UNESCO COMMITTEE

The Ethics of Artificial Intelligence

A Simulation by Project Delegate

Dear Delegates,

Welcome to the UNESCO committee hosted by the Project Delegate! My name is Dina Adhikari, and I'm so excited to be serving as your Chair for this year's session. We'll be focusing on The Ethics of Artificial Intelligence, a topic that's not only timely and complex, but one that will definitely have a major impact on the future of global society. In this committee, it's up to you to help shape that conversation.

To start off with a bit about me: I'm a rising senior at The Derryfield School, and I've been involved in Model UN since my 8th grade and have attended numerous conferences, including three hosted by Harvard. Over time, MUN has helped me grow as a speaker, researcher, and thinker, but even more than that, it's taught me how to understand global challenges from different perspectives. Outside of MUN, I'm fascinated by the growth of technology and international policy, which is why I chose this topic. AI is changing how our world works, from how we apply for jobs and get medical care to how governments make decisions; we are just starting to understand what the long-term effects might be.

In this committee, we'll look at some of the biggest ethical questions surrounding AI: How do we make sure AI systems are fair and unbiased? What role should governments play in regulation? Can international guidelines be created when so many countries are at different levels of technological development? These are not easy questions, and they don't have direct answers, but that's why this topic belongs in the UN.

The background guide will give you an overview of AI's history, ethical challenges, and how different countries and regions are approaching the issue. Your research and ideas, though, are what will drive this committee forward. As you prepare, I encourage you to dig into your country's policies and think critically about what solutions they might support. What values are behind those policies? What compromises would your country be willing to make in international negotiations? The more you understand those details, the more confident you'll feel when debate begins.

If you have any questions about the topic, committee procedures, or anything else, feel free to email me at secretariat@projectdelegate.org. I'm happy to help however I can. I'm really looking forward to meeting all of you and hearing the ideas you bring to this topic. Good luck with your research, and get ready for some great discussions!

Sincerely,

Dina Adhikari

Chair – UNESCO, Project Delegate

Hello Delegates!

My name is Christina Nguyen, your assistant chair for this session, and I am honored to welcome you to the UNESCO Committee on the Ethics of Artificial Intelligence. Here, your debates and decisions can influence the discussion of the use of artificial intelligence worldwide in the near future due to the urgency of the subject.

I am a rising senior at The Derryfield School, and Model UN has been an integral part of my life since my first Harvard Model UN (HMUN) conference in my freshman year of high school. My most recent conference was at HMUN 2025 in late January to early February, and it was my third Model UN conference. Before Model UN, I considered myself to be a “behind-the-scenes” type of individual. However, after my first conference, I learned that the reward for projecting your interests—alliances, negotiations, and much more—is much greater than letting the ideas of others take over your own. Model UN also taught me that one should take in multiple perspectives and not expect a solution without disagreements and negotiations.

I chose this topic because in my previous experiences with Model UN, I have found a greater interest in topics that pertain to the near future, such as AI. With the birth and rise of AI being so recent, including rising issues and regulations that concern it, it is urgent that this topic be discussed worldwide and in a formal setting, which is where the UN comes in. If you would like a more detailed overview of the committee topic, please refer to Dina Adhikari’s cover letter.

I am so excited to meet everyone and hear about your fabulous ideas! If you have any questions, feel free to email our chair, Dina Adhikari, and me at secretariat@projectdelegate.org. Good luck researching and debating, and I hope you have fun at the conference!

Sincerely,

Christina Nguyen

Assistant Chair – UNESCO, Project Delegate

How to Use This Background Guide

This background guide is designed to provide delegates with a foundational understanding of the issue at hand and the role of international governance in addressing it. It serves not only as a source of information but as a strategic tool to help delegates prepare for debate, draft effective resolutions, and represent their assigned countries with nuance and accuracy.

Delegates are encouraged to:

- **Read the guide in full** to understand the historical context, institutional responses, and key challenges outlined.
- **Pay close attention to the Key Terms** section, which includes critical vocabulary that will strengthen the clarity and precision of speeches and resolution language.
- **Analyze the Positions of Major Blocs and Stakeholders** to anticipate likely alliances or conflicts in the committee.
- **Reflect on the Guiding Questions** as a framework for crafting opening speeches, clauses, and negotiation strategies.
- **Use the Policy Recommendations** section as a springboard for solution-building, adapting proposals to align with national priorities.

While this guide offers a starting point, it is not exhaustive. Delegates are expected to conduct independent research on their country's specific stance, current policies, and international commitments to meaningfully contribute to the committee's work. Ultimately, the more research you do, the more prepared you will be; before you can make a change, you have to understand it.

