

Literature Review on Ethical Issues in Challenges and Regulatory Considerations in AI-Driven Healthcare, Robotics, and the Metaverse

Course Title:	Research Methodology in Computer Science
Course Code:	AI800
Weighting:	30%
Lecturer:	Dr. Mohammed Anbar
Student Name:	Sahab AlHumaed
Initiated By/On:	20 /03/2025

TABLE OF CONTENTS

<i>1. Problem Stattment:</i>	3
<i>.2 Research Objective</i>	3
<i>4. References</i>	6

1. Problem Statement:

Artificial Intelligence (AI) implementation in healthcare creates significant ethical issues around transparency and accountability as well as governance.

The deployment of AI-driven diagnostic tools alongside robotic surgery improves healthcare precision and effectiveness but faces ongoing reliability and bias issues as well as ethical compliance challenges.

The "black-box" nature of AI models conceals diagnostic reasoning, reducing trust among healthcare professionals and patients (Wang, 2018). Algorithmic bias exacerbates healthcare disparities, as evidenced by IBM Watson's diagnostic errors and racial discrepancies in pulse oximeter readings (Farhud & Zokaei, 2021; Ness-Teo et al., 2024). Such biases jeopardize fair treatment and patient safety.

Data privacy and security risks remain critical, as demonstrated by the UK NHS-Google DeepMind privacy breach, which exposed regulatory gaps in AI governance (Baihakkia & Ba Qutayanaa, 2023). While frameworks like Responsible Research and Innovation (RRI) and embedded ethics exist, fragmented regulations hinder AI standardization across institutions and regions (Stahl & Coeckelbergha, 2016). This lack of uniform governance weakens public confidence and reduces AI system reliability.

This research critically analyzes existing AI ethics frameworks to identify their limitations and develop a structured governance model that enhances transparency, fairness, and accountability in healthcare AI applications.

2. Research Objective

1. To analyze the challenges of trust and transparency in AI-driven healthcare:

- Aims to uncover the basic transparency problems, accountability issues, and biases present in AI applications within healthcare systems.
- To analyze the consequences of AI systems with non-transparent decision-making on patient trust as well as clinician confidence and medical decisions.
- Aims to investigate specific real-world situations where the lack of transparency in AI systems created ethical problems and diminished trust in healthcare settings.

2. To evaluate existing AI governance models and regulatory frameworks for ethical AI integration:

- The evaluation will cover worldwide AI governance models such as Responsible Research and Innovation (RRI), GDPR and HIPAA together with additional international healthcare-related AI policies.
- Compare AI governance strategies from different nations and institutions to discover regulatory discrepancies and voids.
- The evaluation of AI developers, policymakers, and healthcare providers focuses on their responsibilities for deploying AI systems ethically while managing data privacy and accountability alongside fairness issues.

3. To propose a comprehensive framework for ethical AI integration in healthcare:

- The proposed governance model structures ethical AI principles into medical applications to maintain transparency and accountability while ensuring fairness.
- The development of an interdisciplinary method that brings together AI ethics principles with healthcare regulatory requirements and legal standards as well as technology developments is necessary.
- This research focuses on overcoming AI model explainability issues while maintaining accurate and practical healthcare applications.

4. To validate the proposed framework using real-world AI case studies:

- To analyze ethical dilemmas in AI healthcare through real-world cases such as IBM Watson's diagnostic bias and the NHS-Google DeepMind data privacy controversy.
- The comparison of AI governance policies from various healthcare systems allows for the identification of best practices essential for ethical AI deployment.
- This evaluation assesses the performance of ethical AI models in risk mitigation, compliance assurance, and fairness maintenance within healthcare environments.

5. To provide practical recommendations for ethical AI adoption in healthcare:

- Develop clear step-by-step guidelines which policymakers, AI developers and healthcare institutions can employ for responsible AI implementation.
- To support the establishment of global AI governance standards which ensure fairness alongside transparency and accountability.
- To seeks to sustain AI ethics research activities and policy adjustments to tackle new technological and ethical dilemmas.

3. Literature Review:

Artificial intelligence (AI) represents a transformative technology in healthcare that enhances diagnostic precision and surgical accuracy while making patient treatment more efficient. Without clear transparency and proper accountability and ethical oversight in AI decision-making processes healthcare professionals face significant concerns about system reliability and fair treatment and patient protection. The comprehensive literature analysis assesses current research on AI healthcare ethics while synthesizing multiple perspectives to reveal governance shortcomings.

Trust and Transparency in AI Healthcare : The successful implementation of AI systems in healthcare depends entirely on trust from users. The opaque functioning of advanced AI models creates distrust since medical professionals, patients, and regulators cannot understand the decision-making processes within these models (Farhud & Zokaei, 2021). Explainability problems in AI systems used for robotics and diagnostics prevent healthcare professionals from confirming AI recommendations according to Wang (2018). IBM Watson's AI system received negative feedback for patient misdiagnoses stemming from biased training data according to Ness-Teo et al. (2024). The incident of UK NHS partnering with Google DeepMind showcased ethical risks when AI systems accessed medical data without obtaining proper consent which led to data privacy and regulatory compliance concerns (Baihakkia & Ba Qutayanaa, 2023). The examples demonstrate an urgent need for AI frameworks which explain medical decision-making processes to ensure transparency.

