United Nations A/80/78

General Assembly

Distr.: General
5 June 2025
English
Original: Chinese/English/French/
Russian/Spanish

Eightieth session
Item 101 of the preliminary list*
General and complete disarmament

Artificial intelligence in the military domain and its implications for international peace and security

Report of the Secretary-General

Summary

The present report provides a consolidated summary of elements from the submissions received from Member States and observer States pursuant to resolution 79/239, without prejudice to their individual positions. It includes opportunities and challenges related to artificial intelligence in the military domain; a catalogue of existing and emerging normative proposals; a survey of initiatives in the field of artificial intelligence in the military domain; considerations on next steps; and the observations and conclusions of the Secretary-General.

* A/80/50.

25-06526 (E) 110625 040825

Contents

I. Introduction

1. In paragraph 7 of its resolution 79/239 on artificial intelligence (AI) in the military domain and its implications for international peace and security, the General Assembly requested the Secretary-General to seek the views of Member States and observer States on the opportunities and challenges posed to international peace and security by the application of AI in the military domain, with specific focus on areas other than lethal autonomous weapons systems, and to submit a substantive report summarizing those views and cataloguing existing and emerging normative proposals, with an annex containing these views, to the General Assembly at its eightieth session, for further discussion by States. In paragraph 8 of the same resolution, the Assembly also requested the Secretary-General to invite the views of international and regional organizations, the International Committee of the Red Cross, civil society, the scientific community and industry and to include these views in the original language received in the annex to the aforementioned report. The present report is submitted pursuant to those requests.

2. On 12 February 2025, the Office for Disarmament Affairs sent a note verbale to all Member States and observer States, drawing their attention to paragraph 7 of General Assembly resolution 79/239 and seeking their views on the matter. Notes verbales and letters were also sent to the entities specified in paragraph 8 of the same resolution, drawing their attention to that paragraph and seeking their views on the matter. The views received by 11 April 2025 are reproduced in the annexes to the present report. Any views received after that date will be posted on the website of the Office in the original language of submission.

3. Sections II to VI of the present report provide a consolidated summary of elements from the submissions received from Member States and observer States, without prejudice to their individual positions. The observations and conclusions of the Secretary-General are set out in section VII.

II. Background

4. States referred to rapid advances in science and technology in general, and in AI in particular, noting their widespread impact on society. More specifically, States noted that AI had the potential to transform every aspect of military affairs and have a significant impact on international peace and security.

5. Several States referred to current applications of AI in the military domain, as well as their own efforts to use AI in defence operations. While recognizing the importance of discussions surrounding lethal autonomous weapons systems, States noted that the issue of AI in the military domain was broader, encompassing a wider range of capabilities.

III. Opportunities and challenges

6. It was noted that AI presented both opportunities and challenges, which should be addressed in a realistic manner. It was acknowledged that the pace of AI development meant that the totality of opportunities and challenges could not be predicted at present. It was suggested that the technology itself should not be stigmatized.

A. Opportunities

7. Speed was recognized as a chief advantage of AI, including in information analysis and in decision-making. Scale was also noted as an advantage, such that AI could act as a "force multiplier". Several States referred to the potential for AI to enhance efficiency, accuracy and precision, leading to a lower error probability in comparison with humans. Additional characteristics noted were reliability, safety and robustness.

Applications

8. Several States referred to the applications of AI in the field of intelligence, surveillance and reconnaissance, where it could be used to efficiently analyse large data sets, facilitate the detection of threats, enable increased situational awareness and more accurate operations. It was noted that these same characteristics enabled AI to support decision-making and command and control, potentially leading to more precise operations and reducing risks to civilians and providing greater protection to civilian objects. It was also stressed, however, that AI tools should not replace human decision-making.

9. Several States indicated that AI could be integrated into uncrewed systems. It was noted that AI could improve coordination and communication between military actors and between military actors and others, such as providers of humanitarian assistance. In general terms, it was observed that AI could reduce the burden of routine or repetitive tasks and augment human capabilities in complex tasks.

10. According to some States, AI could be used to enhance information and communications technology security by detecting intrusions or other malicious activities, including to protect critical infrastructure. It was noted that AI could be used to detect AI-generated content used for misinformation and disinformation, as well as to identify hate speech, propaganda or changes in public sentiment.

11. Reference was made to other applications of AI that were not directly related to combat, including optimizing logistics, predictive maintenance, procurement, resource allocation, administration, simulation and training.

International peace and security

12. Several States considered that AI could contribute to the maintenance of international peace and security, for example, AI-supported situational awareness could help to mitigate risks and contribute to the de-escalation of conflicts. It was noted that the use of AI could lower risk to military personnel, for example, by replacing humans in certain dangerous tasks, such as the disposal of unexploded ordnance, or by supporting search-and-rescue operations in remote locations.

13. It was suggested that AI could improve the implementation of international humanitarian law, in particular its fundamental principles of distinction, proportionality and precautions in attack, and the protection of civilians and civilian objects. In that connection, several States noted the ability of AI to improve situational awareness in general, and understanding of the civilian environment in particular, as well as its ability to increase accuracy and reduce the risk of human error. It was also noted that AI could facilitate investigations into civilian casualties and thus ensure that those responsible were held accountable.

14. Several States suggested that AI could help to monitor and verify the implementation of disarmament, non-proliferation and arms control agreements. Reference was made to the potential of AI to support peacekeeping missions, including facilitating planning, logistics and ceasefire monitoring. Other related

applications of AI identified included border security, combating terrorism, the detection of illegal weapons programmes and optimizing humanitarian assistance and disaster response.

B. Challenges

15. Several States noted that the rapid developments in emerging technologies in general and in artificial intelligence in particular posed challenges for international peace and security. While it was important to understand these challenges, it was not currently possible to fully foresee all of them.

16. The following concerns related to AI were highlighted:

- An acceleration in the observe, orient, decide, act loop, compressing the time available for decision-making
- Increasing autonomy and loss of human control, especially in the context of the use of force
- The potential for misuse or malicious use
- Excessive trust by humans in AI applications
- Deepening technological asymmetries between States

International peace and security

17. Several States noted that the integration of AI in the military domain could pose challenges to international peace and security. The use of AI could increase the risk of misunderstanding, miscalculation and unintended escalation, including as a result of the increased speed and scale of AI-supported operations or because of technical failures. These causes could also lower the threshold for the use of force. Several States expressed concern regarding the emergence of an arms race in this field. It was suggested that the use of AI could shift the balance from defensive to offensive actions and that increasing imbalances between States could lead to increased instability, thereby undermining international peace and security.

18. Several States expressed concern regarding the potentially destabilizing effect of the proliferation of AI capabilities, including to non-State actors. It was noted that there was currently no multilateral framework to control the proliferation of weapons that integrated AI capabilities.

Technological considerations

19. States considered risks arising from technological considerations, which included:

- Technical failure and malfunction
- Design flaws
- Unintended behaviour, diverging from design parameters
- Vulnerability to cyberattacks and data poisoning
- Algorithmic and data biases, including gender bias
- Automation bias, resulting from insufficiently trained human operators
- Privacy concerns arising from the collection and processing of large volumes of personal data to train AI models

- Problems caused by poorly trained AI models
- Problems emanating from poor testing, evaluation, validation and verification procedures
- Target selection errors
- Excessive energy consumption
- Excessive reliance on external providers

20. Several States expressed concern regarding the transparency and explainability of complex AI capabilities, which are often referred to as “black boxes”. Concern was also expressed regarding the use of civilian AI applications, such as generative AI, which could add complexity and uncertainty to a conflict situation. Several States also raised concerns regarding the convergence of AI and other technologies.

Legal and humanitarian considerations

21. Several States noted that AI posed challenges for adherence to international law, in particular international humanitarian law and international human rights law. The use of AI could lead to indiscriminate use of force and raised questions of responsibility and accountability in the case of illegal or wrongful acts. Related issues raised included the protection of civilians and civilian infrastructure, as well as the potential to increase the intensity and lethality of conflicts for combatants.

22. Several States raised ethical concerns, noting that the use of AI could diminish the scope for compassion, moral reasoning and human judgment.

Potential areas of misuse

23. Several States noted the potential for AI to be used for cyberattacks by both State and non-State actors, including against critical infrastructure. AI could also potentially be used for misinformation and disinformation campaigns, including for the production of false information and deepfakes, as well as for dissemination by AI-driven bots. The use of such misinformation and disinformation, for instance to influence elections, could be destabilizing.

Weapons of mass destruction

24. Several States stressed the importance of maintaining human control over nuclear weapons and their delivery systems and expressed concern over the possibility of AI being integrated into nuclear command, control and communication systems. Reference was made to the commitment of certain nuclear-weapon States to maintain human control and involvement for all actions critical to informing and executing sovereign decisions concerning nuclear weapons employment. Potential repercussions for strategic stability and escalation were noted.

25. Several States expressed concern that AI could facilitate the proliferation of weapons of mass destruction, including to non-State actors. In that context, one particular concern was that AI could be used to develop and produce biological weapons. It was stressed that, under the provisions of existing treaties, AI must not be used to that end. The view was also expressed that AI could be used to curb the proliferation of weapons of mass destruction.

IV. Existing and emerging normative proposals

26. Several States expressed that AI should be used for peaceful purposes, including the peaceful settlement of disputes. States also stressed the importance of addressing

and mitigating the risks arising from AI in the military domain, with some noting that the challenges arising from military AI should be addressed collectively.

27. In addressing AI in the military domain, States called for an approach that was:

- Flexible, balanced, realistic and incremental, and thus able to adapt to technological advancements
- Precautionary
- Focused on the entire life cycle of AI, including pre-design, design, development, evaluation, testing, deployment, use, sale, procurement, operation and decommissioning
- Based on the applications and use of AI, rather than on the technology itself
- Reflective of existing obligations

It was suggested that efforts in this area should clearly distinguish between lethal and non-lethal uses.

Legal considerations

28. States recalled resolution 79/239, in which the General Assembly affirmed that international law, including the Charter of the United Nations, international humanitarian law and international human rights law, applied to matters governed by it that occur throughout all stages of the life cycle of AI, including systems enabled by AI, in the military domain. It was noted that international law in general, and international humanitarian law in particular, did not categorically prohibit the use of AI capabilities.

29. States affirmed that they were complying with international law in their use of AI in the military domain. It was suggested that compliance with legal obligations, particularly those deriving from international law, must be a key consideration in the governance, design and deployment of AI in the military domain. Moreover, the view was expressed that AI should be designed to enhance compliance with international humanitarian law. Several States stressed the importance of conducting legal reviews of new weapons, means or methods of warfare in that regard.

30. Several States stressed the importance of taking into account ethical considerations, in addition to legal frameworks.

International peace and security considerations

31. Several States noted that AI in the military domain should enhance international peace and security and be used in a way that does not lead to instability or escalation. The view was expressed that States should refrain from seeking absolute military advantage through AI and should ensure that such technology would not become a tool for launching an invasion and pursuing hegemony.

32. Several States noted that AI should not undermine existing disarmament, non-proliferation and arms control agreements. There were calls for efforts to prevent the proliferation of AI technology to non-State actors. The importance of avoiding arbitrary international oversight mechanisms or discriminatory export control was stressed.

Responsible use of artificial intelligence in the military domain

33. Several States were of the view that AI should be applied in a responsible manner throughout its life cycle. It was expressed that the concept of responsibility should be linked to legality and accountability.

34. Several States stressed the importance of a human-centric approach to AI. Many States stressed the importance of human control and responsibility at all times. Reference was made to the importance of concepts such as “context-appropriate human control and judgment” and “meaningful human control”. By contrast, according to other States, these concepts were insufficiently defined. The view was expressed that use of the concept of “meaningful human control” could hamper legitimate research or unduly restrict the use of AI in the military domain.

35. States stressed the importance of ensuring human responsibility and accountability, including within a responsible chain of human command and control, in accordance with international law.

Technological considerations

36. States considered governance principles from a technological perspective, including the following:

- Security, to ensure the robustness of AI systems against external threats
- Safety, including by incorporating guardrails to minimize harm
- Reliability, to prevent unintended consequences and malfunctions
- Clear operational boundaries and constraints to prevent unintended behaviour
- Well-defined use cases
- Governability, by ensuring appropriate human-machine interaction and bias mitigation
- Equity and fairness
- Privacy protection
- Explainability, understandability and traceability
- Transparency

37. The use of training data that enables full compliance with international law was highlighted. Several States emphasized the importance of testing throughout the life cycle to uncover errors and ensure reliability. It was also stressed that there was a need for adequate training of personnel working with AI to mitigate risks and to ensure compliance with international humanitarian law. The importance of monitoring system performance throughout its life cycle, and of measures to securely disable systems at the point of retirement, was underlined.

V. Initiatives in the field of artificial intelligence in the military domain

International forums

38. Several States noted ongoing discussions in the United Nations, as well as the Pact for the Future (General Assembly resolution 79/1) and the Global Digital Compact annexed thereto, and the General Assembly resolution on artificial intelligence in the military domain and its implications for international peace and security (resolution 79/239). The Arria-formula meeting of the Security Council on harnessing safe, inclusive, trustworthy AI for the maintenance of international peace and security, held on 4 April 2025, was also noted.

39. States referred to multilateral discussions on topics related to AI in the military domain, such as those in the Disarmament Commission under the agenda item entitled

"Recommendations on common understandings related to emerging technologies in the context of international security", the work of the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems, established under the Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to Be Excessively Injurious or to Have Indiscriminate Effects, as well as the General Assembly resolutions on lethal autonomous weapons systems (resolutions 78/241 and 79/62).

40. States also referred to their participation in activities related to AI in the military domain organized by the Office for Disarmament Affairs and the United Nations Institute for Disarmament Research.

State-led initiatives

41. Several States noted that they had initiated or participated in initiatives related to AI in the military domain, including:

- The Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy and its subsequent implementation process
- The Responsible Artificial Intelligence in the Military Domain process, which included conferences in the Kingdom of the Netherlands in 2023, at which a Call to Action was endorsed, and in the Republic of Korea in 2024, at which a Blueprint for Action was endorsed. A Global Commission on Responsible Artificial Intelligence in the Military Domain was expected to publish a report prior to the next conference, to be held in Spain in 2025
- The Artificial Intelligence Action Summit that was held in France in 2025, at which the Paris Declaration on Maintaining Human Control in AI-enabled Weapon Systems was adopted
- The AI Safety Summit that was held in the United Kingdom of Great Britain and Northern Ireland in 2023, at which the Bletchley Declaration was adopted
- Work on AI taking place in the context of the Group of Seven
- The AI Partnership for Defence
- The Global AI Governance Initiative proposed in 2023

42. The view was expressed that these initiatives, while useful, could result in fragmentation. Concern was also expressed that the outcomes of these initiatives did not take into account the views of all States concerned and could undermine inclusive work in this area.

Regional initiatives

43. States noted the importance of regional initiatives to foster inclusive and context-specific discussions on AI in the military domain. Examples in this regard include the following:

- The joint statement on cooperation in the field of AI in the defence sector, adopted at the ASEAN Defence Ministers' Meeting Retreat in 2025
- The sixteenth Conference of Defence Ministers of the Americas, held in 2024, at which the Mendoza Declaration was adopted
- Activities in the context of the North Atlantic Treaty Organization, including its AI strategy, last revised in 2024, and its principles of responsible use, elaborated in 2021

- Regional consultations in the context of the Responsible Artificial Intelligence in the Military Domain process in 2024, which have been held in Chile, Kenya, Netherlands (Kingdom of the), Singapore and Türkiye

Domestic initiatives

44. States referred to their domestic efforts, including extant AI legislation, regulations, strategies and bodies, as well as efforts to develop these.

VI. Next steps

45. States called for dialogue on AI in the military domain. Several States called for further study of the impact of AI in the military domain on international peace and security.

46. Numerous States noted that the goal of further dialogue should be to mitigate the risks posed by AI in the military domain. It was suggested that the goal of dialogue should be the development of regulatory or governance frameworks. Several States called for the development of norms, rules and principles to govern the life cycle of AI in the military domain. While some States expressed a preference for the development of a legally binding framework, others did not consider the adoption of new legal measures to be necessary at present. The view was also expressed that norms, rules and principles could form the basis of legal commitments at a later stage. Several States expressed opposition to the concept of norms, rules and principles of responsible development, deployment or use, noting that the concept did not enjoy consensus. The view was expressed that the premature introduction of regulations should be avoided.

47. The importance of avoiding duplication and fragmentation in governance was stressed. It was considered that a discussion on governance should take into account humanitarian, security and development considerations in a balanced manner. States highlighted the importance of avoiding restrictions that would stymie legitimate innovation and technological progress. Several States considered that the peaceful uses of AI, especially by developing nations, should not be impeded.

48. It was suggested that any governance approach should take into account that States were at different stages of integrating AI into military capabilities and had varying security environments. The importance of the participation of all States in discussions on the governance of AI in the military domain was emphasized. Many States considered that future discussions should take a multi-stakeholder approach, including international and regional organizations, civil society, the scientific community and industry. It was stressed, however, that decision-making should remain the sole prerogative of States.

49. States considered various priorities for future dialogue on AI in the military domain, including:

- Ensuring compliance with international law, in particular international humanitarian law
- Protecting human dignity and human rights
- Seeking common understandings on definitions and terminology
- Considering transparency and confidence-building measures
- Addressing autonomy in the use of force
- Addressing AI systems that directly support combat operations

- Ensuring adequate data governance mechanisms
- Strengthening international cooperation and assistance
- Supporting capacity-building, including through knowledge-sharing, technology transfer and the sharing of good practices, so as to bridge the digital divide and the AI divide
- Promoting continued regional dialogue
- Promoting national regulation, including in order to ensure private sector compliance with international law

50. Several States suggested that consideration of lethal autonomous weapons systems should form part of any discussion of AI in the military domain. The view was also expressed that the ongoing discussions on such systems were complementary to discussions on AI in the military domain. Several States recalled their positions on lethal autonomous weapons systems.[1] While some expressed the view that the Group of Governmental Experts established under the Convention on Certain Conventional Weapons was the optimal forum for discussions on AI in the military domain, others stated that, given its specific mandate and non-universal membership, that Group was not an appropriate forum for such discussions.

51. Several States called for discussions on AI in the military domain within United Nations forums. It was suggested that the present report could form the basis for such discussions. States indicated that future discussions should be complementary to ongoing processes, such as the open-ended working group on security of and in the use of information and communications technologies.

52. Several States were of the view that the United Nations disarmament machinery represented an effective and inclusive platform and should play a central role in future discussions of AI in the military domain. It was suggested that the Conference on Disarmament should discuss AI, in particular in relation to nuclear weapons. The view was expressed that discussions could also be held in the First Committee of the General Assembly, which could mandate regular reports by the Secretary-General on the status of the technological development of AI in the military domain. Several States suggested that discussions could be held in the context of the Disarmament Commission.

53. It was also suggested that discussions could be held in the context of the Security Council.

54. Several States suggested that a dedicated process, such as an open-ended working group, should be established. The view was also expressed that the creation of a new process within the United Nations would be inappropriate at this time. The view was expressed that any United Nations process on the subject should be guided by consensus.

VII. Observations and conclusions of the Secretary-General

55. AI has the potential to impact every facet of our lives. When used for peaceful purposes, it can play a significant role in facilitating the achievement of development commitments and objectives, including the Sustainable Development Goals.

56. In the military domain, AI has the potential to bring benefits both to the militaries employing it and to civilian populations, by increasing the accuracy of operations and reducing the scope for human error. At the same time, AI in the

[1] For more detail, see A/79/88.

military domain raises serious challenges, chief among them the maintenance of human responsibility and accountability.

57. The affirmation by the General Assembly in its resolution 79/239 that international law, including the Charter of the United Nations, international humanitarian law and international human rights law, applies throughout the life cycle of AI is an important baseline. However, important questions on how the law applies remain to be resolved.

58. The use of military AI in situations involving the use of force requires particular attention. While there are potential benefits for the protection of civilians and combatants, reported uses of AI in present-day conflicts raise concerns regarding human control and the role of AI in facilitating hostilities in densely populated areas. Machines that have the power and discretion to take human lives are politically unacceptable and morally repugnant.

59. The risks posed by nuclear weapons will not be eliminated until the weapons themselves are eliminated. Pending the total elimination of nuclear weapons, I urge all States that possess these weapons to agree that any decision on nuclear use be made by humans, not machines.

60. AI can lower the barrier for State and non-State actors to developing or acquiring chemical and biological weapons. I therefore urge States to fully meet their obligations under relevant disarmament, non-proliferation and arms control frameworks and systematically evaluate and be well prepared to respond to the challenges and the impact of AI on these frameworks.

61. The potential integration of civilian AI applications into the military domain is a growing cause for concern. The inherently repurposeable nature of AI technologies presents challenges for oversight, transparency and accountability. I urge States to carefully examine the blurring of lines between developments in civilian AI applications and their potential use in the military domain.

62. There is significant value in pursuing additional cooperative mechanisms on AI, especially at the regional and subregional levels. Regional and subregional organizations are uniquely equipped to develop and implement transparency and confidence-building measures as a means to mitigate risk. I therefore encourage States to consider elaborating transparency and confidence-building measures at the regional and subregional levels tailored to the unique characteristics and challenges of AI.

63. Inclusive discussions on the peaceful uses of AI and its governance for the benefit of humanity are ongoing under the auspices of the United Nations, particularly in the context of the implementation of the Global Digital Compact. Nevertheless, Member States' consideration of AI in the military domain has largely taken place outside United Nations forums. General Assembly resolution 79/239 and the present report are notable first steps in bringing this important discussion to the United Nations. I encourage States to conduct these deliberations in an inclusive and constructive manner, with a view to advancing shared understandings and strengthening international cooperation to mitigate risks.

64. States are encouraged to explore efforts, including capacity-building, to ensure meaningful participation by all States in United Nations processes on this subject, which is essential for fostering a shared understanding, developing common approaches and mitigating potential risks.

65. The General Assembly has proved adept at mandating processes that foster inclusive discussions on issues related to emerging technologies and international security, while also fostering input from stakeholders, including international and

regional organizations, civil society, the scientific community and industry. This multi-stakeholder approach is particularly important in the field of AI, where innovation is largely driven by the private sector and much expertise resides outside governments, in academia and the scientific community.

66. **I recommend that States study the ideas contained in the present report and, at the eightieth session of the General Assembly, take concrete steps with a view to the establishment of a dedicated and inclusive process to comprehensively tackle the issue of AI in the military domain and its implications for international peace and security.**

Annex I

Replies received

A. Member States

Argentina

[Original: Spanish]
[10 April 2025]

The following report is submitted in relation to resolution 79/239, entitled "Artificial intelligence in the military domain and its implications for international peace and security", adopted by the General Assembly on 24 December 2024.

General approach

The Argentine Republic recognizes that the rollout of artificial intelligence (AI) in the military environment has major strategic impacts. The use of AI gives rise to tangible benefits for various non-lethal functions, while introducing risks that require a response from the perspectives of international law, ethics and operational responsibility. In that context, respect for international humanitarian law and human rights must be upheld in the development and use of such technologies, and there must be assurance that human responsibility for and control over critical decision-making are maintained at all times.

Opportunities

AI, in particular its non-lethal applications, is a legitimate and valuable tool for enhancing national defence capabilities. Priority uses include the following:

- Logistical and operational optimization
- Support for intelligence processing
- Strengthening of cyberdefence
- Simulation, training and strategic planning

These capabilities contribute to more efficient and safer operations that are better adapted to current scenarios, strengthening defensive effectiveness without undermining humanitarian principles or the international obligations of the State.

Challenges

The accelerated development of AI in military settings poses challenges that must be addressed collectively, including the following:

- Lowering of the threshold for the use of force, and shortening of the time frame for human decision-making
- Possibility of undetected algorithmic bias
- Proliferation of autonomous systems to non-State actors
- Risk of entrenching technological asymmetries between States

These risks underscore the need to establish common principles, verifiable safeguards and cooperative frameworks.

Governance, international cooperation and technological inclusion

It is our understanding that any policy-setting process in this area should be built upon the following principles:

- General or premature regulations that limit the independent development of legitimate defensive technologies should be avoided.
- A clear distinction should be made between lethal and non-lethal uses.
- Significant human control should be guaranteed as an integral operational and policy condition.
- Inclusive international cooperation focused on strengthening capabilities and bridging technological gaps between States should be promoted.

Argentina has reaffirmed these principles at recent multilateral forums, emphasizing the importance of working towards common standards for the responsible use of AI in the military domain, in particular in relation to cyberdefence and cybersecurity.

To give an example of an initiative at the regional level, as part of the sixteenth Conference of Defence Ministers of the Americas, held in Mendoza, Argentina, in 2024, the working group on the responsible development, application and governance of AI in the military domain met to work collectively on the development of international standards.

Reference to the Pact for the Future

Lastly, let it be noted for the record that the Argentine Republic has formally disassociated itself from the Pact for the Future, which is cited in the preamble of General Assembly resolution 79/239. The reference to the Pact therefore does not denote a commitment or adherence to, or support for, the Pact by the Argentine State.

Austria

[Original: English]
[11 April 2025]

Pursuant to the request in paragraph 7 of General Assembly resolution 79/239, Austria would like to share the following reflections and observations on a national basis.

Artificial intelligence related to cybersecurity and cyberdefence

Cybersecurity software enabled by artificial intelligence (AI) is already widely used to help detect intrusions and other malicious activities in computer networks. Such AI tools will likely enable the increasingly automated protection of information technology systems by searching for vulnerabilities and suspicious activities to raise the resilience of software and hardware.

At the same time, AI tools are increasingly used to enhance the sophistication of cyberattacks and create novel computer viruses in a race between offensive and defensive cybersecurity AI models. In addition, AI-enabled software, including large language models, lower the entry barrier for malicious actors who can increasingly create malware without the need for extensive programming skills.

Artificial intelligence related to disinformation campaigns as an element of hybrid strategies

AI-enabled software that can create and disseminate falsified content is increasingly used to enhance disinformation campaigns. Methods used include utilizing generative AI to create tailor-made and localized content on a large scale. Furthermore, AI-driven deepfake audio and video software is rapidly improving and already widely used. Such falsified content can be disseminated using massive AI-driven social media bot networks to create the appearance of shifting public opinion. AI therefore lowers the barriers to conducting large-scale disinformation campaigns as the amount and quality of fake content created is no longer limited by the number or skills of human operators.

However, AI algorithms can also be employed to uncover AI-generated content and astroturfing campaigns, while deceivingly real deepfake audio and video can best be exposed using specialized AI tools. It is necessary to employ such AI-driven tools to counter the ill effects of AI used for the purpose of disinformation campaigns.

Artificial intelligence related to the proliferation of weapons

AI can lower the barrier to acquiring weapons, including weapons of mass destruction. Due to their ability to provide expertise at the push of a button, large language models and the applications based on them could make it easier for malicious actors to manufacture weapons. Use cases range from access to blueprints or printing components for small arms and light weapons to the modification of pathogens for biological warfare. If readily available knowledge reduces the scope and size of weapons programmes, it will be more difficult to detect, prevent and prepare for these threats.

At the same time, machine-learning algorithms can also be used to combat the proliferation of weapons. Due to their anomaly detection and pattern recognition capabilities, they can help to identify malicious activities, including through detecting illicit money flows for weapons programmes or analysing patterns in satellite data.

Artificial intelligence related to arms control verification and decision-making in crisis situations

AI can help with the verification of arms control agreements. This is due to its ability to analyse large amounts of data – from sources such as satellite images, for example – and to classify different objects. This makes it possible to identify military equipment, such as tanks, missiles and barracks, or military activities, such as troop movements and exercises. In addition, as already mentioned, illegal weapons programmes could be detected more easily by AI. Violations of arms control agreements would therefore be much more difficult to commit and the States Parties could be sure that everyone is complying with the agreement provisions.

More and better information based on the ability of AI to analyse and classify sensor data can not only facilitate the implementation of arms control agreements, but also contribute to better decision-making in situations where military tensions between States are particularly high. Political and military leaders could benefit from AI-supported improved situational awareness to de-escalate crises.

Peace and security and the Charter of the United Nations

A particular challenge relating to applications of AI in the military domain is the potential risks to peace and security through unintended escalation and misunderstandings created through the use of AI. The use of machine learning adds

an additional layer of complexity as the functioning of a system might not be fully understood by all actors.

Measures and guardrails to ensure accountability and responsibility and to mitigate algorithmic bias are needed also for the use of AI in decision support systems with regard to human-machine interaction and the necessity of human agency.

All these risks have to be mitigated through oversight and measures that take into account the specific challenges that come with these technologies.