Key Terms

Not all key terms appear in the background guide, but delegates should utilize any key terms that may apply to their country's policies on Artificial Intelligence.

Accountability: Holding individuals, organizations, and/or governments responsible for the outcomes of AI systems, particularly in cases where the AI causes harm.

AI Literacy: The ability to understand, use, and evaluate AI systems in a safe and ethical fashion

Artificial Intelligence (AI): The ability of a machine to perform tasks that would usually require human intelligence, such as learning, problem-solving, and decision-making.

Artificial Neural Network (ANN): A computational model inspired by human neurons that is made of interconnected nodes that process information and learn from data. ANNs are used in deep learning to recognize and solve problems.

Autonomous systems: AI systems that perceive their environment and make their own decisions, functioning without human intervention.

Bias in AI: The presence of systematic prejudice in AI systems due to biased data, often leading to discriminatory outcomes in AI usage.

Cloud AI: AI services and platforms that run on cloud infrastructure, allowing developers to access and deploy AI systems and models without having to build the systems themselves.

Deep Learning: A subset of AI systems that uses ANNs with many layers to analyze numerous elements in data. Deep learning is utilized in complex AI tasks such as image and speech recognition.

Edge AI: The use of AI on devices at the edge of networks, such as sensors, instead of centralized cloud servers, which reduces latency and enhances privacy.

Fairness: A moral principle that emphasizes the equal treatment of all individuals and groups without discrimination when creating and using AI systems.

General AI (Strong AI): AI that understands, learns, and functions across a variety of tasks, similar to a human. This level of AI does not exist yet.

Machine Learning: A subset of AI systems where algorithms learn from data to improve over time without human intervention or explicit programming. Machine Learning focuses on pattern recognition and prediction.

Narrow AI (Weak AI): AI that is designed for specific tasks. Most AI in the world today is narrow AI.

Reinforcement Learning: A type of machine learning where the AI system learns by interacting with its environment and receiving feedback from users.

Singularity: A hypothetical future point when AI systems become so complex that they surpass human intelligence and radically change human society.

Superintelligence: A hypothetical AI system that surpasses all human intelligence, whether that be in creativity, problem solving, or social intelligence.

Transparency: Ensuring that AI systems, especially those used in the government and other critical areas, can be inspected by both users and regulators so that decisions can be understood.

Unsupervised Learning: A type of machine learning where the AI system learns from data without fixed labels or outcomes. Unsupervised learning is used for clustering and finding hidden data patterns.

Introduction to the Ethics of Artificial Intelligence

Origins and Mandate of the Committee

The ethical governance of artificial intelligence (AI) has emerged as one of the most pressing global challenges of the 21st century. In recent years, the widespread adoption of AI systems across diverse sectors, including law enforcement, education, healthcare, finance, and military operations, has raised profound questions about the design and oversight of these technologies. At the heart of this challenge lies the necessity to ensure that AI systems respect fundamental human rights, promote fairness, protect privacy, and contribute to sustainable development.

The Committee on the Ethics of Artificial Intelligence, as envisioned in this Model UN simulation, is inspired by the structure and work of existing multilateral institutions such as UNESCO, the United Nations General Assembly, and specialized agencies like the International Telecommunication Union (ITU). It has been convened to develop normative guidelines, foster multilateral dialogue, and guide the international community toward the creation and enforcement of ethically sound AI practices.

One of the most pivotal moments in the history of AI ethics at the global level came in November 2021, when UNESCO unanimously adopted its *Recommendation on the Ethics of Artificial Intelligence*. This landmark document laid out a comprehensive set of values- human dignity, environmental stewardship, gender equality, cultural diversity, and global solidarity- and translated them into a concrete roadmap of 11 policy action areas. These included data governance, ethical impact assessments, digital education, and mechanisms for transparency and accountability. UNESCO's recommendation was especially notable for its inclusivity: it was the first global AI ethics instrument to be endorsed by all 193 Member States, signaling widespread recognition that ethical guidance must accompany AI development. Importantly, the Recommendation introduced the Readiness Assessment Methodology (RAM), an innovative tool for Member States to assess their capacity to implement ethical AI policies in practice. This practical dimension marked a shift from purely aspirational principles to a results-oriented approach.

