

Artificial intelligence is a major technological force that will reshape the 21st century and the future of humanity. But its reverberating effects are not nor will be confined to the technological; the evolution of AI over the course of the next years and decades will greatly influence the global economy, interstate relations, and society in general. Importantly, the increasing intelligence of AI- which is manifest given its growing ability to address ever more complex tasks and the imminent emergence of artificial “general intelligence”- will accelerate the timing and intensity of its impacts.

Given what is at stake and the global nature of this technological actor, it is vital that the international community unite to establish global norms, regulations, and- when necessary- institutions for the safe and responsible development and use of AI. Thus, proactively shaping the development of AI, rather than reacting haphazardly to its effects after the fact, will allow humankind to leverage the bountiful benefits of AI and minimize the threats it poses to a peaceful and stable society. First, what are some of the key benefits of AI?

The benefits of safe and responsible AI development and use are many. In his acclaimed book, [The Sentient Machine](#), AI expert and entrepreneur Amir Husain describes how artificial intelligence can enhance the efficiency and effectiveness of health care, financial investing/trading, and cybersecurity operations. Furthermore, AI systems can even bolster activities related to conflict resolution and strongly enhance economic productivity.

In the realm of healthcare, computation and dynamic modeling, two key AI tools, can help identify the specific causes and treatment options for “bucket illnesses” like inflammation of the intestines, which can be symptomatic of myriad diseases. Furthermore, AI can aid physicians by continuously monitoring peoples’ physical state- thus providing more extensive health information that can enable physicians to more proactively identify and prevent the emergence of illnesses. Gene editing technology like “CRISPR” also [leverages AI](#) to strengthen the body’s response to diseases like cancer.

In the financial sphere, AI systems employ a counterintuitive approach to identifying financial trading and investment opportunities. AI can also enhance the security and transparency of the transfer of financial information through systems like blockchain. Regarding cybersecurity, AI systems’ capacity to employ [machine learning](#) to quickly identify and contain cybersecurity threats through counterintuitive strategies- as demonstrated by its ability to effectively contain the Adylkuzz virus in April of 2017- far surpasses the cybersecurity abilities of traditional antivirus products.

Another potential benefit of AI is in peacebuilding and conflict resolution. AI [systems like cogSolv](#) can drive successful conflict resolution efforts by conducting in-depth analysis of conflict factors like the local context and relevant psychological and cultural worldviews. Thus, AI-powered programs like cogSolv can supply human experts and peacebuilders on the ground with valuable information they may be lacking.

Why should the international community prioritize the global governance of AI? Critically, [“Our Common Agenda”](#) calls for a “New Global Deal” that recognizes the idea of a global commons. Furthermore, the report identifies the “digital commons” as a global public good. As a [recent report](#) released by the Stimson Center points out, public goods demand collective management since they are non-rival and non-excludable. And, given its penetrating international effects, AI will constitute an increasingly vital part of a “global digital commons.” Furthermore, enhancing the global governance of AI will streamline AI regulation by reducing administrative and logistical issues stemming from dozens of distinct, national AI policies.

Since the threats relative to AI are vast and interconnected, it is vital that we fully leverage the [tools](#) global governance provides us. One key, AI-related threat is in the military sphere, where militaries worldwide- especially among the major powers- have been integrating AI into their military strategies, conventional weapons, and even their nuclear command structure. [One example](#) of the military use of AI is in March 2020 when, according to a UN report, a “lethal autonomous weapons system” was deployed in Libya.

According to Michael Klare, a Senior Fellow at the Arms Control Association (ACA), the [greatest threats posed](#) by the military use of AI reside in the development of autonomous weapons systems that do not require human oversight and the increased use of automated battlefield decision-making systems, which are vulnerable to manipulation from adversaries. Automated decision making is particularly risky [in the context](#) of nuclear weapons. Furthermore, the [military use of AI may degrade](#) international peace and stability given the risks of accidents in the AI software, unintentional conflict due to concerns about how the AI systems will be used, and inadvertent escalation of conflict stemming from the inflexibility of AI systems and human overreliance on them. In sum, the increased deployment of autonomous weapons systems and automated battlefield decision-making systems could enhance great power conflict and [greatly undermine strategic stability](#), potentially driving a country to launch nuclear weapons.

