

**Forum:** Crisis: Special Conference on Artificial Intelligence

**Issue:** Implementing boundaries on weaponization of artificial intelligence to decrease its potential risk to civilians

**Chair:** Claire Lee (Head Chair), Shania Hsu (Deputy Chair), Ayden Chin (Crisis Director), Terry Park (Crisis Director)

---

## Introduction

Artificial Intelligence (A.I.) can draw its roots back to the 1950s, with the publication of accredited mathematician Alan Turing's "Computing Machinery and Intelligence". Though not especially revolutionary, with the turn of the 21st century, artificial intelligence and technology as a whole have rapidly developed beyond the scope of Turing's original papers. AI has found itself in a vastly disruptive position with its highly transformative power in local relations and security. Like the steam engine, electricity and other general purpose technologies, AI stands crucially as a general enabler of other technologies, with a wide array of different uses in different industries. From manufacturing, to finance, to social media, AI currently has incredible potential to aid various sectors, but as its use shifts into real life policy, and as its limits are pushed more and more, practical issues arise.

As these technologies grow more essential to daily life, important dilemmas arise with their implementation and use. Problems with data control, rights to privacy, ethics of AI implementation, what can or cannot be replaced by AI, and accountability of AI fueled weapons are a few of questions that underline this issue. The recognition of the negative potential of AI use is imperative to the understanding of how to establish boundaries on it.

One key issue looming over the massive industry is the growing weaponization of AI by both state and non-state actors. Militarization of AI also has massive implications — it poses a significant threat to the international community, disrupting the entire world of warfare as a whole, and dramatically shifting global power dynamics. Its viability in all kinds of warfare highlights its strength in accomplishing tasks humans cannot. AI has the ability to quickly adapt

and perform actions and learn in ways both similarly and drastically different from humans. Additional issues arise with the threat of AI arms racing, and the potential of these powerful technologies falling into the hands of non-state actors. The limits of AI and machine learning are still unknown, making it ever so important to find common ground to limit its negative effects and to protect civilian lives.

## Definition of Key Terms

### **Artificial intelligence (AI)**

Development of computer systems that can perform tasks that humans typically do.

### **Weaponization**

Something adapted for use as a weapon

### **Autonomous Weapons Systems (AWS)**

AI-powered systems designed to operate without direct human intervention meaning they are allowed to make decisions and carry out actions on their own. An example of a AWS is drones and other military-related vehicles that can be run without the help from humans.

### **Ethical**

In the context of AI and the implementation of boundaries to AI to decrease its risk to civilians. For example, ethical considerations include considering human rights, transparency and accountability, and respecting IHL.

### **International Humanitarian Law (IHL)**

Law of armed conflict that seeks to ensure and to limit the loss of civilians and regulate the methods of warfare. IHL can be used in ensuring the safe and ethical deployment of AI weapons.

### **Human In The Loop (HITL)**

Design approach where human operators are involved in the decision-making process of AI systems, this is designed so that humans will still have control over AI.

### **Explainable AI (XAI)**

A system that can provide clear explanations for the decisions of AI and actions to help transparency and accountability for AI employment

### **Dual use technology**

Dual use technology is tech or equipment that have purposes in both civilian and military applications, in the case of AI, this could include AI algorithms.

## **Arms control**

International agreements aimed to limit the use of weapons, including AI technologies

## **Strategic stability**

An equal balance of power and avoidance of actions that lead to unintended escalations or conflicts

# **Background Information**

## **Historical context**

The past development of nuclear weapons during World War II shows the importance of establishing ethical and moral boundaries that are internationally agreed upon before any further development of advanced weapons. This is needed to prevent further harm to the civilian population as well as to uphold international humanitarian law

### **Emergence of AI in warfare**

AI recently has increasingly helping with military operations and offering help in autonomous weapons systems, surveillance and decision making algorithms. Due to the increase of AI in the military, the question of the ethical use of autonomous lethal forces has been brought up.

