POLICY PROPOSAL: Building Ethical Guardrails for AI in Healthcare

Title: *Healing with Integrity: A Policy Framework for Ethical AI in Healthcare*
Author: AI Module Group 17
Date: 12/07/2025
Theme: *Designing Responsible and Fair AI Systems in Healthcare*

1. Vision

As Artificial Intelligence (AI) becomes increasingly embedded in healthcare—from triage bots to diagnostic tools—it carries the power to **save lives**, **reduce inequities**, and **transform care delivery**. But this potential can only be realized if ethical principles guide every line of code and decision the system makes. This proposal outlines a high-impact policy to ensure AI in healthcare is **trustworthy, equitable, and patient-centered**.

2. Core Ethical Pillars

A. Consent is Care: Enforcing Patient Autonomy

- Patients must have full control over their personal data and how AI systems use it.

- No AI-driven decision should bypass the **informed consent** of the individual involved.

- Interfaces must clearly explain AI's role in care, with options to opt out or seek human clarification.

B. Fairness by Design: Eliminating Bias Before It Harms

- Bias in healthcare AI isn't a glitch—it's a silent threat. Systems must be trained on **inclusive, demographically representative datasets**.

- Institutions must conduct **routine fairness audits** to detect discrepancies in treatment suggestions or outcomes by race, gender, or socioeconomic status.

- Transparency reports should publish fairness metrics such as disparate impact and error rates across subgroups.

C. Radical Transparency & Human Oversight

- AI must never be a "black box" in medicine. Systems must offer **clear, interpretable explanations** of their outputs.

- Clinicians must retain the authority to **overrule AI recommendations** and escalate ambiguous cases.

- Technical documentation—including model training data, decision logic, and update history—must be accessible to oversight bodies.

D. Privacy at the Core

- All patient data must be **anonymized**, encrypted, and processed according to regulations (GDPR, HIPAA, POPIA).

- Only essential data should be collected, and data usage logs must be maintained to protect against misuse or unauthorized access.

3. Policy Implementation Roadmap

- Set up an **AI Ethics Committee** at institutional level, including ethicists, clinicians, data scientists, and patient advocates.

- Mandate **external pre-deployment audits** and ethical certification for all healthcare AI tools.

- Require continuous **feedback loops** from end-users and real-time monitoring of AI performance post-deployment.

4. Closing Call to Action

AI in healthcare should **augment human compassion—not replace it**, **empower clinicians—not pressure them**, and **protect patients—not profile them**. We stand at a pivotal crossroads. Without robust ethical guardrails, the very tools designed to heal may harm. This policy proposal champions a future where AI earns its place in healthcare by aligning with **integrity, inclusion, and human dignity**.