



AI and Robotics in Militarized Conflicts – Paper by Joachim von Braun, PAS President

Emerging challenges, rights, ethics, regulation

#TheSpiritofGeneva:
The Impact of AI on International Humanitarian Law
75 Years of the Geneva Conventions
7 November 2024
Aula Magna, Pontifical Gregorian University

Eminence Cardinal Parolin, Excellencies, Ladies and Gentlemen, esteemed colleagues,

It is an honor to address you today on a truly significant occasion: the 75th anniversary of the signing of the Geneva Conventions, a cornerstone of International Humanitarian Law (IHL). These conventions were established in the aftermath of immense global conflicts and suffering, setting rules of war to protect civilians, prisoners of war, and the wounded. Today we face a new challenge that the original drafters of the Geneva Conventions could not have imagined: the rise of Artificial Intelligence (AI) and its profound implications on modern and future warfare. As we commemorate this anniversary, we must consider how to adapt to ensure that humanitarian principles prevail in the digital age with generative AI. It is crucial to ask: Is IHL still fit for purpose, or does the rise of AI necessitate additional prohibitions to protect civilians in war?

1. The Pontifical Academy of Science (PAS) on War, Peace and AI

The Pontifical Academy of Science has long engaged with issues of peace, security, and AI, exploring the connections between them. The Holy See is a signatory to the Conventions, and consistent with Pope Francis' encyclicals, particularly *Fratelli Tutti*, the Pontifical Academy of Sciences urges nations to create structures that promote peace, justice, and reconciliation on a global scale. The Academy criticizes the arms race and prioritization of military expenditure over social and humanitarian needs. PAS has consistently warned against modern weapons of mass

destruction. We did so, for instance, in 1983.[1] More recently, PAS articulated its deep concerns in its 2022 Statement, “Preventing Nuclear War and War Against Civilian Populations: A Task for the Sciences”,[2] following the Russian attack on Ukraine. In 2021, PAS highlighted the risk of robotics and artificial intelligence in warfare, proposing, “States should agree on concrete steps to reduce the risk of AI-facilitated and possibly escalated wars and aim for mechanisms that heighten the barriers of development or use of autonomous weapons, no systems should be deployed that function in an unsupervised mode. Human accountability must be maintained so that adherence to internationally recognized laws of war can be assured and violations sanctioned.”[3]

Clearly, AI raises fundamental questions about the ethics of warfare, international humanitarian law, and the preservation of human rights. I will describe the emerging technologies, address key challenges, and discuss potential regulatory needs surrounding AI and robotics in militarized conflicts.

2. On International Humanitarian Law

AI technologies must be assessed against the backdrop of the existing framework of the Geneva Conventions and their Additional Protocols, which are universally recognized. This IHL framework regulates the conduct of war by codifying clear prohibitions to protect civilians, civilian objects and civilian infrastructure.[4] While IHL does not address the prevention of conflicts, its framework indirectly contributes to the prevention of escalation and violations during armed conflicts. Violations of the prohibitions enshrined in the Geneva Conventions can amount to war crimes punishable by international criminal tribunals as well as national jurisdictions. Article 8 of the Rome Statute codifies war crimes for the International Criminal Court (ICC) and this provision is implemented in all member states to the Statute, further strengthening the legal framework of IHL. By imposing restrictions on how wars are fought and emphasizing accountability for violations, IHL plays a key role in preventing escalation, protecting human dignity, and fostering a culture of restraint.

The prohibitions of the Geneva Conventions are applicable to the challenges posed by emerging technologies such as autonomous weapons systems (AWS), artificial intelligence (AI), and cyber warfare. I only refer to the latter in passing here. We must ask ourselves *whether we are facing a compliance issue or whether the legal framework needs to be extended as well*.

3. Key AI and Robot Technologies Under Development or Used in Armed Conflicts

Before delving getting into that pivotal discussion, let me briefly outline the types of technologies already in use and those under development in this rapidly evolving field.

