

Policy Proposal: Ethical AI Use in Healthcare

Overview

As AI systems become increasingly integrated into healthcare decision-making, it is essential to ensure their deployment upholds patient rights, clinical safety, and societal trust. This guideline outlines key ethical protocols for patient consent, bias mitigation, and transparency.

1. Patient Consent Protocols

Informed Consent:

Patients must be clearly informed when an AI system is used in diagnosis, treatment planning, or decision support. The explanation should include the role of AI, its limitations, and potential risks.

Opt-in by Default:

AI-assisted services should be opt-in, not opt-out. Patients should explicitly agree to AI usage and have the option to refuse or request human-only evaluation.

Data Usage Consent:

Consent must be obtained for the use of personal health data in AI model training. Data should be anonymized and used strictly within agreed boundaries.

2. Bias Mitigation Strategies

Representative Datasets:

AI models must be trained on datasets that include diverse populations (e.g., age, gender, race, socio-economic status) to prevent skewed outcomes.

Bias Auditing Tools:

Use fairness tools (e.g., IBM AI Fairness 360, Fairlearn) to identify and reduce bias in model predictions, particularly in sensitive conditions like cancer detection or mental health.

Clinical Review of Outputs:

All AI-generated recommendations should be reviewed by a human clinician before implementation — especially when the risk is high (e.g., treatment rejection).

3. Transparency Requirements

Explainable AI (XAI):

AI systems must provide clear reasoning behind their decisions. For example, image-based diagnostics should highlight relevant regions (e.g., heatmaps).

Documentation & Disclosure:

Healthcare providers must maintain records of AI model specifications, data sources, performance metrics, and known limitations.

Error Reporting Mechanism:

There should be a system in place for healthcare workers and patients to report incorrect or harmful predictions, triggering a review process.

Summary

Ethical AI in healthcare requires consent, fairness, and transparency. This guideline ensures that AI systems benefit all patients safely and equitably, while preserving human oversight and trust.