How can we leverage global governance to effectively address the dangers of the military use of AI? Michael Klare advocates for an insightful [framework](#) that suggests starting with non-binding, Track-2 diplomacy (among scientists, arms control experts, etc.) and unilateral/bilateral initiatives and then advancing toward “strategic stability” talks and formal, binding treaties. The “strategic stability” talks would resemble the current US-Russia Strategic Stability Dialogue, but would also include powers like China. Furthermore, this framework stresses the importance of confidence-building measures (CBMs) to enhance trust between the relevant parties, a concept also stressed by Horowitz and Kahn. While they recommend that the US assume leadership to [propose CBMs](#) such as a “Dialogue on AI Safety and Strategic Stability” and standard-setting for the military use of AI, all countries should participate equally in this process. While CBMs are critical, it is important to note that many civil society groups, [including the ACA](#), stress the importance of binding, enforceable international agreements that regulate the use of AI in military contexts.

What about international institutions to govern the use of autonomous weapons and autonomous battlefield decision making? There is precedent regarding international cooperation

to address the spread of dangerous technology. [Security Council Resolution 1540](#) prevents states from supporting non-state actors like terrorist organizations that are seeking to acquire and deploy nuclear, biological, or chemical weapons, and requires them to enhance border protection and export controls regarding these weapons. Given the current lack of global institutions that are able to pass binding resolutions- like the more democratic UN General Assembly- it is incumbent on organizations like the Security Council to act quickly. The Security Council could pass a binding resolution sensibly regulating the use of autonomous weapons and battlefield decision making systems among nonstate actors and national governments.

There is also the [idea of establishing an “International Artificial Intelligence Agency”](#) that would act like the International Atomic Energy Agency (IAEA). The World Federalist Movement/Institute for Global Policy (WFM/IGP) has noted similarities between the mission of the IAEA and a potential International AI Agency in that both agencies are/would be tasked with controlling and monitoring a particular technology and pursuing enforcement when necessary. However, the ability to effectively monitor AI technology could prove more difficult than nuclear technology, and adequate enforcement remains a persistent issue.

Another key threat associated with the proliferation of AI is its negative effects on human rights and social equity and inclusion. [AI systems are likely](#) to perpetuate patterns of social bias and exclusion given that the data that they are trained on to enhance their pattern recognition and decision making capabilities mirrors perverse social norms. For example, automated “risk assessments” tools that some courts are using to make bail and sentencing decisions run the risk of reinforcing racial and ethnic prejudice in the criminal justice system, especially since the data that is used to train the algorithm itself is biased against minority groups. Furthermore, automated recruitment and hiring systems can- instead of providing impartial analysis of candidates’ qualifications- perpetuate existing biases, specifically regarding gender. Recently, [Amazon abandoned](#) a project aimed at designing an AI algorithm to streamline the hiring process since it discriminated against women.

Another important factor to consider regarding inclusion and equity is the underrepresentation of particular cultural contexts and languages, which Seth Dobrin, President of the Responsible AI Institute, pointed out in a recent [National Public Radio \(NPR\) interview](#) in the US. Specifically, languages like Arabic are severely underrepresented in AI systems. Finally, as the Berkman Klein Center report stresses, [AI poses substantial risks to privacy](#) given that AI algorithms are trained on massive amounts of private data that contains personal, identifiable information, and subsequent decision making can further violate individuals’ privacy rights.

The myriad AI-related risks to human rights and privacy emphasize the need to create compassionate and ethical AI systems whose algorithms adhere to ethical and life-affirming principles. Accordingly, many organizations have proposed ethical and sensible AI-principles that address human rights and safety concerns. In May of 2019, the [OECD released its AI Principles](#) that include “human-centered values and fairness.” Furthermore, a [brilliant study](#) from the Berkman Klein Center in 2020 analyzed ethical AI principles prepared by 36 organizations- including the OECD- and found common themes, including privacy, non-discrimination, and

promotion of human values. As over 60% of the ethical frameworks analyzed stressed the importance of human rights, with five frameworks centered around human rights promotion, the report suggests framing a AI global governance regime around international human rights law.

An international human rights framework vis-à-vis AI also would have the advantage of being flexible and able to adapt to the continual evolution of AI and other emerging technologies since, as [Paul Nemitz](#) has stressed the importance of, it is not “tech-specific.” To center AI global governance around human rights, we could [embed the principles of documents](#) like the [Universal Declaration of Human Rights](#) and the [International Covenant on Civil and Political Rights](#) into AI’s algorithms, constituting an effective safeguard against discriminatory behavior. The [Berkman Klein Center](#) also recommends requiring AI developers to abide by the [United Nations Guiding Principles on Business and Human Rights](#).