1) Autonomous Weapons (Lethal Autonomous Weapon Systems - LAWS)

- Drones, such as Unmanned Aerial Vehicles (UAVs), are widely used for surveillance and precision airstrikes. Swarming drones, controlled by AI, operate in coordinated “swarms,” overwhelming adversaries by sheer numbers. Autonomous Ground Vehicles and AI-Controlled Naval Systems are used for tasks ranging from reconnaissance to direct combat.

2) Intelligence, Surveillance, and Reconnaissance (ISR)

- AI-enhanced Surveillance systems operate on data from satellites, UAVs, and ground sensors, analyzing vast amounts of visual, auditory, and signal data. Machine learning models provide predictive analytics, helping to predict enemy movements or logistical needs by analyzing historical data and behaviour patterns.

3) AI for Strategic Planning

- Intelligent Decision-Making Systems assist in strategic planning by analyzing geopolitical situations, assessing risks, and offering data-driven recommendations for military and political leaders. AI for Tactical Decision Aids supports real-time battlefield decisions by providing a clearer picture of the battlefield, integrating data from various sources and suggesting optimal courses of action.

4) Autonomous Decision Support

- Military Simulations are used to simulate war scenarios, helping military planners evaluate strategies, troop deployment, logistics, and potential conflict outcomes. AI-Driven War Gaming helps human commanders assess complex battlefield scenarios through virtual exercises.

5) Logistics and Supply Chain Management

- AI for Logistics Optimization predicts supply chain needs, ensuring military forces are supplied in combat zones. Autonomous Convoys navigate through dangerous areas to deliver supplies.

6) Human Augmentation

- Robotic Exoskeletons enhance soldier performance on the battlefield. AI-Enhanced Prosthetics and Medical Systems are used for battlefield surgery by robots or AI-assisted diagnosis of injuries.

7) Cyber Warfare and Defense

- AI-based Cyber Defense detects, prevents, and responds to cyber-attacks by analyzing network traffic for unusual patterns and identifying potential threats before they cause damage. Offensive Cyber Operations help design and execute cyber-attacks by automating tasks like identifying vulnerabilities and exploiting network weaknesses.

8) AI in Space Warfare

- AI for Space Monitoring analyzes satellite orbits, tracks space debris, and detects potential

threats to national security in space. Autonomous Satellites can adjust their orbits and perform tasks without human intervention, including counter-satellite actions.

9) AI Systems Related to Biosafety and Biosecurity.

- These include algorithms for nucleic acid synthesis screening, construction of high-assurance software foundations for novel biotechnologies, screening of complete orders or data streams from cloud labs and biofoundries, and development of risk mitigation strategies such as medical countermeasures.[5]

In summary, these AI and robotics tools are transforming warfare for combat, decision-making, surveillance, logistics, and cyber operations, leading to a significant shift in military strategy and the future of conflicts, possibly lowering the barriers of starting or escalating wars.

4. Key Challenges

The abovementioned new and emerging technologies pose unique challenges in terms of ethics and the law. I will focus on two areas of mayor concern: **Lethal Autonomous Weapons Systems (LAWS)** and the **use of AI in military decision-making**.

4.1 LAWS

Perhaps the most alarming development is the rise of LAWS. These systems can select and engage targets without human intervention. If machines are allowed to make life-and-death decisions, the fundamental principle of human accountability in warfare is in question. We need to ask:

- Who is responsible when a machine violates the conventions?
- How can we ensure that these systems are programmed to select military targets?
- How can a machine respect the principle of proportionality when it does not consider humanitarian issues?

It is worth recalling that under IHL, humans must determine the lawfulness of military actions, and humans remain accountable, whether technology is involved or not. AWS pose a significant risk for IHL violations, including the inability to distinguish properly between military and civilian targets in a fluid combat zone. Further, removing human agency from life-or-death situation means removing human dignity and restraint from decision making. Against this backdrop I propose three measures:

1) LAWS should never be allowed to function in unsupervised modes. Human accountability must

be maintained to ensure that violations of international law can be sanctioned. AI-based systems should complement and reinforce sound moral judgment, not replace it.

