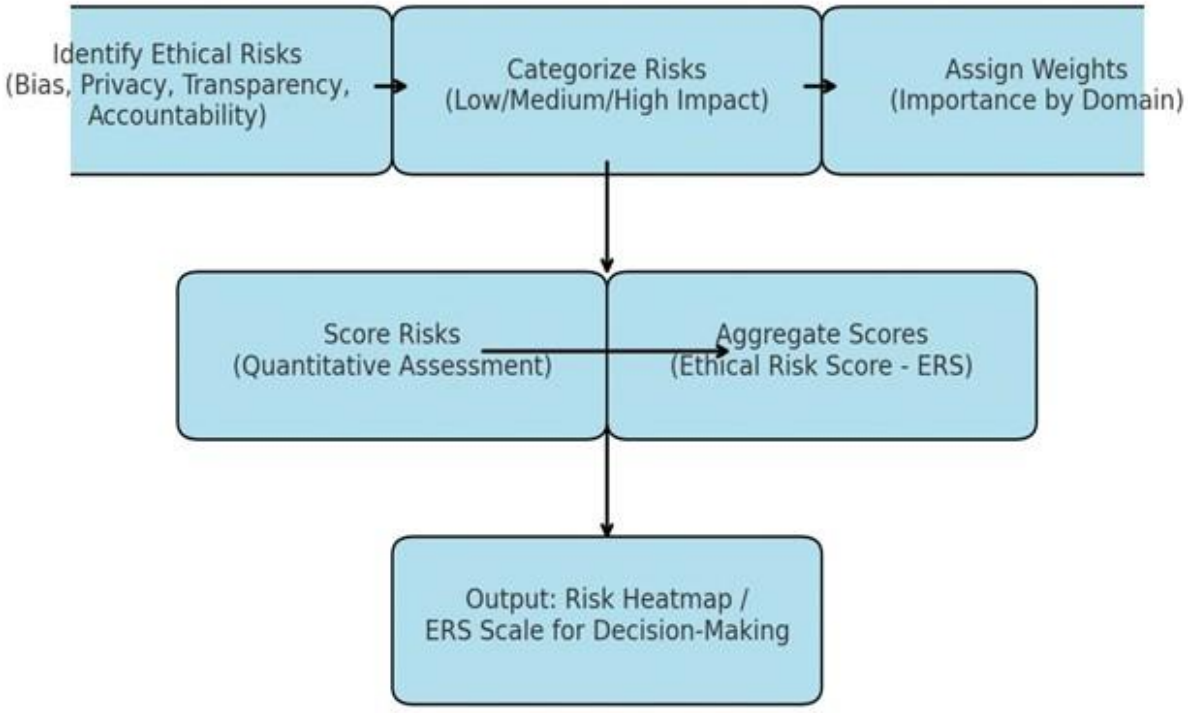

Figure 2: Ethical Risk Score (ERS) Process

3.4 Framework Representation

Table 2. Ethical risk management integration into project management phases

Project Phase	Ethical Risk Considerations	Key Activities	Expected Outputs
Initiation	Identification of ethical context, stakeholders, and societal impact	Define ethical objectives, conduct stakeholder mapping	Ethical risk register (initial)
Planning	Anticipating ethical risks in design and deployment	Ethical impact assessment, bias risk	Ethical risk management plan

		analysis, data governance plan	
Execution	Embedding safeguards in system development and deployment	Implement fairness checks, privacy-preserving mechanisms, accountability structures	AI system with ethical safeguards
Monitoring & Control	Continuous oversight of ethical performance	Ongoing audits, stakeholder feedback loops, bias detection tests	Updated ethical risk register
Closure	Evaluation of ethical outcomes and lessons learned	Post-project ethical review, transparency report	Final ethical compliance report

4. Discussion

I suggest that the findings of this study establish a new framework for integrating ethical risk management into AI project lifecycles, contributing to practice and theory. The framework harmonizes ethical checkpoints in with conditions of project management, which allows a bridge to be created from an abstract sets of principles around AI ethics to project management.