It is noted that article 36 of the Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts (Protocol I) imposes the obligation to review the legality of all new weapons, means or methods of warfare before they are used in an armed conflict.

AI can also be applied in a way to support the effective implementation of international humanitarian law obligations, in particular when it comes to the protection of civilians, as a positive obligation and affirmative action, including through projects, research and applications specifically designed for this task.

Frameworks for multilateral cooperation and information exchange

As the issue of AI in the military domain is rapidly developing and presents challenges to all States, multilateral discussions and formats to exchange experiences and best practices are highly relevant. In this regard, Austria endorsed the Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy. As Co-Chair of the oversight working group of this Declaration, Austria, together with Germany, has been facilitating the sharing of best practices in addressing challenges and in formulating policies in this field. Austria also endorsed the Blueprint for Action of the Responsible Artificial Intelligence in the Military Domain Summit, as well as the Paris Declaration on Maintaining Human Control in AI enabled Weapon Systems.

Relation between the work of the international community on artificial intelligence in the military domain and its work on autonomous weapons systems

Within the broader scope of the application of AI and autonomy in the military domain, there is the specific issue of autonomous weapons systems to be highlighted. Autonomous weapons systems raise particular concerns from a legal, ethical and security perspective. This issue is not the focus of General Assembly resolution 79/239 as discussions in the United Nations framework have already been going on since 2013, with a growing majority of States having expressed their wish to establish rules for, and limits on, autonomous weapons systems at the international level. For this report, Austria would therefore limit its comments to emphasizing its position in favour of a legally binding instrument on autonomous weapons systems and refer here to the important work being undertaken currently by the Group of Governmental Experts in the framework of the Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to Be Excessively Injurious or to Have Indiscriminate Effects, as well as the complementary efforts undertaken in the framework of the first-ever General Assembly resolution on lethal autonomous weapons systems (resolution 78/241), pursuant to which a report of the Secretary-General (A/79/88) was issued, and the follow-up resolution (resolution 79/62), which set up informal consultations on lethal autonomous weapons systems, to be held in New York on 12 and 13 May 2025.

Considerations related to legal frameworks on artificial intelligence

The European Union Artificial Intelligence Act establishes a legislative framework for the European Union for AI systems across various sectors and aims to

foster trust in AI applications and to harness the benefits of AI while safeguarding human rights, fundamental freedoms and democratic values. It emphasizes the importance of transparency, accountability and human oversight in the development and deployment of AI systems while promoting legal certainty, innovation and competitiveness. The Act does not apply to AI systems developed for military, defence or national security activities. However, the AI Act does apply a risk-based approach, which might prove useful when dealing with the wide range of potential AI applications in the military domain.

Way forward

Austria values the work undertaken in the various formats and forums mentioned in its contribution regarding AI applications in the military domain and is confident that they will contribute to an emerging set of internationally agreed norms and standards to ensure the responsible use of AI in the military domain in accordance with international legal obligations and ethical principles.

Chile

[Original: Spanish]
[11 April 2025]

Chile has stated previously that the rapid development of new and emerging technologies is an important issue as regards international security and poses a challenge for all countries. These new technologies, in particular artificial intelligence (AI), may produce enormous benefits for the development and well-being of societies, but at the same time, they raise significant questions about the ramifications of their use in the field of security and defence. New technologies can generate important benefits but also risks and difficulties.

In that regard, Chile considers that it would be advisable to develop a common understanding on the responsible use of AI in the military and security domain and on the development and use of so-called lethal autonomous weapons systems. Chile supports multilateral efforts to establish and strengthen forums for dialogue and discussion among countries, with the aim of finding areas of mutual understanding and consensus on the use of these new technologies.

Chile has taken on a leading role in the field of AI owing to the significant headway that the country has made in developing enabling conditions for the deployment of such technology, as well as its groundbreaking progress with respect to policies and regulatory discussions on AI. In October 2021, Chile launched its first national policy on AI, which was developed in collaboration with various public and private stakeholders. The policy is focused on three essential pillars: enabling factors; technology use and development; and the establishment of regulatory and ethical frameworks to ensure the responsible and safe use of AI.

In 2024, Chile released an updated version of its national policy on AI, which includes new subtopics, such as international coordination, environment and the climate crisis, inclusivity and non-discrimination, children and adolescents, and culture and heritage preservation. The policy is complemented by an action plan that contains more than 100 measures to be completed by 2026, addressing such areas as education, health, environment and culture. The principles set out in the Recommendation on the Ethics of Artificial Intelligence issued by the United Nations Educational, Scientific and Cultural Organization (UNESCO) were also incorporated into the new national policy on AI, so that the policy would be aligned with the most up-to-date international frameworks.

Chile was the first country in the world to apply the readiness assessment methodology, a tool developed by UNESCO to determine a country's readiness to implement AI in an ethical and responsible manner. Chile thereby reaffirmed its commitment to implementing the UNESCO Recommendation on the Ethics of Artificial Intelligence in its national regulations. Chile has promoted the ethical and responsible development of such technology, as reflected in the country's participation in the summits on AI organized by the United Kingdom (2023), the Republic of Korea (2024) and France (2025).

With respect to legislation, a bill to regulate AI systems using a risk-based approach is currently under debate in Chile with the aim of promoting the development and implementation of such systems while upholding democratic principles and the fundamental rights of individuals.

In the field of defence and security, Chile has supported and participated actively in the Summits on Responsible Artificial Intelligence in the Military Domain, held in The Hague (2023) and Seoul (2024). Chile endorsed the documents adopted at the two summits (Call to action (2023) and Blueprint for action (2024)). Chile also supports the work of the Global Commission on Responsible Artificial Intelligence in the Military Domain.

On 13 and 14 June 2024, a regional workshop on the responsible use of AI in the military and broader security domains was held in Chile. The workshop was organized by the Ministry of Foreign Affairs of Chile and the Ministry of Foreign Affairs of Costa Rica and sponsored by the Kingdom of the Netherlands and the Republic of Korea. The Centre for the Study of Law, Technology and Society of the University of Chile and the Centre for Humanitarian Dialogue, based in Geneva, also supported and helped to organize the event. Representatives of Argentina, Brazil, Chile, Colombia, Costa Rica, the Dominican Republic, El Salvador, Jamaica, Mexico, Paraguay, Trinidad and Tobago and Uruguay, as well as the Kingdom of the Netherlands and the Republic of Korea participated in the workshop. Representatives of the Ministry of National Defence and the Armed Forces also participated on behalf of Chile.

Chile believes that AI applications in the military and security domain can yield opportunities and benefits, such as enhanced decision-making and strategic analysis, more efficient logistical operations, enhanced capabilities in cyberdefence and cybersecurity – thereby strengthening the security of critical infrastructure – as well as facilitation of planning for complex peacekeeping and humanitarian aid missions. AI applications can also improve verification and monitoring capabilities for arms control and enforcement of arms control regimes.

Chile believes that AI technologies must be developed, deployed and used in accordance with international law, including, where applicable, the Charter of the United Nations, international humanitarian law, international human rights law and other relevant legal frameworks.

For Chile, it is important to establish control and security measures to prevent irresponsible actors from obtaining and misusing potentially harmful AI capabilities in the military domain, including systems enabled by AI, while bearing in mind that any such measures should not undermine equitable access to the benefits of AI capabilities in other, non-military domains.

Similarly, Chile believes that it is important to join efforts to prevent AI technologies from being used to facilitate the proliferation of weapons of mass destruction among State and non-State actors, including terrorist groups, and to emphasize that AI technologies should be used to support, rather than hinder, disarmament, arms control and non-proliferation efforts. It is particularly crucial to

maintain human control and involvement in all actions critical to informing and executing sovereign decisions concerning the use of nuclear arms, without prejudice to the ultimate goal of a nuclear-weapon-free world.

Chile advocates for the development of confidence-building measures, such as information exchange and consultations among States on good practices and lessons learned. In that regard, Chile considers it important for countries to develop and institute national strategies, principles, standards, policies, frameworks and laws to ensure the responsible use of AI in the military domain. Confidence-building measures can be an effective tool for developing containment, control and credibility mechanisms at the national and international levels, thereby fostering transparency.

Similarly, Chile believes that it is essential to reduce digital and AI-related gaps between developed and developing countries and considers it necessary to enhance understanding and awareness of the implications of AI in the military domain, including knowledge exchange and the sharing of good practices and lessons learned among all States.

In that regard, Chile believes that it is essential to develop initiatives and programmes aimed at fostering capacity-building, in particular in developing countries, to promote their full participation in debates on AI governance in the military domain. Chile recognizes that capacity-building can also help countries to gain a deeper understanding of AI in the military domain and facilitate the responsible and lawful development, deployment and use of military AI capabilities. Capacity-building will also equip countries to engage more effectively in international dialogue and discussion.

Chile considers it important to strengthen international cooperation for capacity-building, promoting dialogue and debate at the national, regional, subregional and interregional levels, including through training programmes, conferences, workshops and seminars for diplomatic, political and technical officials, with a view to bridging the knowledge gap concerning the responsible development, deployment and use of AI in the military domain.

Chile appreciates and considers it important to promote regional and subregional discussion and dialogue on AI in the military domain. Notable efforts to that end include the sixteenth Conference of Defence Ministers of the Americas, held in Argentina from 13 to 16 October 2024, and, in particular, the outcome document of the Conference, known as the Mendoza Declaration, which contains the following recommendations: foster the ethical use of AI in defence; take into account the economic and technological diversity of States members of the Conference; and promote mechanisms for strengthening mutual trust and hemispheric and regional cooperation through which States members of the Conference can share knowledge and good practices, develop consensus-based standards and build technological capabilities for applying AI in the field of defence.

Lastly, Chile believes that, in discussions and dialogue on AI in the military domain, the participation of all interested parties, including civil society, academia, industry, the private sector, the technical community and regional and international organizations, is essential.

China

[Original: Chinese]
[11 April 2025]

The rapid development and widespread application of artificial intelligence (AI) in the military domain is reshaping future warfare paradigms while posing potential

challenges to international peace and security. As the world faces multiple challenges to peace and security, all parties should seek consensus, through dialogue and cooperation, on the regulation of military applications of AI, promote the development of an open, fair and effective governance framework for AI security and minimize risks to ensure that AI technologies remain safe, reliable and controllable and always develop in a way that benefits the progress of human civilization.

China has always engaged in the global governance of military applications of AI in a responsible and constructive manner. We advocate adhering to the concept of a "people-centred approach in military applications of AI" and upholding the vision of common, comprehensive, cooperative and sustainable security, in a bid to build a community with a shared future for humankind. In 2021, China submitted a position paper under the Convention on Certain Conventional Weapons on regulating military applications of AI, which proposed systematic views and recommendations on the responsible development and use of AI in the military domain in terms of strategic security, military policies, law and ethics, technological security, research and development operations, risk management and control, rules-making and international cooperation. In 2023, China proposed the Global AI Governance Initiative, calling on all countries, especially the major Powers, to adopt a prudent and responsible attitude to the research, development and application of AI technologies in the military field. Our specific proposals include the following:

First, a prudent and responsible approach should be taken. While developing their legitimate national defence capabilities, all countries, especially the major Powers, should refrain from seeking absolute military superiority through AI and undermining the legitimate security interests of others. Efforts should be made to avoid misunderstandings and miscalculations and prevent an arms race in this field.

Second, a people-centred approach should be upheld. It is essential to always regard human beings as the ultimate subject of responsibility and to ensure that relevant weapons systems are under human control. Military applications of AI should respect and protect human dignity and human rights and honour the common values of humanity.

Third, the basic principle of "AI for good" should be observed. The application of AI in the military domain should contribute to maintaining peace, comply with international humanitarian law and other applicable international law, and be aimed at reducing collateral casualties.

Fourth, agile governance should be implemented. We should strengthen forward-looking risk assessment and personnel training on AI, take necessary risk mitigation measures and reduce proliferation risks, while not hindering innovation and peaceful uses of technologies.

Fifth, multilateralism should be upheld. We should support the United Nations in fulfilling its due role, welcome the development of inclusive platforms for discussion by all parties, and make efforts to establish governance frameworks based on universal participation and broad consensus.

China believes that the significance of AI in the military domain should be objectively assessed. It is essential to guide the development of military AI in a proper direction while preventing unregulated growth. In the next phase, the international community should collaborate to maximize benefits while minimizing harms. China proposes the following ideas and suggestions:

First, establish clear guidelines. Security and development must be given equal attention. It is imperative to abide by the purposes and principles of the Charter of the United Nations, observe the basic norms governing international relations, and ensure that AI technology will not become a tool for invading other countries and pursuing

hegemony. China is willing to engage in further exchanges with all parties on the concept of a "people-centred approach in military applications of AI" and build consensus continuously.

Second, improve governance measures. In the context of the current state of AI development and application, we should promote the establishment of a testing and assessment system, implement agile governance and carry out tiered and categorized management for rapid and effective response. All countries should, based on their national conditions, establish and improve domestic legal and regulatory systems, refine relevant ethical guidelines, and strengthen education and training, so as to enhance the safety, reliability and controllability of AI technologies.

Third, strengthen international cooperation. All countries should adhere to the principles of openness and inclusiveness, engage in dialogue and exchange to enhance mutual understanding, and strengthen policy coordination and capacity-building cooperation regarding AI governance in order to continuously improve the level of governance.

Egypt

[Original: English]
[11 April 2025]

Pursuant to General Assembly resolution 79/239, the Government of the Arab Republic of Egypt would like to share its views on the opportunities and challenges posed to international peace and security by the application of artificial intelligence (AI) in the military domain.

General Assembly resolution 79/239 represents an important step to foster multilateralism on the topic of AI in the military domain and towards putting it higher on the political agenda. This comes in the wake of the Secretary-General's call to develop norms, rules and principles around the design, development and use of the military applications of AI with the involvement of all relevant stakeholders.

With the understanding that the aforementioned resolution, pursuant to which these views are presented, aims to place specific focus on areas other than lethal autonomous weapons systems, it is key to reiterate Egypt's steadfast position that any meaningful discussion on the subject matter can never overlook the priority of addressing all ethical, legal and security dimensions surrounding lethal autonomous weapons systems, which represent the most pressing threat to the maintenance of international peace and security as far as the military applications of AI are concerned.

Agreeing on a legally binding prohibition of lethal autonomous weapons systems that function without human control or oversight and that cannot be used in compliance with international humanitarian law, as previously suggested by the Secretary-General, is the most effective and realistic course of action. Pursuing a two-tiered approach of prohibition and restriction and/or regulation – comprising the prohibition of weapon systems that function without human control and the regulation of other systems – is essential to establish the necessary universal legal architecture that would provide an enabling environment for maximizing the benefits of the new opportunities offered by AI military applications while tackling the relevant challenges in a realistic, effective and timely manner.

The international policy landscape surrounding AI in the military domain is far from being unified. Egypt follows closely the multiple international initiatives thereon, which demonstrate the increasing awareness of the associated risks. Nevertheless, deliberations during these initiatives have revealed diversions in views, threat perceptions and priorities and we, accordingly, have to caution against the peril

of creating a fragmented policy framework or competing processes, as has been the case with other domains of new and emerging technologies.

There is a clear need for streamlining these initiatives and to bring them under the United Nations umbrella to ensure their inclusivity and effectiveness. The United Nations and its disarmament machinery represent the only effective and inclusive platform to develop the necessary international rules and normative framework, especially as technological developments continue to starkly outpace the necessary regulation at the international level.

Hence, it is imperative to develop a universal, independent, single-track and trusted platform under the auspices of the United Nations to discuss the future governance of AI in the military domain. The envisaged United Nations-led process shall be tailored to avoid certain emerging counterproductive dichotomies. One such dichotomy is that of the legitimate efforts to ensure legal compliance and ethics versus the tendency to further military interests without due regard to humanitarian implications.

It shall be highlighted as well that, while appreciating the discussions conducted within the Group of Governmental Experts on lethal autonomous weapons systems under the umbrella of the Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to Be Excessively Injurious or to Have Indiscriminate Effects, this platform cannot serve as a substitute for the envisaged United Nations process on the applications of AI in the military domain, given that the Group of Governmental Experts is not universal in nature and does not have the mandate to tackle a topic of such versatility and diversity. It is also regrettable that progress within the Group of Governmental Experts remains quite minimal and no tangible results therein have been reached yet.

In conjunction with the opportunities offered by AI technologies, there are a host of risks inherent in the characteristics of such technologies, which can function unpredictably and unexplainably. These risks include disinformation, inadvertent escalation and cyberrisks, as well as misuse and proliferation by non-State actors. The risks can be novel or can make existing ones more complex.

It is widely acknowledged that there is a broad array of possible military applications of AI. However, meaningful efforts to elaborate on their future governance shall establish the right order of priorities in terms of their inherent risk and their impact on peace and security. This aims to ensure focused and structured discussions, while avoiding undue distraction. With that said, Egypt holds a very strong view that the emphasis shall be placed, aside from the issue of lethal autonomous weapons systems, on other autonomous or semi-autonomous system capabilities that enable the use of force and/or lower the threshold for the use of force and, accordingly, may lead to additional arms race dynamics spanning both conventional and non-conventional weapons. The potential for increased autonomy of nuclear weapons and advanced conventional weaponry, such as hypersonic missiles, would create unknown risks and transform the future of conflict in unpredictable ways.

Emphasis shall be placed also on command and control and target selection activities rather than on logistics planning and intelligence, surveillance and reconnaissance, given their less disruptive impacts. Similarly, more focus is to be put on offensive rather than defensive capabilities.

The envisaged deliberations within the desired United Nations-led process shall aim first at reaching a common understanding on the main elements underpinning the development, deployment and use of AI in the military domain. These elements include:

- Full compliance with applicable international law, including the cardinal principles of international humanitarian law, such as necessity, proportionality and distinction, as well as other ethical considerations throughout the life cycle and stages of AI applications in the military domain.
- The centrality of preserving the human element throughout the whole life cycle of AI military applications, including human judgment, intervention, oversight and control as the key enablers to maintain accountability. It is necessary to ensure that all software, algorithms and designs involving the utilization of AI applications in the military domain remain subject to critical human revision and the principle of explainability. While governments claim that human control over AI-enabled systems is maintained from a doctrinal standpoint, some may be more tempted to increasingly make their weapons systems more autonomous to further military interests.
- The balance between mitigating proliferation risks to non-State actors and curbing malicious use versus maintaining the rights of States to acquire AI and dual-use technologies. It is critical to avoid introducing any arbitrary international oversight mechanisms or imposing any type of discriminatory export controls.
- A capacity-building component with the aim of ensuring proper investment in human capital, technology transfer and sharing of knowledge and best practices in a way that preserves the right of developing countries to benefit from the potential benefits of the various AI military applications, and with the aim of bridging the digital divide.
- The boundaries of AI in the military domain and its interplay with other new and emerging technologies. It is pertinent to discuss ways to ensure complementarity with other United Nations-led processes, including the open-ended working group on security of and in the use of information and communications technologies, given, for example, the intersections between AI and cyberoperations. In addition, discussions shall mainly focus on the military domain aside from the wider security domains.

Finally, it is important to ensure inclusivity and equitability in elaborating governance pathways for a responsible, accountable and human-centric AI within the United Nations multi-stakeholder perspectives providing key inputs that feed into policy discussions. However, their participation shall be without prejudice to sovereign prerogative of States in the policymaking process.

El Salvador

[Original: Spanish]
[10 April 2025]

Background

In recent years, the use of artificial intelligence (AI) in the military domain has played a very important role. It has been established in numerous reports that these new technologies are increasingly sophisticated and widespread, which has made it possible to put such computing tools to use in military planning and decision-making processes, including those concerning who or what to attack. This gives rise to many questions regarding the overall impact, legal implications and risks for civilians resulting from the use of these technologies. One example has been the debate during multilateral negotiations regarding the political, legal and humanitarian implications

of such technologies in respect of autonomous weapons systems. The range of military applications of AI, however, is much broader.

For that reason, we need to expand our understanding of the use and applications of AI in the military setting, in particular in relation to the specific tasks of military targeting and use of force.

The issue of responsible use of AI in the military domain has become particularly significant since the discussions held at the first Summit on Responsible Artificial Intelligence in the Military Domain, held in the Kingdom of the Netherlands in February 2023. The issue has also started to gain greater importance in the meetings of the Group of Governmental Experts on lethal autonomous weapons systems, which meets in Geneva.

While it is important to note that, to date, AI applications and uses have been debated primarily in the context of discussions of autonomous weapons systems, the use of AI in military applications is a much broader issue which has taken on new dimensions and applies not only to applications focused on the autonomy of weapons systems but also, in particular, to applications aimed at automating certain military functions.

In broad terms, the discussion of AI is an emerging issue that is still being explored and is evolving very rapidly, to the extent that initiatives are being developed at the national, regional and multilateral levels to address its impact. It is clear that the countries of Latin America and the Caribbean are not at the same level as developed countries in terms of the technology and capacity-building that facilitate the identification and understanding of opportunities and risks stemming from the use of AI. It is therefore important for these countries to develop a national position that equips them to engage actively in the discussions emerging in international forums and settings and, in so doing, to secure cooperation for capacity-building and specific aspects so that they can be at the forefront of the issue and understand the opportunities and potential security-related risks at the national, regional and global levels.

Initiatives in which El Salvador has participated

- El Salvador participated in the Summit on Responsible Artificial Intelligence in the Military Domain that was held in the Kingdom of the Netherlands in 2023 and endorsed the declaration issued at the Summit (February 2023).
- El Salvador participated in the Latin American and the Caribbean Conference on the Social and Humanitarian Impact of Autonomous Weapons, at which the Belén Communiqué was adopted (February 2023).
- El Salvador is part of the “Group of 16” in the context of the discussions of the Group of Governmental Experts on lethal autonomous weapons systems. While this issue has a distinct focus, it is related to the use of AI in the military domain.

National position

- Some AI applications may have certain benefits in the military field, in particular those applications that are not associated with the functions of identifying and recognizing military targets or are not associated explicitly with the use of force, which involves risks to civilians. Such applications concern other, administrative tasks, such as data analysis and automated learning not connected to human interaction in military operations.
- Misuse of those applications, however, could have adverse impacts, in particular in relation to the protection of civilians and civilian infrastructure, which are

categories that receive special protections under the rules of international law, including international humanitarian law and international human rights law.

- It is important to have a risk-based approach through which certain AI functions can be regulated or prohibited, in particular functions in which limitations are placed on significant human control over the use of force and those that reproduce algorithmic biases arising from the use of databases that are unrepresentative or that contain historical data. Such functions pose risks related to human rights and, in the long term, to international security, especially when the power to decide whether a human being lives or dies is left to a machine or when such tools contain highly sophisticated technological elements, such as self-learning, that may have grave humanitarian, social, economic, political or even environmental consequences.
- At present, there is an urgent need to introduce adequate regulation in the field of AI, as regulation is essential for ensuring that such technology is developed in an ethical and safe manner. That will help to protect users and society from potential abuses and risks and will also encourage innovation by providing a clear and safe environment for developers and researchers.
- While the ultimate goal is to develop binding legal instruments, these technologies are progressing at a rate that is outpacing the evolution and development of international law in this field. For that reason, we believe that it would be appropriate to take an approach focused on responsible behaviour, which could then be used to build a foundation of comprehensive legal commitments to better address the issue.
- It is important to consider the challenges that emerging technologies pose with respect to security issues. For example, materials technologies, such as three-dimensional printing, could be used for the manufacture of small arms and light weapons; robotics could be used for the development of robots with autonomous capabilities in the military field; and the dual nature of certain uses and applications of AI means that they might replicate biases in command and control functions in armed conflict, thereby posing elevated risks to civilians.
- Loss or surrogacy of control in the military domain may give rise to unintended risks. AI can be used to increase human capabilities, but a lack of control in the military context may pose other risks that must be fully explored. AI-related support in the military domain should be used to strengthen or inform decision-making in specific contexts but should never replace human decision-making and reasoning.
- The use of AI in the military domain must comply with international law, international human rights law and international humanitarian law and must serve the public good.
- Countries must be able to strengthen their capacities to identify risks arising from the misuse of AI and the associated linkages with international law.
- Other actors involved in the creation and development of this type of technology, such as private enterprises and academia, should be included in multilateral discussions, and international cooperation between stakeholders should be encouraged in order to unlock the benefits offered by peaceful uses of AI to support the development of countries.

Finland

[Original: English]
[11 April 2025]

Finland is pleased to submit its views on General Assembly resolution 79/239 on artificial intelligence (AI) in the military domain and its implications for international security, adopted on 24 December 2024, in which the Assembly requests the Secretary-General to seek the views of Member States on "the opportunities and challenges posed to international peace and security by the application of artificial intelligence in the military domain, with specific focus on areas other than lethal autonomous weapons systems".

The adoption of international principles or regulations on the application of AI in the military domain is fundamental to ensure compliance with international law, to increase security and to reduce potential risks of conflicts. At the same time, it is necessary to enable the development of national defence capabilities that do comply with international law. Finland has committed to developing, deploying and using AI capabilities in the military domain in a responsible manner, in accordance with international law, in particular international humanitarian law, and in a way that does not undermine international peace, security and stability, while pursuing research, development, experimentation and innovation efforts in the area of AI technology.

It has become increasingly important to identify the foreign, security and defence policy implications of disruptive technologies and develop means for addressing them. Finland actively participates in global debates on technology regulation, advocating fundamental and human rights and addressing related risks in the development and application of AI and relevant policies.

In addition to identifying the risks of disruptive technologies, it is also important to recognize the opportunities they offer for security, defence capability development, economic growth, productivity, sustainable development, technological competence and sectoral investments.

Opportunities

Disruptive technologies present significant opportunities for advancing various sectors, driving the clean transition, fostering sustainable economic growth and enhancing efficiency and productivity. They also have the potential to enhance security, education, well-being and health at a global level.

AI and other emerging technologies offer opportunities for advancing defence capabilities while fundamentally shaping the future of battlefields and the means and methods of warfare. Technological advancements enable more efficient information collection and data processing, heightened situational awareness, faster decision-making and more precise and longer-reaching engagement. The importance of remotely operated and autonomous unmanned systems is growing in modern warfare, and such systems will change the future of war, operations and the battlefield. Anticipating advancements in technology, integrating emerging technologies into defence systems and making use of the unexpected will become increasingly important, as the pace of technological development picks up speed in the future. Technological edge can also compensate for numerical inferiority.

Challenges

At the same time, it is important to establish a wide understanding of the security threats, potential for misuse, human rights issues and interdependencies related to the development of disruptive technologies, such as AI. As they are developed, they will

pose new challenges for the defence and security sectors, in particular. The development of AI makes cyberattacks, information influence activities and, one of their instruments, disinformation, more targeted and effective. Furthermore, AI is already being used to influence elections. In such an environment, increased focus must also be placed on keeping confidential information secure.

International law, in particular the Charter of the United Nations, international human rights law and international humanitarian law, fully applies to cyberspace. Respect for and adherence to the framework of responsible State behaviour in cyberspace remain essential to maintaining international peace, security and stability. Technological development raises new issues. These issues are related, for example, to the cyberenvironment, the use of AI, new weapons technologies and the exploitation of critical raw materials. Hybrid influence activities may include practices aimed at hindering the realization of accountability under international law. Finland advocates taking fundamental and human rights and the risks related to them strongly into account when developing and applying AI and drawing up relevant regulations. Establishing national principles, standards and norms, policies and frameworks is important to ensure responsible AI applications in the military domain, in compliance with international law.

Technological development has provided hostile actors with new opportunities to engage in hybrid influence activities exercised below the threshold of open conflict. Hostile cyberoperations have become an established part of power politics and of the range of instruments available for influence activities conducted by State actors. Cyber, hybrid and information operations are also conducted under normal conditions, which may, for their part, obscure the boundaries between war and peace. Despite the increasingly technological nature of warfare, conventional warfare capabilities remain important, in particular in large-scale and long-term conflicts.

Many countries are facing intense information influence activities that also deploy AI. The harmful use of information has become an everyday part of broad-spectrum influencing, and the competition in the information environment has increased.

Developments in infrastructure and technology and the growing number of users offer greater opportunities for hostile actions in the cyber domain. Many countries are constantly facing intelligence-gathering on information networks, cyberespionage and cyberattacks by hostile actors, who also strive to have physical impacts on critical infrastructure. Alongside that of State actors, the role of politically motivated or State-led non-State actors as orchestrators of hostile activities is growing.

France

[Original: French]
[11 April 2025]

I. Artificial intelligence in the military domain and its implications for international peace and security

Opportunities to leverage

Help with planning and decision-making. The French armed forces are harnessing their databases on events related to ammunition and explosives in order to develop tools for predicting potential threats in specific areas.