Challenges and Limitations of Current AI Systems: The present state of AI in healthcare contains multiple challenges which foster distrust and create ethical dilemmas. AI algorithms that exhibit bias and discrimination deepen healthcare disparities especially among marginalized populations. AI medical devices like pulse oximeters show racial bias which results in wrong readings for patients who are not white (Baihakkia & Ba Qutayanaa, 2023). The absence of consistent ethical regulations worldwide splits AI governance into different parts which stops the development of joint accountability systems (Farhud & Zokaei, 2021). AI systems also struggle with accountability dilemmas—when AI-assisted decisions lead to medical errors, the question remains: Responsibility for AI-assisted medical errors lies with either the AI system itself or the developer or the physician. (Stahl & Coeckelbergha, 2016).

The highlighted problems demonstrate why we need urgent implementation of ethical governance systems to manage AI healthcare risks.

Explainable AI (XAI) offers a viable solution to implement AI systems ethically: XAI frameworks have been introduced to improve decision-making transparency and interpretability within AI systems.

While trust and transparency issues remain major concerns, Explainable AI (XAI) has been proposed as a potential solution. XAI frameworks aim to bridge the gap between AI predictions and human understanding by making decision-making processes more interpretable.

The objective of these frameworks is to connect AI predictions with human comprehension in the context of critical medical decisions (Farhud & Zokaei, 2021). XAI methods enhance transparency but fail to completely address trust and accountability challenges. Research demonstrates that non-experts find XAI models too complex to understand which creates barriers to clinical implementation (Stahl & Coeckelbergha, 2016). Embedded ethics functions as an alternative strategy where ethical principles become part of AI system development from its earliest stages instead of being addressed afterwards. The proactive approach encourages joint work between ethicists, medical experts, and AI engineers to create AI innovations that uphold responsibility (Baihakkia & Ba Qutayanaa, 2023).

Regulatory Frameworks for Ethical AI Governance: Through interdisciplinary collaboration Responsible Research and Innovation (RRI) governance frameworks promote ethical development of AI technologies (Baihakkia & Ba Qutayanaa, 2023). Multiple jurisdictions show regulatory fragmentation because they operate without a shared AI ethics framework (Ness-Teo et al., 2024). The absence of standardized regulations generates inconsistency in worldwide AI governance initiatives which results in divergent AI accountability and compliance standards. The EU has implemented robust data privacy protection through its GDPR regulations yet other parts of the world lack comparable protective measures (Farhud & Zokaei, 2021). Global harmonization of AI ethics policies plays a vital role in promoting responsible AI implementation within healthcare systems.

Implications for Future AI Ethics Research: The analyzed literature shows an immediate requirement for ethical oversight in healthcare systems driven by AI technology. XAI frameworks and regulatory models present potential solutions yet face ongoing issues with scalability and standardization along with practical application problems. Future research should focus on: The medical field requires strong AI governance systems which follow medical ethical principles and abide by regulatory requirements. Develop technological advancements that enhance AI transparency while maintaining model performance levels. The application of fairness mechanisms within AI systems serves to reduce both algorithmic bias and discriminatory outcomes. Global AI policies must be standardized to achieve fair governance in all regions.

4. References

1. Almusaylim, Z. A., Alhumam, A., Mansoor, W., Chatterjee, P., & Jhanjhi, N. Z. (2020). Detection and mitigation of RPL rank and version number attacks in smart Internet of Things. *Preprints*. <https://doi.org/10.20944/preprints202007.0476.v1>
2. Alsukayti, I. S., & Singh, A. (2022). A lightweight scheme for mitigating RPL version number attacks in IoT networks. *IEEE Access*, 10, 111115–111133.
3. Baihakkia, M. A., & Ba Qutayanaa, S. M. S. (2023). Ethical issues of artificial intelligence in healthcare. *Journal of Science and Technology in Policy*, 9(1), 32–45. <https://jostip.utm.my/index.php/jostip/article/view/129/89>
4. Mayzaud, A., Badonnel, R., & Chrisment, I. (2017). A distributed monitoring strategy for detecting version number attacks in RPL-based networks. *IEEE Transactions on Network and Service Management*, 14(2), 472–486. <https://doi.org/10.1109/TNSM.2017.2705290>
5. Nandhini, P. S., Srinath, P., Veeramanikandan, P., & Malliga, S. (2021). Version attack detection using claim algorithm in RPL-based IoT networks: Effects and performance parameters evaluation. In *Proceedings of the 2nd International Conference on Smart Electronics and Communication (ICOSEC 2021)* (pp. 209–215). <https://doi.org/10.1109/ICOSEC51865.2021.9591735>
6. Osman, M., He, J., Mokbal, F. M. M., Zhu, N., & Qureshi, S. (2021). ML-LGBM: A machine learning model based on light gradient boosting machine for the detection of version number attacks in RPL-based networks. *IEEE Access*, 9, 83654–83665. <https://doi.org/10.1109/ACCESS.2021.3087175>
7. Rouissat, M., Belkheir, M., Mokaddem, A., & Bachir, N. (n.d.). Parent supervision lightweight solution against version number attacks for IoT networks. *Preprints*.
8. Tay, Y. C., & Tan, J. P. (2022). Ethical AI in medical technology: A review. *Singapore Management University Research Collection*. https://ink.library.smu.edu.sg/cgi/viewcontent.cgi?article=10404&context=sis_research
9. Wei, X., & Zhou, L. (2022). A systematic review of AI ethical challenges in healthcare. *ScienceDirect*. <https://www.sciencedirect.com/science/article/abs/pii/S0967586819300256>
10. World Health Organization. (2021). Artificial intelligence and healthcare: Ethics and policies. *International Journal of Public Health*, 50(i). <https://pmc.ncbi.nlm.nih.gov/articles/PMC8826344/pdf/IJPH-50-i.pdf>