4.1 Comparison with Other Initiatives

There are a number of international initiatives - EU Ethics Guidelines for Trustworthy AI, OECD AI Principles, IEEE Ethically Aligned Design, that provide high level ethical standards, however software frameworks and implementation pathways remain more conceptual and often do not guide project managers on how to operationalize projects. Conversely, our framework:

- Provides an ethics based approach to the creation of a project, rather than undertaking as a post project compliance measure
- Employs a risk based approach that allows project managers to quantify, prioritize, and track ethical risks in the same way they would track technical or financial risks
- Provides accountability as the benchmarks for ethics are embedded within the normal deliverables of the project (risk registers, compliance reports, audits)
- Develops a practical link to existing ethics frameworks aligned for project management.

Principles / Requirements	EU AI Act	OECD AI Principles	IEEE Ethically Aligned Design	Proposed Project Mgmt Framework
Transparency	✓	✓	✓	✓
Accountability	✓	✓	✓	✓
Risk Management	✓	✓	✓	✓
Human Oversight	✓	✓	✓	✓
Fairness / Non-Discrimination	✓	✓	✓	✓
Sustainability / Societal Benefit	✘	✓	✓	✓

Figure 3: Comparative Alignment of Framework with Global AI Governance Guidelines

4.2 Implications for Project Management Practice

From a managerial standpoint, the framework presents some benefit, including:

- Operational Clarity: Project teams will easily be able to see if, when, and how they can conduct ethical assessments.
- Stakeholder Engagement: Involving various stakeholders at various stages increases transparency and fairness.
- Scalability: The framework can be used by different setups across sectors including healthcare, financial services, and autonomous systems.
- Integration with Agile and Traditional Approaches: Ethical checkpoints may be worked into sprints (Agile) or gates (Waterfall/PRINCE2).

This would suggest that project managers in AI-heavy industries or even other industries can implement parts of the framework without “blowing up” existing practices entirely.

4.3 Contributions to Theory

The study offers contributions to the literature by:

- Understanding responsible innovation theories for the first time in terms of project management practice.
- Creating an ethical risk lens into managing projects throughout the project lifecycle.
- Offering a replicable process creating framework from literature, establishing a framework, and validating framework informed by case studies.

4.4 Limitations

Our analysis demonstrates that it has practical value, this study has limitations. To begin with, the framework only has been validated in case-based scenarios, as opposed to a full-length live project, which is key to offering insight into longevity and applicability of the framework. The scope of the framework is also limited by the project management lens, as contexts that impact application and adaptation (e.g., organizational, cultural, and regulatory dynamics) have not been assessed. Moreover, ethical risks involving AI technologies rapidly shift and change, thus, the framework will need a continuous basis of review to ensure it meets these changing technologies. Whilst comparisons to leading global guidelines were made, there does not exist an empirical cross-validation process that covers whole sectors or geographic regions, demonstrated in evidence of the construct being established beyond the test cases cited.

In spite of the contributions listed above, some limitations need to be identified:

- The validation of the framework was limited to a case validation, rather than live project application.
- The ethical landscape is fluid, and ethical risks in AI technologies change rapidly, thus ethical codes and related documentation must be regularly revised and updated.
- Specific ethical safeguards may require adaptations, or reductions, to account for the particulars of a given industry or culture.

4.5 Future Directions

Some avenues for further research include:

- Empirical testing of the framework in a live project within a multi-industry context.
- Toolkits (e.g., dashboards, checklists, etc.) assist project managers in the application of ethical checkpoints as documentation.
- To establish a linkage of the respective organizational policy and governance structure and the ethical checkpoints applied during the course of the project, i.e., the ethical checkpoint must align with regulations.