## **Tech development**

Advancements in AI— including machine learning and natural language processing — have the possibility of allowing the development of autonomous weapons to be capable of independent decision making when encountering combat scenarios

### **Dual use**

Dual use is when something can be used for both civilian uses and military purposes. For example, in regards to AI, dual use technology could be algorithms that are used for both purposes. This technology could be detrimental as the proliferation of AI and the weaponization of AI also exacerbates the risk of harm to civilians.

## **Policies**

### **International legal frameworks**

Including the Geneva Convention of 1949, protocols have been put in place to prohibit the targeting of civilians, mandating their protection during armed conflicts. However the use of AI weapons poses a question of accepting accountability this could happen if AWS are in use, meaning the AI makes decisions by itself with human intervention.

### **Ethical guidelines and principles**

This can take place in ways such as transparency, accountability and human control, all important to help guide the development of AI enabled weapons. Ethical guidelines also are essential in providing rules to ensure the responsible and moral use of AI.

### **Multilateral diplomacy and cooperation**

The importance of collaboration between countries, especially those considered major global powers, need to show diplomacy in negotiations surrounding the use of AI in warfare. These conventions are crucial for developing the consensus of regulations in order to ensure safety.

### **Challenges With AI**

With the rise of AI powered weapons, questions about its potential risk to civilians is also brought up, for example, with the use of AI weapons, specifically AWS or autonomous decision making weapons, which means these weapons can be employed with human intervention. This raises concerns as decisions made by AI will be hard to link with human operators. The further development of AI weapons also brings up questions of bias. Although AI algorithms claim to be fair, they may inadvertently perpetuate bias, if AI is in charge of big decisions, the biased data can further social inequalities. Finally AI powered weapons will also be susceptible to cybersecurity attacks including hacking. If these AI weapons systems are manipulated, it will pose even more risk to civilians.

## **Major Countries and Organizations Involved**

### **The United States of America (USA)**

The USA which has been deeply entwined with the rise of AI. Their use of AI in the expansion of their military raises many ethical concerns; one of which is the killings of civilians via unmanned drone strikes and what party must take culpability for those deaths. Other issues arise with their development of various weapons and technologies and the rapid integration into defense systems like autonomous vehicles, surveillance, and cyber warfare, and their potential harm to civilians. The USA has worked to draft treaties like "Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy", signed by 50 other states, which aims to

ensure the responsible use of AI technologies and development. In spite of this, the US still views AI as a major strategic tool and holds plans to continue expansion of generative learning programs for various military and governmental needs. The USA is notably against centralized AI control as pushed by the EU but is instead pushing for state controlled regulation nationally.

### **The People's Republic of China (PRC)**

China has rapidly expanded the use of AI both domestically and internationally. Through the development of surveillance technologies, China has rapidly integrated AI into both daily life and its military. Full capabilities of their “unmanned” systems and drones remain unknown but are likely to have significant implications domestically and geopolitically; dramatically affecting its strategic standpoint globally. Chinese AI researchers and military professionals have begun talks on the issue of safety and boundaries on AI. China, which is ahead of the vast majority of countries with both its public and private sector China as a whole still continues plans on further investment and researching on the capabilities of these technologies despite their potential civilian risk.

### **The United Kingdom (UK)**

As of 2023, the UK currently follows the US and China as the third leading producer of AI technologies. With heavy governmental funding of AI and enterprises, the UK has grown as a massive global AI powerhouse in addition to its investment in domestic AI infrastructure. Its goals with AI warfare remain ambitious with its continual development of Autonomous Weapon Systems (AWS), as they must continue to live up their quoted goal as stated by the UK parliament of being "Ambitious, safe and responsible" while developing their military might. The UK, instead of supporting larger, more centralized regulations on AI, has instead pushed a series of proposed principles: Safety, security and robustness, appropriate transparency and explainability, Fairness, Accountability and governance, and finally contestability and redress. Still the UK has also not pushed any large reform movements towards regulation of AI.

### **The State of Israel**

Israel has been noted to be pioneering the use of AI in civilian and military settings. The Hamas attacks call into question the efficacy of their AI surveillance, but Israel looks to respond with a various array of AI tools and technologies such as autonomous drones. Their innovations on cyberwarfare autonomous driving and weapon systems vastly strengthen the nation's military capability. Along with their use of police predictive AI, Israel's advancements pose many issues

both within its borders and in a larger regional setting. Israel is among one of the nations against further UN regulation of autonomous weapons.