The momentum generated by UNESCO's Recommendation carried over into the broader United Nations system. In March 2024, the UN General Assembly adopted its first-ever AI-focused resolution, A/RES/78/265, titled *Seizing the Opportunities of Safe, Secure and Trustworthy Artificial Intelligence Systems for Sustainable Development*. This solution acknowledged AI's potential to drive progress in sectors like health, climate adaptation, and education. It also warned of the dangers of its unregulated spread, especially the risks of digital surveillance, algorithmic discrimination, labor displacement, and geopolitical destabilization. The resolution reasserted the UN's commitment to human rights and equitable development, emphasizing that AI must serve all humanity, not just those in the most technologically advanced nations. It also called for sustained multilateral cooperation and global investment in AI capacity-building to bridge the digital divide.

The creation of the Committee on the Ethics of Artificial Intelligence can be seen as a response to the urgent need for an institutional forum dedicated specifically to ethical AI governance. The committee's mandate includes facilitating international consensus on key ethical principles, advising on national and regional policies, providing implementation support, especially in the Global South, and fostering inclusive, pluralistic debates on the future of AI. It also provides a mechanism for states to share best practices and align their approaches with globally endorsed guidelines. The committee does not seek to impose one-size-fits-all rules but aims to craft adaptable solutions that recognize both universal rights and the diversity of its Member States.

Past Actions and Institutional Capabilities

The committee's work builds on the foundation laid by regional, national, and multilateral efforts to regulate and ethically manage the AI revolution. One of the most comprehensive examples is the European Union's progression from soft law to binding regulation. In 2019, the European Commission's High-Level Expert Group on Artificial Intelligence published its *Ethics Guidelines for Trustworthy AI*, articulating seven key requirements such as human agency and oversight, technical robustness, privacy and data governance, and societal well-being. These principles were incorporated into the legally binding 2024 AI Act, which introduced a tiered risk classification system: banning applications deemed to pose unacceptable risks (e.g., social scoring), strictly regulating high-risk uses (e.g., in employment or criminal justice), and applying minimal restrictions to low-risk systems (e.g., spam filters or AI generated art). This combination of aspirational ethics and enforceable law represents a pioneering model that many states are now considering adapting to their regulatory environments.

In 2019, a landmark study by Anna Jobin, Marcello Ienca, and Effy Vayena reviewed 84 AI ethics guidelines from a broad spectrum of issuers—governments, corporations, research institutions, and civil society organizations. They found a surprising degree of global consensus on five foundational ethical principles: transparency, justice and fairness, non-maleficence, responsibility, and privacy. Yet, they also noted that these concepts were interpreted in significantly different ways, depending on the cultural, political, and institutional context. For example, transparency could refer to algorithmic explainability in one setting, while in another, it might refer to clear data usage disclosures or the traceability of decision-making processes. Similarly, fairness might be interpreted through technical lenses like statistical parity or more philosophical frameworks focused on historical injustice and systemic inequality. This divergence complicates efforts to create uniform ethical standards, especially without guidance on how such principles should be enforced.

Further complicating this landscape is the lack of capacity in many national governments to design and implement ethics infrastructure. Lionel Tidjon and Foutse Khomh's 2023 study provides compelling evidence of how limited this capacity can be, even in countries that have publicly endorsed AI ethics frameworks. They analyzed 14 countries and discovered that only four had practical tools aligned with their stated principles. Across the board, barriers included a lack of ethics training, insufficient technical expertise, the absence of auditing tools, and the challenge of cross-sector coordination. For example, India has promoted inclusivity as a core AI principle but lacks national guidelines on ensuring accessible interfaces or equitable language representation in AI tools. France, while advocating fairness, has not yet standardized bias detection protocols for public-sector machine learning systems. These gaps reveal the limitations of ethics as a rhetorical device without operational capacity.

To address this, scholars and policymakers have proposed the creation of systems like EthicsOps—a governance framework modeled after the DevOps lifecycle in software development. EthicsOps integrates ethical review into every phase of the AI lifecycle, from design to deployment to monitoring and iteration. Such frameworks recommend internal documentation standards, ethics checkpoints, stakeholder consultations, and independent audits. The Canadian government has begun piloting these ideas through its Algorithmic Impact Assessment (AIA), which evaluates the potential consequences of public-sector AI systems before they are deployed. EthicsOps holds promise as a replicable model that can be scaled to different national contexts depending on institutional maturity and resource availability.

The Committee on the Ethics of Artificial Intelligence is uniquely positioned to guide the development of these implementation tools on a global scale. Rather than duplicating regulatory efforts, the committee can synthesize best practices, foster knowledge-sharing, and promote international alignment on auditing methods, impact assessments, and ethics review protocols. It can also provide model frameworks that can be adapted by states with different levels of digital development. Beyond tools and templates, the committee can act as a convener: bringing together engineers, policymakers, philosophers, legal experts, and civil society actors to co-develop robust and context-sensitive implementation strategies. It may also recommend the integration of AI ethics into national education systems, support the creation of independent AI ethics boards, and help develop indicators for measuring ethical compliance over time.

