Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Views of Switzerland pursuant to resolution 79/239 "Artificial intelligence in the military domain and its implications for international peace and security" adopted by the General Assembly on 24 December 2024, in accordance with the request of the UN Secretary-General contained in Note Verbale ODA/2025-00029/AIMD

Executive Summary:

Switzerland recognizes the transformative potential of AI in the military domain including to enhance operational efficiency and decision-making - but also underscores legal, humanitarian, and strategic risks involved.

AI-enabled systems in the military domain must only be developed, deployed, and used in compliance with international law, in particular International Humanitarian Law (IHL). Human responsibility, accountability, and context-appropriate control and judgment must always be preserved. Such systems must be reliable, explainable, and robust to ensure safe and lawful use, with appropriate guardrails to minimize harm and allow for adequate human intervention. Lifecycle management is essential: legal, ethical, and technical considerations must be integrated from design through deployment to decommissioning. Switzerland stresses the importance of avoiding new pathways of escalation and supports using AI for risk reduction, arms control, and de-escalation.

Switzerland fully supports further work in the UN, including the establishment of processes to address the governance of AI in the military domain. Such governance processes must be inclusive, aimed at developing shared understandings, transparency and confidence-building measures, must be embedded in the existing legal frameworks and should be cognizant ongoing pertinent processes, notably on Autonomous Weapons Systems. Switzerland supports as part of such governance processes the developing of principles, best practices, and new international norms where needed to address emerging risks and potential legal gaps.

1. Opportunities and Risks

Artificial Intelligence (AI) is likely to transform many aspects of military affairs. It promises to support military tasks and operations, for instance by enhancing reliability, efficiency, accuracy, safety and robustness. Key areas include situational awareness, decision-making, intelligence, surveillance and reconnaissance, logistics and supply chains, training and simulation and command and control by analyzing large datasets and enabling faster, more informed decisions. For instance, in **surveillance and reconnaissance**, AI can analyze drone and satellite imagery to detect movements more quickly than human analysts. AI could also **support target recognition** by processing sensor data to distinguish between friendly and hostile forces. In logistics, AI can optimize supply chains, predict equipment failures, and ensure that resources reach the right place at the right time. For **decision support**, AI simulations can provide commanders with predictive insights and potential outcomes to guide strategic planning. Training and simulation systems powered by AI offer realistic and adaptive environments that better prepare soldiers. Finally, AI can **support command and control** by streamlining information flow, improving decision-making, and enhancing coordination across units. AI can also aid in threat detection, cybersecurity, peacekeeping, arms control verification, and conflict de-escalation through early warning systems, predictive analytics, and monitoring mechanisms, helping to promote stability and security. However, if these developments may bring benefits to the armed

forces, the integration of AI into military operations also presents several important concerns and potential risks that need to be considered/addressed.

When used responsibly in armed conflict, AI holds the potential to contribute to bolstering compliance with international humanitarian law (IHL) and strengthening the protection of civilians and civilian objects, for instance by improving risk assessments or increasing targeting precision to reduce collateral damage. However, several military applications of AI in armed conflict, especially involving high-risk applications, also raise serious legal, humanitarian, ethical, security, and strategic stability concerns that must be addressed. For instance: **Target Selection Errors:** While AI may technically identify objects or individuals based on its training data, contextual understanding and value judgements necessary for compliance with international law pose a particular challenge, which could lead to misidentification of objects or persons as military targets, and thus to unlawful or unintended strikes. **Escalation Risks**: In a fast-moving crisis, a black-box decision-support tool could recommend aggressive action without offering clear reasoning. Without explainability, commanders may either blindly follow flawed guidance or waste critical time questioning it. **Misinterpretation of Intentions**: An AI system assessing the risk associated with actions of persons and/or of objects may raise (legal and security) concerns, especially when assessments are based on patterns derived from past behaviors and contexts without context-appropriate human control and judgement. For instance, an AI system monitoring opponent's behavior may misclassify routine troop movements as hostile, due to flawed data, potentially prompting preemptive action and unintended escalation.