Support people. The artificial intelligence (AI) system for flight crew training has been established to improve the training of French pilots by analysing data collected from flights and simulations. AI can also help people when faced with large

quantities of data. For example, the Oreille d'or acoustic analysis system processes vast amounts of acoustic data in order to help French operators to focus their attention on value-added signals only.

Counter our vulnerabilities in the area of information and communications technology. AI technologies can be used to support cybersecurity and address the proliferation of false information. The French armed forces rely on deepfakes detection systems.

Promote the implementation of international humanitarian law and the protection of people and property. AI can contribute to the implementation of the cardinal principles of international humanitarian law, such as distinction, proportionality and precaution. AI can also be used to protect people by helping to clear landmines using drones with AI-based sensors.

Enhance arms control. AI can be used to better monitor and detect clandestine launches, changes in weapons production sites and the testing of chemical and biological weapons. AI could also improve the traceability of arms exports to enhance the control of such exports.

Strengthen prevention, peacekeeping and peacebuilding. AI would enable peacekeeping operations to be better adapted and therefore more effective. The Resistance instantaneous translation system put forward by the French armed forces is designed to enable, offline and without a network connection, communication with the local population, thereby combating disinformation.

Risks to mitigate

Technology-specific risks. Machine learning techniques present various bias-related risks stemming from unintended biases, intended biases, biases related to the reconstruction of particularly sensitive data, and results that are opaque or difficult to explain. There is also the issue of the exponential consumption of energy resources.

Increasing risks to international security and stability. In the wrong hands, AI can exacerbate certain risks to international security and stability (escalation scenarios, an arms race, proliferation to non-State actors, extension of information operations and hostile acts in the cyber domain), which will need to be addressed using adapted risk mitigation measures. The risk of a lack of accountability posed by the dependence on technology makes it necessary to ensure human responsibility.

II. Key principles and measures for "responsible artificial intelligence" throughout the life cycle

Develop artificial intelligence that ensures respect for international humanitarian law

Adapt legal reviews. While such reviews are fully applicable to military AI, the precise methods underpinning them will need to be adapted to the specific characteristics of this technology.

Carry out appropriate follow-up reviews. Such reviews must be carried out, as needed, during the various phases of a weapon system's life cycle. They must be performed when a device undergoes innovations or when new components are added that may significantly modify its effects.

Develop reliable and secure artificial intelligence

Evaluate, classify and certify systems. Such systems must be evaluated and classified at the appropriate level (depending on the criticality of the relevant

functions) through risk analysis during the design phase. They must be associated with defined use cases. These verifications should be repeated at intervals commensurate with the issues at stake.

Rely on controlled, sovereign data. Countermeasures and appropriate bans should be implemented to address the risk of data breaches.

Correct and retrain systems. It is important to identify and characterize any errors encountered (during testing or operational use), make operators aware of the need for feedback and continually verify the system's compliance with our international obligations.

Subject artificial intelligence to appropriate human control and a responsible chain of command

Ensure that decisions and actions comply with the law. The operator or military leader must be able to exercise his or her own judgment in verifying whether or not the results obtained comply with the orders given and with relevant legal obligations.

Guarantee human responsibility. Human responsibility in the design, deployment and use of AI technologies is an inalienable principle requiring the formalization of chains of responsibility among those in charge of command, control and execution functions.

Adapt human control. Analysing and characterizing appropriate human control without restricting the capabilities of the AI technology-based system is a complex issue that must take into account various human, technical and contextual factors.

Train military leaders and personnel to use these systems proficiently. A training and practice phase must be introduced prior to use in order to make personnel aware of the relevant benefits and risks.

Develop sustainable artificial intelligence

Protect research. The purpose and scope of research programmes must be open-ended, without overly broad prohibitions being automatically imposed.

Consider the ethical implications of research. France has a standing body, the Defence Ethics Committee, which considers the ethical challenges posed by new technologies in the area of defence.

Develop frugal artificial intelligence. Promoting frugal behaviour means thinking about the use of AI and improving the resilience and sustainability of relevant systems, while controlling costs.

III. A dedicated process to implement global governance aimed at operationalizing the principles of responsible artificial intelligence

A universal and inclusive process. The relevant discussions must include all stakeholders, specifically States – in particular, the active participation of States that develop and use these systems is absolutely essential; therefore, during the decision-making process, the various relevant positions will need to be taken into account and corresponding rules will need to be adopted to ensure consensus – as well as industrial, scientific, academic and civil society actors, in order to guarantee that the discussions are connected to reality and to preserve innovation. The First Committee of the General Assembly may be an appropriate forum.

Streamlined and coherent governance architecture. A single framework should make it possible to streamline efforts to achieve efficiency gains and enhance

the impact of results. It will be essential to ensure complementarity with the discussions of the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons System, which must be able to continue its work after 2026, under a new mandate.

A process with an operational focus, centred on issues specific to the military sector. Governance must be anchored in the legal corpus applicable to armed conflict, and therefore primarily in international humanitarian law. The priority of any international process must be to ensure respect for existing legal norms by discussing the establishment of guiding principles and their means of implementation by States (facilitating the exchange of best practices and fostering international cooperation and assistance using methods adapted to military affairs), while promoting appropriate confidence-building and risk reduction measures.

Germany

[Original: English]
[11 April 2025]

I. Introduction

In recent years there has been an unprecedented evolution in artificial intelligence (AI) technologies, including in the development of applications based on disruptive technologies, such as generative AI. It is indispensable for States to be able to leverage the opportunities arising from these technological developments and ensure that technological progress will not be hampered. At the same time, States need to ensure that AI applications in the military domain will be developed and used responsibly and in full compliance with international law, including international humanitarian law. International exchange is of utmost importance, in order to master this balancing act.

Against this background, Germany contributes actively to international processes on questions related to the responsible use of AI in the military domain. Inter alia, Germany promoted General Assembly resolution 79/239 on artificial intelligence in the military domain and its implications for international peace and security as part of the core group of co-sponsors and fully supports the efforts of the Secretary-General to submit a substantive report on the views of Member States on "the opportunities and challenges posed to international peace and security by the application of artificial intelligence in the military domain".

Germany welcomes the opportunity to examine the views of Member States and other stakeholders in more depth and to share elements of its own considerations in addressing these important questions.

II. Principles and working assumptions

Germany's approach to ensuring the responsible military use of AI builds on the following fundamental principles identified in the framework of various international forums and discussions.

Germany actively contributed to the elaboration of the North Atlantic Treaty Organization (NATO) principles of responsible use in 2021 and remains fully aligned with those important standards: lawfulness in developing and using AI applications; human responsibility in order to ensure accountability in the design and operation of AI in military systems; explainability and traceability of AI applications in the military domain; reliability, safety, security and robustness throughout the entire life

cycle of systems with AI and autonomy; and governability by ensuring appropriate human-machine interaction and bias mitigation.

In addition, Germany endorsed the outcome documents of the two summits on Responsible Artificial Intelligence in the Military Domain held in The Hague in 2023 (Call to Action) and in Seoul in 2024 (Blueprint for Action), as well as the Political Declaration on Responsible Military Use of AI and Autonomy initiated by the United States of America in 2023, and is actively engaged in the implementation of the Declaration.

Furthermore, Germany is also part of the AI Partnership for Defence initiative, in which like-minded nations promote the responsible use of AI, advance shared interests and best practices on AI ethics implementation, establish frameworks to facilitate cooperation and coordinate strategic messaging on AI policy.

In February 2025, Germany endorsed the Paris Declaration on Maintaining Human Control in AI-enabled Weapon Systems, highlighting the importance of safeguarding human control in the application of AI in the military domain.

III. Key aspects concerning the use of artificial intelligence in Germany's Federal Armed Forces

Germany's Federal Armed Forces (Bundeswehr) are examining the possibility of using AI both to fulfil their core mission and to gain superiority in terms of information, decision-making and effectiveness, as well as to optimize administrative and logistical processes and those involved in the predictive maintenance of complex systems. AI is also used to support specialist personnel in the context of civil-military early crisis detection across different remits in the analysis of mass data and to make projections for deployments. AI is an integral part of major defence projects, which are also being implemented in a European context, contributing to maintaining and fostering European technological excellence. In terms of national and technological developments in the international armaments sector, AI serves to ensure the capabilities required for national and allied defence in the future. The development of possibilities to deploy AI, in particular for the protection of national security and for military purposes, is carried out within the remits and responsibilities of the respective ministries and departments. Without prejudice to the foregoing, AI technologies and AI applications of security relevance are embedded in the AI strategy of the German Federal Government.

The Bundeswehr makes the highest ethical demands of and sets the highest legal standards on the use of AI in weapon systems. In particular, the Bundeswehr follows the provisions of international humanitarian law with regard to armed conflicts and the guidelines of the Data Ethics Commission of the Federal Government and NATO, in particular the above-mentioned six principles of responsible use for the military use of AI, for the duration of those systems' life cycles.

IV. Essential considerations

In order to maintain necessary defence and deterrence capabilities, Germany remains determined to seize the opportunities that are related to AI in the military domain and is convinced that technological progress must not be hampered, in particular given the inherent dual-use character of the technologies at stake.

At the same time, Germany will continue to expand the knowledge base by assessing and addressing the risks associated with the use of AI in the military domain, including those related to unintended biases, such as those based on gender. In this context, Germany attaches high importance to the essential role of academia and the valuable contributions by research institutes and think tanks working in this

area. In order to foster relevant research, Germany supports relevant research organizations, including the United Nations Institute for Disarmament Research (UNIDIR), by financially contributing to goal-oriented research projects.

Ensuring the inclusivity of the discussions, both geographically and by taking into account the views not only of Member States, but also of industry, civil society and academia, is of utmost importance for Germany.

In addressing the opportunities and risks associated with AI-based weapons systems, Germany attaches particular importance to the concept of human control and considers the existence of an effective framework of human control a necessary condition for ensuring that all weapons systems are in compliance with international humanitarian law. This implies not only technical control, but also an element of judgment. Germany's concept of a framework of human control encompasses a set of technologically possible steps and actions that set clear boundaries within which the system's algorithm is allowed to operate. International law, and in particular international humanitarian law, is a central element within these boundaries. When it comes to the actual use of AI on the battleground, context is of utmost importance. Germany considers the concept of a framework of human control to be an appropriate way to take this into account adequately.

Specific attention is necessary when the use of AI is related to nuclear weapons, an area in which the scientific and political debate is still in its early stages. The possible use of AI in nuclear weapons' command and control systems might have serious repercussions for strategic stability or nuclear escalation. At the same time, AI might open up new avenues for containing the spread and use of weapons of mass destruction. Germany sought to contribute to these debates by hosting a conference on artificial intelligence and weapons of mass destruction as part of its well established conference series on the theme of "Capturing technology – rethinking arms control" held in Berlin on 28 June 2024.

The Convention on the Prohibition of the Development, Production and Stockpiling of Bacteriological (Biological) and Toxin Weapons and on Their Destruction and the Convention on the Prohibition of the Development, Production, Stockpiling and Use of Chemical Weapons and on Their Destruction prohibit entire categories of weapons of mass destruction. Applications such as (generative) large language models can facilitate the proliferation of dual-use knowledge that might be misused to develop, produce or use biological and chemical weapons. The convergence of AI applications, such as AlphaFold, and synthetic biology can enable malign actors to design novel proteins that, due to changes of the DNA sequence, can escape detection. AI can be used to analyse big data clouds, such as human genome data, and have great benefits for the development of individual medical therapies, but could also be misused to develop biological weapons that target specific ethnic groups.

In close cooperation with our international partners, Germany will therefore continue to identify possible lines of actions for assessing the impact of AI applications on the development and production of prohibited weapons and introduce possible regulations. At the same time, Germany will leverage the benefits of AI for verification, bioforensics and risk reduction.

V. Germany's commitment to international processes

Since its inception, Germany has actively contributed to the Responsible Artificial Intelligence in the Military Domain process and will continue to do so. Germany was among the core group of co-sponsors of General Assembly resolution 79/239 on artificial intelligence in the military domain and its implications for

international peace and security. Germany highly commends the interregional and multi-stakeholder approach of this important initiative and is looking forward to its continuation in Spain in September 2025.

In full complementarity, Germany has contributed to the United States-initiated Political Declaration on Responsible Military Use of AI and Autonomy, including by co-chairing the working group on oversight (jointly with Austria).

Furthermore, Germany is actively engaged in the AI Partnership for Defence and participates in the UNIDIR Expert Network on the Governance of Artificial Intelligence in the Military Domain.

Germany supports the Chair of the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems in Geneva, Ambassador Robert in den Bosch, and remains actively engaged in the process, including by coordinating several Member States' positions in the framework of the so-called two-tier group. In close cooperation with our international partners, Germany will continue to work towards the timely fulfilment of the Group's mandate, preferably by the end of 2025.

In the context of NATO, Germany recognizes the potential of AI for the further development of the armed forces and defence capability of the alliance, as well as the challenges that the use of AI will pose to the interoperability of the armed forces of the nations comprising the alliance. Multinational AI developments and AI standardization aspects within NATO, the European Union and Germany's partner countries must be fully taken into account, in order to ensure the interoperability of the Bundeswehr as a military force in the context of international operations. Consequently, Germany welcomed the fact that that NATO countries agreed on the principles of responsible use in the context of the NATO strategy on AI.

VI. The way ahead

As emerging disruptive technologies will continue to evolve and shape our world, Germany considers inclusive international coordination on the responsible military development and use of AI to be indispensable. Existing international processes provide an excellent framework for addressing the meaningful aspects involved and for taking into account the views of a variety of relevant stakeholders. Germany will continue to contribute actively to these efforts in order to implement and broaden the support for political commitments on the responsible military use of AI, such as the United States-led Political Declaration and the Responsible Artificial Intelligence in the Military Domain process. Germany looks forward to examining the results of the report of the Secretary-General on AI in the military domain. Germany will continue to contribute actively to the process on lethal autonomous weapons systems in the framework of the Group of Governmental Experts in Geneva.

Greece

[Original: English]
[10 April 2025]

The integration of artificial intelligence (AI) into the defence sector has fundamentally influenced the means and methods of conducting military operations. AI military applications have provided significant operational benefits, including improved decision-making speed, enhanced threat detection and prediction, real-time situational awareness and assessment, optimized resource allocation and planning, logistical support, augmented human capabilities in complex tasks and efficient processing of large-scale intelligence data.

However, despite these advancements, it is essential to recognize that technological progress also introduces complex, multidimensional challenges that require careful scrutiny to ensure that they do not undermine peace, security and stability, both regionally and globally.

In this regard, a key area of concern, as far as Greece is concerned, is the use of military systems with machine learning capabilities, which raises several challenges – including transparency and explainability – since complex models may operate as "black boxes" with undefined decision-making processes, in particular given the constantly evolving battlefield environment.

In addition, the potential use of generative AI in military equipment introduces an important layer of complexity and uncertainty, as these systems might autonomously generate novel solutions and adapt to changing battlefield conditions by continuously analysing and learning from new data – capabilities that are of paramount concern to Greece. To address these challenges, it is crucial to impose clear operational boundaries and constraints on the use of such systems, in order to prevent unintended behaviour.

Given the aforementioned context, one of the most alarming challenges regarding the use of AI in military contexts lies in its integration into command and control and decision support systems for the use of nuclear weapons. The prospect of delegating decisions related to nuclear deterrence, or even the initiation of relevant protocols for their use, to AI-enabled systems requires careful consideration to ensure both human oversight and involvement in those decisions and the establishment of appropriate cybersecurity safeguards to prevent unintended escalation.

Equally concerning, within the current challenging geopolitical environment, is the effort by States to maintain military superiority – an effort that could fuel an arms race that is characterized by a lack of transparency and mutual suspicion. This competition can exacerbate geopolitical instability and pose significant challenges to global security, as the balance of power is disrupted and the technological gap between advanced and developing States becomes increasingly pronounced.

Moreover, the increasing development and deployment of AI-enabled capabilities by armed forces has the potential to lower the threshold of armed conflict. The accelerated pace of decision-making and the growing reliance on unmanned systems in operational theatres heighten the risk of unintended escalation, as the human element on the battlefield is increasingly replaced by unmanned systems.

In this context, another parameter that requires appropriate consideration is the proliferation and diversion of AI-enabled capabilities to States that disregard the rules-based international order and to non-State actors, including terrorist organizations. As AI technologies become more accessible, there is a significant risk that such actors could acquire and deploy them to pursue destabilizing objectives, further challenging international security.

Military AI applications also create risks and challenges related to psychological operations and misinformation, as they enable the mass production of false information, deepfakes and falsified data aimed at deceiving the public and destabilizing institutions. Automated accounts (bots) and targeted propaganda algorithms strengthen psychological operations, influencing public opinion and electoral processes and creating social tensions, including by undermining populations' trust in peacekeeping operations through disinformation campaigns. Social biases, such as those related to gender, age, race and disability, also create concerns, and it is essential to implement risk assessments and mitigation measures to prevent unintended bias and discrimination in algorithms.

Furthermore, AI applications in cybersecurity can be used either to protect critical infrastructures or for malicious purposes, such as cyberattacks and data interception. Hybrid threats combining traditional military operations with offensive intelligence tactics require increased vigilance and coordination between State and international actors to avoid escalation and preserve regional and international peace and security.

In the light of the above, Greece strongly supports international efforts to ensure the responsible use of AI in the military domain, as, despite the challenges described above, it can enhance the implementation of international humanitarian law and contribute to the protection of civilians by improving target accuracy, enhancing surveillance and optimizing humanitarian assistance.

In this spirit, on 4 April 2025, together with France and the Republic of Korea, and with the valuable support of Armenia, Italy and the Kingdom of the Netherlands, Greece organized an Arria-formula meeting of the Security Council on the theme "Harnessing safe, inclusive, trustworthy artificial intelligence for the maintenance of international peace and security". This meeting provided valuable insights into the ways in which the United Nations can contribute to the maintenance of international peace and security, especially through regulation, non-proliferation and the prevention of the diversion of AI capabilities in the military domain, enhancement of rule of law, democratic values, social cohesion and economic development.

In addition, as part of its international engagement, Greece has supported the joint statements issued at the two summits on Responsible Artificial Intelligence in the Military Domain, held in The Hague (15 and 16 February 2023) and Seoul (9 and 10 September 2024) on actions for the responsible development and use of AI in the military domain. Greece has also endorsed both the United States of America-led Political Declaration on Responsible Use of AI and Autonomy and the Paris Declaration on Maintaining Human Control in AI-enabled Weapon Systems.

Furthermore, Greece has established a high-level advisory committee[1] on AI to develop a comprehensive national AI strategy, alongside the necessary structures within the Ministry of National Defence to address the technological, legal, ethical and political challenges arising from the applications of AI and autonomy in the military domain.

Last but not least, in order to constructively contribute to the international dialogue on the responsible use of AI in the military domain, Greece is organizing an international conference on the theme "Armed conflicts and crisis management in the era of AI", which will be held in Athens on 22 and 23 May 2025.

India

[Original: English]
[1 April 2025]

Artificial intelligence (AI) is a transformative technology that is substantially affecting every aspect of human life. It is being developed at an unprecedented scale and speed and being adopted and deployed rapidly for a range of applications. AI can have transformational effects on reducing poverty and improving the lives of people. This is particularly relevant in the case of developing countries like India.

[1] The Committee's landmark study entitled "A blueprint for Greece's AI transformation" provides guiding principles and flagship projects to drive artificial intelligence advancements in Greece, with priorities including the safeguarding and enhancement of democracy, climate mitigation and adaptation and support for security.

There is a need for collective global efforts to establish governance and standards for AI that uphold our shared values, address risks and build trust. AI governance and standards should: take into account deep interdependence across borders; promote innovation; be deployed for the global good; and promote access and equity to ensure that the benefits of AI are available to all, especially countries in the global South. India is committed to open discussions about innovation and governance.

Discussions on military AI need to be anchored in military reality, where there is rapid integration of AI into military doctrines and operations. Ongoing conflicts around the world have demonstrated both the risks and opportunities flowing from the growing adoption of these technologies.

The development, deployment and use of AI in the military domain poses ethical, legal and security challenges. Without downplaying these challenges, India supports the view that has been expressed about the potential of AI to improve compliance with international humanitarian law.

India supports collective global efforts to appropriately regulate the development, deployment and use of AI in the military domain. These efforts should address legal and ethical concerns and enable the identification and mitigation of risks associated with AI in the military domain.

Any collective efforts to appropriately regulate AI in the military domain should be focused on applications and use, and not on the technology and its constituent components. Stigmatization of technology should be avoided. Access to technologies for developmental uses must not be restricted.

AI should be used lawfully in the military domain, in accordance with the inherent right of individual or collective self-defence under international law. International humanitarian law continues to apply fully to AI in the military domain. The cardinal principles contained in international humanitarian law, namely distinction, proportionality and precaution, apply to all means and methods of warfare in the past, present and future.

Human judgment and oversight in the use of AI in the military domain is essential to mitigate risks and ensure compliance with international humanitarian law.

Any collective effort or appropriate regulations with regard to AI in the military domain should take into account existing legal obligations and respect national jurisdiction and competence, as well as relevant national capacities.

India is committed to the responsible use of AI in the military domain.

India is developing a framework for evaluating trustworthy AI in the defence sector to address the complex challenges posed by modern AI technologies. The framework is centred on five key principles: (a) reliability and robustness; (b) safety and security; (c) transparency; (d) fairness; and (e) privacy. These principles provide a foundation for further discussions on appropriately regulating the development, deployment and use of AI in the military domain.

Indonesia

[Original: English]
[11 April 2025]

Indonesia welcomes the discussion on opportunities and challenges posed to international peace and security by the application of artificial intelligence (AI) in the

military domain, with a specific focus on areas other than lethal autonomous weapons, pursuant to paragraphs 7 and 8 of General Assembly resolution 79/239.

As AI in the military domain includes a wide range of systems and applications to devise an inclusive multilateral deliberation on the subject at the United Nations, such discussion should extend beyond kinetic capabilities (such as lethal autonomous weapons systems) to non-kinetic ones, which can be adversarial (e.g. autonomous cyberwarfare systems, adaptive radar-jamming or electronic warfare capabilities) or supportive (e.g. logistics, medevac or tactical surveillance) in military roles. It should also cover other capabilities that may have a direct effect on strategic balance, such as improved sensing (e.g. satellite or anti-submarine), intelligence or war planning.

Indonesia remains firmly committed to the maintenance of international peace and security, as enshrined in the preamble of the Indonesian Constitution. Guided by this commitment, Indonesia believes that the use of AI in the military domain must be governed in a manner that promotes peace, security and sustainable development goals. AI must be a force for peace and security, not a driver of insecurity, conflict or strategic rivalry.

While AI itself is not a weapon, Indonesia recognizes that it serves as both a force multiplier and a threat amplifier, capable of generating significant benefits and serious risks for international peace and security. The use of AI in the military domain brings up various ethical, legal, moral and technical questions, which should be carefully considered and deliberated in relation to compliance with international law, including international humanitarian law and international human rights law.

On the one hand, AI is believed to offer a wide range of potentials. It can augment data processing; increase operational efficiency, precision and accuracy; and potentially improve compliance with international humanitarian law, such as by supporting proportionality assessment and precautionary measures to reduce harm to civilians. AI can also enhance intelligence, surveillance and reconnaissance capabilities, support logistics and planning, and improve personnel management.

On the other hand, AI raises a range of risks and consequences, including the potential to fuel arms races, proliferate to non-State actors, enable criminal and irresponsible misuse, exacerbate imbalance in military power through technological superiority and increase instability, miscalculation, escalation and legal ambiguity. Technical risks also include cybervulnerabilities, system malfunctions, data bias, target misidentification and other operational uncertainties.

Indonesia is particularly concerned about the existential risks arising from the potential integration of AI into nuclear command, control and communication systems. Indonesia reaffirms its principled position that the use and threat of use of nuclear weapons violates international law and that we need urgent and decisive actions to uphold and strengthen the norms against nuclear weapons. The introduction of AI into nuclear weapons systems exacerbates the existential risks of nuclear weapons use, be it intentional, inadvertent or accidental, and increases nuclear dangers. This is a threat to the security of all nations. Indonesia urges all nuclear-armed States to reassess their dependence on nuclear weapons and reaffirm our collective commitment to a world free of nuclear weapons. Pending the total elimination of nuclear weapons, nuclear-armed States must maintain meaningful human control, responsibility and accountability over nuclear weapons and their delivery systems in the context of the development of AI.

Given these considerations, Indonesia urges a precautionary approach in addressing the challenges of the use of AI in the military domain. Indonesia emphasizes that the development, application and use of AI in the military domain must be governed to harness its benefits and mitigate its risks. Such governance must

serve the collective peace, security and prosperity of all nations. Accordingly, Indonesia puts forward the key points below.

Firstly, Indonesia affirms that international law must be upheld throughout the life cycle of AI technologies. This includes the Charter of the United Nations, international humanitarian law, international human rights law and disarmament and non-proliferation treaties. States should conduct legal reviews at all stages, from procurement to evaluation. States must guarantee accountability for their development and application of AI in the military domain, including the legality of AI applications in the conduct of warfare or hostilities. In the absence of such laws regulating the use of AI in the military domain, it is important to underscore that the usage shall be governed by the laws of humanity and the dictates of public conscience.

Beyond international law, ethical considerations should complement the legal frameworks in guiding the governance of the use of AI in the military domain. Principles such as traceability, accountability, responsibility, explainability, humanity, transparency, equity and fairness must be promoted in the development and application of AI.

Secondly, Indonesia stresses the essential role of the human element in ensuring accountability and responsibility at all levels, be it at the State, corporate or individual level, in the design, development, deployment and use of AI in the military domain.

The development, application and use of AI in the military domain must remain human-centred and be governed to serve the interests of humanity. Effective and meaningful human control must be preserved and strengthened through training, particularly in decisions involving the use of force. Critical decisions must involve human judgment, intervention, oversight and control. Further, while “meaningful human control” has been increasingly accepted in governing the use of AI in the military domain, Indonesia is of the view that this concept has yet to satisfy the legal, moral, technical and regulatory questions associated with such use. There needs to be an agreement on what “meaningful” human control entails in practice.

While AI governance will primarily regulate State conduct, it must also address civilian stakeholders, particularly technology companies involved in the use of AI in the military domain. States must ensure that the private sector is compliant with international law and ethical standards while still supporting the growth of the AI innovation ecosystem. Researchers and companies bear responsibility for ensuring that their AI technologies are reliable, safe, secure, accountable and under accountable human control. They should also be responsible for monitoring, communicating and addressing the risks entailed in their product.

Thirdly, Indonesia underscores the urgent need for multilateral, inclusive and comprehensive legal and regulatory governance frameworks. These must reflect the interests of all States, irrespective of their level of AI development. All States must have an equal voice in shaping the rules and norms governing the use of AI in the military domain to ensure fair representation and foster global trust.

Broad stakeholder involvement is critical, given the multifaceted ethical, legal and technical dimensions of AI. Engagement from diverse disciplines and cultures is also necessary to ensure that AI systems align with international law, humanitarian law, human rights and disarmament commitments before their application in the military sphere.

Fourthly, it is critical to remain cognizant and foster meaningful discussion on the risks, challenges and implications stemming from the development, deployment and use of AI in the military, be it technological or non-technological. Indonesia highlights the importance of continuously assessing the broader implications of

military AI for international peace and security, particularly in the context of non-proliferation and disarmament. More comprehensive studies are needed to understand these impacts, which remain underexplored.

Identification of risks associated with AI development, deployment and use in the military domain will support evidence-based forecasting, risk assessments and the eventual development of risk mitigation measures.

Enhancing understanding and raising awareness of the risks associated with the use of AI in the military domain is also crucial. In this regard, transparency should be promoted by, among other things, sharing national policy and strategy, especially to identify, evaluate and mitigate risks; sharing AI capabilities in the military domain, where appropriate, to increase accountability and confidence-building measures; and sharing lessons learned and best practices across borders, industries and sectors.

Fifthly, AI governance must not hinder technological development or limit access to AI by developing nations. Frameworks should avoid imposing conditionalities or barriers that restrict equitable access. A balanced approach is needed that addresses risks such as proliferation while ensuring AI accessibility to States with limited resources.

Lastly, AI governance must place a strong emphasis on bridging the digital and AI divide. Developing countries face significant constraints, not only in terms of AI capabilities but also in their ability to govern these technologies effectively. If this gap remains unaddressed, global governance efforts will be undermined, as numerous States remain ill-equipped to tackle the complex and cross-border challenges that AI presents.

Indonesia emphasizes the urgent need to address the stark digital and AI divides between and within nations, particularly regarding access to financial, human and technical resources. These divides risk deepening global inequalities and heightening the potential for conflict.