### **The Russian Federation**

Russia looks to meet the west in their development of AI, and aims to continue expanding its AI military capabilities but notably lags behind that of China and the US. They have rapidly expanded their use of AI in domestic surveillance and national security. Their goal to reach a similar technological standing as the US and China has been made difficult with sanctions on Russia significantly curtailing Russia's ability to continue research and push boundaries on these technologies. Despite Russia's vying, they have also noted the importance of boundaries when it comes to the implementation of these AIs, although the extent of which they utilize AI is still largely unknown. Russia also holds back on pushing wider regulations on the development of AI.

### **The Republic of Korea (ROK)**

Korea's AI development capabilities can largely be attributed to a key collaborative effort between the Korean private and public sectors. Besides its focus on data and AI infrastructure aimed at national security and protection, Korea's implementation of AI within its national defense has been categorized into four distinct efforts: AI based command systems, unmanned underwater vehicles on a large scale, AI in aviation training systems, and AI object tracking systems. Their race for AI dominance has raised the serious threat of a consequential "leapfrogging" and dramatic weapons race with AI technologies within the Korean peninsula.

### **The European Union (EU)**

The EU hopes to set a precedent for standardization of boundaries on AI, working to push legislation to limit and control the quickly growing industry and market. They hope to have a larger extraterritorial effect, attempting to reduce things like AI fraud and discrimination. Through the passing of its AI Act (AIA), the EU looks to have a larger global effect by leading the way for more boundaries and limitations to be placed on AI within the international community beyond its member states. One issue arising with their push for legislation is the conflict with various member states and private enterprises' current legislation, views, and allocated research and budget.

### **Stop Killer Robots**

A campaign group started by a coalition of international NGOs, formed in April 2013, Stop Killer Robots works to push international law forwards on the grounds of artificial intelligence(A.I.), and more specifically its use in autonomous weapon systems. They propose the only solution to stop civilian harm from lethal autonomous weapons is through multilateral action and a preemptive ban. With its creation, they hope to create a form of legally binding legislation to ensure the thorough stop to all lethal autonomous weapons.

## Timeline of Events

<b>Date</b>	<b>Description of event</b>
1979	Robert Williams, an employee at a Ford Motor Company, was killed by an automated robot. This was the first recorded death by a human being caused by a robot malfunctioning.
1991	Dynamic Analysis and Replanning Tool (DART) for military supplies transportation was implemented in the U.S. military. This was the first time that artificial intelligence was utilized for military purposes.
July 28th, 2015	The Future of Life Institute submitted an open letter, endorsed by over 1,000 robotics and artificial intelligence (AI) researchers. The content of this letter was to advocate for the United Nations to consider implementing a ban on the development of weaponized AI systems.
July 20th, 2017	China issued the “New Generation Artificial Intelligence Development Plan” (AIDP). The basis of AI implementation on matters in China was established.
March 21st, 2021	In the UN Convention on Certain Conventional Weapons (CCW), the USA, the United Kingdom, Israel and Russia blocked a proposed ban on the Autonomous Weapon System (AWS). Some nations such as Austria, Brazil, and China, have suggested that the development of AWS risks setting in motion a potentially catastrophic military arms race that could lead to further instability around the world.

April 21st, 2021	The first international law regarding AI usage was officially proposed. This act named "The Artificial Intelligence Act"(AI Act) was proposed European Union regulation in the European Union. However, the proposal is not enacted yet.
November, 2021	The first ever global standard on AI ethics was established and adopted by 193 member states. The recommended standard, "Recommendation on the Ethics of Artificial Intelligence" was officially established by UNESCO.