2) AI driven systems deployed in military contexts must be programmed to comply with IHL principles, such as distinction, proportionality, and necessity. If it is not possible to reliably embed ethical guidelines within the algorithms, ensuring they do not violate international laws, the technology should be prohibited from use without complete human control.

3) AI systems must be designed with transparency and be auditable, so that decision-making processes can be understood and regulated. This ensures that in the event of violations, those responsible for designing and deploying the AI can be held accountable, including in a court of law.

To achieve these aims, collaboration between governments, international organizations, and tech companies is essential to create ethical AI frameworks that focus on conflict prevention, peacebuilding, and respect for human dignity. The PAS could assist in bringing such actors together.

4.2 Decision Support Systems

The second area of major concern I wish to highlight, alongside LAWS, is the **use of AI in military planning and decision making**. Today, this is already one of the most prominent areas for AI application in so-called “decision support systems”, which integrate various forms of intelligence data to aid analysis. In light of these risky developments, I propose the following:

1) Each AI-supported process must be transparent, reviewable by a human at every stage, and sufficiently *reliable* to support a decision. It is crucial to maintain the principles of human oversight, including the necessity and proportionality of military decisions. However, generative AI trained on large data sets is inherently opaque.

2) The requirement for decision support systems to be reviewable by human agents is becoming increasingly challenging as generative AI progresses autonomously. Therefore, human review necessitates **capable international agencies** of high sophistication.

3) AI systems should be trained to uphold IHL principles. Additionally, **innovative peace research, supported by public oversight**, might engage generative AI to actively foster peace.

4) AI for moral decision-making and peace: Considering the above actions, we should remember how often immoral human actions occur in armed conflicts. Therefore, we should not exclude new pathways to solutions: we should be open to explore new opportunities with science to address the moral and ethical issues of humans and AI in armed conflicts. We must **ask how AI could**

strengthen International Humanitarian Law. Until very recently we thought that AI could not make decisions based on empathy and moral principles because they are ‘machines’ and ‘silicone chips have no empathy’. However, science-based experiments comparing human empathy with the empathic reactions of advanced AI suggests that ethical AI may be possible, depending on its design.[6] **AI that takes the Geneva Conventions seriously could** confront AI that facilitates effective but immoral human decision making in armed conflicts.

Returning to my initial question – Is IHL still fit for purpose? Broadly, it is, and a fundamental rethinking of the legal principles is not required. However, new prohibitions are needed in the area of LAWS and accountability of AI based war strategizing. Science advances in AI that assist moral judgment and peace should be pursued.

5. Call to Action

Seventy-five years ago, the Geneva Conventions set the standards for protecting human life in times of war. Today, as we look to the future, we face continuous violations of these principles in different parts of the world and new challenges through technology and AI. But the principles that guided the drafters of the conventions —the sanctity of human life, the importance of accountability, and the protection of the vulnerable—are as relevant as ever. As we strive for global compliance with these principles, we must ensure that the technology used is compliant as well. I call for the following actions and initiatives:

1) Safeguards against misuse: IHL can only address AI challenges with additional safeguards. One such safeguard would be to develop prohibitions of AWS/LAWS as an addition to the conventions without reopening the entirety of the treaties.

2) Multilateralism for risk reduction: The impact of generative AI on International Humanitarian Law depends on its design, regulation, and deployment. Without adequate safeguards, it will lead to legal, ethical, and security risks that undermine global stability and increase the likelihood of significant IHL violations and even global disasters. The role of the UN needs strengthening to address the growing AI risks.

3) Multi-actor Initiatives: We should expand initiatives that foster collaboration. I hope the initiative taken with this dialogue today continues in various formats.[7] PAS will be supportive of that.

In closing, I stress that AI and robotics in warfare present grave risks. *I repeat and emphasize the call by PAS in 2021 for regulatory action:* The responsibility lies with policymakers, military leaders, and society at large to ensure that these technologies are developed and used in ways that align with our ethical values and legal standards. International cooperation is essential to prevent the misuse of AI in warfare. We must work together to create regulatory frameworks that slow the rush to autonomous weapons and enforce human accountability. AI should not lower the

threshold for conflict, nor should it diminish the moral responsibility of human agents in war. Instead, let us use these powerful technologies to strengthen peace, uphold human rights, and foster a future where war is prevented, not perfected and perpetuated, by the machines we create.