Just as importantly, the committee can anticipate emerging ethical dilemmas not yet fully addressed by current policy frameworks. The rise of generative AI, the use of AI in warfare, biometric surveillance, synthetic media, and algorithmic content moderation all raise questions that challenge the boundaries of existing legal and ethical norms. Addressing these frontiers requires not only normative clarity but also institutional agility, the ability to adapt as new technologies emerge. The committee's value lies in its ability to offer precisely that combination: grounded ethical vision and practical implementation support in a world where AI is evolving faster than regulation can keep pace.

Current Challenges

I. Background on the Ethics of Artificial Intelligence: Causes and Consequences of the Challenge

The ethical challenges surrounding artificial intelligence are deeply rooted in both the nature of the technology and the global context in which it is developing. At the core of these challenges are two overlapping phenomena: the unprecedented pace of AI advancement and the uneven global capacity to govern it responsibly. This section offers a detailed overview of the historical causes and modern consequences of the ethical dilemmas associated with AI.

A. Technological Acceleration and Governance Lag

AI development has accelerated rapidly in the last decade due to major advancements in machine learning, big data analytics, and computational power. Breakthroughs in natural language processing, image recognition, and autonomous decision-making have enabled machines to perform increasingly complex cognitive tasks. However, this progress has outpaced the ability of policymakers to regulate or even fully understand the implications of these technologies. The result is a governance lag, which is an absence of timely, coherent, and enforceable standards to ensure that AI is used ethically.

This lag is exacerbated by the fact that AI systems often function as “black boxes,” meaning their internal logic is difficult to explain or interpret even for the engineers who designed them. As AI begins to make decisions that affect human lives (e.g., in policing, hiring, healthcare), the lack of transparency creates risks of bias, abuse, and error without adequate avenues for accountability or redress.

B. Structural Inequities and Global Digital Divides

A major cause of ethical concern is the deep global inequity in access to AI development and governance tools. The majority of advanced AI research and applications are concentrated in a handful of wealthy nations and multinational corporations. This centralization raises fears of digital colonialism, where the technological standards, data flows, and ethical norms of a few countries dictate the experiences of the global majority.

Meanwhile, many developing countries lack the infrastructure, expertise, and financial resources needed to participate in the AI revolution. Without targeted support, they risk being marginalized or exploited by AI systems designed without their input. For example, AI translation tools often fail to support less widely spoken languages, while facial recognition systems exhibit significantly higher error rates for people of color, which are problems that stem from a lack of diverse representation in training data and design teams.

C. Commercial Incentives Vs. Public Good

Another major cause of the AI ethics crisis is the dominance of profit-driven motives in AI development. Many of the world's most powerful AI systems are designed by private companies whose primary obligation is to shareholders, not to the public. This can lead to ethical shortcuts by using opaque data collection practices, manipulating user behavior, or prioritizing speed over safety in deployment.

The concentration of AI expertise within the private sector also limits democratic oversight. Proprietary systems are often shielded from scrutiny under intellectual property laws, making it difficult for regulators or the public to evaluate their fairness or safety. This commercial opacity creates ethical blind spots that public institutions are ill-equipped to address without new mechanisms for oversight and accountability.

D. Consequences of Ethical Neglect

The consequences of failing to address these challenges are far-reaching. Algorithmic bias can reinforce structural racism and gender inequality, such as when AI systems used in hiring penalize applicants based on biased data patterns. AI in law enforcement has led to wrongful arrests and surveillance overreach. In the healthcare sector, algorithms have misallocated resources or denied treatments based on flawed input data. Furthermore, autonomous weapons systems raise grave concerns about human control over the use of AI.

At a societal level, AI has the power to shape discourse, influence elections, and control access to economic opportunity, If deployed irresponsibly, it could exacerbate inequality, erode trust in institutions, anc concentration power in the hands of a few unaccountable actors these risks are especially acute in environments lacking legal safeguards or independent media scrutiny.

E. Guiding Questions for Debate and Resolutions

The following guiding questions are designed to steer delegates toward the critical issues that must be addressed in this committee. These questions aim to provoke thoughtful debate, clarify conflicting priorities, and lay the groundwork for innovative and actionable resolutions. A strong resolution will address many or most of the following dimensions:

1. How can international bodies define a baseline set of ethical principles for AI that accommodates cultural and political diversity?
2. Should there be a global oversight mechanism for AI systems? If so, what form should it take- a treaty, regulatory agency, voluntary framework?
3. How can transparency be mandated in privately developed AI systems without stifling innovation or violating proprietary rights?
4. What role should the United Nations and affiliated bodies play in auditing or accrediting AI technologies?
5. In what ways can AI design and deployment be made more inclusive of marginalized and underrepresented populations?
6. How can global cooperation prevent an AI arms race between states, particularly in the context of military and surveillance technologies?
7. What obligations do AI developers have to address the unintended harms caused by their algorithms? What mechanisms should be in place to redress?
8. Should data used to train AI models be subject to ethical review in the same way human research subjects are? What standards would apply?