## Relevant UN Resolutions and Treaties

- First Committee of the UN General Assembly, 1 November 2023 (A/78/409 DR XXXIII)
- Seventy-eighth session First Committee, 12 October 2023 (A/C.1/78/L.56)

## Possible Solutions

Historically, there have been attempts to address this issue, including the Convention on Certain Conventional Weapons adopted by the United Nations in 1980. 89 countries, including the US, the UK, and China, all claimed that they would adhere to the stipulations proposed by this treaty. It aimed to restrict specific weapons causing undue suffering or harm to civilians. The convention debated the question of banning autonomous weapons at its once-every-five-years review meeting. While the convention did indeed provide a convenient platform for discussing emerging technologies like AI weaponry through protocols, it ultimately failed to present a viable approach. This happened for several reasons, with a primary factor being the inability of the GGE (Group of Governmental Experts) to reach a consensus on categorizing weapons and these systems naturally lacked transparency. Overall, in order to enhance the treaty's effectiveness, other measures, such as engineering safeguards for autonomous weapons, are needed.

In response to the ongoing concerns, a contemporary solution, the Global Partnership on Responsible Artificial Intelligence (GPAI), consisting of 29 countries such as Italy and Japan, aims to enforce these limits. The GPAI is an international multi-stakeholder initiative established in 2020 to guide the responsible development and use of artificial intelligence. Although the



partnership itself doesn't specifically focus on regulating weapons, its emphasis on responsible development and promotion of ethical principles indirectly contribute to preventing harmful AI practices. This is attested by past achievements such as crafting five fundamental principles for responsible AI and launching pilot projects. These initiatives effectively shifted attention away from weaponizing AI toward its positive applications. There are limitations, though. GPAI doesn't hold any legal authority to ensure adherence to its principles in respective countries. Therefore, its effectiveness relies merely on voluntary commitment of member states, potentially leading to further repercussions.

**A possible solution that would effectively complement GPAI while simultaneously addressing AI weaponization is the creation of an independent authorized AI body.** This centralized organization could set legally binding regulations, assess and mitigate risks, foster international cooperation, and ensure transparency. Nevertheless, monitoring and enforcing regulations across diverse governments and technologies can be costly and time-consuming. Therefore, a joint-cooperative effort among nations and industries becomes crucial in maintaining this robust regulatory framework.

## Questions for Further Research

- What role should the international community play in setting boundaries on the weaponization of AI and how the UN or other large IGOs play in restricting or passing legislation on this matter?
- How can countries work to prevent technological stockpiling and weapons racing when it comes to autonomous systems?
- What hand should the international community play in affecting the private sectors in regards to AI?
- How can restrictions help prevent monopolization of AI technologies ?
- How can countries work to de-escalate tensions regarding the use of AI technologies in warfare?
- How can countries shift from more weaponized AI to promote forms of positive peace?
- How can systems be established to ensure the protection of civilians from potentially dangerous autonomous weapons?

## Bibliography

“The Global Partnership on AI (GPAI) - OECD.AI.” *OECD AI Policy Observatory*,

<https://oecd.ai/en/gpai>.

Noor, Ousman. “Convention on Conventional Weapons runs out of road as states adopt meaningless report.” *Stop Killer Robots*, 22 May 2023,

<https://www.stopkillerrobots.org/news/states-adopt-meaningless-report-after-civil-society-excluded-from-un-discussions-on-autonomous-weapons-systems/>.

“Responsible AI Working Group Report.” *Global Partnership on Artificial Intelligence - GPAI*,

<https://www.gpai.ai/projects/responsible-ai/gpai-responsible-ai-wg-report-2022.pdf>.

“Untitled.” *Global Partnership on Artificial Intelligence*,

<https://gpai.ai/projects/responsible-ai/areas-for-future-action-in-responsible-ai.pdf>.

Copeland, B.J.. “artificial intelligence”. *Encyclopedia Britannica*, 26 Jan. 2024,

<https://www.britannica.com/technology/artificial-intelligence>. Accessed 22 January 2024.

Dresp-Langley, Birgitta. “The weaponization of artificial intelligence: What the public needs to be aware of.” *Frontiers in artificial intelligence* vol. 6 1154184. 8 Mar. 2023,

doi:10.3389/frai.2023.1154184

Heyns, Christof. “Human Rights and the Use of Autonomous Weapons Systems (AWS) During Domestic Law Enforcement.” *Human Rights Quarterly*, vol. 38, no. 2, 2016, pp. 350–78. *JSTOR*,

<http://www.jstor.org/stable/24738054>. Accessed 25 Jan. 2024.

