THIS MUN VI

March 2026

MODEL UNITED NATIONS VI

TSINGHUA INTERNATIONAL SCHOOL

General Assembly

ESTABLISHING A GLOBAL FRAMEWORK FOR THE ETHICAL DEVELOPMENT AND GOVERNANCE OF ARTIFICIAL INTELLIGENCE

UN GENERAL ASSEMBLY BACKGROUND GUIDE

Topic: Establishing a Global Framework for the Ethical Development and Governance of Artificial Intelligence

Difficulty: Intermediate to Advanced

Statement of Problem

Artificial Intelligence (AI) is no longer a futuristic concept; it is a present-day reality transforming every facet of human society. From healthcare diagnostics and autonomous vehicles to algorithmic trading and automated weapons systems, AI's potential to drive progress is immense. It promises solutions to global challenges like climate change, disease, and poverty.

However, this rapid advancement brings profound and unprecedented challenges. Issues of ethics, bias, privacy, security, and human rights are at the forefront of international discourse. The "black box" nature of some AI systems, their potential to perpetuate and amplify societal biases, the risk of mass surveillance, and the existential questions posed by autonomous weapons (LAWS) demand a coordinated global response.

The current international governance landscape is fragmented, with various nations, corporations, and organizations proposing different guidelines and principles. This patchwork approach risks creating a "race to the bottom" in regulatory standards and significant gaps in oversight. The central question before this committee is: How can the international community harness the benefits of AI while mitigating its risks through a cohesive, equitable, and effective global framework?

History of the Problem & Past Actions

The global community has begun to address AI governance through various initiatives:

- UNESCO's Recommendation on the Ethics of AI (2021): The first global standard-setting instrument on AI ethics, adopted by 193 countries. It outlines common values and principles but is non-binding.
- Office of the Secretary-General's Envoy on Technology: Advocates for a global digital compact and promotes cooperation on AI and other tech issues.
- ITU (International Telecommunication Union): Hosts the annual AI for Good Global Summit, focusing on how AI can achieve the SDGs.
- Regional Frameworks:
    - European Union's AI Act: A landmark, comprehensive legal framework that adopts a risk-based approach, banning certain AI applications (e.g., social scoring) and imposing strict regulations on high-risk ones (e.g., CV-scanning tools).
    - OECD AI Principles: Adopted by over 50 countries, these principles promote AI that is innovative, trustworthy, and respects human rights and democratic values.
- National Approaches:
    - United States: A more fragmented, sector-specific approach guided by executive orders (e.g., the Blueprint for an AI Bill of Rights) and emphasizing innovation and competition, particularly against China.
    - China: Has implemented aggressive regulations focused on algorithmic recommendation systems, generative AI, and data security, often emphasizing social stability and state control.
    - Global South: Many developing nations are concerned about being left behind in the AI race ("digital divide") and about AI systems developed in the Global North being imposed upon them without consideration for their local contexts and needs.

Multi-Stakeholder Efforts: Initiatives like the Global Partnership on AI (GPAI) and the Partnership on AI bring together experts from industry, civil society, and academia.

Key Issues and Questions for Debate

Delegates should consider the following subtopics and guiding questions:

1. Ethical Principles and Human Rights:
    a. How can the principles of transparency, fairness, accountability, and privacy be universally defined and implemented in AI systems?
    b. How can we ensure AI does not infringe upon fundamental human rights, including the right to non-discrimination and the right to privacy?
    c. Should there be a universal moratorium or a complete ban on Lethal Autonomous Weapons Systems (LAWS)?

2. Global Governance and Regulatory Models:
    a. Should a new UN treaty or a specialized agency (e.g., a "International AI Agency") be established, or should governance be handled through softer law and existing bodies?
    b. How can a global framework balance the need for innovation with the necessity of precautionary regulation?
    c. How can regulations be enforced across borders, especially against powerful multinational corporations?

