

Ethical AI Use in Healthcare: Policy Proposal

Title: *Guidelines for Responsible, Transparent & Fair AI Deployment in Healthcare Settings*

Author: Kerry Koech

Date: 24th June 2025

1. 📄 Introduction

Artificial Intelligence (AI) holds immense potential in healthcare—from diagnostics to treatment recommendations and patient monitoring. However, without proper safeguards, it can also cause harm or reinforce inequalities. This policy outlines ethical guidelines to ensure AI is used safely, fairly, and transparently in clinical settings.

2. ✅ Patient Consent Protocols

- **Informed Consent:** Before deploying AI-driven decisions, patients must be clearly informed that AI is involved, what data is being used, and how it impacts their care.
- **Opt-In/Opt-Out Rights:** Patients must be allowed to opt in or out of AI-assisted diagnostics or data usage.
- **Data Anonymization:** Personal data must be anonymized before being processed or shared, ensuring compliance with privacy laws like HIPAA or GDPR.

3. 🧑 Bias Mitigation Strategies

- **Diverse Training Data:** AI models must be trained on datasets representative of various genders, ethnicities, age groups, and socioeconomic backgrounds.
- **Regular Fairness Audits:** Implement AI fairness tools (e.g., AI Fairness 360) to monitor bias in predictions and outcomes.
- **Oversight Committees:** Establish an interdisciplinary ethics review board to assess AI tools pre-deployment and periodically thereafter.

4. 🔍 Transparency & Explainability

- **Explainable AI (XAI):** AI models used in healthcare must offer explainable outputs so clinicians and patients can understand the rationale behind recommendations.
- **Model Documentation:** All AI systems should include clear documentation of the algorithm's purpose, limitations, and update history.
- **Real-Time Feedback:** Systems should provide traceable logs and decision confidence scores, especially in high-stakes use like diagnosis.

5. Continuous Improvement

- **Post-Deployment Monitoring:** Continuously monitor AI performance on real-world data. Adapt to new patterns (e.g., disease variants) via periodic retraining.
- **Incident Reporting:** Establish a protocol for clinicians and patients to report adverse effects or suspicious AI behaviours.

Summary

This policy ensures AI in healthcare aligns with ethical principles—respecting autonomy, promoting justice, avoiding harm, and maintaining transparency. These guidelines must evolve alongside AI advancements to protect the dignity, safety, and trust of all patients.