485. “United Kingdom Artificial Intelligence Market 2023.” *International Trade Administration | Trade.Gov*,

[www.trade.gov/market-intelligence/united-kingdom-artificial-intelligence-market-2023](http://www.trade.gov/market-intelligence/united-kingdom-artificial-intelligence-market-2023). Accessed 29 Jan. 2024.

7. Artificial Intelligence and Military Advances in Russia - *JSTOR*, [www.jstor.org/stable/resrep24532.13](http://www.jstor.org/stable/resrep24532.13). Accessed 29 Jan. 2024.

Bajak, Frank. "Pentagon's AI Initiatives Accelerate Hard Decisions on Lethal Autonomous Weapons." AP News, AP News, 26 Nov. 2023, [apnews.com/article/us-military-ai-projects-0773b4937801e7a0573f44b57a9a5942](https://apnews.com/article/us-military-ai-projects-0773b4937801e7a0573f44b57a9a5942).

Chapter 22 the Development of Artificial Intelligence in Russia from Artificial Intelligence, China, Russia, and the Global Order on JSTOR, [www.jstor.com/stable/resrep19585.28](https://www.jstor.com/stable/resrep19585.28). Accessed 29 Jan. 2024.

Engler, Alex, et al. "The EU AI Act Will Have Global Impact, but a Limited Brussels Effect." Brookings, 4 June 2023, [www.brookings.edu/articles/the-eu-ai-act-will-have-global-impact-but-a-limited-brussels-effect/](https://www.brookings.edu/articles/the-eu-ai-act-will-have-global-impact-but-a-limited-brussels-effect/).

Government Warned to Proceed with Caution on AI in Autonomous Weapons ..., [committees.parliament.uk/committee/646/ai-in-weapon-systems-committee/news/198762/government-warned-to-proceed-with-caution-on-ai-in-autonomous-weapons/](https://committees.parliament.uk/committee/646/ai-in-weapon-systems-committee/news/198762/government-warned-to-proceed-with-caution-on-ai-in-autonomous-weapons/). Accessed 29 Jan. 2024.

"Homepage." Autonomous Weapons Systems, 25 Oct. 2023, [autonomousweapons.org/](https://autonomousweapons.org/).

Israel's AI Revolution: From Innovation to Occupation, [carnegieendowment.org/sada/90892](https://carnegieendowment.org/sada/90892). Accessed 29 Jan. 2024.

MacCarthy, Mark. "'Ai Weapons' in China's Military Innovation." Brookings, 7 Sept. 2023, [www.brookings.edu/articles/ai-weapons-in-chinas-military-innovation/](https://www.brookings.edu/articles/ai-weapons-in-chinas-military-innovation/).

"Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy - United States Department of State." U.S. Department of State, U.S. Department of State, 17 Jan. 2024, [www.state.gov/political-declaration-on-responsible-military-use-of-artificial-intelligence-and-autonomy/](https://www.state.gov/political-declaration-on-responsible-military-use-of-artificial-intelligence-and-autonomy/).

Putin Says West Cannot Have AI Monopoly so Russia Must up Its Game, [www.reuters.com/technology/putin-approve-new-ai-strategy-calls-boost-supercomputers-2023-11-24/](https://www.reuters.com/technology/putin-approve-new-ai-strategy-calls-boost-supercomputers-2023-11-24/). Accessed 29 Jan. 2024.

"UK, US and Russia among Those Opposing Killer Robot Ban." The Guardian, Guardian News and Media, 29 Mar. 2019, [www.theguardian.com/science/2019/mar/29/uk-us-russia-opposing-killer-robot-ban-un-ai](https://www.theguardian.com/science/2019/mar/29/uk-us-russia-opposing-killer-robot-ban-un-ai).

“United Nations, Main Body, Main Organs, General Assembly, Resolutions, Official Documents.”  
United Nations, United Nations, [www.un.org/en/ga/78/resolutions.shtml](http://www.un.org/en/ga/78/resolutions.shtml). Accessed 29 Jan.  
2024.