9. What should be the ethical limits of AI in sensitive domains such as healthcare diagnostics, judicial sentencing, or predictive policing?
10. How should the international community ensure that developing countries have a voice in shaping AI ethics frameworks?
11. To what extent can AI be used as a tool to promote sustainable development and reduce global inequality, rather than exacerbate it?
12. How can the committee promote capacity-building and knowledge-sharing among member states with unequal access to AI expertise and infrastructure?
13. What legal or ethical liabilities should apply when AI systems malfunction or produce discriminatory outcomes?
14. Should there be mandatory human oversight or intervention protocols for AI systems?
15. How do we ensure that ethical considerations remain central as AI continues to evolve, especially with the emergence of new domains such as neurotechnology, emotion recognition, and generative AI?

These guiding questions are not meant to be exhaustive but rather to help frame the discussion. Delegates are encouraged to raise additional questions and explore innovative approaches that reflect both the urgency and the opportunity of this moment in global technological governance.

F. Role of the Committee in Bridging the Gap

This committee plays a central role in addressing these causes and mitigating their consequences. It provides a platform to elevate ethical considerations alongside technical and economic concerns. Through cross-cultural dialogue, it helps harmonize diverse perspectives into shared principles. By highlighting implementation challenges, it directs attention to capacity-building and institutional support. And by anticipating future dilemmas—such as those posed by generative AI, neurotechnology, or predictive policing—it promotes proactive, not reactive, governance.

Ultimately, this background section serves to equip all delegates with a common understanding of the stakes, challenges, and structural roots of the AI ethics debate. With this foundation, delegates are better prepared to propose meaningful, implementable, and globally inclusive solutions in the sessions that follow.

Positions of Major Blocs and Stakeholders

This section outlines the general positions and perspectives of key blocs of countries and relevant actors regarding the ethical governance of artificial intelligence. While each nation may have its specific priorities, these groupings help delegates identify broad alignments and understand how different stakeholders view the global AI ethics agenda.

Technologically Advanced States

Highly developed economies with robust AI research ecosystems, including those in North America, Western Europe, and parts of East Asia, tend to prioritize innovation-friendly regulation. These states are concerned with preserving competitive advantages, protecting intellectual property, and enabling cross-border AI collaboration. While they often express support for ethical AI, their policies tend to emphasize self-regulation by the private sector and soft law approaches, such as voluntary ethical guidelines. For example, the OECD Principles on AI, adopted by many of these countries, promote trustworthy AI through non-binding commitments.

These countries may resist strict international regulatory frameworks that could constrain national innovation ecosystems or subject private firms to global audits. However, many also support global cooperation on preventing misuse of AI in areas such as autonomous weapons, algorithmic bias, and surveillance overreach, as seen in G7 declarations on AI safety.

Emerging Economies and Middle-Income Countries

Many countries in Latin America, Southeast Asia, Eastern Europe, and parts of the Middle East face a dual challenge: they seek to reap the developmental benefits of AI while avoiding digital dependency on foreign firms. These nations often call for capacity-building, technology transfer, and inclusion in global AI standard-setting. Brazil and Indonesia, for instance, have both advocated at UNESCO for more equitable access to AI infrastructure.

They may support ethical regulation of AI as a way to level the playing field, particularly around data sovereignty, equitable access to AI tools, and protection of vulnerable populations. These countries are often interested in a stronger role for the UN and multilateral bodies to counterbalance private corporate dominance.

Least Developed Countries (LDCs)

LDCs face unique challenges in AI governance due to limited infrastructure, low digital literacy, and weak institutional capacity. Many lack domestic AI industries altogether and are more likely to import AI technologies from abroad. This creates significant concerns around fairness, data privacy, and algorithmic harm.

LDCs tend to advocate for stronger international guardrails, inclusive development frameworks, and funding for AI literacy and institutional development. In UNESCO negotiations, several African member states have called for a global AI ethics observatory to monitor real-world impacts in underrepresented regions.