These risks underscore the obligation to ensure compliance with existing international law, particularly international humanitarian law, but also the urgent need for further dialogue and study of this issue to better understand risks and challenges, possible necessary measures as well as to consider the necessity, added value and feasibility of developing additional normative governance structures. This could include national legislation, the elaboration of best practices, international norms, standards or instruments, or the establishment of operational guidelines.

2. Legal Framework

The development and use of AI, as well as any other technology, does not take place in a legal vacuum. AI-enabled systems in the military domain must be developed, deployed, and used in full compliance with existing international law, particularly the UN Charter, IHL and human rights law, and other relevant legal frameworks. No technology must ever challenge the validity of international law. **International law, particularly the United Nations Charter in its entirety, international human rights law and international humanitarian law (IHL), apply and must be observed and complied with.**

States and parties to a conflict must **respect and ensure respect for IHL in all circumstances**, including when using AI in military operations. Hence, AI-enabled systems in the military domain should be designed to enhance compliance with IHL and the protection of civilians and civilian objects. This could be achieved, for instance, by ensuring that AI systems prioritize accuracy, harm minimization, and accountability - such as through strict target selection, validation, and verification processes. Moreover, AI should be used in a way to enhance the implementation of the legal obligation to take all feasible precautions in military operations, including to avoid or at the very least minimize incidental harm, by supporting commanders in protecting civilians and civilian objects throughout the conduct of hostilities, for example by improving risk assessments.

A key area of action is to ensure that AI systems in the military domain are designed with, and trained on, datasets that enable their use in full compliance with international law. Beyond the

conduct of hostilities, military applications of AI must comply with all relevant rules and principles of IHL, should they be used to perform other tasks governed by IHL, for instance in relation to detention and internments of persons or with regard to crowd control and public security measures in occupied territories.

In developing and using AI in the military domain, there is a risk that overly permissive legal interpretations - such as broadening the definition of lawful targets or raising thresholds for acceptable incidental harm - may become embedded in system design or training data. If applied at scale, such interpretations could gradually undermine the protective purpose of IHL and significantly increase harm to civilians. This risk underscores the importance of safeguarding the integrity of legal norms, which must remain a central consideration in the governance, design, and deployment of military AI going forward.

3. Understandings and Principles

Building on, and flowing from, this legal framework outlined above, and also taking into account the humanitarian, ethical, security, and strategic stability concerns, the **following understandings and principles are relevant and should be further developed:**

a) Human Responsibility, Accountability and Involvement

- **Responsibility and Accountability:** States must ensure that humans remain responsible and accountable at all times, in accordance with applicable international law, for decisions involving AI-enabled systems in the military domain.
- **Context-appropriate Human Control and Judgement:** Critical military decisions - from the board room to the battlefield - and especially those involving the use of force, must always be made with context-appropriate human control and judgement. AI-enabled systems can assist in decision-making but should not replace legal and ethical considerations and judgements, such as cognitive autonomy for decisions. States must only integrate these systems into a chain of command and control in which humans are able to maintain judgment and can exercise appropriate levels of control. Unintended biases should be addressed to the extent possible.

b) Reliability, Predictability/Explainability, Robustness

- **Reliability:** AI-enabled systems must be reliable to prevent unintended consequences or malfunctions, especially if they could have a negative impact or harm civilians and civilian objects. AI in the military domain must only be used if the effects and consequences can be reasonably foreseen.
- **Predictability/Explainability:** The decision-making processes of AI should be predictable and explainable to those responsible for their deployment, allowing them to understand and anticipate system behaviors.
- **Robustness:** AI-enabled systems must also be robust – technically and operationally – in order to remain secure and safe when deployed and used.

c) Risk Mitigation

- **Enhancing Situational Awareness:** AI should be used to improve battlefield awareness by, inter alia, detecting civilian presence with a view to reducing the likelihood of harm.