As global public goods, peace and security require international cooperation among all countries, both developed and developing, to address shared challenges and seize collective benefits, including those related to the development, application and use of AI in the military domain. In this context, Indonesia calls for enhanced and balanced international cooperation and assistance to promote global AI capacity and governance frameworks. Such cooperation must be pursued on an equitable and mutually agreed basis, taking into account the specific needs and contexts of developing countries. This includes, but is not limited to, initiatives in capacity-building, education, technology transfer, lifelong learning, technical training, joint research and knowledge-sharing.

Such cooperation must be multi-level, not only among States and international organizations but also across sectors within countries. Public-private partnerships should be encouraged in order to promote responsible innovation and to raise awareness within the industry of the implications that their technologies may have for international peace and security.

International cooperation is critical not only to resolve the digital and AI divides, but also to create an enabling environment for confidence-building between States. It can help reduce geopolitical division and competition in the AI field. International cooperation must be rooted in principles of equality, trust, mutual benefit, respect for sovereignty and solidarity to pave the way to meaningful collaboration, including technological transfer and knowledge-sharing.

Indonesia also recognizes the value of strengthening regional cooperation mechanisms that take into account local and regional specificities. These mechanisms

can serve as foundational building blocks towards a broader global consensus, while also providing space for more granular and context-sensitive deliberations.

Iran (Islamic Republic of)

[Original: English]
[12 March 2025]

In response to the request made to the Secretary-General under paragraph 7 of resolution 79/239, in which the General Assembly sought the views of Member States on the opportunities and challenges posed to international peace and security by the application of artificial intelligence (AI) in the military domain with specific focus on areas other than lethal autonomous weapons, the Islamic Republic of Iran hereby submits its views.

Artificial intelligence is becoming one of the major drivers of change in today's world, leaving an indelible mark on how the military industry will evolve in the near future, thereby affecting international peace and security at its very core. State and non-State actors actively advance their contending AI agendas, which cannot be left unregulated. Considering the leading role of non-State actors, and the need for striking a balance between regulatory and innovative procedures and trends, it is vital that the regulatory authority remain the sovereign prerogative of Member States.

From a substantive point of view, as has been the case for other technologies used in cyberspace and outer space, the Islamic Republic of Iran supports the exclusively peaceful application of AI, bearing in mind that, under the proper circumstances, military entities can also peacefully benefit from AI dividends.

Given the varying levels of development across nations, it is of paramount importance to ensure that the digital divide does not evolve into an AI divide. Inclusivity of all AI-related regulatory procedures may be guaranteed only within the consensus-based framework of the United Nations. This approach safeguards the sovereignty of Member States, fosters an environment of equitable AI development for all and provides innovative flexibility for the AI industry to flourish. The centrality of the United Nations in AI-related regulatory matters impedes exclusivist national approaches to the matter. Inclusivity and a consensus-based approach to this vitally important matter must reign supreme.

Despite ongoing discussions on AI in various international forums, our grasp of the issue and its implications for international peace and security remains incomplete. It is premature to assert the full applicability of international law, humanitarian law and international human rights law to AI. Facing the enormity of this new and fast-evolving phenomenon, the international legal framework might need adaptation and evolution of its own kind.

As regards international regulatory efforts, the Islamic Republic of Iran supports the establishment of legally binding arrangements between Member States as its preferred course of action, as opposed to norm-setting or political instruments.

Within the framework of its principled position on disarmament, the Islamic Republic of Iran rejects any politically motivated, discriminatory or conditional approach, or double standards. Thus, the terminology utilized by the General Assembly must reflect a sense of unity and consensus. In this vein, concepts such as "responsible application" are too abstract to regulate a field defined by concreteness and exactitude. Such an abstract notion would lend itself to misinterpretation and open the door to a politicized approach. The Islamic Republic of Iran expresses its strong objection to the use of such subjective terminology. It supports and proposes

replacing the term “responsible application” with “peaceful application” in any future instrument.

Israel

[Original: English]
[10 April 2025]

Israel notes the adoption of General Assembly resolution 79/239 and, in accordance with paragraph 7 of the resolution, has the honour to submit its national contribution to the report of the Secretary-General to the Assembly at its eightieth session for further discussion by member States.

Israel holds the opinion that the concept of artificial intelligence (AI) is currently subject to a range of possible interpretations that may be refined over time.

It is clear that the use of AI in the military domain is becoming more common and frequent than ever before. Israel voted in favour of the aforementioned General Assembly resolution and encourages States and all stakeholders to engage in a discussion, while maintaining a professional and non-politicized nature, that takes into account the legitimate considerations of all States, including security, humanitarian, economic and developmental considerations.

In order to conduct a serious and responsible discussion on AI in the military domain that may also ultimately have a meaningful effect, we believe that a pragmatic, balanced and incremental approach must be adopted.

As technology brings a wide variety of opportunities to almost every field, including in the military domain, we welcome exploration of the benefits that these developments can bring and ways to materialize them, as well as potential risks and ways to mitigate them. It is Israel’s view that emerging technologies, such as AI technologies, may also serve to advance adherence to existing international humanitarian law. These potential opportunities mandate that such technologies should not be negatively stigmatized.

Israel remains a constructive voice in the global discourse on AI for military use. It has recently endorsed the Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy, led by the United States. We look forward to joining future meetings of this initiative and continue promoting the responsible military use of AI and autonomy.

As part of the Declaration, as well as in other contexts, States have in recent years contemplated guidance on the development and use of artificial intelligence in the military domain, either domestically or internationally. Some of the more basic and commonly shared principles within such guidance, which may also be relevant to discussions in the context of resolution 79/239, seem to be that:

- The military use of AI must be in compliance with applicable international law.
- It should be responsible and enhance international security.
- States shall ensure accountability with regard to the use of AI capabilities in accordance with applicable international law, including by operating them within a responsible chain of human command and control.

Among the practical measures that States should take to effectuate these principles are the following:

- States should take appropriate steps, such as legal reviews, to ensure that their military AI capabilities will be used consistently with their respective obligations under international law, in particular international humanitarian law.
- States should take appropriate measures to ensure the responsible development, deployment and use of military AI capabilities. These measures should be implemented at relevant stages throughout the life cycle of military AI capabilities.
- Relevant personnel should exercise appropriate care in the development, deployment and use of military AI capabilities, including weapon systems incorporating such capabilities.
- Senior officials should effectively and appropriately oversee the development and deployment of military AI capabilities with high-consequence applications, including, but not limited to, weapon systems involving such capabilities.
- States should support appropriate efforts to ensure that military AI capabilities are used responsibly and lawfully, and pursue continued discussions with other States on how military AI capabilities are deployed and used as such.

Israel sees value in inclusive multilateral discussions on AI in the military domain, and its implications for international security, that would strike the right balance between military necessity and humanitarian considerations.

Italy

[Original: English]
[11 April 2025]

Italian presidency of the Group of Seven

Artificial intelligence (AI) was placed at the heart of political and technical discussions throughout the Italian presidency of the Group of Seven in 2024. The Apulia leaders' summit recognized the impact of AI on the military domain and the need for a framework for its responsible development and use.

From 18 to 20 October 2024, the first-ever Group of Seven ministerial meeting on defence took place in Naples. On that occasion, the ministers of defence of the Group of Seven reaffirmed their determination to address security challenges in a cohesive and concrete manner, at a time in history marked by great instability. Moreover, they stressed the need for a more cooperative approach in defence-related research and development, including in terms of sharing and leveraging expertise and knowledge, while fostering a safe environment to prevent malign access, in order to maintain competitive advantage, including in the field of emerging and disruptive technologies.

Finally, in their statement, the Group of Seven Non-proliferation Directors Group recognized the profound impact of emerging disruptive technologies, such as AI, on arms control, non-proliferation and disarmament, as well as on the future of military operations.

I. Responsible Artificial Intelligence in the Military Domain process

Italy values the Responsible Artificial Intelligence in the Military Domain process that was launched by the Netherlands and the Republic of Korea in 2023 with the aim of providing a platform to discuss key opportunities, challenges and risks associated with military applications of AI. At the second Responsible Artificial Intelligence in the Military Domain summit, held in Seoul in 2024, Italy endorsed the

Blueprint for Action, a document outlining key principles for responsible AI governance, including the importance of compliance with international law, human responsibility and accountability, the reliability and trustworthiness of AI systems, and appropriate human involvement in the development, deployment and use of AI in the military domain.

The States that have endorsed the Blueprint stress the need to prevent AI technologies from being used to contribute to the proliferation of weapons of mass destruction, and the importance of not undermining arms control, disarmament and non-proliferation efforts. Moreover, in order to develop a common understanding of AI technology and its applications in the military domain, the Blueprint calls on States to commit to further discussions and develop effective legal review procedures, trust- and confidence-building measures and appropriate risk reduction measures. In this framework, the exchange of information and good practices, as well as the active involvement of other stakeholders, is crucial to progress in the debate.

II. Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy

Italy also values the Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy. Affirming that the military use of AI could and should be ethical, responsible and enhance international security, the endorsing States recognized that a set of measures should be implemented in the development, deployment and use of military AI capabilities. In particular, States committed to: minimizing unintended bias in military AI capabilities; ensuring that their safety, security and effectiveness were subject to appropriate and rigorous testing; and implementing appropriate safeguards to detect and avoid unintended consequences and to respond effectively in such cases. Moreover, it is important that a responsible human chain of command and control is defined and that military AI capabilities are used in a manner consistent with international obligations.

III. Pact for the Future

In September 2024, world leaders adopted the Pact for the Future, reaffirming their global commitments and enabling States to address new and emerging challenges and opportunities. In action 27, States are encouraged to seize the opportunities associated with emerging technologies, including AI, while at the same time addressing the potential risks posed by their misuse. In particular, Member States will continue to assess such risks in the military applications of AI and the potential opportunities throughout their life cycle, in consultation with relevant stakeholders.

IV. Paris Declaration on Maintaining Human Control in AI-enabled Weapon Systems

Italy has also recently endorsed the Paris Declaration on Maintaining Human Control in AI-enabled Weapon Systems, which was adopted on the margins of the Artificial Intelligence Action Summit held in Paris from 6 to 11 February 2025. Underscoring that responsibility and accountability can never be transferred to machines, the endorsing States committed to a human-centric approach to the development, deployment and use of AI applications in the military domain. They also committed to ensuring that the deployment of AI in the military sector was fully in accordance with international law and international humanitarian law, while fostering research, development and innovation with AI technology.

V. Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems

The rapid advancements in artificial intelligence and machine learning also have significant implications for the role of autonomy in weapons systems. In Italy's view, the Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to Be Excessively Injurious or to Have Indiscriminate Effects, which combines the diplomatic, legal and military expertise of representatives from Governments, international organizations and specialized institutions, is by far the most suitable forum to address current and emerging issues relating to the development and use of weapons systems. Italy actively contributes to the discussions held within the Group of Governmental Experts on Emerging Technology in the Area of Lethal Autonomous Weapons Systems, which was launched under the auspices of the Convention and is committed to advancing discussions on the development of elements of a future instrument in accordance with the mandate agreed at the 2023 Meeting of the High Contacting Parties to the Convention.

In Italy's view, that instrument should set out clear prohibitions and regulations, so as to be eventually adopted as an Additional Protocol to the Convention. In accordance with that approach, lethal autonomous weapons systems that cannot be developed and used in accordance with international humanitarian law would be ipso facto prohibited. On the other hand, systems featuring decision-making autonomy in critical functions, which can be developed and used in full compliance with international humanitarian law, would be regulated. The human element is in fact, in Italy's view, crucial for the entire life cycle of lethal autonomous weapons systems, namely for their design, development, production, deployment and use. An appropriate level of human judgment and control should be retained also to ensure responsibility and accountability under international humanitarian law.

Japan

[Original: English]
[11 April 2025]

In resolution 79/239, the General Assembly requested the Secretary-General to seek the views of Member States and observer States on the opportunities and challenges posed to international peace and security by the application of artificial intelligence (AI) in the military domain, with specific focus on areas other than lethal autonomous weapons systems, and to submit a substantive report summarizing those views and cataloguing existing and emerging normative proposals, with an annex containing those views, to the General Assembly at its eightieth session, for further discussion by States. Japan hereby submits its views on this subject for the purpose of contributing to the preparation of the report and to the furthering of the discussion on this topic.

I. General views

Japan is committed to maintaining and strengthening a free and open international order based on the rule of law so that all people can enjoy peace, stability and prosperity, and to promoting diplomacy to realize a safe and secure world in which human dignity is protected. In line with these goals, Japan has actively engaged in efforts to enhance international peace and security as well as arms control and disarmament.

Japan is of the view that the application of AI in the military domain should be examined in a comprehensive manner, with a sufficient understanding of its risks and benefits, and taking into account both humanitarian considerations and security

perspectives. It is useful to deepen the understanding of the application of AI in the military domain, and to promote realistic and practical efforts toward its responsible use in order to maximize its benefits while reducing its risks.

Regarding the application of AI in the military domain, Japan supports the view that, firstly, existing international law applies to matters governed by it that occur throughout the life cycle of AI; secondly, AI capabilities should be applied in a responsible manner; and, thirdly, humans remain responsible and accountable for their use and effects. Japan also emphasizes the need for enhanced transparency as an important confidence-building measure for maximizing benefits while reducing risks.

II. Japan's views and approach regarding opportunities for and challenges to international peace and security owing to the application of artificial intelligence in the military domain

Opportunities

Views

Rapid advances in science and technology, including AI, are fundamentally changing the paradigm of security. Countries are striving to develop cutting-edge technologies that could dramatically alter the character of warfare and thus prove to become game changers, and it has become extremely difficult in practice to distinguish between technologies for civilian use and those for security purposes. AI holds extraordinary potential to transform every aspect of military affairs, including military operations; command and control; intelligence, surveillance and reconnaissance activities; training; information management; and logistical support. Considering the varied usage of AI in the military domain, the application of AI may bring benefits such as improvement of precision, accuracy and efficiency; enhanced situational awareness and understanding; facilitation of rapid information analysis; reduction of human errors; and labour-saving. Its proper application could contribute to better protection of civilians in conflicts and post-conflict peacebuilding.

Japanese approach towards the utilization of "opportunities"

In the application of AI in the military domain, it is necessary to consider whether such application is effective in overcoming the issues identified by humans, while keeping in mind the functions and limitations of AI. The application of AI in itself should not be the goal, and it should not be considered in isolation from its functions and limitations. Therefore, States should ensure that military AI capabilities have explicit, well-defined uses and that they are designed and engineered to fulfil those intended functions. With this in mind, it is important to foster a common international understanding of AI and its functions and limitations in the military domain, as well as a common understanding of the potential application of AI in the military domain. As for the application of AI by the defence authorities, the Ministry of Defence of Japan published the Ministry of Defence Basic Policy on Promoting the Utilization of AI in July 2024, which set out its current thoughts on the functions and limitations of AI in the military domain and the areas for the application of AI which it prioritizes. In the Basic Policy, in the light of the current capabilities and limitations of AI, the Ministry of Defence has identified the following seven fields in which it focuses on the application of AI:

- Detection and identification of targets
- Intelligence collection and analysis
- Command and control
- Logistics support operations

- Uncrewed assets
- Cybersecurity
- More efficient administrative works

The Basic Policy also indicates that it is necessary to keep in mind that AI is applied to support human decision-making, and that human involvement is essential when applying AI.

Challenges

Views

The application of AI in the military domain can present risks of misuse or malicious use, and of escalation and lowering the threshold of conflict, which may originate from bias, unintended consequences and other factors. In this regard, Japan stresses the need to prevent AI from being used to contribute to the proliferation of weapons of mass destruction by States and non-State actors, and emphasizes that AI should support, not hinder, disarmament, arms control and non-proliferation efforts.

The Japanese approach toward addressing "challenges"

In the light of risks such as bias, misuse and malicious use, the Ministry of Defence of Japan will work to reduce the risks posed by AI, taking as a reference the concepts of human-centric AI, safety, fairness, privacy protection, ensuring security, transparency and accountability, as set out in the AI Guidelines for Business in Japan published in April 2024, while also paying attention to discussions in the international community and with the defence authorities of other countries.

In addition, Japan is closely following the possible impact of emerging technologies such as AI on nuclear disarmament and non-proliferation. In this regard, Japan welcomes the commitment made by the United States, the United Kingdom and France at the 2022 Review Conference of the Parties to the Treaty on the Non-Proliferation of Nuclear Weapons to maintain human control and involvement for all actions critical to informing and executing sovereign decisions concerning nuclear weapons employment and calls on other nuclear-weapon States to follow suit. Furthermore, the International Group of Eminent Persons for a World without Nuclear Weapons, in its recommendation to the 2026 Review Conference, stressed the need to work together to address challenges and opportunities associated with emerging technologies.

III. Views on the future of discussions and international cooperation

A flexible, balanced and realistic approach is necessary for the governance of AI in the military domain in order to keep pace with the rapid development and advancement of technologies. Japan stresses that efforts for responsible AI in the military domain can be taken in parallel with, and do not hamper, efforts for research, development, experimentation and innovation in AI technology.

It should be noted that discussions problematizing specific AI technologies may lead to a hindrance of technological development and innovation in the civilian sector, with the possibility of producing a chilling effect. Furthermore, the application of AI in the military domain should be discussed in an inclusive manner with the involvement of and exchange among stakeholders.

In the light of the above considerations, Japan strongly supports the outcomes of the Responsible Artificial Intelligence in the Military Domain summits and the

Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy and expects more States to join these initiatives.

As for lethal autonomous weapons systems, it should be noted that Japan strongly supports the continuation of the discussions under the Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to Be Excessively Injurious or to Have Indiscriminate Effects and expects that the discussions on the application of AI in the military domain will complement and strengthen discussions in the Group of Governmental Experts on Emerging Technology in the Area of Lethal Autonomous Weapons Systems established under the Convention.

Japan recognizes that transparency in the application of AI in the military domain is important as a confidence-building measure leading to risk reduction, as well as effective collaboration and cooperation among countries. Japan also recognizes the importance of capacity-building to facilitate the responsible approach in the development, deployment and use of AI in the military domain, and commits to strengthening international cooperation on capacity-building aimed at reducing the knowledge gap regarding such an approach. In this regard, methods such as the exchange of good practices and lessons learned will be useful, and Japan will make use of opportunities to exchange views with other countries.

Finally, regarding the application of AI in the military domain, Japan will continue to actively and constructively participate in international discussions with the aim of achieving a common understanding in the international community through balanced discussions that take into account humanitarian considerations and security perspectives.

Lithuania

[Original: English]
[9 April 2025]

Lithuania appreciates the opportunity to make a submission to the report of the Secretary-General in accordance with General Assembly resolution 79/239. Lithuania was pleased to support the resolution, which was adopted by the General Assembly on 24 December 2024.

Lithuania notes that the development and use of artificial intelligence (AI) in the military domain presents both opportunities and challenges for international peace and security. Lithuania places great importance on the development of norms and principles of responsible use, which would allow States to harness the benefits and mitigate the potential risks of AI in the military domain. Lithuania firmly believes that all responsible States have an interest in ensuring the responsible application of AI in the military domain. Lithuania is convinced that addressing the implications of military AI requires global action and a multi-stakeholder approach, involving the public and private sectors, civil society and academia.

Lithuania is highly supportive of the Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy, which Lithuania joined on 13 November 2023. The Political Declaration consists of non-legally binding principles and best practices for ensuring responsible and lawful use of AI in the military context. The Political Declaration takes into account measures such as legal reviews, appropriate oversight, minimizing unintended bias and ensuring that military AI capabilities have explicit, well-defined use cases. Lithuania strongly encourages more States to sign on to the Political Declaration.

Furthermore, Lithuania subscribes to the AI strategy of the North Atlantic Treaty Organization, which was adopted in 2021 and revised in 2024. The strategy sets out six principles of responsible use for AI in the military domain, namely: lawfulness, responsibility and accountability, explainability and traceability, reliability, governability and bias mitigation. These non-legally binding principles, to which Lithuania is committed, are intended to apply across the whole life cycle of an AI application.

Finally, Lithuania is pleased to elaborate on its views on the opportunities and challenges posed to international security by the application of AI in the military domain. Lithuania believes that military AI could and should be used responsibly to, first and foremost, strengthen the State's national security, and to contribute to the implementation of international law, including international humanitarian law and the fulfilment of the State's various obligations concerning the protection of civilians. Apart from strengthening the protection of civilians in armed conflict, responsible AI offers opportunities to improve decision-making, logistics, planning and other efficiency-increasing operations.

As regards the potential risks of AI in the military domain, Lithuania would highlight challenges that include, without being limited to, cybersecurity, unintended bias in military AI capabilities, and the unintended behaviour of AI-enabled systems. Lithuania believes that such potential risks are best addressed by implementing principles of responsible use, as well as capacity-building and the proper training of personnel on the use of AI applications and AI-enabled systems. Lithuania emphasizes that, in order to access the benefits of AI in the military domain and use AI as a critical defence capability, States should avoid placing unnecessary excessive restrictions that hinder AI innovation, especially if irresponsible States refuse to accept any such constraints on military AI.

Mexico

[Original: Spanish]
[10 April 2025]

Artificial intelligence, autonomous weapons systems and the challenge facing the world in terms of their regulation

Mexico submits the present document pursuant to General Assembly resolution 79/239, entitled "Artificial intelligence in the military domain and its implications for international peace and security".

Mexico recognizes that the application of artificial intelligence (AI) in the military domain may offer benefits. However, it also poses significant challenges to international peace and security, which require the urgent and coordinated attention of the international community.

Mexico values the multilateral exchanges carried out within the framework of the United Nations, such as the inaugural Military AI, Peace and Security Dialogue, entitled "Opportunities, Risks and International Peace and Security", coordinated by the Office for Disarmament Affairs, which contribute to a common understanding of emerging risks and shared responsibilities. We agree that the integration of AI into military functions poses fundamental challenges to international peace and security, including unintended escalation of conflict, strategic ambiguity and increasing autonomy in the use of force.

Mexico considers that priority should be given to strengthening international cooperation, promoting transparency, sharing best practice and building capacities

that support a culture of regulatory compliance and respect for international law, as well as advancing the elaboration of regulatory frameworks that ensure that the development and deployment of AI in military contexts is governed by ethical, legal and humanitarian principles, preventing this technology from deepening asymmetries or eroding international stability.

International peace and security

Mexico believes it is imperative to take action to prevent the proliferation and misuse of these technologies, including by non-State actors and outside of clear legal frameworks.

The incorporation of new and emerging technologies in the military domain must not take precedence over international peace and security. Such incorporation must take into consideration human development and social empowerment, particularly for the benefit of developing countries. Peaceful uses and the resolution of disputes, rather than the pursuit of a more efficient military machine, should thus guide the aims of this type of technology.

The increasing sophistication of digital threats and the potential for emerging technologies to be used as means of conducting State-to-State attacks, as well as the difficulty of ensuring the reliability and accuracy of autonomous systems in military contexts, exposure to vulnerabilities throughout the AI life cycle, algorithmic biases, data poisoning and the use of generative models for malicious purposes, underscore the need for proactive risk mitigation.

Scientific and technological advances, especially in AI, autonomous systems and digital and quantum technologies, surpass the current capacity of regulatory frameworks to manage such risks. Mexico therefore reiterates the need to develop comprehensive governance frameworks, foster international cooperation and multilateral dialogue, and prioritize transparency, accountability and meaningful human control throughout the life cycle of these technologies, including rigorous testing and ethical safeguards for their deployment.

In the absence of clear international legal frameworks and the necessary multilateral consensus, the use of the term “responsible” in this context should not be interpreted as a tacit endorsement or acceptance of the use or development of autonomous, AI-enabled military capabilities. The principle of responsibility must necessarily be linked to legality and accountability.

In this respect, Mexico considers it essential to establish governance and regulatory mechanisms that reduce the likelihood that AI and other disruptive technologies are used for hostile purposes, recognizing that risks are not present only during their operational deployment, but rather that they arise from the initial stages of design and development.

Operational context

Mexico notes that, given the different military operational contexts in which this technology could be incorporated, AI may have differentiated impacts.

In the context of armed conflict, there must be assurance that any AI-based technology is used in accordance with international humanitarian law, in particular the principles of distinction, proportionality, precaution and humanity.

In the field of peacekeeping operations and disaster response, AI can make positive contributions to logistical coordination, risk prediction and care for affected populations, provided that the human rights framework is fully respected.

With regard to border security, Mexico recognizes that AI can strengthen monitoring capabilities; it stresses, however, the importance of guaranteeing respect for the dignity of all persons and avoiding automated decisions that perpetuate discriminatory practices.

Lethal autonomous weapons systems

Mexico considers that a key aspect of this discussion is lethal autonomous weapons systems, which are a matter of particular concern as regards international peace and security. In this respect, Mexico stresses that multilateral discussions on the incorporation of new technologies in the military domain should not be fragmented and believes that lethal autonomous weapons systems should be an integral part of these exchanges.

Mexico believes it is urgent that the international community establish clear prohibitions and regulations on lethal autonomous weapons systems, owing to their incompatibility with international humanitarian law and their ethical, legal and security risks.

Mexico has promoted and co-sponsored resolutions 78/241 and 79/62 on lethal autonomous weapons systems in the General Assembly, with a view to establishing a legitimate multilateral arena to address these challenges.

Mexico supports the call of the Secretary-General and the International Committee of the Red Cross to initiate negotiations on a legally binding instrument establishing the necessary prohibitions and regulations on lethal autonomous weapons systems by 2026, as provided in the New Agenda for Peace.

Mexico has manifested its political commitment to this issue through its participation in the San José Conference (2023), its adherence to the Belém Communiqué and its active involvement in the conference entitled “Humanity at the crossroads: autonomous weapons systems and the challenge of regulation” (Vienna, 2024), the outcome report of which was endorsed by Mexico.

Mexico believes that lethal autonomous weapons systems pose a number of risks, including:

- They exclude human judgment from critical decisions on the use of force.
- They replace indispensable contextual assessment in military operations.
- They weaken mechanisms for accountability and attribution of responsibility.

Responsibility for the use of force must never be transferred to machines. Decisions on the deployment, activation or override of armed systems must always remain with human persons, who are subject to legal responsibility.

Mexico reiterates that all military technology, including that based on AI, must respect the international obligations emanating from:

- The Charter of the United Nations
- International humanitarian law
- International human rights law
- International criminal law
- The law of international responsibility

In this regard, Mexico considers it critical to prohibit those weapons systems whose technology:

- Cannot distinguish between military and civilian targets
- Cannot apply the principle of proportionality to collateral damage
- Does not have cancellation mechanisms if an attack is found to be unwarranted
- Causes unnecessary suffering or superfluous injury

Mexico insists on the urgency of initiating negotiations on a legally binding instrument that establishes specific prohibitions and regulations on lethal autonomous weapons systems; guarantees that meaningful human control is maintained over critical activities; and includes effective implementation, monitoring and accountability mechanisms.

Benefits and risks

With respect to the specific uses of AI, Mexico recognizes both benefits and risks in the following areas:

- Command and control: under certain conditions, AI might improve the efficiency of operational decisions, but these must remain under significant human control, especially when they pertain to the use of force. AI has the ability to process and analyse large volumes of data and information, far exceeding human capabilities, making it possible to speed up, facilitate and streamline the prediction of future trends and inform strategic decisions in real time.
- Cyberoperations: AI offers valuable capabilities for predicting and responding to cyberincidents, but also increases the risks of escalating tensions, including automated offensive use without adequate oversight.
- Information management and logistics: massive data processing through AI can facilitate real-time decisions, but it must carried out in line with protocols that ensure ethical, explainable and responsible use of AI.

Notwithstanding the above, Mexico underscores the technological risks associated with the integration of AI in military contexts, given that the evidence suggests that technical failures or unforeseen errors, which can escalate a conflict, persist.

Netherlands (Kingdom of the)

[Original: English]
[7 April 2025]

The Kingdom of the Netherlands welcomes the opportunity to submit its views, in accordance with resolution 79/239, adopted by the General Assembly on 24 December 2024, on the challenges and opportunities posed by artificial intelligence (AI) in the military domain to international peace and security.

The Netherlands recognizes the potential military applications of AI and is committed to the responsible development, deployment and use of AI in the military domain. The fundamental position of the Netherlands is that the application of AI in the military domain must be in accordance with international law, including the Charter of the United Nations, international humanitarian law and international human rights law.

