AI Ethics Toolkit

Planetary Health Accord Implementation Framework
Section: Guides - AI Ethics Protocols
Revision: Version 1.0 (2025-05-30)
Estimated Reading Time: 10 minutes

In this document:

- Purpose and Scope

- Core Ethical Principles

- AI Ethics Protocols

- Implementation Guidelines

- Accountability Mechanisms

- Cultural and Contextual Adaptation

- Case Studies

- Call to Action

Purpose and Scope

The **AI Ethics Protocols** provide a comprehensive framework to ensure the ethical development, deployment, and governance of artificial intelligence (AI) within the Planetary Health Accord's health systems. These protocols aim to safeguard communities, respect cultural sovereignty, and promote planetary well-being by embedding justice, transparency, and accountability in AI applications. They address risks such as algorithmic bias, data exploitation, and cultural insensitivity, while fostering trust through community oversight and integration with traditional healing practices. The protocols apply to all stakeholders, including health workers, technologists, policymakers, and Community Health Legions, across global and local health contexts.

Objectives:

- Ensure AI systems prioritize health equity and community empowerment.

- Protect cultural and indigenous knowledge from appropriation or misuse in AI development.

- Mitigate algorithmic bias and ensure transparency in AI decision-making.

- Align AI applications with One Health principles, supporting human, animal, and environmental health.

- Establish community-driven governance and accountability mechanisms for AI systems.

Core Ethical Principles

The protocols are grounded in the following ethical principles, aligned with the Planetary Health Accord:

1. **Justice and Equity**: AI systems must prioritize marginalized communities, ensuring equitable access to benefits and protection from harm.

2. **Community Sovereignty**: Communities have authority over AI systems, including data ownership and decision-making processes (see Governance Structure).

3. **Cultural Respect**: AI respects traditional healing practices, indigenous cosmologies, and local cultural norms, avoiding appropriation or erasure.

4. **Transparency**: AI processes, including algorithms and data sources, are open and understandable to communities and stakeholders.

5. **Accountability**: Developers and users of AI are accountable for outcomes, with mechanisms for redress and system correction.

6. **One Health Alignment**: AI supports interconnected human, animal, and environmental health, contributing to planetary resilience (see Pandemic and Climate-Preparedness).

7. **Do No Harm**: AI systems are designed to prevent harm, with safeguards like kill switches and bias audits to mitigate risks.

AI Ethics Protocols

The following protocols outline specific requirements for ethical AI use in health systems:

1. **Community Consent and Control**

   - Obtain free, prior, and informed consent from communities before deploying AI systems.
   - Establish Community AI Oversight Boards within health assemblies to govern AI use.

- Ensure communities can veto or modify AI applications that conflict with cultural or ethical values.

2. **Data Sovereignty and Privacy**

- Store health data under community-controlled, federated learning networks to protect privacy (see Technology and Data Infrastructure).

- Prohibit commercialization or unauthorized sharing of health data, with quantum-resistant encryption for security.

- Respect indigenous data sovereignty, ensuring traditional knowledge is used only with explicit community agreements.

3. **Bias Mitigation and Fairness**

- Conduct quarterly bias audits with community input, using the AI Bias Audit Framework.

- Train AI models on diverse, representative datasets, including data from marginalized populations.

- Provide visual explanation interfaces and confidence scores to make AI decisions interpretable.

4. **Cultural Competency**

- Adapt AI interfaces to local languages, cultural health concepts, and traditional practices, co-designed with traditional healers.

- Prohibit AI systems that undermine or replace indigenous healing without community consent.

- Include cultural competency modules in AI development, aligned with Cultural Competency Guidelines.

5. **Transparency and Explainability**

- Publish open-source AI algorithms and data sources, excluding sensitive cultural data, via the Global Knowledge Commons.

- Provide plain-language explanations of AI processes to communities, available in multiple formats (e.g., audio, text).

- Disclose all AI failures or errors transparently, with immediate corrective actions.

6. **Harm Prevention Mechanisms**

- Implement kill switch protocols to halt AI systems causing harm, accessible to Community AI Oversight Boards.

- Establish AI Red Teams to proactively identify risks, including bias, security, and cultural insensitivity.
- Ban AI systems that automate decisions without human oversight in critical health contexts (e.g., diagnostics, triage).

7. **Planetary Health Integration**

- Use AI to monitor One Health connections, such as zoonotic disease risks and environmental health trends.
- Align AI applications with climate-adaptive care, supporting ecosystem restoration and resilience.
- Ensure AI systems minimize environmental impact, using energy-efficient algorithms and hardware.