Private Sector and Technology Companies

Private AI developers play a dominant role in shaping the global AI landscape. While many companies have adopted internal ethical principles, such as Google's AI Principles or Microsoft's Responsible AI Standard, the degree of transparency and enforcement varies. Most private actors prefer voluntary codes of conduct and oppose rigid international regulation.

Some tech companies are supportive of multi-stakeholder governance models that include industry voices alongside governments and civil society. However, others lobby against external oversight, particularly around data use, algorithmic transparency, and AI audits. The controversy surrounding OpenAI’s GPT models and data disclosure practices illustrates the tension between innovation and accountability.

Civil Society and Academic Institutions

Non-governmental organizations, research institutes, and human rights advocates have played a critical role in drawing attention to the societal risks of AI. They often emphasize the need for enforceable global norms, transparent governance, and protection of vulnerable groups. Initiatives like the Algorithmic Justice League and Access Now advocate for bans on certain AI applications, such as facial recognition in public spaces.

They tend to call for public participation in AI governance processes, ethical review of algorithmic systems, and stronger limitations on state and corporate surveillance. Their advocacy shapes both grassroots movements and formal policy debates.

Regional and Cultural Perspectives

Different regions of the world may emphasize distinct ethical frameworks based on their cultural, historical, and political traditions. For example, European states often highlight individual rights and data protection, as reflected in the General Data Protection Regulation (GDPR), while some Asian states stress collective welfare, innovation, and state-led governance, with countries like China pursuing state-centric AI planning through national strategies.

Delegates should be aware of these regional differences when negotiating shared ethical standards. Successful resolutions may need to strike a balance between universal principles and culturally grounded applications.

By understanding the broad coalitions and stakeholder dynamics at play, delegates will be better equipped to draft inclusive, realistic, and effective approaches to the ethical governance of artificial intelligence.

Pathways Forward: Policy Solutions and Recommendations

The ethical challenges posed by artificial intelligence are neither abstract nor inevitable; they are a consequence of how institutions, governments, and developers choose to design and deploy these systems. Therefore, this committee has the opportunity not only to highlight problems but also to propose practical, scalable, and rights-centered solutions. The following section outlines several strategic pathways that delegates may consider when drafting resolutions.

I. Strengthening Global Norms and Legal Guidelines

The current international landscape lacks any binding legal frameworks governing AI ethics, leaving regulation to voluntary principles and national-level efforts. Delegates may propose the development of a Global Framework Convention on the Ethical Use of AI, similar in structure to the UN Framework Convention on Climate Change, that could eventually evolve into a treaty with enforceable protocols.

Such guidelines would establish shared commitments to transparency, human rights, algorithmic accountability, and equitable benefit-sharing. It should include a peer-review mechanism similar to the Universal Periodic Review used by the Human Rights Council, where states periodically report and receive feedback on AI practices. In parallel, delegates can build upon existing documents like the UNESCO Recommendation on the Ethics of AI (2021), which includes 11 core principles and calls for banning social scoring systems. Proposals could seek to institutionalize its monitoring tools, such as the AI Readiness Assessment methodology, and extend them into formal compliance systems.

II. Promoting Equitable Access and Capacity-Building

Many developing states face significant barriers in participating in global AI governance due to a lack of infrastructure, investment, and technical expertise. Resolutions should support the creation of an International AI Capacity Fund under UNDP, modeled after the Green Climate Fund. Delegates may also advocate for a Digital Solidarity Tax, a small levy on global AI revenues (especially from transnational tech companies), to fund this effort. The African Union has previously raised the need for Global South participation in AI research; these initiatives would respond directly to such appeals.

III. Enhancing Algorithmic Transparency and Redress Mechanisms

Transparency and accountability remain among the most urgent ethical concerns. To protect citizens, resolutions should also recommend:

1. The creation of regional AI ethics authorities or grievance redress units
2. Minimum explainability standards for all AI used in government procurement
3. Access-to-information guarantees for users affected by automated decisions

Case studies like Amazon's scrapped AI hiring tool, which penalized resumes with the word “women’s”, demonstrate the consequences of opaque, unreviewed systems. Tools to address these issues would empower communities to challenge discriminatory outcomes, particularly in criminal justice and healthcare, two sectors where algorithms have already inflicted measurable harm.

IV. Limiting Harmful Applications of AI

Resolutions should seek to address immediate risks by restricting or banning unethical AI uses. Proposed prohibitions might include:

1. Facial recognition in public spaces, as seen in bans already enacted in San Francisco and parts of the EU
2. Predictive policing systems, which often rely on biased historical crime data
3. Fully autonomous weapons, expanding on the work of the CCW and responding to UN Secretary-General António Guterres’ 2018 call for a global ban.