On 15 and 16 February 2023, the Netherlands hosted the first Responsible Artificial Intelligence in the Military Domain summit. Since then, the Responsible Artificial Intelligence in the Military Domain process has provided a multi-

stakeholder platform for representatives of Governments, knowledge institutions, think tanks, industry and civil society organizations to discuss the key opportunities and challenges associated with military applications of AI. Discussion takes place annually at global level and throughout the year during regional Responsible Artificial Intelligence in the Military Domain events hosted, so far, by Singapore, Kenya, Türkiye, Chile and the Netherlands.

At the 2023 summit, the Netherlands and 57 other countries agreed on a joint call to action on the responsible development, deployment and use of AI in the military domain. In 2024, the Netherlands endorsed the Blueprint for Action, which was agreed during the 2024 Responsible Artificial Intelligence in the Military Domain summit, hosted by the Republic of Korea and co-hosted by the Netherlands. In addition, the Netherlands has endorsed the Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy.

During the 2023 summit, the Netherlands launched the Global Commission on Responsible Artificial Intelligence in the Military Domain and tasked it with identifying short- and long-term recommendations for Governments and the wider multi-stakeholder community. The Netherlands is awaiting the publication of the Commission's strategic guidance report in September 2025.

The section below further summarizes the Dutch position and sets out key issues requiring further consideration.

Opportunities for international peace and security

From a military perspective, the primary benefits of AI are speed and scale. AI technology enables much faster processing and analysis of data. AI-driven scenario development and decision support systems also help commanders formulate courses of action. This improves strategic insight and the ability to respond to threats promptly and effectively.

The Netherlands believes that AI can also contribute to international peace and security by providing greater insight, improving connectivity, enhancing protection of civilians and reducing risks during front-line operations:

- AI-driven analysis and decision support systems enhance commanders' information position, with regard to both the situation on the ground and long-term strategic developments. This helps improve insight into civilian population dynamics within conflict zones, climate security challenges, gender-based violence and the behavioural patterns of terrorist organizations. This information can in turn be used to improve risk and conflict management, thereby contributing to international peace and security.
- The Netherlands sees value in the application of AI in the military domain to improve connectivity between defence forces, and between defence forces and other actors such as humanitarian assistance actors, monitoring organizations and local governments. Data can be exchanged among a large number of users, creating "single sources of truth" with "smart" sensors that operate in a secured, networked environment. AI agents can also be used to share data at increasingly high speeds. Improved connectivity through improved higher-speed data-sharing benefits international peace and security by enhancing communication, information-sharing and international cooperation, for example on early warning systems and crisis management.
- The Netherlands attaches great importance to the potential of AI for protecting civilians. AI can recognize patterns and deviations in large volumes of data, which can provide a more comprehensive understanding of the civilian

environment. This increased understanding can reduce the risk of misidentification, collateral damage and civilian casualties. More broadly, AI offers the potential to improve the identification of possible threats to civilians and civilian objects, enabling armed forces to respond quickly and appropriately. AI can also assist in optimizing humanitarian assistance efforts such as providing food, shelter and medical care in areas of conflict. Lastly, AI can improve investigations into civilian casualties by gathering and analysing data and evidence, in order to determine the cause of harm and ensure that those responsible can be held accountable.

- AI reduces risks for front-line military personnel, since AI-driven autonomous systems may replace humans in certain activities in difficult or dangerous terrain. Examples include underwater surveillance and supporting search-and-rescue operations under extreme weather conditions. AI may also help reduce medical and rehabilitation costs by reducing the exposure of military personnel to high-risk environments.

Challenges for international peace and security

The Netherlands identifies various risks to international peace and security arising from the application of AI in the military domain:

- The Netherlands is concerned that AI could be used to amplify, improve and automate cyberattacks and the manipulation of information, both of which undermine international peace and security. With the rise of generative AI, information manipulation and automated cyberattacks are easier to carry out. When deployed in the military domain, they disrupt operational communication lines and complicate decision-making. In the long term, widespread dissemination of disinformation and automated cyberattacks could erode trust in military lines of communication. They could also affect trust between States, thereby potentially damaging fragile relationships, especially between nations that are already on the brink of potential conflict.

- The risks associated with the application of AI in the military domain could lead to systems that potentially violate international law. These inadequacies could occur due to insufficient adaptation to context, data and military jargon, and in turn lead to an oversimplification of military decision-making or disregard for specific operational contexts, for example. States could also potentially violate international legal obligations if an application behaves unpredictably, produces discriminatory outcomes based on irrelevant characteristics, or proposes unlawful courses of action. Due to the increased prevalence of AI, the impact of automation bias, bias in data sets and human decisions based on inadequate AI systems could create significant challenges for assigning responsibility and ensuring accountability and appropriate remediation. Importantly, AI applications cannot be expected to reason or function in the same way humans do.

- The risk of AI-driven escalation poses potential risks to international peace and security. As AI accelerates the “Observe, orient, decide and act” loop by increasing speed and scale capabilities, misperceptions may arise due to discrepancies between military intentions and the analyses produced by AI-driven systems. Therefore, AI could unintentionally contribute to escalation. Because AI systems are capable of identifying possible targets at greater speed and scale than humans, their use may also increase the intensity and lethality of conflicts.

- As a consequence, the creation of robust defensive systems is an increasingly significant challenge. The speed at which new AI applications are emerging makes it difficult to implement strategies and tactics for effectively countering and defending against them in a military context. This specific consequence of the increasing use of AI systems could potentially favour offensive actions, and therefore negatively influence international peace and security.
- As terrorist organizations, organized crime networks and other non-State actors gain access to military AI capabilities, destabilization is a further concern. In this context, the Netherlands is concerned that AI could make the production of chemical, biological, radiological and nuclear weapons more accessible to these actors.

Given the rapid evolution of AI technologies, the Netherlands acknowledges that challenges and opportunities around international peace and security cannot be entirely foreseen at present. Some are entirely new, while others exist already but may be exacerbated by the application of AI. Ongoing international dialogue on this issue is essential in order to ensure responsible application of AI in the military domain by all States.

Responsible application of artificial intelligence in the military domain

In order to ensure that AI is applied responsibly in the military domain, context-appropriate human judgment and control must be retained. Humans must remain responsible and accountable. However, it is important to note the points set out below.

More human control does not ensure more responsible artificial intelligence

The Netherlands believes that there is no one-size-fits-all method to integrate sufficient human judgment and control in AI applications. Human judgment and control range from direct human control to higher levels of automation and autonomy, depending on a number of factors. The required degree of human judgment and control humans should exercise over AI-driven applications and systems must therefore be decided on a case-by-case basis. This is the only way to account for multiple factors such as the operational context, the impact on the technology's ability to operate autonomously in hostile environments, system parameters and human-machine interaction.

Research and development is key for the responsible deployment of artificial intelligence applications in the military domain

The Netherlands believes in the importance of research and development. States must adequately assess whether their AI applications act in the way they are designed to and can be deployed in a specific-use context. This is especially necessary during combat and in other high-stakes environments. Through research and development in a general sense, and through proven and reliable testing, evaluation, verification and validation procedures for specific AI applications, potential issues can be discovered and eliminated or mitigated before deployment. In addition, it is important that military personnel be adequately trained on and familiarized with AI applications before the applications are deployed, to ensure that they understand the applications' capabilities and limitations. This is particularly important given the rapid technological developments around AI applications and the fact that they are becoming less expensive to use.

International governance of military artificial intelligence should be flexible, inclusive and realistic

With regard to international governance around AI in the military domain, the Netherlands recognizes the need for a flexible, balanced and realistic approach. Firstly, governing frameworks need to be flexible in order to keep up with rapid technological and battlefield developments. Secondly, parties need to work towards a shared understanding of AI in the military domain and the opportunities, risks and potential solutions that accompany it. This will require an inclusive global dialogue and the active involvement of all stakeholder groups, including States, knowledge institutions, civil society and industry. Thirdly, States should focus on establishing safeguards for the responsible application of AI in the military domain, for example, with a focus on issues such as ensuring traceability or understandability. Fourthly, international governance of military AI deployment must take account of States' different views on regulation. Within the parameters of existing legal obligations, international governance of AI in the military domain should not hamper States' abilities to innovate.

Discussion on autonomous weapon systems

As AI has a significant potential for operating autonomous weapon systems, there are clear parallels between the broader discussion on its use in the military domain and the discussion about the regulation of autonomous weapon systems. The Netherlands regards the international discussions on these two topics as complementary and mutually beneficial.

New Zealand

[Original: English]
[11 April 2025]

The present national submission from New Zealand responds to the note verbale dated 12 February 2025 from the Office for Disarmament Affairs and should be read alongside New Zealand's response to the Office's note verbale dated 1 February 2024.[1]

Position of New Zealand on artificial intelligence in the military domain

New Zealand recognizes that the potential and existing applications of artificial intelligence (AI) in the military domain will have far-reaching and multifaceted impacts.

As yet, while it is unclear what the nature and extent of many of these impacts will be, AI is already being applied in a wide range of military functions by some military organizations, including for intelligence, planning, logistics, navigation and communication. Although it has certain risks, AI in the military domain can give users significant advantages, including greater speed, efficiency, accuracy and situational awareness. Like other militaries, the New Zealand Defence Force intends to pursue the opportunities presented by AI for improving its operations and maintaining interoperability with its partners.

We reiterate paragraph 1 of General Assembly resolution 79/239, namely, "that international law, including the Charter of the United Nations, international humanitarian law and international human rights law, applies to matters governed by

[1] Available at www.mfat.govt.nz/assets/Peace-Rights-and-Security/Disarmament/New-Zealand-submission-to-the-UN-Secretary-General-on-autonomous-weapon-systems.pdf.

it that occur throughout all stages of the life cycle of artificial intelligence, including systems enabled by artificial intelligence, in the military domain". In addition to binding legal obligations, relevant ethical standards should be taken into account throughout the life cycle of AI in the military domain.

New Zealand recognizes that AI is relevant to the development and use of some weapon systems, for instance in elevating levels of autonomy. New Zealand's position on autonomous weapon systems is detailed in its response to the note verbale of the Office for Disarmament Affairs dated 1 February 2024.

It is conceivable that AI could be applied to the development of weapons of mass destruction. Biological and chemical weapons are clearly prohibited under international law, and New Zealand affirms that the general-purpose criterion in both the Convention on the Prohibition of the Development, Production and Stockpiling of Bacteriological (Biological) and Toxin Weapons and on Their Destruction and the Convention on the Prohibition of the Development, Production, Stockpiling and Use of Chemical Weapons and on Their Destruction would apply were AI to be used to develop such weapons, which means, inter alia, that AI must not be used to this end. In addition, as has been noted by States Parties to the Treaty on the Prohibition of Nuclear Weapons, including New Zealand, it is essential that meaningful human control be maintained over nuclear weapons and their delivery systems, pending their elimination and the achievement of a nuclear-weapon-free world.

Existing and emerging normative proposals

Reaching common understandings and building norms are important aspects of promoting the responsible military use of AI. In 2024, New Zealand joined the United States-led Political Declaration on Responsible Military Use of AI and Autonomy, along with many other countries. The Declaration affirms that "military use of AI can and should be ethical, responsible, and enhance international security". New Zealand has also engaged in the Responsible Artificial Intelligence in the Military Domain summits.

New Zealand sees value in multilateral discussions, including through the United Nations, dedicated to developing and agreeing norms around AI in the military domain. The participation of non-State stakeholders, including civil society, international and regional organizations, and industry in these discussions is important throughout these processes.

Norway

[Original: English]
[11 April 2025]

Norway welcomes the opportunity to submit its views on the opportunities and challenges posed to international peace and security by the application of artificial intelligence (AI) in the military domain, with specific focus on areas other than lethal autonomous weapons systems, pursuant to General Assembly resolution 79/239, entitled "Artificial intelligence in the military domain and its implications for international peace and security".

As recognized in the Secretary-General's policy brief of July 2023 on A New Agenda for Peace, AI is both an enabling and a disruptive technology that is being increasingly employed in a wide array of civilian, military and dual-use applications. The increasing ubiquity of AI, coupled with rapid scalability, lack of transparency and pace of innovation, presents potential risks to international peace and security and poses governance challenges.

As a consistent advocate of international law, multilateralism and responsible innovation in the defence sector, Norway supports efforts to promote common understandings, strengthen governance and develop adequate regulation of AI in the military domain. As a minimum starting point, AI applications in the military domain must be developed, deployed and applied in a responsible manner throughout their entire life cycle and in compliance with applicable international law, in particular international humanitarian law.

Importantly, in resolution 79/239, the General Assembly affirmed the applicability of international law, including the Charter of the United Nations, international humanitarian law and human rights law in the use of AI in the military domain and stressed the importance of responsible, human-centric AI use.

AI as an enabling technology holds extraordinary potential to transform every aspect of military affairs, including procurement, hardware, software, operations, command and control, strategic communications, surveillance, intelligence, training, information management and logistical support. The application of AI in the military domain presents foreseeable and unforeseeable opportunities and risks on both the tactical and strategic level. As a general-purpose technology, AI represents a force multiplier with a capacity to reshape the conduct of warfare. Technological convergence between artificial intelligence, neurotechnology, synthetic biology and quantum computing adds further complexity.

It is foundational that AI is developed, deployed, used and governed responsibly, in line with fundamental ethical principles, in strict compliance with States' obligations under international law, including international humanitarian law and human rights law, and with risk identification and mitigation at the very core.

The Norwegian Strategy for Artificial Intelligence in the Defence Sector (2023) outlines key areas where AI may contribute constructively to areas other than lethal autonomous weapons systems:

- **Enhanced situational awareness and decision support**. The utilization of AI is both a possibility and a necessity in intelligence, surveillance and reconnaissance, as large and increasing volumes of data cannot be analysed manually. AI can be used for filtering out relevant data, for example, by pre-processing data, automatic translation or detection of special objects in images, detecting anomalies and repetitions, as well as cross-checking information to detect attempts at disinformation. Improvements in this area can lead to more effective, precise operations and reduced loss of life.
- **Cyberdefence**. Digitalization and increased dependence on information and communications technology introduce vulnerabilities along with the benefits. The digital space provides threat actors with the opportunity to commit data breaches, engage in espionage and sabotage and conduct influence campaigns. AI can support the defence sector's ability to detect, monitor, report, manage and counter digital threats. The use of AI can, among other things, more quickly provide a more complete picture of goals and complex relationships, collect information from relevant sources and streamline the use of analysis. Knowledge and expertise development relating to how AI can constitute a digital threat are essential to being able to detect and avert digital attacks in the future. AI therefore has to be a central element in the further development of the sector's defence against digital threats, both through existing and future instruments.
- **Logistics**. Successful, effective military operations depend on effective logistics support. By streamlining logistics using systems that adopt AI, better operational capability and greater preparedness can be ensured. Applications of

AI in the civilian logistics sector has already progressed far. Many of these could likely be easily adapted for use in the military sector.

- **Support activities**. Many military support activities could likely be improved and streamlined using AI. These include tasks that support and strengthen operational capability, such as operating and maintaining materiel, procuring, managing and disposing of materiel and buildings, recruiting, training and managing personnel and delivering common services, such as accounting and archiving. AI has the potential to strengthen support activities through improved utilization of data for analyses and decision-making support, automation of tasks and improved ability to handle information and knowledge. This could make it possible to switch to a model of predictive maintenance, improved information flow, introduce new and better support systems for human resources management and improved modelling of cost trends for materiel and buildings. A successful introduction of AI technology in support activities could therefore lead to reduced time consumption and increased efficiency.

Additionally, AI applications in the military domain have the potential to enhance the implementation of international humanitarian law and assist in efforts to protect civilians and civilian objects in armed conflicts. They can be beneficial to peacebuilding and peacekeeping activities, and enhance verification and monitoring capabilities for arms control, disarmament and other compliance regimes.

AI in the military domain also introduces unprecedented challenges. AI has inherent vulnerabilities that can have unintended consequences and lead to the degradation of meaningful human control, responsibility and accountability. The use of deep learning has the potential to make AI models hard to understand, explain and predict. Lack of understanding can, for instance, render conflict escalation dynamics more opaque and unpredictable.

Effective safeguards must be in place to ensure that humans retain meaningful control and oversight over the development, deployment and use of AI. This is particularly important the closer the application is to combat operations and the use of force, e.g. decision support systems. Accountability and responsibility for the use and effects of military AI must always remain with humans.

AI systems may be highly sensitive to the quality and representativeness of training data. Possible biases, dis- and misinformation, or incomplete training data can lead to models that generate inaccurate or discriminatory results. Automation bias can cause overreliance by the human user on the outputs of the system.

Highly automated or autonomous response capabilities in the cyber domain – particularly those without adequate human-in-the-loop mechanisms – may lead to unintended responses and rapid escalation.

Increased reliance on cybertechnology for tasks that previously were performed manually or with basic automation also comes with the risk of malicious exploitation of vulnerabilities in that technology. Increasing reliance on commercial systems raises concerns about dependency on external providers, loss of control over updates and other vulnerabilities related to proprietary systems.

The aforementioned are mere examples of potential risks associated with the application of AI in the military domain. There are also many unknown unknowns. In a military context, these factors can, combined or by themselves, undermine mission outcomes and pose fundamental legal, ethical, humanitarian and military risks.

The Norwegian Strategy for Artificial Intelligence in the Defence Sector (2023) also outlines key principles for the responsible development and use of AI:

- **Lawfulness**. AI applications must be developed and used in accordance with international law, including international humanitarian law and human rights law. In the study, development, acquisition or adoption of a new AI-reliant weapon, means or method of warfare, each State is under an obligation to determine whether its employment would, in some or all circumstances, be prohibited by international human rights law or any other rule of international law applicable to the State.
- **Responsibility and accountability**. Human responsibility and accountability for the use of AI must be ensured. Decision-making authority over the use of an AI system and responsibility for its actual use must be unambiguously determined.
- **Explainability, understandability, traceability**. AI applications must be sufficiently explainable, understandable, transparent and traceable.
- **Training**. AI operators must have the necessary training to understand the behaviour of the AI application, including how to identify abnormal behaviour.
- **Reliability, safety and security**. AI applications should have explicit and well-defined scopes of use. The resilience, reliability and security of AI applications must be subject to testing and verification throughout the entire life cycle within their respective scopes of use. AI applications must have adequate levels of security and be protected against digital threats.
- **Control**. Meaningful human control must be ensured. AI systems must include an interface for human-machine interaction that is adequate for its intended use, which provides the capacity to identify and mitigate unintended consequences, as well as the means to take necessary corrective action if the system operates in an unintended way.

There is need for the international community to deepen the dialogue on the military applications of AI and their implications for peace and security, including on measures to ensure responsible AI in the military domain. Particular attention should be given to systems supporting combat operations, including the use of AI for situational awareness and decision support, where undesired outputs and behaviours in the AI application, and loss of meaningful human control, can have particularly harmful consequences. There is also a need to address AI in hybrid warfare, including but not limited to AI in cyberoperations, AI in electronic warfare and AI in information operations.

Norway is committed to strengthening international cooperation on information-sharing and capacity-building. By developing a shared knowledge base, States would promote common understanding, close gaps, enhance transparency and build trust. To this end, Norway would encourage the development and publication of national strategies and policy documents related to military applications of AI. Attention should be given to risk reduction and confidence-building measures.

The timely development of adequate international AI governance, with flexibility to respond to the rapid technological advancements, can help to prevent technology-driven arms races while ensuring that innovation supports global security.

Pakistan

[Original: English]
[9 April 2025]

The rapid advancement and integration of artificial intelligence (AI) technologies in the military domain are poised to fundamentally transform warfare.

AI is increasingly integrated into military operations through applications in autonomous weapon systems; command and control; decision support systems; intelligence, surveillance and reconnaissance; training; logistics; and cyber/information warfare. While these advancements offer operational efficiencies, they also pose significant risks to international peace and security.

Challenges associated with artificial intelligence in the military domain

Strategic risks: interplay with nuclear weapons

The integration of AI with nuclear weapons systems introduces strategic risks, particularly in nuclear command, control and communications. When AI capabilities are integrated with nuclear force posture and employment policies, they can lead to miscalculations, accidents and catastrophic consequences.

The concept of nuclear deterrence relies heavily on human rationality, perception and political decision-making. The integration of AI potentially removes or significantly reduces these critical human factors, increasing the risk of automated or accidental escalation. Recognizing these profound concerns, some States have publicly committed to retaining meaningful human control over nuclear weapons employment decisions – a principle Pakistan supports and urges all nuclear-weapon States to endorse.

In regions with nuclear weapons, reliance on AI-driven decision support systems and fully autonomous weapons systems in the conventional domain can also lead to escalatory risks. Completely eliminating human control during crises could make it difficult to control the magnitude and duration of conflicts. Automating responses in volatile, high-stakes scenarios, particularly in regions with tense nuclear dynamics, can compound conventional-nuclear entanglement and adversely impact strategic stability.

The use of AI for data assessment and intelligence, surveillance and reconnaissance can engender a false sense of confidence for States considering pre-emptive, destabilizing counterforce strikes or targeting second-strike capabilities, posing serious risks to regional and global stability

Operational risks: loss of human agency

AI-driven autonomy in military operations risks diminishing human oversight, complicating crisis management. As warfare accelerates to “machine speed”, human decision-making becomes severely compressed, reducing opportunities for crisis mitigation and diplomatic intervention.

Humans may overly trust AI-generated recommendations from decision support systems, even if flawed or incomplete, resulting in automation bias. Critical military decisions might become overly reliant on machine outputs, causing commanders to overlook human intuition, context or caution, potentially escalating conflicts unintentionally.

AI-enabled capabilities, driven by the allure of increased operational efficiency and the race for decisive advantage, could result in an increased propensity for use, thus lowering the threshold for armed conflict. In times of crisis, a low threshold for the use of force would be highly destabilizing.

Technical risks

Military applications of AI may entail technical vulnerabilities, including algorithmic bias, data poisoning and susceptibility to cyberattacks. Conflicts could erupt due to the malfunction or manipulation of early warning systems or data

poisoning attacks. AI capabilities often function as “black boxes”, producing decisions lacking transparency or explainability, complicating validation and accountability. Such vulnerabilities can lead to unpredictable outcomes, system failures and significant risks to operational integrity. AI capabilities tested in one environment with specific data sets may not perform reliably in completely different environments with more complicated dynamics.

Normative, legal and ethical risks

The use of AI in the military domain poses ethical, normative and legal challenges, particularly concerning compliance with international humanitarian law. The essence of international humanitarian law relies fundamentally on human judgment, discretion and context-sensitive decision-making – qualities inherently difficult for AI systems to replicate. Delegating critical functions, such as target selection and engagement, including lethal force decisions, to autonomous systems risks violating the core international humanitarian law principles of distinction, proportionality, precautions in attack and military necessity. AI systems that produce unpredictable, unreliable or unexplainable outcomes further complicate adherence to international humanitarian law, potentially leading to unlawful or unintended harm.

Additionally, the absence of direct human decision-making or overreliance on AI-driven decision support systems raises critical questions of accountability and responsibility, making attribution and liability in cases of illicit or wrongful acts extremely challenging. If something goes wrong, commanders might deflect responsibility onto AI, complicating legal accountability and potential war crimes investigations.

Ethical concerns further arise from delegating life-and-death decisions to autonomous systems, potentially diminishing compassion, moral reasoning and human judgment, thus exacerbating the risk of unjustified violence and civilian casualties.

Proliferation and global security risks

The proliferation of military AI technologies presents significant risks to international security. The spread of advanced AI capabilities, particularly autonomous weapons, risks initiating new arms races and destabilizing regional and global security environments. The ease of proliferation and potential acquisition by non-State actors further exacerbate these concerns.

Proposed international response: central role of the United Nations machinery

AI technologies are general-purpose and their peaceful uses are integral to achieving the Sustainable Development Goals. At the same time, the implications of AI in the military domain are cross-cutting and can significantly impact international peace and security, thus necessitating a coordinated international response.

Pakistan acknowledges the value of AI governance initiatives outside the United Nations but remains cognizant of their limitations, particularly regarding universal participation and formal multilateral legitimacy. While these initiatives can complement United Nations efforts by fostering dialogue and political will, pursuing them in isolation risks fragmentation. Therefore, discussions on military applications of AI should be brought within United Nations forums to ensure inclusivity, legitimacy and a coherent global framework reflecting the interests of all States.

For these reasons, the United Nations must remain central to any international response. The United Nations disarmament machinery should play a central role in developing an international governance framework for military AI and preventing the

fragmentation of the normative landscape. The scale and novelty of the military implications of AI require a multifaceted, holistic multilateral response. The universal membership of the United Nations uniquely positions it as the ideal forum where all States – both developed and developing – have a voice.

No single forum or instrument will suffice. A structured strategy utilizing multiple United Nations disarmament bodies is needed, with each forum addressing the issue from its unique angle and mandate, in a complementary manner. We propose leveraging all relevant forums, from the General Assembly and its First Committee to the Disarmament Commission, the Conference on Disarmament and the Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to Be Excessively Injurious or to Have Indiscriminate Effects. Such an approach would comprehensively address strategic, humanitarian, legal and technical dimensions, avoiding gaps and redundancies. Each forum's work should inform the others, creating synergies towards the common goal of mitigating military AI risks while preserving the peaceful use of AI.

Conference on Disarmament

The Conference on Disarmament should prioritize addressing the strategic risks associated with military AI, particularly in the nuclear domain, aligning directly with its agenda items 1 and 2 ("Cessation of the nuclear arms race and nuclear disarmament" and "Prevention of nuclear war, including all related matters"). In 2023, Pakistan proposed establishing a new agenda item in the Conference on Disarmament on this subject (CD/2334).

Under this new agenda item, the Conference on Disarmament should establish a subsidiary body or an ad hoc group specifically mandated to examine stability-related risks of military AI, assess how it contributes to nuclear risks and pursue negotiations on concrete measures. These measures could include:

- Making a commitment to maintaining human control and not replacing human judgment in decisions regarding nuclear weapons employment
- Prohibiting the use of AI capabilities to manipulate data or target nuclear command, control and communications systems
- Developing restraint measures on deployment and use of certain AI capabilities, which can initiate pre-emptive strikes and contribute to escalatory nuclear risks

The Conference on Disarmament is uniquely suited for these discussions, bringing all militarily significant States together on an equal footing and operating by consensus, thereby safeguarding all States' vital security interests. Addressing this issue could revitalize the work of the Conference on Disarmament, demonstrating responsiveness to new and emerging threats.

Disarmament Commission

With its universal membership and deliberative mandate, the Disarmament Commission is ideally positioned to develop practical guidelines and recommendations on the responsible military use of AI. Historically, the Disarmament Commission has effectively developed similar guidelines (e.g. confidence-building measures in 1988 and regional approaches to disarmament in 1993).

Within its Working Group II, the Disarmament Commission could develop guidelines and recommendations on confidence- and security-building measures related to military AI applications at both the global and regional levels. Key elements could include reaffirming normative foundations, recommending operational and technical risk mitigation measures, developing military AI risk reduction strategies

and addressing proliferation concerns while ensuring equitable access to peaceful AI uses.

First Committee of the United Nations General Assembly

The First Committee of the General Assembly should institutionalize regular assessment reports by the Secretary-General of the United Nations and maintain a catalogue of technological development of military AI capabilities and associated risks based on voluntary information shared by the Member States. These periodic assessments would provide authoritative insights into evolving capabilities, offering timely information and facilitating informed international policy responses.

The First Committee, in reviewing such reports, could hold dedicated debates on AI and possibly establish an open-ended working group under the General Assembly, if needed, to negotiate for a more institutional platform, e.g. a United Nations register on military applications of AI (though for now leveraging existing forums remains preferable).

These reports could also identify areas where consensus is emerging or further work is needed, guiding agendas of forums like the Conference on Disarmament, the Disarmament Commission and the Convention on Certain Conventional Weapons.

Convention on Certain Conventional Weapons

The Group of Governmental Experts established under the Convention on Certain Conventional Weapons remains essential for addressing the humanitarian, ethical and legal implications of lethal autonomous weapon systems. Its inclusive nature (engaging civil society and the International Committee of the Red Cross as observers) is an asset.

Pakistan values the work accomplished by the Group of Governmental Experts since 2017, notably the 11 guiding principles established in 2019. However, progress under the Convention has been slow and largely principle-based rather than focused on concrete regulations. Pakistan agrees with assessments that discussions under the Convention have given “insufficient and declining attention” to the security dimensions of AI-enabled weapons, highlighting the need for complementary actions in the Conference on Disarmament and other forums. Nonetheless, on the humanitarian front, the Group of Governmental Experts established under the Convention should continue and intensify its work.

Pakistan advocates concluding negotiations on a legally binding protocol to the Convention prohibiting lethal autonomous weapon systems from operating without human control or incapable of complying with international humanitarian law. The current mandate of the Group of Governmental Experts allows Member States to develop elements of such an instrument for presentation at the seventh Review Conference of the High Contracting Parties to the Convention, potentially initiating formal negotiations thereafter.