5. Conclusions

This research provides a new vision for the conversation around AI ethics, moving beyond principle-based governance approaches to an ethical risk management framework-based on project-management expertise. Rather than creating a 'one-size-fits-all' set of mechanisms and processes, we make ethics a concern for every phase of implementation in the process of legitimacy-building. The framework provides concrete

mechanisms and tools, for instance the Ethical Risk Score (ERS), for practitioners to develop ethically competent action plans to address ethical issues systematically. Further, pairing the ethical risk management framework with existing reference points (the EU AI Act, OECD AI Principles, IEEE standards) provides a means of compliance and increases the use of existing resources towards practical implementation.

As limitation of this study, the framework has primarily been developed through expert consultation and comparative analysis rather than applying the ethical risk management framework to live AI projects. Furthermore, the framework has also been developed with reference to international reference points, and as such may require adjustments to integrate and comply with regional regulatory environments. The fast-changing nature of AI ethics field will also require revisions to the framework be made continuously.

Moving forward, we will focus on practical validation that studies the framework in action and is inclusive of case studies and pilots within industry with a potential for scalability across sectors such as healthcare, finance, and public administration. By working in this way we can create a link between theory and practice to support responsible AI innovation and further most effective integration of governance principles into future processes of development.

Author Declaration on AI Use

I certify that I only used artificial intelligence (AI) technologies to help with language editing, grammar correction, readability, and structural clarity when writing this manuscript. No discoveries, interpretations, or conclusions were unique because of AI. As the author, I am solely responsible for all research concepts, analytical effort, arguments, and intellectual contributions. AI was just utilized as a writing tool, and I maintain the work's originality and academic integrity.

REFERENCES

1. Dogru, A. K., & Keskin, B. B. (2020). AI in operations management: applications, challenges and opportunities. Journal of Data, Information and Management, 2(2), 67-74.

2. Tariq, B., Ashraf, M. R., & Rashid, U. (2025). Ethical Imperatives in AI Design: A Comprehensive Framework for Risk Mitigation and Responsible Innovation. Ubiquitous Technology Journal, 1(2), 61-73.

3. Li, W., Yigitcanlar, T., Browne, W., & Nili, A. (2023). The making of responsible innovation and technology: An overview and framework. Smart Cities, 6(4), 1996-2034.

4. Smith, Michael & Brown, Christopher & Asante, Godfred. (2025). The Rise of Intelligent Threats: AI-Driven Cyber Attacks in Multi-Cloud Environments.

5. Thomas, John & Wilson, Sarah & Asante, Godfred. (2025). Exploiting Artificial Intelligence for Cyber Attacks in Multi-Cloud Hosted Applications.

6. Clarke, Keron & Asante, Godfred. (2025). The Deficiency of Artificial Intelligence in the Cybersecurity Sphere.

7. Li, W., Yigitcanlar, T., Browne, W., & Nili, A. (2023). The making of responsible innovation and technology: An overview and framework. Smart Cities, 6(4), 1996-2034.

8. Xue, L., & Pang, Z. (2022). Ethical governance of artificial intelligence: An integrated analytical framework. Journal of Digital Economy, 1(1), 44-52.

9. De Almeida, P. G. R., Dos Santos, C. D., & Farias, J. S. (2021). Artificial intelligence regulation: a framework for governance. Ethics and Information Technology, 23(3), 505-525.

10. Yigitcanlar, T., Corchado, J. M., Mehmood, R., Li, R. Y. M., Mossberger, K., & Desouza, K. (2021). Responsible urban innovation with local government artificial intelligence (AI): A conceptual framework and research agenda. Journal of Open Innovation: Technology, Market, and Complexity, 7(1), 71.

11. Canals, J., & Heukamp, F. (2020). The future of management in an AI world. Palgrave Macmillan.

12. Weber-Lewerenz, B. (2021). Corporate digital responsibility (CDR) in construction engineering—ethical guidelines for the application of digital transformation and artificial intelligence (AI) in user practice. SN Applied Sciences, 3(10), 801.

13. Hickman, E., & Petrin, M. (2021). Trustworthy AI and corporate governance: the EU's ethics guidelines for trustworthy artificial intelligence from a company law perspective. European Business Organization Law Review, 22(4), 593-625.