Delegates could also mandate Human Oversight Protocols for high-risk applications, ensuring human responsibility in domains like medical diagnosis and judicial sentencing.

V. Upholding Human Rights and Democratic Values

Delegates must anchor all AI ethics work in internationally recognized human rights standards. That includes:

1. Safeguards against mass surveillance, particularly in authoritarian contexts
2. Anti-discrimination clauses modeled after frameworks like the International Convention on the Elimination of All Forms of Racial Discrimination
3. Whistleblower protections for developers who report unethical AI behavior

A resolution might suggest convening a Special Rapporteur on AI and Human Rights, under the Human Rights Council, to investigate abuses and issue annual reports.

Concluding Reflections: Defining the Ethical Future of AI

This committee does not operate in a vacuum. It is meeting at a moment when countries are already grappling with the consequences of unregulated AI: wrongful arrests in Detroit caused by facial recognition, public school surveillance software with racist outcomes, and discriminatory healthcare algorithms that misallocate resources away from minority patients.

Yet it is also a moment of opportunity. The 2021 UNESCO Recommendation on AI was adopted by 193 countries, showing that a global consensus is possible. But implementation lags behind commitment. Delegates must now propose mechanisms that translate ideals into action.

True AI ethics will not emerge from technical tweaks but from rebalancing power. Delegates are urged to:

1. Craft resolutions that mandate inclusive decision-making
2. Propose systems that prevent harm rather than simply respond to it
3. Embed justice and equity, not just efficiency, into the very foundations of AI governance

Further Research Recommendations

I. International Norms and Multilateral Guidance
   A. UNESCO Recommendation on the Ethics of Artificial Intelligence (2021) https://unesdoc.unesco.org/ark:/48223/pf0000381137 . A foundational document that outlines 11 principles and policy actions. Delegates may study its language and explore national commitments. It includes a section on banning social scoring systems and an AI readiness assessment framework.
   B. OECD AI Policy Observatory https://oecd.ai/en/ . A comprehensive international database tracking national AI strategies, ethical frameworks, and performance metrics. Particularly helpful for comparing country-level implementation and alignment with ethical principles.
   C. UN High Commissioner for Human Rights Report on Artificial Intelligence and Privacy (2021) https://www.ohchr.org/en/documents/thematic-reports/ahrc4831-right-privacy-digital-age-report-united-nations-high . A UN human rights analysis of the challenges AI poses to civil liberties and democratic values, including facial recognition and data surveillance.

II. Governance Gaps and Global Inequality
   A. Access Now: "A Human Rights-Based Approach to Artificial Intelligence" (2018) https://www.accessnow.org/wp-content/uploads/2018/11/AI-and-Human-Rights.pdf . This civil society report discusses how AI can exacerbate existing inequalities and offers recommendations for centering rights in AI policy. It is particularly relevant for understanding Global South advocacy.
   B. Brookings Institution: "Artificial Intelligence and Emerging Technology Initiative" https://www.brookings.edu/project/artificial-intelligence-and-emerging-technology-initiative/ . Features ongoing analysis of global AI developments, with frequent focus on governance gaps between high-income and low-income countries.

III. Case Studies of Ethical Harms

A. ProPublica: "Machine Bias" (2016) https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing A landmark investigative piece on algorithmic bias in U.S. criminal sentencing software. Essential reading for understanding structural discrimination in automated systems.

B. The Markup: "Do Algorithms Decide Your Life?" series https://themarkup.org/2020-in-review/2020/12/15/algorithms-bias-racism-surveillance An accessible and detailed journalistic series investigating how algorithms affect housing, credit, employment, and education. Useful for grounding policy debates in real human impact.

C. Nature: "Dissecting racial bias in an algorithm used to manage the health of populations" (2019) https://www.science.org/doi/10.1126/science.aax2342. A peer-reviewed study revealed racial disparities in a widely used U.S. healthcare algorithm. Demonstrates the systemic effects of skewed data in life-critical sectors.

IV. Emerging Technologies and Ethical Frontiers

A. Electronic Frontier Foundation: "Who's Watching You Now? Facial Recognition Technology and the Law" https://www.eff.org/pages/face-recognition. Explores the rapid deployment of facial recognition across borders and the absence of legal safeguards, especially in protest surveillance.

B. Future of Life Institute: "Autonomous Weapons: An Open Letter from AI and Robotics Researchers" https://futureoflife.org/open-letter/autonomous-weapons-open-letter/ . Over 20,000 experts—including Elon Musk and Stephen Hawking—signed this letter calling for a ban on lethal autonomous weapons. It can serve as a precedent for multilateral agreement language.