Conclusion

Pakistan emphasizes the need for coordinated, inclusive international action to mitigate substantial military AI risks. It envisions a governance approach balancing security and development, ensuring stability while enabling beneficial AI development. Through a structured, multi-forum strategy within the United Nations, the international community can establish robust normative guardrails, uphold international security and preserve equitable, non-discriminatory access to the peaceful uses of AI.

Peru

[Original: Spanish]
[11 April 2025]

In paragraph 7 of its resolution 79/239, adopted on 24 December 2024, with Peru voting in favour, the General Assembly requests the Secretary-General to seek:

> the views of Member States and observer States on the opportunities and challenges posed to international peace and security by the application of artificial intelligence in the military domain, with specific focus on areas other than lethal autonomous weapons systems, and to submit a substantive report summarizing those views and cataloguing existing and emerging normative proposals, with an annex containing these views, to the General Assembly at its eightieth session, for further discussion by States.

In this regard, Peru presents below some aspects of its position with a view to contributing to the preparation of the aforementioned report of the Secretary-General.

I. Significance of artificial intelligence in the military domain

Peru recognizes the rapid and dynamic evolution of emerging technologies in the military field, in particular the potential applications of artificial intelligence (AI). Peru is closely monitoring developments in this field – including the ways in which AI appears to be transforming military operations, from the use of autonomous drones to decision support systems – and considers it critical to advance a sustained, multilateral dialogue aimed at establishing principles that ensure the ethical and responsible use of such tools.

Given that AI can be integrated into both weapons systems and systems to support military operations, Peru considers it essential to address the challenges and concerns raised by AI use from the humanitarian, legal, security, technological and ethical perspectives, including risks linked to algorithmic biases. Such concerns are compounded by the possible impacts the use of this technology could have on international stability and security.

This is all the more worrying given the implications AI use has for nuclear weapons and other weapons of mass destruction. The principle of meaningful human control must therefore be emphasized.

II. Views

Compliance with international law

The development, implementation and use of AI-based technologies in the military domain must comply with international law, including international human rights law and international humanitarian law, as well as with the fundamental principles enshrined in the Charter of the United Nations.

In this respect, any normative activities to regulate AI in the military domain must ensure its responsible and ethical use and also guarantee the non-proliferation of AI-based military technologies and equitable access to knowledge and technological capabilities.

This is to ensure that any use of AI respects human dignity, protects the civilian population and guarantees international stability and peace.

Recognition of benefits and risks

AI offers valuable opportunities for a better understanding of operational situations and thus for improving implementation of international humanitarian law and protection of civilians and civilian objects.

However, its use may entail foreseeable and unforeseeable risks in the military domain, for example, those arising from algorithmic biases, design flaws, misuse or malicious use, among others. Furthermore, AI can have an impact on complex regional and global dynamics given that it influences the risks of escalation, miscalculation, lowering the threshold for conflict and emergence of an arms race.

Responsible development

The use of AI in the military domain should promote peace, the protection of the civilian population and the concept that technological advances should complement rather than replace human capabilities.

Consistent with the principles applicable to autonomous weapons systems, the application of AI in the military domain must ensure that responsibility and accountability can never be transferred to machines. In this regard, Peru emphasizes the need to preserve meaningful human control over all decisions involving the use of force.

All risks and challenges related to this technology must be addressed in a comprehensive manner throughout its life cycle.

Controls and safeguards to prevent misuse of this technology in the military domain can be established without hindering AI-related research, development, experimentation and innovation in other fields.

Implementation and transparency

Defining strategies, principles, standards and norms, as well as national policies and legal frameworks that guarantee the responsible use of AI in the military sphere, is a matter of priority.

Establishing confidence-building and risk reduction measures, as well as mechanisms for the exchange of good practices, in the interest of transparency and cooperation among States, is also important.

Format of discussions

Peru considers it essential to maintain an ongoing dialogue at the global, regional and inter-State levels on the development of measures to ensure responsible AI in the military domain.

It also calls for inclusive participation in this area, whereby the views of States, particularly developing States, as well as the contributions of other stakeholders, such as industry, academia, civil society and regional and international organizations, are considered.

It is important to take into account the fact that different States and regions are at different stages of integrating AI capabilities into the military domain and operate in different security environments.

This underscores the value of promoting capacity-building in developing countries and strengthening international cooperation with a view to reducing existing gaps and enhancing the participation of such countries in discussions on the use of this technology.

Participation of Peru in international discussions

Responsible Artificial Intelligence in the Military Domain Summits

- Peru participated in the 2023 and 2024 Summits and in the related regional workshop
- Endorsement of the final declaration of the 2024 Responsible Artificial Intelligence in the Military Domain Summit ("Blueprint for Action")

Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy:

- Peru participated as an observer in the inaugural plenary meeting on this initiative and subsequently formalized its endorsement

AI Action Summit – Military Talks (Paris, 2025)

- Peru attended, with high-level participation, and signed the Paris Declaration on Maintaining Human Control in AI-Enabled Weapon Systems.

Republic of Korea

[Original: English]
[11 April 2025]

As an enabling technology, artificial intelligence (AI) holds potential to fundamentally transform multiple dimensions of military affairs – from decision-making and intelligence gathering to logistics, surveillance, and command and control systems. With the rapid development of AI, there is a growing interest among States to leverage this technology in the military domain.

AI capabilities and AI-enabled systems, as they become increasingly integrated into military operations, present both opportunities and challenges, particularly for international peace and security. These developments raise important questions from humanitarian, legal, security, technological and ethical perspectives.

For the purpose of the present submission, the views set out below specifically focus on areas other than lethal autonomous weapons systems.

Opportunities of artificial intelligence in the military domain

AI capabilities and the systems integrated with AI, including those used in intelligence, surveillance and reconnaissance and decision support systems, enable increased situational awareness, enhanced precision and accuracy and improved efficiency by processing large-scale data, supporting optimization and generating predictive insights. These capabilities and systems can contribute to maintaining and promoting international peace and security.

1. Enhancing the implementation of international humanitarian law and assisting the protection of civilians and civilian objects in armed conflicts

AI-enabled intelligence, surveillance and reconnaissance and decision support systems can enhance the implementation of the fundamental principles of international humanitarian law – distinction, proportionality and precautions in attack – by enabling more accurate battlefield assessments and improving situational awareness. AI can help to distinguish between combatants and non-combatants, and assess the potential collateral damage, using timely and well-informed information. By improving the battlefield awareness, including the presence of civilians, AI assists

the necessity and appropriateness of taking precautionary measures to protect civilians and civilian infrastructure.

2. Supporting peacekeeping operations

AI can support the monitoring of ceasefire agreements and peace accords. It can also facilitate early warning mechanisms to detect potential violations, strengthening mission effectiveness and safety. The Republic of Korea has launched a smart camp pilot project in the Hanbit unit in the United Nations Mission in South Sudan to enhance the safety, efficiency and operational capabilities of United Nations peacekeeping camps through the application of AI and other emerging technologies.

3. Enhancing verification and monitoring capabilities for arms control and compliance regimes

AI can enhance the capabilities of international verification mechanisms to monitor compliance with arms control and non-proliferation agreements. The International Atomic Energy Agency may leverage AI to increase the efficiency of safeguards processes, in particular for those that involve classifying data, finding patterns and identifying outliers in the data. AI-enabled systems can also help to identify early indicators of chemical or biological weapons use and uncover increasingly sophisticated sanctions evasion tactics, reinforcing international non-proliferation regimes.

In addition to the opportunities outlined above, AI can help to mitigate strategic risks – such as miscalculation, misunderstanding and unintended escalation – by improving the analysis of actors' behaviour and enhancing the capacity to detect and respond proactively. Furthermore, AI capabilities can facilitate the development of capacities to enhance cyberdefence posture, protect critical national infrastructure and combat terrorism, among others.

Challenges of artificial intelligence in the military domain

The military application of AI could give rise to novel challenges or exacerbate existing ones if not developed, deployed and used responsibly.

Challenges may stem from the technical and operational characteristics of AI. For instance, its black box nature makes it difficult to understand how and why specific outputs are generated, resulting in limited explainability and traceability. Design flaws and unintended biases in data, algorithms or system architecture can lead to malfunctions or outputs that deviate from intended objectives. Overreliance on AI systems, such as automation bias, or insufficient training may raise issues related to the lack of appropriate human judgment and involvement. These factors could increase the likelihood of miscalculation, misinterpretation or unintended escalation in conflict, thereby posing a challenge to international peace and security.

The dual-use nature of AI technologies could increase the risk of misuse or abuse by irresponsible actors with malicious intent. For example, in the cyber domain, AI-driven disinformation campaigns and cyberattacks, such as data poisoning and spoofing, may be accelerated. Furthermore, irresponsible actors may exploit AI technologies to facilitate the development of novel chemical or biological weapons, raising proliferation concerns and amplifying risks to international peace and security.

Implementation of responsible artificial intelligence in the military domain

In order to harness the benefits and opportunities of AI while addressing its associated risks and challenges, AI capabilities and the systems enabled by them in

the military domain must be developed, deployed and used responsibly throughout their entire life cycle.

The Republic of Korea is committed to ensuring and promoting responsible application of AI in the military domain. This includes the following key principles and measures:

- AI should be ethical and human-centric.
- AI capabilities in the military domain must be applied in accordance with applicable international law, including international humanitarian law and international human rights law.
- Humans remain responsible and accountable for the use and effects of AI applications in the military domain, and responsibility and accountability can never be transferred to machines.
- The reliability and trustworthiness of AI applications need to be ensured by establishing appropriate safeguards to reduce the risks of malfunctions or unintended consequences, including from data, algorithmic and other biases.
- Appropriate human involvement needs to be maintained in the development, deployment and use of AI in the military domain, including appropriate measures that relate to human judgment and control over the use of force.
- Relevant personnel should be able to adequately understand, explain, trace and trust the outputs produced by AI capabilities in the military domain, including systems enabled by AI. Efforts to improve the explainability and traceability of AI in the military domain need to continue.

The Republic of Korea supports discussions and dialogues on further developing measures to ensure responsible AI in the military domain, including through international normative frameworks; rigorous testing and evaluation protocols; comprehensive verification, validation and accreditation processes; robust national oversight mechanisms; continuous monitoring processes; comprehensive training programmes and exercises; enhanced cybersecurity; and clear accountability frameworks.

Establishing robust control and security measures is crucial to prevent irresponsible actors from acquiring and misusing potentially harmful AI capabilities in the military domain, including systems enabled by AI.

The Republic of Korea encourages the development of effective trust and confidence-building measures and appropriate risk reduction measures, as well as the exchange of information and consultations on good practices and lessons learned among States.

The Republic of Korea stresses the need to prevent AI capabilities from being used to contribute to the proliferation of weapons of mass destruction by State and non-State actors and emphasizes that AI capabilities should not hinder arms control, disarmament and non-proliferation efforts. It is crucial to maintain human control and involvement for all actions critical to informing and executing sovereign decisions concerning nuclear weapons employment, without prejudice to the ultimate goal of a world free of nuclear weapons.

AI capabilities and AI-enabled systems in the military domain should be developed, deployed and used in a way that maintains and does not hamper international peace and security.

Future governance of artificial intelligence in the military domain

In envisioning future governance of AI in the military domain, it is critical to foster a common understanding of AI technology – its capabilities and limitations – and a shared understanding of the possible applications of AI in the military domain, as well as of its implications for international peace and security.

Capacity-building is also important, especially for developing countries, to promote their full participation in governance discussions and to facilitate the responsible approach to, and shared understanding of, the development, deployment and use of AI in the military domain. The exchange of knowledge, good practices and lessons learned can also facilitate a common understanding.

Given the rapid advancement of AI, governance mechanisms should be flexible enough to adapt to its advancement. Also, the Republic of Korea supports a balanced approach that addresses both opportunities and risks. Overly risk-centric or restrictive governance discourses may stifle innovation and obscure the potential of AI to support international peace and security. Future governance should not serve as a barrier to innovation, but rather support it and play a role as an enabler for the responsible application of AI in the military domain.

As the international community is in the early stages of understanding the implications of AI in the military domain for international peace and security and considering the current state of technological and policy development, it would be premature to narrowly define the trajectory of AI governance or to establish legally binding instruments or norms without a common and shared understanding of what constitutes responsible AI in the military domain. The Republic of Korea believes that governance discussions should be realistic and proceed incrementally, guided by continued dialogue.

Recognizing that AI innovation is being driven by the private sector, the Republic of Korea believes that future governance efforts must adopt an open and inclusive approach engaging with multiple stakeholders, including industry, academia, civil society and regional and international organizations.

The Republic of Korea acknowledges national, regional and global efforts to address the opportunities and challenges of AI in the military domain, including the development of relevant national strategies, legislation, principles, norms, policies and measures, and recognizes the importance of promoting dialogue at all levels.

To ensure the responsible application of AI in the military domain, the Republic of Korea newly established the Data Policy Division and the Defence AI Policy team within the Ministry of National Defence in 2022 and 2025, respectively. In 2024, the Ministry launched the Defence Data and AI Committee as the highest-level deliberative and decision-making body.

In order to promote dialogue, the Republic of Korea, together with the Netherlands, Singapore, Kenya and the United Kingdom, hosted the second Responsible Artificial Intelligence in the Military Domain Summit in September 2024 in Seoul. The Responsible Artificial Intelligence in the Military Domain Summits and a series of regional consultations on responsible AI in the military domain in 2024 have served as an incubator for exchanging expertise, promoting inclusive dialogue and fostering mutual understanding. Looking ahead, the third Responsible Artificial Intelligence in the Military Domain Summit, to be held in Spain in September 2025, along with upcoming regional consultations on responsible AI in the military domain in 2025, will continue to guide the international community’s efforts towards the responsible application of AI in the military domain.

The Republic of Korea believes that discussions on the responsible application of AI in the military domain within the United Nations framework, including the First Committee of the General Assembly and the Disarmament Commission, should work in a complementary manner with other relevant initiatives outside the United Nations, including the Responsible Artificial Intelligence in the Military Domain Summits process, the Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy, and the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems. The Republic of Korea holds the view that these initiatives are mutually reinforcing and complementary.

Data governance is also crucial. As data play a central role in training, deploying and evaluating AI systems, relevant stakeholders must engage in further discussion on adequate data governance mechanisms, including clear policies and procedures for data collection, storage, processing, exchange and deletion, as well as data protection.

Russian Federation

[Original: Russian]
[10 April 2025]

The Russian Federation welcomes the adoption of General Assembly resolution 79/239 of 24 December 2024 and, in accordance with paragraph 7 thereof, has the honour to submit its national contribution to the report of the Secretary-General to the General Assembly at its eightieth session for further discussion by Member States.

Introduction

The Russian Federation attaches great importance to matters relating to the application of artificial intelligence (AI) in the military domain. We are interested in further substantive discussion of this issue in specialized international forums.

We consider the Group of Governmental Experts on lethal autonomous weapons systems, established by the High Contracting Parties to the Convention on Certain Conventional Weapons, to be the best forum for such a discussion. It is precisely the Group of Governmental Experts that is called upon to maintain a reasonable balance between humanitarian concerns and the legitimate defence interests of States in relation to such weapons, and to take consensus-based decisions. The Group's consideration of the military applications of artificial intelligence is broad in scope, is not limited to the issue of lethal autonomous weapons systems, and touches upon a number of important aspects (including legal, technical and military) related to the use of the technology for military purposes.

We note the discussion of this topic in the context of the existing regimes in the area of arms control, disarmament and non-proliferation. This paper is focused on analysing the risks and opportunities that artificial intelligence presents in terms of the fulfilment by States Parties of their obligations under relevant international legal instruments.

We welcome the readiness of Member States to begin discussing the topic of military applications of artificial intelligence in the Disarmament Commission as part of the discussion on emerging technologies in the context of international security. The purpose of this exchange of views is to agree upon recommendations on aspects of "military" AI that are not addressed in other forums.

In the course of work in the above-mentioned international forums, special attention should be paid to the development of a common terminology, the application

of existing international law, human control, accountability, and the risks and opportunities posed by the technology.

Definition

There is no consensus definition under existing international law of AI-based weapons systems and military equipment, which makes it difficult to address the issue. The development of a common working understanding of such tools and, in general, of the terminology associated with the application of such technology for military purposes will provide a clearer picture of the subject and of the prospects for the discussion of the topic.

The working definition should:

(a) Include a description of types of AI-based weapons systems and military equipment and the specific key features of their use;

(b) Not be limited to the existing understanding of such tools, but rather account for how such systems might evolve in the future;

(c) Be universally understood by the expert community, including scientists, engineers, technicians, military personnel, lawyers and ethicists;

(d) Not be construed as limiting technological progress or undermining research in the field of peaceful robotics and AI;

(e) Not define AI-based weapons systems and military equipment solely by describing their functions.

Categorizing these tools as either "bad" or "good" should be avoided; in other words, they should not be categorized on the basis of the political preferences of a particular group of States.

Existing highly automated military systems should not be placed in a "special" category requiring urgent restrictions and prohibitions. It is precisely this level of automation that enables such systems to operate effectively in dynamic combat situations and in various environments, and that guarantees an adequate degree of specificity and accuracy, thus ensuring that they conform to the principles and norms of international law, including international humanitarian law.

Artificial intelligence-based weapons systems and military equipment in the context of international law

It is generally accepted that existing international law, including international humanitarian law, applies fully to AI-based weapons systems.

The Russian Federation believes that there are currently no convincing grounds for imposing any new restrictions or prohibitions on AI-based weapons systems, or for updating or adapting international law, including international humanitarian law, to address such tools. The discussions towards agreeing on some kind of "rules of conduct" or norms and principles for "responsible" use of AI-based weapons systems and military equipment are also premature. The concept of "responsible" use of AI promoted by Western countries is based on highly controversial criteria that are not known to international law (including international humanitarian law), raises many questions and does not enjoy consensus support from the international community.

The principles of humanity, the dictates of the public conscience and the human rights component cannot be used as the absolute and sole sufficient condition for imposing restrictions and prohibitions on certain types of weapons and military equipment. Concerns regarding AI-based weapons systems and military equipment

should be addressed through the good-faith implementation of existing international legal norms.

Strict compliance with the norms and principles of international law, including international humanitarian law, in situations of armed conflict remains one of the priorities of the Russian Federation. The Armed Forces of the Russian Federation adhere strictly to the norms of international humanitarian law enshrined in federal and departmental legal instruments. Issues relating to compliance with international humanitarian law, including those connected with the use of new types of weapons, are reflected in regulations and training programmes for all categories of military personnel. In 2022 a concept paper of the Armed Forces of the Russian Federation on the development and use of AI-based weapons systems was adopted.

Russian law takes full account of the guidelines on AI-based weapons systems approved by consensus in 2019 by the States Parties to the Convention on Certain Conventional Weapons. We view the further exchange of information on concrete practical measures to implement these guidelines at the national level as a way of building confidence and enhancing transparency.

Control of artificial intelligence-based weapons systems and military equipment

We consider an important limitation to be that humans should have control over the operation of AI-based weapons systems and military equipment. The control loop for such tools should therefore allow for a human operator or an upper-level control system to intervene to change the operating mode of such systems, including to partially or completely deactivate them.

The Russian Federation believes that humans always remain responsible for decisions to use force. The control exercised is based on all information available at the time the decision is made. However, the specific forms and methods of human control should be left to the discretion of States, and direct control need not be the only option.

Control over such systems and equipment can be exercised by:

(a) Increasing their reliability and fault tolerance;

(b) Limiting the types of targets;

(c) Limiting the time frame of their operation, their geographical coverage and the scale of their use;

(d) Making prompt interventions and deactivating them;

(e) Testing them in realistic operational environments;

(f) Allowing people who have successfully mastered the procedures for the use of AI-based tools to operate (control) them;

(g) Monitoring the manufacture of individual elements and the device as a whole;

(h) Monitoring the dismantling and disposal of individual elements and the device as a whole.

We consider it inappropriate to bring into the discussion the concepts of “meaningful human control”, “forms and degrees of human involvement”, “context-appropriate human control and evaluation” and “predictability, reliability, traceability, explainability”, which are promoted by certain States, since such notions generally have no legal bearing and lead only to the politicization of discussions.

Responsibility

The Russian Federation believes that States and individuals (including developers and manufacturers) at any time bear responsibility under international law for their decisions to develop and use AI-based weapons systems and military equipment. Responsibility for the use of such tools lies with the official who assigns them a task and gives the order for their use. To use AI-based weapons systems and military equipment, that official should possess the required knowledge and skills related to their functioning and operation, and should be responsible for taking the decision on the appropriateness of their use and planning the forms and means of their use.

Opportunities and limitations of artificial intelligence-based weapon systems and military equipment

It is commonly known that AI-based weapons systems and military equipment can be more effective than a human operator in performing assigned tasks and can reduce the likelihood of errors. In particular, such tools are capable of significantly reducing the negative impacts – in the context of international law, including international humanitarian law – that are associated with mistakes by operators, their mental or physical state or their moral, religious or ethical beliefs. The use of such tools can ensure greater accuracy in the targeting of weapons against military facilities and help to reduce the risk of unintentional strikes against civilians and civilian objects.

An assessment of the potential risks related to the use of AI-based weapons systems and military equipment and measures to mitigate them should be part of the process of designing, developing, testing and deploying new technologies in any kind of military system.

The risks associated with such tools could be minimized by:

(a) Ensuring effective life cycle management;

(b) Conducting comprehensive tests at all stages of the life cycle, including in near-real-life environments;

(c) Increasing their reliability and fault tolerance;

(d) Setting readiness criteria;

(e) Ensuring maximum protection against unauthorized access;

(f) Training operators;

(g) Prioritizing the use of AI technologies in the gathering and processing of information to support military decision-making;

(h) Facilitating continuous monitoring of the operations of such systems by the operator and enabling the emergency termination of a combat mission at the operator’s command;

(i) Preventing such tools from falling into the hands of non-State actors, who could use them for illegal purposes.

These measures may be taken at all stages of the life cycle (development, production, operation, disposal) of weapons and military and special equipment.

Next steps

We believe it would be useful for States to continue the consideration of issues related to the application of AI for military purposes in the Group of Governmental

Experts on lethal autonomous weapons systems, as the best international forum for such a discussion in the context of the existing regimes of arms control, disarmament and non-proliferation, as well as in the Disarmament Commission. At the same time, the discussion in one forum should not duplicate the exchange of views that is already taking place in parallel forums.

We oppose the fragmentation of efforts in this area. It seems counterproductive to transfer the issue of the use of AI for military purposes to any other international platforms, to establish additional forums for its consideration or to discuss it in a narrow forum without the participation of the overwhelming majority of States Members of the United Nations (including the main developers of AI-based weapons systems, including the Russian Federation).

In particular, the discussions on this topic in the context of the non-inclusive "summits on responsible use of AI for military purposes" organized by a group of Western States, and summits on AI in general, are not constructive. These events and their outcome documents do not take into account the views of all stakeholders and cannot be considered a basis for further work that would reflect a common understanding of the subject. They are divisive and are not conducive to the pooling of efforts in this area.

Attempts to "consolidate" unilateral approaches to these issues in alternative forums, including such "summits", bypassing the specialized multilateral forums, will have extremely negative consequences. They have the potential to seriously undermine ongoing constructive and inclusive work on "military" AI and fragment efforts to develop common understandings and recommendations in this area.

In the course of the discussion in the above-mentioned international forums, we believe it is necessary to focus mainly on agreeing upon common specialized terminology and approaches with regard to the application of existing international law, including international humanitarian law, to AI-based weapons systems and military equipment, on ensuring human control over such tools and on the risks and opportunities created by this technology.

The Russian Federation requests the Secretary-General to take into account the above proposals in his substantive report pursuant to paragraph 7 of General Assembly resolution 79/239 and to include the present document in the annex to that report.

Serbia

[Original: English]
[4 April 2025]

The development and application of artificial intelligence is an important factor of change in the way military operations are conducted in today's world. It provides for new possibilities, bringing about, at the same time, new challenges for international stability and peace and security in the military domain. It is, therefore, necessary to initiate the creation of an appropriate international framework to regulate its application.

1. Possibilities and advantages of the use of artificial intelligence in the military domain

The application of artificial intelligence in the non-lethal military context may improve many areas of military operations:

(a) Raise the level of operational awareness;

(b) Improve the process of decision-making with respect to quality and speed;

(c) Upgrade the quality of intelligence data and reconnaissance by rapid data processing and provide for a quick detection of threats;

(d) Support protection of civilians and non-combatants in military conflicts;

(e) Support peace operations and missions by monitoring ceasefires and predicting conflict dynamics;

(f) Improve processes and procedures of predictive maintenance and logistics optimization by reducing costs and saving resources.

2. Key challenges and threats from the use of artificial intelligence in the military domain

The development and integration of artificial intelligence in combat and non-combat systems pose a critical challenge to international peace and stability, as well as for international humanitarian law, primarily in the following areas:

(a) Technical risks and function failures due to application errors in a dynamic environment, which can threaten human life, cause material damage and affect the implementation of international humanitarian law;

(b) Legal and ethical risks regarding compliance with international law, particularly with respect to the implementation of its principles, such as distinction, proportionality and precautionary measures in targeting;

(c) Lack of explicit rules to establish responsibility for acts and activities operated by artificial intelligence;

(d) Algorithms' imperfection may, presumably, lead to bias, mistakes in the process of decision-making and discrimination since the application of non-representative data groups may lead to erroneous identification of civilians or a threat to ethnic or national groups;

(e) The application of the algorithms of artificial intelligence may create false impressions of reduced responsibility of individuals included in the process of conducting operations;

(f) Strategic risks in making decisions via artificial intelligence, based on faulty premises;

(g) Non-selective convergence and integration with new technologies, particularly in the areas of information and cyberoperations or the application of nuclear, chemical and biological means;

(h) Lack of professional staff for the development, organization and responsible application of artificial intelligence systems in conflicts;

(i) Misusing artificial intelligence in information operations by creating and distributing disinformation, which can instigate conflicts and worsen tensions.

3. Creating a legal and ethical framework

Bearing in mind the assessed risks and challenges, it is necessary to create, within the international community, mandatory legal and ethical frameworks to:

(a) Promote and work on starting a dialogue within the United Nations with the aim of increasing the compliance with the norms of international humanitarian law, including the establishment of international legal norms, rules and principles that would ensure that the development and application of artificial intelligence systems

are in accordance with the principles of international humanitarian law (distinction, proportionality and precautionary measures to protect individuals who are not participating in armed conflicts);

(b) Initiate a legality assessment process in the application of systems and weapons vis-à-vis the approved applicability of artificial intelligence;

(c) Ensure the protection of the life and freedom of individuals during armed conflicts and their privacy in peacetime, in particular in the context of monitoring;

(d) Strengthen United Nations mechanisms by introducing the mandatory consideration of the risk of the application of artificial intelligence for military purposes, upgrading the Conference on Disarmament, harmonizing the work of the Disarmament Commission, establishing new specialized bodies of the United Nations and expanding the existing United Nations initiatives for the responsible use of artificial intelligence;

(e) Start a United Nations dialogue to define the responsible use of artificial intelligence in the military domain and establish security protocols for its application (testing, evaluation, validation and verification);

(f) Develop measures for harmonizing the private sector enlistment with the principles of international humanitarian law during the development, establishment and application of the systems and services of artificial intelligence for military domains;

(g) Expand the existing United Nations institutes and documents on recommendations regarding the ethics of the development and application of artificial intelligence to include specific guidelines for conducting conflicts.

The application of artificial intelligence systems in the context of international armed conflicts calls for wide multilateral action by the international community with the aim of promoting responsibility for their use. The United Nations should have the leading role in instigating a dialogue, norms and international community capacity-building in order to prevent fragmentation and bring about proper management.

Singapore

[Original: English]
[11 April 2025]

As a small State, Singapore has always supported the rules-based multilateral system and the role of the United Nations. The United Nations provides the foundation for international law and norms. Multilateral institutions, systems and laws are critical for the survival of all States, in particular small States.

Singapore believes that artificial intelligence (AI) capabilities in the military domain, including systems enabled by AI, should be applied in a responsible manner throughout their entire life cycle and in compliance with applicable international law, in particular, international humanitarian law.

AI has the potential to bring about benefits in the military domain in terms of enhancing precision and situational awareness, and consequently reducing collateral harm to civilians and/or civilian objects. However, AI can also pose risks of conflict escalation and miscalculation in the absence of appropriate governance frameworks. In this regard, Singapore believes that it is important for the international community to engage on this topic.