Another widely recognized threat is AI-related job displacement. While AI will [enhance productivity and economic growth](#)- although [mainly for a select group of wealthy countries](#)- experts worry that many jobs will be displaced by AI and automation, driving long-term, structural unemployment. According to NPR, [one study](#) conducted by MIT and Boston University found that, in the United States, robots could replace 2 million workers in the manufacturing sector alone by 2025. Forecasts are different across regions; [an OECD study](#) found that about 9% of jobs in general in the US and other OECD countries are threatened by automation. While AI and automation may open up new job opportunities and enhance real income for consumers by reducing the cost of production through productivity gains, the “[displacement effect](#)” of automation is significant.

To address the adverse economic effects of AI and automation, think tanks and experts worldwide like [The Millennium Project](#) have proposed programs such as Universal Basic Income (UBI) to address long-term, structural unemployment. The idea of [taxing the enormous wealth](#) created by the adoption of automation and AI, in addition to taxing robots, has also been suggested by tech leaders like [Bill Gates](#). [Other suggestions](#) include employment subsidies and guaranteed employment. One pertinent international institution in this respect is the [International Labour Organization](#) (ILO), which works with countries to establish [legally binding conventions and protocols](#) through international treaties. Some conventions, such as collective bargaining and [equality of opportunity/discrimination](#) (regarding humans competing with robots and AI), are particularly relevant in the context of automation-related job displacement. One problem is persuading countries to ratify the treaties that establish these binding conventions.

Finally, it is important to briefly mention the role AI might play in international relations. Importantly, AI does not only have the potential to enhance dangerous [great power competition](#), but it could also greatly [endanger the international system](#) by strengthening authoritarian leaders’ ability to censor and control their citizenry, further destabilizing authoritarian societies. This could have a cascading effect on democratic countries and the entire international order. Promoting the [widespread adoption](#) of ethical AI principles and organizing multilateral forums between states and civil society could be a starting point to address this issue.

Where do we go from here? One idea endorsed by WFM/IGP is to negotiate a [UN Framework Convention on AI](#) (similar to the [UNFCCC](#)) which would drive further international dialogue and negotiations regarding the creation and implementation of ethical AI principles and how to address the socioeconomic effects of AI. Furthermore, an [Intergovernmental Panel for Artificial Intelligence](#) (IPAI), supported by political leaders like Emmanuel Macron, could complement the Convention and enhance inclusivity in AI global governance. Importantly, the Millennium Project is supportive of a [UN Treaty on Artificial General Intelligence](#), which would help set the “initial conditions” for artificial general intelligence.

While creating new institutions like an IPAI may be necessary to strengthen AI global governance, we should also consider existing, flexible institutions and frameworks like the UN, the International Labour Organization, and the Universal Declaration of Human Rights, that could provide a means of sensibly regulating AI according to shared, universal values. Leveraging international human rights frameworks is a logical starting point.

Furthermore, it is vital that everyone in the world plays a role in shaping the future of AI. The UN’s [Global Digital Compact](#) is a good start. However, given the fast-paced evolution of AI and other emerging technologies, it is critical to enhance the opportunities for individual and civil society input in AI global governance; a single dialogue is not enough. One idea to streamline worldwide public participation is to establish an [International Science & Technology Organization](#), again proposed by the Millennium Project. An International S & T Organization would be an “online collective intelligence platform” that would facilitate a continuous dialogue among members of the global community regarding science and technology concepts. Such an online platform could address new AI developments and, in my opinion, could be focused on employment, education, health care, the military, and interstate relations. Another section could be dedicated to the ethical implications of AI. I imagine that this platform could allow policymakers and tech experts worldwide to interact with members of the general public, integrating their input into policymaking and technical design. General themes regarding public input could be extracted at the national, regional, and global level. To establish such a platform equitably and safely, issues like the digital divide and misinformation would have to be fully addressed.

To create the conditions for lasting peace and stability in a world increasingly shaped by emerging technologies like AI, it is crucial to develop and implement these technologies purposefully and in consultation with all of humanity. And developing these technologies “the right way”- in accordance with our values- may necessitate slowing down and prioritizing how we deploy these technologies rather than how quickly we do so. One needs to just look at the [recent US Supreme Court case](#) challenging the immunity that internet and social media companies enjoy. Is it not preferable to establish a strong, just foundation for these technologies from the beginning? Or would we rather suffer the unintended consequences of an untempered global obsession with technological growth?