Implementation Guidelines

To operationalize the protocols, stakeholders should follow these steps:

1. **Establish Governance Structures** (Months 1-3)

- Form Community AI Oversight Boards within health assemblies, including traditional healers and marginalized group representatives.
- Train boards using AI Literacy Certifications to ensure ethical oversight capacity.

2. **Conduct Ethical Assessments** (Months 4-6)

- Assess proposed AI systems against the protocols using the Policy Integration Toolkit.
- Engage communities to validate AI alignment with cultural and equity goals.

3. **Develop and Deploy AI Systems** (Months 7-12)

- Co-design AI tools with communities, ensuring open-source development and cultural adaptation.
- Pilot AI systems in Health Sanctuary nations, with continuous community feedback.

4. **Monitor and Report** (Ongoing)

- Use the Monitoring Dashboard Template to track AI performance, bias, and cultural impact.

- Publish annual AI ethics reports, shared via community assemblies and the Global Knowledge Commons.

5. **Iterate and Improve** (Annual)

   - Update AI systems based on community feedback and emerging ethical standards.

   - Conduct recertification of AI developers and users every three years to maintain ethical competency.

Accountability Mechanisms

To ensure compliance with the protocols, the following mechanisms are established:

- **Community Oversight**: Community AI Oversight Boards have authority to audit, modify, or halt AI systems, with support from Regional Health Hubs.

- **Grievance Redress**: Communities can report AI harms via digital democracy tools or offline channels, with resolution timelines of 30 days.

- **Independent Audits**: Third-party AI ethics auditors, including indigenous and marginalized representatives, conduct annual reviews.

- **Penalties for Non-Compliance**: Violations (e.g., data misuse, bias perpetuation) result in system suspension, developer sanctions, or reparations to affected communities.

- **Recognition Programs**: Ethical AI innovators are honored through Planetary Health Steward Awards (see Recognition and Celebration Programs).

Cultural and Contextual Adaptation

The protocols are designed for global applicability, with adaptations for diverse contexts:

- **Multilingual Access**: Protocols available in UN official languages, with translations for local languages (e.g., Swahili, Quechua, Hindi).

- **Offline Compatibility**: Guidelines provided via paper manuals, community radio, and QR code summaries for low-connectivity areas.

- **Cultural Customization**: Protocols co-developed with traditional healers to reflect local health cosmologies and practices.

- **Disability Inclusion**: Formats include audio, braille, and sign language, with accessible interfaces for neurodivergent users.

- **Funding Support**: Implementation subsidized by the Global Health Solidarity Fund, ensuring free access (see Financing Mechanisms).

Case Studies

The protocols are informed by real-world applications in Health Sanctuary nations:

1. **India: Community-Controlled AI Diagnostics (2027-2029)**

   - Implemented AI for tuberculosis screening with community consent and bias audits.

   - Used culturally adapted interfaces, co-designed with Ayurvedic practitioners.

   - Outcome: 35% increase in early detection rates, with zero reported cultural conflicts.

2. **Pacific Islands: Climate-Linked AI Surveillance (2026-2028)**

   - Deployed AI for zoonotic disease monitoring, aligned with One Health principles.

   - Established kill switch protocols managed by indigenous oversight boards.

   - Outcome: 50% reduction in outbreak response time, with community trust maintained.

3. **South Africa: Transparent AI Governance (2027-2030)**

   - Published open-source AI algorithms for maternal health diagnostics.

   - Conducted quarterly bias audits with community health assemblies.

   - Outcome: 20% reduction in maternal mortality disparities in rural areas.

Call to Action

Stakeholders of the Planetary Health Accord are called to uphold the AI Ethics Protocols:

- **Health Workers**: Advocate for ethical AI use, ensuring community consent and cultural respect.
- **Technologists**: Develop open-source, community-controlled AI systems aligned with the protocols.
- **Policymakers**: Embed the protocols in national health policies using the Policy Integration Toolkit.
- **Communities**: Lead AI oversight through health assemblies, ensuring sovereignty and equity.

- **Start Now**: Access the Planetary Health Starter Kit to launch AI ethics initiatives in your region.

Together, we can ensure AI serves as a force for justice, healing, and planetary health. Join the Accord to build ethical health systems for all.

Access the Protocols: Available in PDF and editable markdown at Tools Library.
Feedback: Share input via [globalgovernanceframework@gmail.com].
Related Resources: Explore AI Literacy Certifications and AI Bias Audit Framework for complementary tools.