Singapore's approach to the governance of artificial intelligence in the military domain

One of the key objectives of Singapore's National AI Strategy 2.0 is to foster a trusted environment that protects users and facilitates innovation. To this end, various government sectors, including defence, are developing frameworks for AI governance to allow for the harnessing of the benefits of AI, while ensuring that the potential harm of its use are mitigated.

Through consultations with defence technologists, military planners, international law experts and policy professionals, Singapore developed national principles on AI in the military domain, which were announced in 2021 and address four key areas of concern:

(a) **Responsible**. First, the risk of emergent AI behaviour must be addressed. AI systems must have well-defined intended uses, and both developers and users are responsible for the outcomes of AI systems;

(b) **Reliable**. Second, the risk of errors or inaccuracies in an AI system's output must be addressed. AI systems should be tested and assured to a level appropriate for their intended use. They should be designed to minimize unintended bias and produce consistent output;

(c) **Robust**. Third, the risks from the exploitation of AI by malicious actors must be addressed. AI systems should be designed with cyber and adversarial AI threats in mind. In order to address the "black box effect", their development process should be well-documented to support explainability;

(d) **Safe**. Fourth, we must focus on the risk of AI failure in safety-critical contexts. AI systems should be safe to use, not only in terms of the deployed platforms, but also the surrounding assets and personnel.

These guiding principles have informed Singapore's governance approach to the development, testing, training and deployment of AI-enabled systems for military purposes.

International and regional initiatives on artificial intelligence in the military domain

Singapore has engaged actively in international initiatives on AI governance in the military domain. In 2023, Singapore endorsed the Responsible Artificial Intelligence in the Military Domain Call to Action, and the Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy. In 2024, Singapore co-hosted the Responsible Artificial Intelligence in the Military Domain Summit in Seoul, Republic of Korea, where we endorsed the Responsible Artificial Intelligence in the Military Domain Blueprint for Action.

Singapore also recognizes the importance of regional initiatives to ensure inclusive and context-specific discussions on AI in the military domain. Singapore co-hosted the 2024 regional consultations on responsible AI in the military domain for Asia, which provided a platform for regional countries to exchange views, including on the opportunities and risks posed by AI in the military domain.

In February 2025, Singapore and other Association of Southeast Asian Nations (ASEAN) member States adopted a joint statement on cooperation in the field of AI in the defence sector at the ASEAN Defence Ministers' Meeting Retreat in Penang, Malaysia. In the statement, the ASEAN Defence Ministers committed to promote the accountable and responsible use of AI, to deepen regional understanding and awareness of the implications of AI in the defence sector through information

exchange, and to share best practices and lessons learned among ASEAN member States.

Way ahead for discussions on artificial intelligence and international peace and security at the United Nations

Singapore believes that any further discussions to build upon the international community's support for this resolution should be of an open and inclusive nature. In that regard, we would be favourably disposed to the creation, within the ambit of the United Nations, of an open-ended working group that is centred on AI in the military domain. If such an open-ended working group is created, it should adopt a multi-stakeholder approach involving, among others, technologists, military planners, international legal experts, and policy professionals. We reaffirm our commitment to working with all Member States to advance the responsible application of AI in the military domain.

Spain

[Original: Spanish]
[11 April 2025]

Introduction

Artificial intelligence (AI) entails a revolution in all fields, including security and defence. Its development and use brings with it great advancements and opportunities, while also posing numerous challenges.

The adoption of this technology by the armed forces is not only redefining the way military operations are conducted but is also transforming the global strategic balance.

The development and incorporation of AI by the Ministry of Defence is based on responsible, ethical and legitimate military use that is in line with international humanitarian law and ensures respect for human rights.

AI is changing the traditional concept of military power and security, providing advanced capabilities for data collection and analysis, decision-making and execution of operations in multidomain settings. This involves a paradigm shift in the way States approach defence and security, facilitating faster and more accurate responses to emerging threats.

In the military domain, AI is having a disruptive impact on unpredictable battlefields, resulting in a paradigm shift in the planning and conduct of military operations. AI is also affecting other areas of the military domain: logistics, training, information management and interpretation, intelligence, surveillance, target acquisition and reconnaissance.

It is worth noting that, in line with its commitment to the responsible use of AI, Spain is the host country of the 2025 Responsible Artificial Intelligence in the Military Domain Summit and endorsed the "Call to Action" (The Hague, 2023) and the "Blueprint for Action" presented at the most recent Summit in 2024.

Conceptual and regulatory framework of the Ministry of Defence

The development, deployment and application of AI by the Ministry of Defence are guided by a set of fundamental principles that ensure its safe and ethical use in compliance with national and international regulations. These principles, which are set out in the "Strategy for the development, incorporation and use of artificial

intelligence by the Ministry of Defence" (developed under resolution 11197/2023 of the State Secretariat for Defence) and are in line with the North Atlantic Treaty Organization (NATO) 2021 AI strategy (revised in 2024), are aimed at maximizing the opportunities AI offers in the field of defence while seeking to mitigate the risks associated with its use in the military domain:

- Lawfulness: AI applications should be developed and used in accordance with applicable national and international law, including the Universal Declaration of Human Rights and international humanitarian law.
- Human responsibility and accountability: any development or use of AI should allow for clear human oversight in order to ensure due accountability and the attribution of responsibility.
- Explainability and traceability: AI applications, including the use of auditable methodologies, sources and procedures, should be understandable and transparent for relevant personnel.
- Reliability and transparency: AI applications should be tailored to precise, well-defined and limited use cases, and information should be provided to foster a general understanding of these applications by all stakeholders. The safety, security and robustness of these capabilities should be subject to testing and guarantees under such use cases throughout their life cycle.
- Governability: AI applications should be developed and used in accordance with the intended design functions, and they should include the ability to detect and avoid unintended consequences. Disconnection or deactivation mechanisms should be enabled when unplanned or undesired behaviour is identified.
- Bias mitigation: all necessary measures should be taken to minimize errors and subjective tendencies in the development and use of AI.
- Privacy: the development, implementation and use of AI-based applications must respect the privacy of individuals, from their design and throughout their life cycle.

With respect to the regulatory framework, a set of standards and best practices is being developed on the development, implementation and use of AI in the military domain by the Ministry of Defence to ensure the responsible and efficient use of AI in accordance with national and international legal frameworks, and, in particular, in strict compliance with international humanitarian law and human rights.

Opportunities

The Ministry of Defence focuses its development of AI capabilities on diverse areas to improve the effectiveness of the armed forces. According to the strategy, the use of AI is focused on the areas of operations, intelligence, logistics, cybersecurity and decision support.

AI will help increase the accuracy, speed and effectiveness of decision-making during military operations, subject to international humanitarian law at all times, with the aim of executing missions more efficiently and reducing risks to troops, as well as helping to enhance the protection of civilians and civilian objects in armed conflicts.

Its ability to analyse large volumes of data in real time improves situational awareness and threat response capabilities, enhancing operational security. All of these improvements to capabilities provide for human control at all times and do not delegate responsibility to machines.

With regard to military training and education, within the framework of the European general staff colleges (“C5 Commandants Group”) (Great Britain, France, Germany, Italy and Spain), work is under way to establish a collaborative space relating to AI in military education.

Spain also collaborates with the NATO Data and Artificial Intelligence Review Board on the responsible use of data and AI in the military domain.

In addition, the Ministry of Defence has announced strategic investments in specific regions with the aim of promoting projects related to AI and other advanced technologies. These investments seek not only to strengthen the industry, but also to promote the industrial revitalization of new regional areas.

Challenges

The development and application of AI in the military domain must be aligned with national and international regulatory frameworks, including respect for the implementation of international humanitarian law, thus consolidating efforts to ensure that there is effective human control over critical decisions associated with the use of AI in military operations.

As concerns privacy and data protection, the massive collection and processing of data required to train AI models poses risks in terms of the protection of personal data and information security.

Safety and reliability

The biggest challenge relates to the safe and reliable use of AI. The primary associated risks are the following:

- The training data of AI algorithms can contain biases, which can lead to wrong decisions or unintended consequences.
- Poor training of AI models can lead to interpretation errors, with potentially catastrophic consequences for military operations.
- AI systems can be targeted by cyberattacks, which can manipulate their behaviour or disable them.
- There is a risk of data poisoning, whereby malicious actors alter training data sets to cause algorithm failures.

In Spain, the development of AI in the military domain is governed by the principles of accountability and continuous monitoring, and risk assessment, auditing and traceability mechanisms are implemented in each phase of the system’s life cycle. Any development or use of AI should allow for clear human oversight in order to ensure proper accountability and attribution of responsibility, leaving clear traceability of human actions related to and taken in parallel with AI activities, without delegating final decisions to machines.

AI must also be reliable and predictable, maintaining a level of autonomy that is controlled and supervised by trained operators.

Any AI solution should be assessed in a different setting than the one in which it was trained, and it should be subjected to non-functional testing – load, stress and performance testing under defined, changing scenarios – to study its behaviour and allowed deviations.

In addition, these AI capabilities should be subject to rigorous testing and constant audits throughout their life cycle, enabling early detection of potential errors and improving their operational reliability. Human oversight and control protocols

should be implemented in all phases of deployment, ensuring that critical decisions are not delegated exclusively to AI. In this regard, efforts are under way to ensure that AI developments are certified by recognized entities.

In order to improve the robustness of AI-based systems and protect them against external actions, integrating security from the design stage is key, thus guaranteeing that such systems are resistant to cyberattacks and adversarial manipulation and ensuring the integrity of the data and models used.

AI can be targeted by attacks such as data poisoning or model manipulation attacks; continuous monitoring of system performance and regular validation testing and audits are therefore required. The development of backup and disaster recovery plans should be promoted, guaranteeing the operability of systems in adverse scenarios.

Collaboration with cybersecurity agencies and AI experts should also be fostered, ensuring that the armed forces have the best tools and strategies to protect these systems from external threats and ensure their operational reliability.

The talents and training of personnel in such technologies are critical and constitute one of the four areas of focus of the Ministry of Defence, ensuring that operators understand the scope and limitations of these systems and can intervene in case of deviations in their behaviour. Training and sensitization of personnel in the legal and ethical use of AI is key to mitigating the risks associated with bias, ensuring that the use of AI in the armed forces is objective, reliable and complies with national and international regulations, especially international humanitarian law.

A guide to good practice is currently being developed, which could serve as the basis of a document that includes contributions from all areas of the Ministry of Defence. The good practices for the responsible use of AI in the military domain proposed by NATO have been disseminated. One example is the NATO responsible AI assessment and toolkit, the purpose of which is to operationalize the principles of responsible use of AI adopted by NATO, which include lawfulness, responsibility, traceability, reliability, governability and bias mitigation.

Switzerland

[Original: English]
[11 April 2025]

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 analysing large data sets and enabling faster, more informed decisions. For instance, in surveillance and reconnaissance, AI can analyse 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 the military domain also presents several important concerns and possible risks.

When used responsibly in armed conflict, AI holds the potential to contribute to bolstering compliance with international humanitarian law 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 forms of AI in the military domain 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 on the basis of its training data, contextual understanding and value judgments 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 behaviours and contexts without context-appropriate human control and judgment. For instance, an AI system monitoring an opponent's behaviour may misclassify routine troop movements as hostile, due to flawed data, potentially prompting pre-emptive 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, do not take place in a legal vacuum. AI in the military domain must be developed, deployed and used in full compliance with existing international law, particularly the Charter of the United Nations, international humanitarian law and human rights law, and other relevant legal frameworks. No technology must ever challenge the validity of international law. International law, particularly the Charter of the United Nations in its entirety, international human rights law and international humanitarian law, apply and must be observed and complied with.

States and parties to a conflict must respect and ensure respect for international humanitarian law in all circumstances, including when using AI in military operations. Hence, AI in the military domain should be designed to enhance compliance with international humanitarian law 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 in the military domain is designed with, and trained on, data sets that enable its use in full compliance with international law. Beyond the conduct of hostilities, AI in the military domain must comply with all relevant rules and principles of international humanitarian law, should it be used to perform other tasks governed by international humanitarian law, 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 international humanitarian law 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 AI in the military domain going forward.

3. Understandings and principles

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

1. *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 in the military domain.
- **Context-appropriate human control and judgment**. 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 judgment. AI in the military domain can assist in decision-making but should not replace legal and ethical considerations and judgments, 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.

2. *Reliability, predictability/explainability, robustness*

- **Reliability**. AI in the military domain must be reliable to prevent unintended consequences or malfunctions, especially if it 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 behaviours.

- **Robustness**. AI in the military domain must also be robust – technically and operationally – in order to remain secure and safe when deployed and used.

3. *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 in the military domain should incorporate safeguards that minimize harm and allow adequate human intervention in case of system failures.

4. *Avoiding new pathways of escalation*

- **Stability**. AI in the military domain 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 in the military domain could support de-escalation and crisis management.

5. *Life cycle management of military artificial intelligence systems*

Responsible military use of AI requires a comprehensive and risk-sensitive approach that addresses the entire life cycle of AI in the military domain. 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 life cycle-based approach is particularly essential for high-risk AI 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 international humanitarian law. For systems with lower risk, such as administrative support tools or logistical planning systems, life cycle management should be applied on the basis of a context-specific risk assessment.

- During the design and development phase, States must ensure that systems are trained on high-quality, representative data sets – 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 judgment 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.

4. International governance

Switzerland underscores the importance of an inclusive and sustained United Nations process to consolidate shared understandings of the benefits, risks and challenges of AI in the military domain and to develop principles for its responsible use. Accordingly, all Member States and relevant stakeholders, as well as scientists and representatives of technology industries, civil society and academia, need to be included, to ensure legitimacy, expertise and broad-based support. Related United Nations processes should be transparent, regularly convened and aligned with other relevant initiatives.

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 international humanitarian law. In addition, humanitarian and ethical concerns, the safeguarding of stability and the reduction of security risks must be at the centre of such efforts. Effective governance frameworks, shared norms and sustained multilateral dialogue should help to prevent unintended escalation, foster transparency and mutual confidence, and strengthen the role of international law in times of technological disruption. By anchoring the governance of AI in the military domain in these principles, States contribute to a more predictable, resilient and peaceful security environment.

Specific efforts could include:

- Promoting common understandings, definitions and terminology and a common scope related to AI in the military domain
- Identifying and better understanding humanitarian, legal, security and ethical opportunities and concerns
- Exploring transparency and confidence-building measures
- Developing principles, norms, best practices and other recommendations
- Providing guidance for their implementation

Ukraine

[Original: English]
[11 April 2025]

Ukraine has been actively developing and applying artificial intelligence (AI) in various areas of activity, including the military domain. Ukraine clearly understands both the potential of this technology to enhance human well-being and military capabilities, and the significant risks of its misuse in the civilian and especially in the military sphere. These risks are especially intense in the context of the Russian Federation’s unprovoked and unjustified full-scale invasion of Ukraine, during which it has systematically violated the laws and customs of war and international humanitarian law.

Ukraine supports and participates in international efforts to build global consensus on the responsible development, deployment and use of civilian and military AI.

To date, Ukraine has, inter alia, signed the Bletchley Declaration in 2023; is one of the endorsing States of the Political Declaration on Responsible Military Use of

Artificial Intelligence and Autonomy, launched at the 2023 Responsible Artificial Intelligence in the Military Domain Summit, held in The Hague; supported the Responsible Artificial Intelligence in the Military Domain Call to Action agreed at the 2023 Responsible Artificial Intelligence in the Military Domain Summit and the Responsible Artificial Intelligence in the Military Domain Blueprint for Action adopted as the outcome document of the 2024 Responsible Artificial Intelligence in the Military Domain Summit; joined the Statement on Inclusive and Sustainable Artificial Intelligence for People and the Planet at the 2025 Artificial Intelligence Action Summit in Paris; and co-sponsored all three General Assembly resolutions on AI adopted to date, including resolution 79/239 on artificial intelligence in the military domain and its implications for international peace and security.

Ukraine stands ready to take an active part in new global initiatives to encourage the safe, ethical and responsible development of AI. It also supports discussions on AI in its different aspects across the United Nations system, including within the Security Council.

Being both a peace-loving nation with no territorial claims against others and a victim of Russian military aggression not recognizing any such claims against itself, Ukraine develops and uses military AI exclusively to strengthen its defence capabilities by exercising the right to self-defence provided by the Charter of the United Nations.

In using AI in the military context, Ukraine identifies the following key risks to international peace and security:

- Competition in integrating AI into combat and weapons systems risks triggering a new, more dangerous round of the global arms race to the detriment of achieving sustainable development goals, and particularly the emergence of fully autonomous weapon systems operating without human intervention.
- As with other digital technologies, with the growing threat of cyberattacks and increasing complexity and expansion of areas of application, AI in the military systems is becoming more vulnerable to cyberinterference and manipulation by an interested party aimed at depriving them of their intended application characteristics and selective use functionality.
- Excessive reliance on AI for decision-making could lead to losing human control over critical military processes.
- Hasty integration of underdeveloped AI into weapons systems, especially with flawed target identification capabilities, may result in indiscriminate effects and increased civilian casualties.
- There is currently no multilateral framework to control the proliferation of weapons with integrated AI.
- The use of AI-integrated weaponry without adherence to the laws and customs of war and international humanitarian law presents serious legal and ethical concerns.

United Kingdom of Great Britain and Northern Ireland

[Original: English]
[11 April 2025]

Artificial intelligence (AI) is a family of general-purpose technologies, any of which may enable machines to perform tasks that would traditionally require human or biological intelligence, especially when the machines learn from data how to do

those tasks. AI technologies are maturing and being adopted at extraordinary pace. As a group of technologies with different systems, methods and applications, they have different developmental trajectories and implications. What is certain is that they have the potential to drive transformational change across all aspects of society, the economy and policy, including defence and security.

The United Kingdom welcomes the opportunity presented by General Assembly resolution 79/239 to consider the implications of AI in the military domain beyond those related to lethal autonomous weapons systems, which have been subject to extensive and valuable discussions, including those ongoing in the Group of Governmental Experts established under the Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to Be Excessively Injurious or to Have Indiscriminate Effects. A rigorous assessment of the broader strategic implications of military AI, bringing together thinking, ideas and good practices discussed in informal and formal international forums on this agenda, will allow for a holistic discussion on how to make the most of the opportunities AI presents in the military domain, while addressing effectively associated risks.

Opportunities of artificial intelligence in the military domain

The integration of AI into the military domain will potentially transform defence, global security dynamics and the character of warfare. Advanced technologies enabled by AI, which can categorize and refine large quantities of data from different sources, faster and more comprehensively, will support greater efficiency and improved decision-making, and accelerate the tempo and rigour of operational planning. AI in intelligence, surveillance and reconnaissance systems can provide a more accurate picture of the operational context and enable planners to reduce the impact on civilians – resulting in greater protection for civilians and civilian infrastructure. Autonomous logistics and unexploded ordnance functions will reduce the need to have military personnel on the ground. AI in the military context could therefore strengthen national and international security and lower the risk to human life and reduce casualties.

Research by the United Kingdom Ministry of Defence on AI and peacekeeping identified ways in which peace operations could benefit from AI-enhanced capacities and systems, including:

- Analytical capability that will improve situational awareness, operational decision-making, scenario planning and sentiment analysis capability.
- Autonomous systems, such as unmanned aerial vehicles, could provide enhanced coverage of large geographical areas or high-risk regions (where it may be risky for peacekeeping personnel to maintain a permanent presence).
- Logistics could improve delivery of healthcare and aid provision to local populations, supporting mission objectives and building community trust.

Such capabilities can be applied to enhance monitoring and verification of arms control and peace agreements, making it easier to detect violations or confirm compliance in a timely and credible manner. AI tools could enable better detection, identification, attribution and verification of hostile sub-threshold operations of various kinds, which would reduce the effectiveness of such activities and potentially deter them in the first place. They can help also to monitor and identify online hate speech, propaganda or changes in public sentiment in real time that might escalate tensions or undermine any peace talks or ceasefire.

Challenges and risks

AI use in the military context may exacerbate existing risks and pose additional threats both above and below the threshold of armed conflict. The rush to adopt AI capabilities to gain strategic advantage could result in countries using AI in ways that are unacceptable on legal, ethical or safety grounds. New risks of AI-induced escalations or accidents caused by malfunctions or the fragility, brittleness, immaturity or insecurity of AI systems will require new protocols and de-escalation mechanisms. Hostile actors may seek to attack national AI systems and undermine confidence in their performance, safety and reliability (e.g. by "poisoning" data sources, corrupting hardware components within supply chains and interfering with communications and commands), which could disrupt systems and skew military decision-making in times of crisis and other operational environments.

In times of conflict, these technologies – and the operational tempo they enable – are likely to compress decision times dramatically, tax the limits of human understanding and may require responses at machine speed. The black box nature of many AI capabilities means that humans are often unable to discern how or why a particular output has been delivered. AI-driven operations may lead to unpredictable and opaque behaviour and make accurate inferences and judgments about the intent of an adversary difficult or could be misinterpreted or provoke unintended consequences. Operators could place excessive confidence in algorithmic outputs without a full grasp of the underlying assumptions, constraints and flaws of AI systems. Without appropriate safeguards, norms and protocols in place, AI-driven systems could exacerbate the risk of misunderstanding, miscalculation and unintended escalation.

The widespread availability of advanced AI capabilities or tools and other dual-use technologies likely increases proliferation risks and development of novel weapons by State and non-State actors. AI could be used also to augment or advance disinformation attempts designed to engender hostility towards countries, which could cause conflict and escalate tensions.

United Kingdom commitment to secure and responsible artificial intelligence in the military domain

The United Kingdom recognizes that AI raises profound concerns about fairness, bias, reliability and the nature of human responsibility and accountability, especially in a military context. While States have a long history of incorporating new technologies and will continue to rely on long-established legal, safety and regulatory regimes, we must recognize the particular challenges arising from the nature of AI and importance of positively demonstrating that we are responsible and trustworthy.

The United Kingdom sets out its commitment to secure and responsible AI through its Defence AI Strategy and associated AI ethical principles. These AI ethical principles, set out in the United Kingdom's "Ambitious, safe, responsible" policy, establish the ethical framework considerations of human-centricity, responsibility, understanding, bias and harm mitigation, and reliability. The Joint Service Publication "Dependable artificial intelligence (AI) in defence", published in November 2024, provides clear direction to the teams within the Ministry of Defence and beyond on how to implement these AI ethical principles to deliver robust, reliable and effective AI-enabled services and capabilities.

Through its AI ethical principles, the United Kingdom seeks to cultivate trust in AI technologies and their applications, realizing the full potential of human-machine teaming, while mitigating the risks associated with its use, misuse or disuse and preventing unintended consequences. This approach allows the United Kingdom to

harness the innovation and creativity found across defence and industry in a way that will enable the ambitious adoption of AI-enabled solutions.

The Government of the United Kingdom is clear that any use by the United Kingdom of AI to enhance defence processes, systems or military capabilities is governed by national and international law. The United Kingdom armed forces always seek to abide by their legal obligations across the full range of activities, from employment law to privacy, procurement and the law of armed conflict, also known as international humanitarian law. They have robust practices and processes in place to ensure that their activities and people abide by the law. These practices and processes are being – and will continue to be – applied to AI-enabled capabilities. Deployment of AI-enabled capabilities in armed conflict needs to comply fully with international humanitarian law, satisfying the four core principles of distinction, necessity, humanity and proportionality. We are clear that use of any system or weapon that does not satisfy these fundamental principles would constitute a violation of international law.

Human responsibility and accountability exercised through context-appropriate human involvement is also crucial. This context-appropriate human involvement is necessary to satisfy our policies, ethical principles and obligations under international humanitarian law. The nature of human involvement will vary depending upon the nature of the capability, operational environment and context of use. The United Kingdom will ensure that human political control of its nuclear weapons is maintained at all times.

United Kingdom contribution to international initiatives

Global stability requires the ambitious, but responsible development of military AI. The international community's understanding of the risks, safeguards and standards related to AI use in the military context continues to evolve. Given that the risks are inherently international in nature, they require a global response.

The United Kingdom has been at the forefront of international efforts in support of secure and responsible development and use of AI. It is proud to have hosted the inaugural AI Safety Summit, which agreed the Bletchley Declaration on AI safety, and to have played a role in commissioning the International AI Safety Report – the world's first comprehensive synthesis of current literature of the risks and capabilities of advanced AI systems, published in February 2025, which builds understandings critical to informing international discussion, such as harnessing AI for peace and security. We support efforts under the Global Digital Compact to close digital divides and enhance international governance on AI for the benefit of humanity.

The United Kingdom actively supports international initiatives to drive action in relation to the military domain. We have supported work by organizations like RAND Europe, University of California, Berkeley and the Global Commission on Responsible Artificial Intelligence in the Military Domain to bring together diverse and widely recognized experts to explore these issues, make sense of the latest thinking and map out ways forward for policymakers with workable recommendations.

The United Kingdom continues to be an active participant in international dialogues on AI-related defence and security issues and continues to share its experiences of developing and operationalizing secure and responsible approaches to AI adoption within the military domain. The United Kingdom welcomes progress made through initiatives like the Responsible Artificial Intelligence in the Military Domain Summits, which the United Kingdom co-hosted in 2024, and United States-led Political Declaration on Responsible Military Use of Artificial Intelligence and

Autonomy to increase understanding of the opportunities and strategic risks, and how to address these through appropriate measures that support secure and responsible AI use. AI ethics and assurance are dynamic fields that require continuous engagement, collaboration and iteration.

Looking ahead

The United Kingdom looks forward to building on progress made to date in existing processes, including through discussions in the United Nations based on the Secretary-General's report and focused on tangible actions. Given the nature of AI in the military context, it will be crucial to have an inclusive and multi-stakeholder approach, informed by technical, military and legal expertise from States, industry, academia and civil society.

While we have an abundance of information, our collective understanding of military applications and implications remains low and there remain substantial knowledge gaps and misunderstandings about the nature and capabilities of AI. Further work is required to build the capacity of States, enhance our collective understanding of the implications and potential risks and challenges of military AI at the strategic level and establish universally agreed terminology to allow for constructive discussions. Discussions should focus on tangible, effective and appropriate measures and practices that could help to address risks, including such things as safeguards and norms of behaviour, new communication channels and transparency mechanisms to reduce the risk of misinterpretation, updated doctrines, confidence-building measures and arms control agreements that reflect the impact of military AI.

B. European Union

[Original: English]
[11 April 2025]

The European Union welcomes this occasion to submit its views on the challenges and opportunities posed to international peace and security by artificial intelligence (AI) in the military domain, in accordance with resolution 79/239, adopted by the General Assembly on 24 December 2024.

First and foremost, the European Union would like recall its long-standing position that the use of AI in the military domain must be in accordance with international law, notably the Charter of the United Nations, international humanitarian law and international human rights law.

Likewise, the European Union wishes to recall another long-standing position, namely, that human judgment and control over the use of force must always be retained. Humans must remain responsible and accountable also when it comes to AI in the military domain so as to ensure that this technology is applied in a responsible manner.

The European Union recognizes that the application of AI to military systems entails opportunities as well as challenges. The development of AI is so fast that not all advantages or risks can be predicted at this point of time.

In this respect, the European Union welcomes the ongoing focus of the United Nations on the matter as well as the discussions within relevant international forums. In this respect, the European Union in particular appreciates the continuation of the Responsible Artificial Intelligence in the Military Domain process, which began in the Netherlands in 2023 with the first Responsible Artificial Intelligence in the Military Domain Summit, followed by the Responsible Artificial Intelligence in the

Military Domain Summit hosted by the Republic of Korea in 2024. The European Union welcomes the continuation of the process, with Responsible Artificial Intelligence in the Military Domain Summit, to be held in 2025 in Spain, and extends its gratitude to Spain for organizing the next Summit.

The European Union notes that the 2023 Responsible Artificial Intelligence in the Military Domain Call to Action, as well as the 2024 Responsible Artificial Intelligence in the Military Domain Blueprint for Action, have been endorsed by all European Union member States. The European Union believes that the Responsible Artificial Intelligence in the Military Domain Summit concept of multi-stakeholder, inclusive processes on the issue of responsible military use of AI is a promising approach. In this respect, the European Union recognizes the value of other recent contributions, such as the International AI Summit and the Artificial Intelligence Action Summit, which was hosted by France on 10 and 11 February 2025. The European Union also acknowledges the work in the framework of the Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy, as a valuable contribution to the broader international debate on consequences of AI for international peace and security.

The European Union believes that the Responsible Artificial Intelligence in the Military Domain outcome documents and the Political Declaration, to which all European Union member States are signatories, are complementary and highly important for further developing global thinking, governance and practical solutions to the responsible military use of AI.