3. Equity and the Global Digital Divide:
    a. How can the international community ensure that AI benefits are distributed equitably and do not exacerbate existing inequalities between and within nations?
    b. What mechanisms can be put in place to support capacity building, technology transfer, and inclusive AI development in the Global South?
    c. How can we prevent "algorithmic colonialism," where AI systems designed in wealthy nations are deployed in developing countries without regard for local culture or needs?

4. Security and Liability:
    a. How can malicious use of AI (e.g., for cyberwarfare, disinformation, surveillance) be prevented?
    b. Who is legally and morally responsible when an AI system causes harm: the developer, the manufacturer, the operator, or the AI itself?

c. How can international norms for the military use of AI be established?

d.

Possible Directions

A successful resolution will be comprehensive and may include:

- A reaffirmation of fundamental ethical principles for AI development (e.g., based on UNESCO's recommendation).
- Mechanisms for international cooperation on AI safety research and standards.
- Provisions for capacity building and technical assistance for developing nations.
- A stance on the development and use of Lethal Autonomous Weapons Systems (LAWS).
- A proposal for a new oversight body or a mandate for an existing UN agency.
- Guidelines for public and private sector accountability and liability.
- Measures to promote transparency and explainability in AI systems (XAI).

Possible Stances

The Innovation Bloc (e.g., USA, Japan, South Korea): Will likely emphasize light-touch regulation to foster innovation and economic competitiveness. May oppose strict, binding treaties in favor of flexible guidelines and public-private partnerships.

The Regulatory Bloc (e.g., EU member states): Will advocate for a strong, rights-based regulatory framework modeled on the EU AI Act. Will push for global bans on certain AI applications and strict rules for high-risk AI.

The State-Sovereignty Bloc (e.g., China, Russia): Will strongly oppose any framework that infringes on national sovereignty. Will emphasize the right of states to regulate AI within their own borders according to their own social and political values. May focus on state security applications.

The Global South Bloc (e.g., India, Brazil, Nigeria, Rwanda): Will focus on closing the digital divide, capacity building, and technology transfer. Will be wary of solutions imposed by technologically hegemonic powers and will demand a seat at the table in shaping global norms. Will highlight concerns about bias and relevance in AI models.

Further Research

Delegates are strongly encouraged to conduct further research:

1. UN Documents: Read the full text of the UNESCO Recommendation on the Ethics of AI.
2. Regional Policies: Study the EU AI Act and the US Executive Order on Safe, Secure, and Trustworthy AI.
3. News: Follow AI governance developments from sources like *Reuters*, *Financial Times*, *MIT Technology Review*, and *Wired*.
4. Academic Sources: Look for papers from institutions like the Future of Life Institute, Stanford's Human-Centered AI Institute, and the Centre for the Governance of AI.

Key Terms and Definitions

Artificial Intelligence (AI): A branch of computer science dedicated to creating systems capable of performing tasks that typically require human intelligence. This includes learning, reasoning, problem-solving, perception, and language understanding.

Machine Learning (ML): A subset of AI that enables systems to learn and improve from experience without being explicitly programmed, often using statistical methods.

Algorithmic Bias: Systemic and repeatable errors in a computer system that create unfair outcomes, such as privileging one arbitrary group of users over others. This often reflects biases present in the training data or the designers.

Autonomous Weapons Systems (LAWS): Often called "killer robots," these are weapons systems that, once activated, can select and engage targets without further human intervention.

The Black Box Problem: The lack of transparency in how complex AI models, particularly deep learning networks, arrive at their decisions. This makes it difficult to understand, trust, or audit the AI's reasoning.

General AI (AGI): A hypothetical form of AI that possesses the ability to understand, learn, and apply intelligence across a wide range of cognitive tasks, at a level equal to or beyond human capability. This is distinct from the Narrow AI that exists today, which is designed for specific tasks.

Data Sovereignty: The concept that digital data is subject to the laws and governance structures of the nation within which it is located.

Explainable AI (XAI): Methods and techniques in the application of AI technology such that the results of the solution can be understood by human experts.