Literature

- Joachim von Braun, Margaret S. Archer, Gregory M. Reichberg, Marcelo Sánchez Sorondo (2021) *Robotics, AI, and Humanity - Science, Ethics, and Policy*. Springer Open Access.
<https://link.springer.com/book/10.1007/978-3-030-54173-6>
- Crawford, Emily; Pert, Alison (2020). *International Humanitarian Law*. Cambridge University Press. [ISBN 978-1-108-57514-0](#).
- Toni Erskine and Steven E. Miller (2024) *AI and the decision to go to war: future risks and opportunities*
<https://www.tandfonline.com/doi/epdf/10.1080/10357718.2024.2349598?needAccess=true>
- Fleck, Dieter (2008). *The Handbook of International Humanitarian Law*. Second Edition. Oxford University Press, USA. [ISBN 978-0-19-923250-5](#).
- Fleck, Dieter (2021). *The Handbook of International Humanitarian Law*. Oxford University Press. [ISBN 978-0-19-258719-0](#).
- Solis, Gary D. (2021). *The Law of Armed Conflict: International Humanitarian Law in War*. Cambridge University Press. [ISBN 978-1-108-83163-5](#).

End notes

[1] The Pontifical Academy of Sciences has previously addressed the risk of nuclear war comprehensively. Notably, an important “Declaration On The Prevention Of Nuclear War” was issued by an assembly of Presidents of scientific academies and other scientists from around the world, convened by the PAS on September 23-24, 1982.

[2] https://www.pas.va/content/dam/casinapioiv/pas/pdf-vari/statements/pas_final_note_on_war_2.pdf

[3] Joachim von Braun, Margaret S. Archer, Gregory M. Reichberg, Marcelo Sánchez Sorondo (2021) *Robotics, AI, and Humanity - Science, Ethics, and Policy*. Springer Open Access.
<https://link.springer.com/book/10.1007/978-3-030-54173-6>

In the meantime, hundreds of artificial intelligence experts, including some members of the

Pontifical Academy of Sciences, signed a Statement on AI Risk which states, “Mitigating the risk of extinction from AI should be a global priority alongside other societal-scale risks such as pandemics and nuclear war.” (May 30, 2023).

[4] The Geneva Conventions are international humanitarian laws consisting of four treaties and three additional protocols that establish international legal standards for humanitarian treatment in war. The singular term Geneva Convention colloquially denotes the agreements of 1949, negotiated in the aftermath of the Second World War (1939–1945), which updated the terms of the two 1929 treaties and added two new conventions. The Geneva Conventions extensively define the basic rights of wartime prisoners, civilians and military personnel; establish protections for the wounded and sick; and provide protections for the civilians in and around a war-zone.

[5] Memorandum on Advancing the United States’ Leadership in Artificial Intelligence; October 24, 2024 Harnessing Artificial Intelligence to Fulfill National Security Objectives; and Fostering the Safety, Security, and Trustworthiness of Artificial Intelligence. Presidential Actions.

MEMORANDUM FOR THE VICE PRESIDENT <https://www.whitehouse.gov/briefing-room/presidential-actions/2024/10/24/memorandum-on-advancing-the-united-states-leadership-in-artificial-intelligence-harnessing-artificial-intelligence-to-fulfill-national-security-objectives-and-fostering-the-safety-security/>

[6] Yuval Noah Harari. 2024. *Nexus: A Brief History of Information Networks from the Stone Age To AI*. Random House

[7] There have been promising recent initiatives. For example, an international summit on ‘Responsible AI in the Military Domain,’ held in the Hague in February 2023, issued a ‘Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy’. This declaration has been endorsed by more than 50 states as of February 2024. It comprises a list of desirable measures to promote the safe and prudent utilization of military AI, including the proposition that its use should occur ‘within a responsible human chain of command and control’ and that AI systems must be programmed to comply with IHL principles (US Department of State 2023).