- **Predictive Analytics:** AI-driven predictive models should be used to assist in assessing risks and developing, inter alia, conflict de-escalation strategies and prevent civilian casualties.
- **Built-in Guardrails:** AI-enabled systems should incorporate safeguards that minimize harm and allow adequate human intervention in case of system failures.

d) Avoiding new Pathways of Escalation

- **Stability:** AI-enabled systems must only be designed, deployed, and used in a way that does not exacerbate international tensions or create new pathways for escalation.
- **Arms Control:** AI could support arms control and must not undermine existing non-proliferation, arms control, and disarmament norms and instruments, or hinder compliance with such norms, particularly concerning biological and nuclear weapons.
- **Crisis Management:** AI-enabled systems could support de-escalation and crisis management.

e) Lifecycle Management of Military AI Systems

Responsible military use of AI requires a comprehensive and risk-sensitive approach that addresses the entire lifecycle of AI-enabled systems. This includes the design, development, testing, deployment, operation, updating, and decommissioning of such systems. At each phase, relevant legal, humanitarian, operational, and technical considerations must be systematically integrated. This lifecycle-based approach is particularly essential for high-risk AI applications in the military domain, such as those involving autonomous weapons, target selection, or decision support risking harm or death to people or damage to objects and more generally where the decisions are governed by IHL. For systems with lower risk, such as administrative support tools or logistical planning systems, lifecycle management should be applied based on a context-specific risk assessment.

- During the design and development phase, States must ensure that systems are trained on high-quality, representative datasets - that are based on a minimum of biases - enabling their use in full compliance with international law, norms and standards, to minimize unwanted bias.
- In the testing and evaluation phase, rigorous validation and verification procedures must be implemented to confirm reliability, legal compliance, and operational robustness under realistic conditions.
- In the deployment and operational use phase, safeguards must be in place to monitor system performance, ensure context-appropriate human control and judgement, and enable adequate human intervention.
- Throughout the updating and learning phases, States must establish strict protocols for system modifications, including version control, re-validation, and formal approval processes.
- For the retirement or decommissioning phase, measures must be in place to securely disable or archive systems to prevent misuse, unintended activation, or re-deployment.
- By integrating lifecycle management tailored to the risk profile of each AI application, States can help ensure that military AI systems remain lawful, ethical, effective, and trustworthy throughout their use.

4. International Governance

Switzerland underscores the importance of an inclusive, broadly carried and sustained United Nations process to consolidate shared understandings of benefits, risks and challenges, and to develop principles for the responsible use of AI in the military domain.

The overarching aim of all international governance efforts for responsible use of AI in the military domain must be to ensure compliance with international law, in particular IHL. In addition, humanitarian and ethical concerns, the safeguard of stability and the reduction of security risks must be at the center of such efforts. Effective governance frameworks, shared norms, and sustained multilateral dialogue should help prevent unintended escalation, foster transparency and mutual confidence, and strengthen the role of international law in times of technological disruption. By anchoring military AI governance in these principles, States contribute to a global security environment that is more predictable, resilient, and peaceful.

Specific efforts to achieve this aim could include:

- promoting common understandings on definitions, scope and terminology related to AI in the military domain;
- identifying and better understanding the humanitarian, legal, security, and ethical opportunities and concerns;
- exploring transparency and confidence-building measures;
- developing principles, norms, best practices, and other recommendations; and
- providing guidance for their implementation.

Consideration should also be given to the fact that a core challenge for governance approaches lies in managing the dual-use nature of emerging technologies, which can serve both civilian and military purposes. Various lines of governance efforts on both civilian and military sides should avoid unnecessarily restricting legitimate, responsible use. Legal, ethical, or security concerns often stem less from the technologies themselves than from the way how they are used - underscoring the need for technology-neutral approaches, as far as possible, that enable harvesting new opportunities that new AI-enabled procedures and technologies can provide.

Switzerland believes that any governance process in the United Nations should be guided by key criteria such as participation open to all UN Member States and other relevant stakeholders, transparency, regular sessions, and articulation with other relevant processes.

While a broader UN-centered process is important and should ensure inclusive multistakeholder participation, it is also a fact that private technology companies and research institutions are the main actors when it comes to technological developments in the field of AI. For this reason, incorporating, science, technology industries, civil societies, and academia representatives is increasingly essential to ensure legitimacy, expertise, and broad-based support.