14. Thiebes, S., Lins, S., & Sunyaev, A. (2021). Trustworthy artificial intelligence. Electronic Markets, 31(2), 447-464.

15. Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. Nature machine intelligence, 1(9), 389-399.

16. Crockett, K., Colyer, E., Gerber, L., & Latham, A. (2021). Building trustworthy AI solutions: A case for practical solutions for small businesses. IEEE Transactions on Artificial Intelligence, 4(4), 778-791.

17. Wu, W., Huang, T., & Gong, K. (2020). Ethical principles and governance technology development of AI in China. Engineering, 6(3), 302-309.

18. Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. Computers and Education: Artificial Intelligence, 2, 100041.

19. Raisch, S., & Krakowski, S. (2021). Artificial intelligence and management: The automation–augmentation paradox. Academy of management review, 46(1), 192-210.

20. Jarke, M., & Radermacher, F. J. (1988). The AI potential of model management and its central role in decision support. Decision Support Systems, 4(4), 387-404.Daly, A., Hagendorff, T., Hui, L., Mann, M., Marda, V., Wagner, B., ... & Witteborn, S. (2019). Artificial intelligence governance and ethics: global perspectives. arXiv preprint arXiv:1907.03848.

21. Fetzer, J. H. (1990). What is artificial intelligence?. In Artificial intelligence: Its scope and limits (pp. 3-27). Dordrecht: Springer Netherlands.

22. Jarrahi, M. H., Askay, D., Eshraghi, A., & Smith, P. (2023). Artificial intelligence and knowledge management: A partnership between human and AI. Business Horizons, 66(1), 87-99.

23. Jiang, Y., Li, X., Luo, H., Yin, S., & Kaynak, O. (2022). Quo vadis artificial intelligence?. Discover Artificial Intelligence, 2(1), 4.

24. Wu, X., Huang, Y., Liu, Z., Lai, W., Long, E., Zhang, K., ... & Lin, H. (2019). A universal artificial intelligence platform for collaborative management of cataracts. The Lancet, 394, S22.

25. Mannuru, N. R., Shahriar, S., Teel, Z. A., Wang, T., Lund, B. D., Tijani, S., ... & Vaidya, P. (2025). Artificial intelligence in developing countries: The impact of generative artificial intelligence (AI) technologies for development. Information development, 41(3), 1036-1054.

26. Almagharbeh, W. T., Alfanash, H. A., Alnawafleh, K. A., Alasmari, A. A., Alsaraireh, F. A., Dreidi, M. M., & Nashwan, A. J. (2025). Application of artificial intelligence in nursing practice: A qualitative study of Jordanian nurses' perspectives. BMC Nursing, 24, 42. https://doi.org/10.1186/s12912-024-02658-6

27. Almagharbeh, W. T. (2025). The impact of AI-based decision support systems on nursing workflows in critical care units. International Nursing Review, 72(1), e13011. https://doi.org/10.1111/inr.13011

28. Oladejo, A. O., Adebayo, M., Olufemi, D., Kamau, E., Bobie-Ansah, D., & Williams, D. (2025). Privacy-Aware AI in cloud-telecom convergence: A federated learning framework

for secure data sharing. International Journal of Science and Research Archive, 15(1), 005-022.

29. Adebayo, M. Deepfakes and Data Privacy: Navigating The Risks in the Age of AI. NDPC–, 106.

About the Author

Rishiraj Kohli

USA

Rishiraj Kohli is a Senior IT Project Manager at Schwarz Partners and a Ph.D. candidate in Information Technology with a concentration in Artificial Intelligence at the University of the Cumberlands (expected 2027). He is also a certified Project Management Professional (PMP) and Certified ScrumMaster (CSM). With extensive experience leading enterprise AI, digital transformation, and compliance-driven initiatives in insurance, logistics, and healthcare, he specializes in bridging governance, risk management, and innovation. His work focuses on making AI adoption accessible and responsible for non-technical professionals while ensuring regulatory alignment. He can be contacted at rkohli5848@ucumberlands.edu or rishirajkohli@gmail.com