V. Interactive Tools and Regional Frameworks

A. AlgorithmWatch: "AI Ethics Guidelines Global Inventory" https://inventory.algorithmwatch.org/. A living database of over 160 national, regional, and corporate AI ethics frameworks. Delegates can use this to research country positions or compare thematic priorities.

B. African Union Development Agency (AUDA-NEPAD): Continental AI Strategy (2022 Draft Overview) https://au.int/en/documents/20240809/continental-artificial-intelligence-strategy. Outlines a pan-African strategy for ethical and sustainable AI. Emphasizes capacity-building, data sovereignty, and inclusive innovation.
C. European Commission: "Ethics Guidelines for Trustworthy AI" (2019) https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai. Presents the EU's official guidelines, including 7 key principles and assessment checklists. A model for binding regional legislation like the forthcoming AI Act.

Works Cited

"A Human Rights-Based Approach to Artificial Intelligence." *Access Now*, 2018, https://www.accessnow.org/a-human-rights-based-approach-to-ai/.

"AI Ethics Guidelines Global Inventory." *AlgorithmWatch*, https://inventory.algorithmwatch.org/.

"Continental Strategy on Artificial Intelligence for Africa (Draft Overview)." *NEPAD*, https://www.nepad.org/publication/continental-strategy-artificial-intelligence-africa.

Andrus, Matt. "Face Recognition Tech: The Consequences for Communities of Color." *American Civil Liberties Union*, 2019, https://www.aclu.org/news/privacy-technology/face-recognition-tech-consequences-communities-color.

Angwin, Julia, et al. "Machine Bias: There's Software Used Across the Country to Predict Future Criminals. And It's Biased Against Blacks." *ProPublica*, 23 May 2016, https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

"Artificial Intelligence and Emerging Technology Initiative." *Brookings*, https://www.brookings.edu/project/artificial-intelligence-and-emerging-technology-initiative/.

Crawford, Kate. *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. Yale University Press, 2021.

"Who's Watching You Now? Facial Recognition Technology and the Law." *EFF*, https://www.eff.org/pages/face-recognition.

"Ethics Guidelines for Trustworthy AI." *European Commission*, 2019, https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai.

"Autonomous Weapons: An Open Letter from AI and Robotics Researchers." *Future of Life Institute*, https://futureoflife.org/open-letter/autonomous-weapons-open-letter/.

Gibney, Elizabeth. "The Battle for Ethical AI at Google." *Nature*, vol. 603, no. 7901, 2022, pp. 565–568. https://doi.org/10.1038/d41586-022-00757-8.

Hao, Karen. "How AI Is Reinforcing the Global Digital Divide." *MIT Technology Review*, 21 Dec. 2020, https://www.technologyreview.com/2020/12/21/1015300/ai-digital-divide-openai-gpt-3-translation/.

Jobin, Anna, et al. "The Global Landscape of AI Ethics Guidelines." *Nature Machine Intelligence*, vol. 1, no. 9, 2019, pp. 389–399. https://doi.org/10.1038/s42256-019-0088-2.

Karger, Howard. "The Algorithmic Welfare State?" *New Labor Forum*, vol. 29, no. 1, 2020, pp. 74–82. https://doi.org/10.1177/1095796019890467.

Latonero, Mark. "Governing Artificial Intelligence: Upholding Human Rights & Dignity." *Data & Society*, 2018, https://datasociety.net/library/governing-artificial-intelligence/.

Obermeyer, Ziad, et al. "Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations." *Nature*, vol. 574, no. 7780, 2019, pp. 447–453. https://doi.org/10.1038/s41586-019-0322-6.

"OECD AI Policy Observatory." *OECD*, https://oecd.ai/.

"Foundation Models and Responsible AI." *Stanford Human-Centered Artificial Intelligence*, https://hai.stanford.edu/research/foundation-models.

"Do Algorithms Decide Your Life?" *The Markup*, https://themarkup.org/series/hello-algorithm.

"Recommendation on the Ethics of Artificial Intelligence." *UNESCO*, 2021, https://unesdoc.unesco.org/ark:/48223/pf0000381137.locale-en.

"The Right to Privacy in the Digital Age: Report of the United Nations High Commissioner for Human Rights." *United Nations OHCHR*, 2021, https://www.ohchr.org/en/documents/thematic-reports/a-76355-right-privacy-digital-age.

Vincent, James. "Amazon Scraps Secret AI Recruiting Tool That Showed Bias Against Women." *The Verge*, 10 Oct. 2018, https://www.theverge.com/2018/10/10/17959854/amazon-ai-recruiting-tool-biased-against-women-report.