The European Union recognizes that there are military advantages of applying AI in the military domain. This goes, in particular, for speed, scale and precision of military operations. AI can provide a tactical advantage from the management and pre-processing of vast data sets stemming from surveillance and weapons systems, drones and satellite images, which can enable human operators to achieve speedier and better decisions. AI applications can reduce costs by improving logistics or the maintenance of equipment via predictive maintenance management. Likewise, AI can provide greater distance military operations and more precision of military operations in uncertain environments.

At the same time, the very advantages of speed and scale by AI applications in the military domain also pose challenges. AI accelerates the observe, orient, decide, act loop. The increase of speed and scale capabilities may give rise to misperceptions due to inconsistencies between military intentions and the analyses produced by AI-driven systems. AI could thus unintentionally contribute to escalation. Speed is also a challenge to the objective of retaining human judgment and control over the use of force.

Against this backdrop, the European Union stresses the importance of international cooperation aimed at studying the impact of AI in the military domain and possible governance frameworks.

Annex II

Replies received from international and regional organizations, the International Committee of the Red Cross, civil society, the scientific community and industry[1]

A. International and regional organizations

African Commission on Human and Peoples' Rights

[11 April 2025]

I. Introduction

The African Commission on Human and Peoples' Rights (the African Commission), as the premier treaty-based human and peoples' rights body of the African Union (AU), is entrusted with the mandate of promoting and protecting human and peoples' rights in Africa under the African Charter on Human and Peoples' Rights (African Charter). In the African Commission's study on Addressing Human Rights Issues in Conflict Situations, the African Commission's Focal Point who led the study observed that 'it is … in conflict and crisis situations that the most egregious violations and abuses of rights are perpetrated…With the changes in the nature of conflicts and the attendant heightened threat to human and peoples' rights, there is a greater need for the human rights system to pay increasing attention to and provide effective responses to the challenges that these new dynamics present to the protection and observance of rights.' In the current context, one of the major new dynamics that carries serious implications for peace and security and therefore human and peoples' rights relate to Artificial Intelligence (AI) and in particular its rapid development and use in the military domain.

During its 1214th meeting, the AU Peace and Security Council (PSC), in requesting the AU Commission to conduct a study to assess the adverse impact of AI on peace and security, underscored the necessity of ensuring African perspectives in shaping global AI governance frameworks. Against this background and having regard to its work on AI and other technologies and human and peoples' rights[2] and human rights in peace and security, the African Commission is pleased to share its views in response to the invitation of the Secretary-General for submission of inputs on AI in the military domain and its implications for international peace and security.[3]

II. AI in the military domain and peace and security

The development and use of AI technologies in the military domain particularly to automate military functions such as surveillance, targeting, and the deployment of

[1] In accordance with operative paragraph 8 of General Assembly resolution 79/239, the replies received from international and regional organizations, the International Committee of the Red Cross, civil society, the scientific community and industry are included in the original language received. The Secretary-General remains committed to multilingualism as a core value of the United Nations.

[2] Resolution ACHPR/Res. 473 (EXT.OS/ XXXI) 2021 on human and peoples' rights and artificial intelligence (AI), robotics and other new and emerging technologies in Africa, available at https://achpr.au.int/en/adopted-resolutions/473-resolution-need-undertake-study-humanand-peoples-rights-and-art.

[3] The Focal Point of the African Commission on its study on human and peoples' rights and AI, robotics and other technologies acknowledges with appreciation the contribution of Professor Thompson Chengeta, who is the consultant providing technical assistance in the development of the study, through the Centre for Human Rights, University of Pretoria.

lethal force have far reaching consequences for peace and security and hence for human and peoples' rights. The AU Continental AI Strategy, endorsed during the 44th Extraordinary Session of the Executive Council of the African Union, highlights AI governance and regulatory challenges, particularly in military applications, warning that AI could exacerbate conflicts through inaccurate predictions or deployment of autonomous weapon systems. Additionally, the framework raises concern about disinformation, misinformation, cybersecurity threats, and military risks.

From the perspective of the development and use of AI in the military domain, peace and security should not be seen just from the perspective only of what it means for stability of states and societies. Beyond its conception under the UN Charter and public international law associated with friendly relations of states, peace and security is also a fundamental right of all peoples. The African Charter thus stipulates that 'All peoples shall have the right to national and international peace and security. The principles of solidarity and friendly relations implicitly affirmed by the Charter of the United Nations and reaffirmed by that of the Organization of African Unity shall govern relations between States.'[4]

The framing of peace and security as a right of peoples compels states to assess and govern the development and deployment of AI technologies in the military domain through a human rights lens that prioritises the prevention of harm, suffering, and injustice. Together international law conception of peace and security, it places an affirmative duty on states to ensure that AI systems do not contribute to conflict, perpetuate structural inequalities, or violate the rights and dignity of individuals and communities. By embedding peace and security within the framework of human rights, states are not only accountable for avoiding direct acts of aggression, but also for proactively creating and maintaining environments in which human flourishing, security, and justice are protected from the potentially disruptive or harmful impacts of emerging military technologies.

The implication of AI in the military domain to peace and security, farmed comprehensively, thus goes beyond how it shapes the obligation of states for non-aggression. It also covers how algorithm-driven systems may dehumanise individuals, introduce bias, and lead to unaccountable or disproportionate harm. It raises critical questions about the erosion of human oversight, the potential for unlawful killings or violations of international humanitarian law, and the targeting of vulnerable or marginalised populations.

By transforming military capabilities, the application of AI in the military domain can also have implications for peace and security by heightening tendencies for engaging in hostilities. The resultant escalation of tension and violence will be inimical not only to stability and peace between and within states but also most importantly carries more adverse consequences for the development needs of the less developed parts of the world such as Africa. While AI may contribute to advancing the development needs of Africa, its development and use in the military domain can have devastating consequences for development detrimental in particular to the right to development enshrined in Article 22 of the African Charter.[5]

This link between peace and development is also central to the Sustainable Development Goals (SDGs), especially SDG 16, which promotes peace, justice, and strong institutions. Without peace and security, sustainable development cannot be achieved. Recognising this link is critical in the governance of military AI, as the

[4] Article 23(1) of the African Charter.

[5] All peoples shall have the right to their economic, social and cultural development with due regard to their freedom and identity and in the equal enjoyment of the common heritage of mankind.

militarisation of AI can aggravate instability, particularly in fragile regions, and undermine Africa's developmental aspirations. By reaffirming the interconnectedness of peace and development, the African Commission calls for a governance approach that upholds peace as both a human right and a developmental imperative.

III. The need for a human and peoples' rights-based regulation of the development and use of AI in the military domain

Given the ways in which the use of AI in the military domain transforms the conduct of hostilities and how the development of AI relies on the extraction of natural resources particularly critical minerals such as rare earth minerals, it is the submission of the African Commission that both the process of extraction of resources in the development of AI in the military domain and the use of AI in the military domain need to be in full compliance with human and peoples' rights standards and international law principles, including international humanitarian law.

First and foremost, it is of paramount significance that the development and use of AI in the military domain complies with the right to peace and security enshrined in Article 23 of the African Charter on Human and Peoples' Rights. As a right that is born out of the recognition of the inseparability of the enjoyment of other human rights states from peace and security, this right entails that the use of AI in the military domain should be consistent with the international law prohibition of the use of force enshrined in the UN Charter and the Constitutive Act of the African Union.

Second, the use of AI technologies in conflict settings need to ensure respect for applicable human and peoples' rights and international humanitarian law principles, including most notably needs to adhere to the principles of precaution, necessity, distinction, proportionality and legitimacy. These requirements apply irrespective of whether the context in which the use of AI in the military domain relates to international armed conflicts or non-international armed conflicts. As established in the African Commission's study,[6] parties to conflict are obliged to observe human rights standards where such conflicts do not meet the IHL threshold of armed conflict. As such, those who use AI technologies in conflict situations that do not meet the IHL threshold of armed conflict are legally obliged to respect and ensure respect for the human and peoples' rights standards established under treaty and customary international human rights law.

Third, the development of AI in the military domain and the use AI technologies in hostilities need to comply with the principle of transparency. This is fundamental because it is the basis for ensuring effective regulation of the development and use of AI in the military domain and for compliance with applicable human rights and international law standards. Additionally, transparency is critical for ensuring compliance with the obligation for respecting the dignity, privacy and data protection of individuals. The principle of transparency is also a pre-requisite for addressing some of the concerns that arise from use of AI in the military domain including bias (owing to the source and type of data used) and explainability. Transparency is also critical not only with the development of AI in the military domain but also with respect to the transfer of AI technologies in the military domain.

Fourth, from the perspective of human and peoples' rights and IHL, the other standard key to human rights and international law-based regulation of the development and use of AI concerns accountability. In the event of the occurrence of violations of human and peoples' rights standards or IHL principles from the development and use of AI in the military domain, there has to be both institutional and individual accountability. Accountability in this instance encompasses not only

[6] ACHPR, Addressing human rights issues in conflict situations, https://achpr.au.int/en/node/895.

the measures that are taken against perpetrators but also the remedial steps that need to be put in place for redressing victims.

Firth, building and sharing of technical knowhow critical to ensuring regulation by states is the other principle. Recent developments including the jamming of GPS systems affecting flights reported in Eastern DRC and the deployment by the Islamic State of West Africa of armed drones, highlight not only the need for effective regulation but also the need for developing the requisite infrastructure and technical capacity for ensuring effective regulation.

IV. The link between the development of AI in the military domain and Africa's natural resources and its implications for peace and security

The African Commission is also of the view that when discussing peace and security, stakeholders must be aware of the link between development of military AI, Africa's natural resources – particularly critical minerals – and the notion of peace and security. Article 21(1) of the African Charter on Human and Peoples' Rights affirms: "All peoples shall freely dispose of their wealth and natural resources. This right shall be exercised in the exclusive interest of the people. In no case shall a people be deprived of it."[7] Article 21(5) further provides that "States parties to the present Charter shall undertake to eliminate all forms of foreign economic exploitation particularly that practised by international monopolies so as to enable their peoples to fully benefit from the advantages derived from their national resources."[8]

This provision is particularly important in the context of military AI, which depends heavily on critical minerals such as cobalt, lithium, and rare earth elements – resources abundantly found in Africa. The 2924 Report of the Chairperson of the African Commission's Working Group on Extractive Industries, Environment and Human Rights Violations, stressed the "significance of critical minerals for new and emerging technologies" and highlighted that Africa has been burdened by a "resource curse phenomenon."[9] The report of the Chairperson noted that "extraction of minerals and other resources not only fuels but also at times becomes the site where contestation over whose control and use triggers conflicts. In some instances, this has created a vicious cycle of insecurity and violence, a condition that not only leads to major human and peoples' rights violations but also the perpetuation of a vacuum of effective governance and the concomitant exploitative, socially and environmentally costly extraction of the resources of the continent."[10]

Therefore, governance of military AI must not only ensure the legal use of force but also address the exploitative chains of extraction that power such technologies. This requires strict oversight, equitable benefit sharing, and regional solidarity to prevent Africa's resources from being used to fuel further conflict and inequality.

V. Conclusion

The African Commission is of the view that the development and use of AI in the military domain carries far reaching consequences for international peace and security in general and for less developed parts of the world such as in Africa that historically suffered violations and remain vulnerable to the adverse impacts of the development and use of AI in the military domain without robust and effective legal regime for such development and use in the military domain. The African

[7] Article 21(1) of the African Charter.
[8] Article 21(5) of the African Charter.
[9] African Commission's Working Group on Extractive Industries, Environment and Human Rights Violations (2024), https://achpr.au.int/en/intersession-activity-reports/extractive-industries-environment-and-human-rights-violations (accessed 08 April 2025).
[10] As above.

Commission affirms that the development and use of AI in the military domain needs to be regulated on the basis of international law, human and peoples' rights and international humanitarian law standards with particular regard to the development and peace and security interests and human and peoples' rights needs of less developed parts of the world.

More specifically, beyond and above the right to peace and security, the governance of AI in the military domain needs to ensure respect for applicable human and peoples' rights and international humanitarian law principles, including most notably needs to adhere to the principles of precaution, necessity, distinction, proportionality and legitimacy, the principles of transparency, accountability and redress for victims and the obligation to build and share technical knowhow necessary for enabling societies to avert the risks that the development and use of AI in the military domain carries for peace and security. Only by ensuring that the development and use of military AI are aligned with international legal standards including those relating to the right to peace and security, the right to development, the right to privacy and protection of personal data, the right to remedy and the responsibility for exercising human control, the right to and control over natural resources and by addressing the structural inequities underpinning global technological advancement, can states uphold their duties to their peoples and advance genuine peace, justice, and security in relation to the development and use of AI in the military domain.

B. International Committee of the Red Cross

[19 March 2024]

Summary

The full submission is available at: https://www.icrc.org/en/article/artificial-intelligence-military-domain-icrc-submits-recommendations-un-secretary-general.

The International Committee of the Red Cross (ICRC) welcomes the opportunity to submit its views for consideration by the United Nations Secretary-General, in accordance with resolution 79/239.

The recommendations that the ICRC makes in this submission are in line with its long-standing mandate and practice of promoting respect for and the development of IHL, including its application to new technologies of warfare. This submission is intended to support States in ensuring that military applications of AI comply with existing legal frameworks and, where necessary, identifying areas where additional legal, policy, or operational measures may be required.

1. Normative proposals: Reaffirming existing IHL as the starting point

The ICRC has consistently emphasized that, while IHL does not explicitly prohibit or regulate the use of AI in military applications, it does restrict its development and use, and places strict constraints on AI when it is integrated into weapon systems or used in some way to conduct warfare.[1]

Existing and emerging normative proposals on the military application of AI should build upon established international legal frameworks and mechanisms, including IHL. Where necessary, these frameworks can be reinforced through the development of additional legal instruments, operational guidance or policy measures to address specific risks or challenges posed by emerging technologies. The form and content of such measures may vary depending on the specific use case. The ICRC

[1] This has also been affirmed by States, including in the UN General Assembly with Resolution 79/239.

encourages the international community to engage in concrete discussions on particular applications of AI in the military domain and to prioritize consideration of those that pose the greatest risks to people affected by armed conflicts.

2. A Human-centred Approach to military AI

In line with the resolution, the ICRC advocates for a human-centred approach to the development and use of AI in armed conflict.[2] This approach has at least two key dimensions: first, ensuring a focus on the humans who may be affected by the use of AI; and second, emphasizing the obligations and responsibilities of the humans using or ordering the use of AI in military operations.

Despite the growing development of AI-related technologies in the military domain, IHL requires individuals to make legal determinations. Humans must, for instance, determine the lawfulness of attacks that they plan, decide upon or execute, and they remain accountable for those determinations. The ICRC considers that human judgement is crucial for reducing humanitarian risks, addressing ethical concerns and ensuring compliance with IHL. Accordingly, while certain technical tasks may be carried out by machine processes, it is not the system itself that must comply with the law, but the humans using it.[3]

This does not mean that commanders and combatants cannot or should not use tools, including AI-decision-support systems. However, these tools must only be designed and used to support, rather than hinder or replace, human decision-making.[4] Further, States and parties to armed conflicts must ensure that human control and judgement are preserved in decisions that pose risks to the life and dignity of people affected by armed conflict. This is essential for ensuring respect for applicable laws, including IHL, and upholding ethical standards.[5]

3. Specific Applications of ai in the military domain

The ICRC has identified three specific applications of AI in the military domain that pose particularly significant risks to those affected by armed conflict:

1. AI in Autonomous Weapon Systems

Resolution 79/239 acknowledges the increasing integration of AI into weapons and weapon systems, a development that raises significant legal and humanitarian concerns. The integration of AI, particularly machine learning (ML) techniques, into autonomous weapon systems (AWS) exacerbates existing challenges posed by AWS in ensuring compliance with IHL. In particular, it increases difficulties for human users to understand, predict, and control the system's functioning and effects.

Users of AWS must be able to, with a reasonable degree of certainty, predict the effects of that weapon in order to determine whether it can be directed at a specific military objective, and take steps to limit those predicted effects, as required by IHL. This entails the ability to understand the functioning of the AWS: the nature and functioning of its sensors, the definition of its target profile and the potential effects in the circumstances of use, including any risk of error or malfunction. This is

[2] ICRC, AI and machine learning in armed conflict: A human-centred approach, 2019 (updated in 2021).

[3] ICRC, International Humanitarian Law and the Challenges of Contemporary Armed Conflicts: Building a Culture of Compliance for IHL to Protect Humanity in Today's and Future Conflicts (IHL Challenges Report), 2024, p. 61.

[4] *Ibid*.; ICRC, IHL Challenges Report – Chapter 2: Contemporary and future challenges in the conduct of hostilities, 2019, p. 32.

[5] ICRC, Decisions, Decisions, Decisions: computation and Artificial Intelligence in military decision-making, ICRC, 2024, p. 8.

particularly relevant for AWS that function in opaque ways (the "black box" challenge), such as AWS relying on AI techniques, which prevent the human user from being able to understand, predict or explain the system's output. This impossibility effectively results in a lack of control over the weapon's effects, rendering it indiscriminate by nature.

In this regard, we reiterate the joint call made by the ICRC President, with the UN secretary-general,[6] for new, legally binding rules prohibiting certain AWS and constraining the use of others.[7] In particular, we recommend a prohibition on

- unpredictable autonomous weapons – those that, due to their design or the circumstances and manner of use, do not allow a human user to understand, explain or predict the system's functioning and effects;
- autonomous weapons designed or used to target humans directly. This is required because of the significant risk of IHL violations and the unacceptability of anti-personnel autonomous weapons from an ethical perspective.[8]

The ICRC supports all efforts by States to urgently adopt a legally binding instrument to regulate AWS, in whichever forum they choose.[9] The integration of AI into AWS should also be considered when discussing normative proposals on military applications of AI. Doing so is essential to ensure a consistent and comprehensive approach to the regulation of military AI, to avoid normative gaps, and to effectively address the serious legal, ethical, and humanitarian risks that are exacerbated by the integration of AI into AWS. In this regard, the ICRC considers it important that binding prohibitions and restrictions on AWS, including AWS that incorporate AI, are integrated into broader discussions on the governance of military AI.

2. *AI in Military Decision-Making*

AI decision-support systems (AI-DSS) are computerised tools that bring together data sources – such as satellite imagery, sensor data, social media feeds or mobile phone signals – and draw on them to present analyses, recommendations and predictions to decision makers.

The use of AI-DSS raise concerns related to system functioning, data quality, and human-machine interaction. These systems risk increasing the rate of unforeseen errors, perpetuating problematic biases – particularly those based on age, gender, ethnicity, or disability, and making it difficult for the users to understand how and why the system generates its output from a given input.

Generally, AI-based systems will perform better when given well-defined goals and access to representative and high-quality data. However, armed conflict environments are marked by uncertainty, volatility, and deliberate deception techniques by adversaries, which makes it extremely difficult to obtain reliable or transferable data. Even where good data exists, it may not reflect the specific operational or humanitarian dynamics of a particular context.[10] Moreover, for AI systems that rely on training data, the utility of those data can rapidly diminish once

[6] ICRC, Joint call by the United Nations Secretary-General and the President of the International Committee of the Red Cross for States to establish new prohibitions and restrictions on Autonomous Weapon Systems, 2023.

[7] ICRC, ICRC Submission on AWS to the UN Secretary-General, 2024, p. 6.

[8] *Ibid*.

[9] *Ibid*.

[10] ICRC, IHL Challenges Report, 2024, pp. 64–65; ICRC, AI and machine learning in armed conflict: A human-centred approach, 2019 (updated in 2021); ICRC, Decisions, Decisions, Decisions: Computation and Artificial Intelligence in Military Decision-Making, ICRC, 2024, pp. 31 and 54.

a conflict begins. Parties to armed conflicts will continuously seek to maintain the initiative and operate in a manner that is not anticipated by their adversary, adapting their strategies and tactics accordingly. This can fundamentally alter the environment in which the system was expected to operate, making the original data no longer representative of the new operational conditions. In such cases, the system's outputs may become unreliable, and the AI model may require re-evaluation or retraining in order to remain fit for purpose.

Human interaction with these systems raises further concerns, such as "automation bias" – a propensity to rely on machine outputs even when other available information may call those outputs into question – which is particularly pronounced in high-pressure or stressful environments like in armed conflicts.[11] Taken together, these factors can hamper a user's ability to scrutinize the information available. The practical consequence might be, for instance, that someone plans, decides upon or launches an attack based solely on an AI-DSS's output, thereby effectively serving as a human rubber stamp rather than assessing the lawfulness of the attack by considering all the information reasonably available including the AI-DSS output.[12]

On the positive side, the careful use of AI-based systems may facilitate quicker and more comprehensive information analysis, which can support decisions in a way that enhances IHL compliance and minimizes risks for civilians. In the context of urban warfare in particular, the ICRC has recommended that online open-source repositories should be used to gather information about the presence of civilians and civilian objects.[13] Importantly, IHL imposes obligations to take constant care to spare the civilian population and to take all feasible precautions in attack. Therefore, in developing and using AI-DSS, armed forces should be considering not only how such tools can assist them to achieve military objectives with less civilian harm, but also how they might be designed and used specifically to protect civilians. However, the important point is that these computer outputs can inform but must not displace the need for legal determinations.

Beyond targeting decisions, militaries are also exploring the use of AI to support other operations traditionally carried out by humans, including detention operations. While technology deployed responsibly and with robust human oversight can contribute to IHL compliance, it also carries risks including bias, lack of transparency, and faulty programming and analysis, all of which can undermine compliance with IHL.[14]

To support efforts by States and other actors to ensure that military uses of AI-DSS remain consistent with IHL and humanitarian principles, the ICRC has formulated a non-exhaustive set of preliminary recommendations relating to the development and use of AI-DSS in armed conflict. They focus on 1) ensuring human control and judgement; 2) system design requirements; 3) testing, evaluation, verification and validation; 4) legal reviews; 5) operational constraints on use; 6) user training; 7) after-action reviews; and 8) accountability, among others. The recommendations are annexed to the full version of this submission.

[11] ICRC and the Geneva Academy, Artificial Intelligence and Related Technologies in Military Decision-Making on the Use of Force in Armed Conflicts: Current Developments and Potential Implications, ICRC, 2024, p. 17.

[12] ICRC, IHL Challenges Report, 2024, p. 65.

[13] *Ibid*., p. 66; ICRC, Reducing Civilian Harm in Urban Warfare: A Handbook for Armed Groups, 2023, p. 15.

[14] ICRC, IHL Challenges Report, 2024, p. 22.

3. AI in Information and Communications Technologies

AI is expected to change how actors defend against and conduct information and communications technology (ICT) activities, including in armed conflict. In particular, States have noted with concern that the use of AI and other emerging technologies in malicious ICT activities may further increase their scale and speed, as well as the harm they may cause.[15] For example, AI enables tools to identify and develop exploits for new vulnerabilities in software or networks, or to conduct harmful ICT activities autonomously, whether in offence or in defence. The ICRC is concerned that this could increase the risks of indiscriminate attacks, incidental civilian harm, including damage to critical civilian infrastructure, as well as the uncontrolled escalation of conflict, particularly in complex and interconnected digital environments.[16]

Similarly, information or psychological operations are not a new feature of armed conflicts; however, AI is changing how information is created and spread. AI-enabled systems, particularly generative AI, have been widely used to produce harmful content – text, audio, photos and video – which is increasingly difficult to distinguish from authentic, original content.[17] The ICRC is concerned about the consequences for civilians that might result from the creation and spread of such information through ICT, including information that contributes to or encourages violence, causes lasting psychological harm, undermines access to essential services or disrupts the operations of humanitarian organizations.

In light of these concerns, the ICRC underlines the importance of applying existing international law, including IHL, to the use of AI in ICT activities. The ICRC urges States to ensure that the development and use of AI-supported ICT activities respect the protections afforded to civilians and civilian infrastructure in armed conflict. Moreover, in light of the emergence of increasingly autonomous ICT capabilities, the ICRC further encourages States to address the serious challenges posed by these tools, particularly by considering whether existing international law, including IHL, provides sufficient safeguards against the harm such tools can cause, or whether additional limits are needed.

4. Conclusion

The ICRC is grateful for the opportunity to share its above views and recommendations on ways to address the challenges and concerns raised by AI for the secretary-general's consideration, and stands ready to contribute further to assist States in taking effective action to address the risks posed by AI applications in the military domain.

C. Civil society

Autonorms

[10 April 2025]

The following is the AutoNorms project's submission pursuant to Resolution 79/239 on "Artificial intelligence in the military domain and its implications for international peace and security" adopted by the United Nations General Assembly on 24 December 2024. The resolution requests the UN Secretary-General to seek

[15] 34th International Conference of the Red Cross and Red Crescent, Resolution 2 "Protecting civilians and other protected persons and objects against the potential human cost of ICT activities during armed conflict", 2024.

[16] ICRC, IHL Challenges Report, 2024, pp. 66–67.

[17] *Ibid*., pp. 58–59.

views, including those of Member States, civil society, the scientific community and industry, on "opportunities and challenges posed by the application of artificial intelligence in the military domain, **with specific focus on areas other than lethal autonomous weapons systems**". The AutoNorms team welcomes the opportunity for representatives of academia to submit their views on this important and timely topic.

The AutoNorms project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 852123). Led by Professor Ingvild Bode and hosted by the Center for War Studies at the University of Southern Denmark, the project examines how the integration of artificial intelligence (AI) technologies into weapon systems and military targeting shapes international norms governing the use of force.[1]

Introduction

Over the past 2-3 years, the international debate about applications of AI in the military domain has been characterized by two significant, near-simultaneous changes. First, there has been a **move away from its predominant focus on autonomous or AI technologies in weapon systems** towards considering AI technologies across a wider range of military decision-making tasks, especially in relation to targeting. To reflect his move, this submission focuses on the employment of **AI-based decision support systems (AI DSS),** or systems that are meant to be used as tools to directly or indirectly inform the complex process of use-of-force decision-making, for example, by analyzing large volumes of data, recognizing patterns within the data, predicting scenarios, or recommending potential courses of action to human decision makers.

Second, there has been a **growing emphasis on human-machine interaction** in the context of using AI in the military domain.[2] This emphasis results from the broad recognition that, even when humans are 'in' or 'on' the loop of targeting decision-making, they need to exercise a sufficient level of oversight, control, and agency over the targeting process. Human oversight is a governance principle featuring prominently across various international initiatives, including A/RES/79/239. However, **dynamics of human-machine interaction as part of the use of AI DSS both introduce new issues and solidify existing sets of challenges that require governance attention.** Our submission highlights these challenges and the need to ensure the exercise of human oversight and agency throughout the full targeting decision-making spectrum. It is structured in three parts, starting with explicating challenges of human-machine interaction, then commenting on the relative under-development of the international debate about AI DSS, and finally, sketching a way forward.

Challenges of human-machine interaction in the use of AI DSS

The use of AI DSS involves various dynamics of human-machine interaction because military personnel such as operators and intelligence analysts routinely and increasingly interact with a network of AI systems throughout the targeting process. These interactions involve multiple challenges **which have the potential to affect**

[1] The members of the AutoNorms team are Professor Ingvild Bode, Dr Hendrik Huelss, Dr Anna Nadibaidze, Dr Guangyu Qiao-Franco, and Dr Qiaochu Zhang. The AutoNorms project is based at the Center for War Studies, University of Southern Denmark, Odense, Denmark. For more information, please visit our website: www.autonorms.eu.

[2] Ingvild Bode and Anna Nadibaidze, "Symposium on Military AI and the Law of Armed Conflict: Human-Machine Interaction in the Military Domain and the Responsible AI Framework," *Opinio Juris*, April 4, 2024, https://opiniojuris.org/2024/04/04/symposium-on-military-ai-and-the-law-of-armed-conflict-human-machine-interaction-in-the-military-domain-and-the-responsible-ai-framework/.

the exercise of human agency, or humans' capacity to understand a system's functions and its effects in a relevant context; deliberate and decide upon suitable actions in a timely manner; and act in a way where responsibility is guaranteed.[3]

Dynamics of human-machine interaction result in **distributed agency between humans and AI systems, where they are not separated into two distinct entities but rather form part of a socio-technical system**.[4] As part of this system, both sides may influence each other in different ways, which then translate into various forms of distributed agency located along a spectrum. In some instances, dynamics of human-machine interaction will offer more opportunities for exercising human agency in targeting decisions. In other instances, however, the humans involved in use-of-force decision-making will be more constrained in their ability to exercise agency.

For example, **humans' ability to exercise agency might be limited by cognitive biases such as automation bias or anchoring bias**. Humans could over-trust AI DSS even when knowing that there might be malfunctions or unintended errors involved, risking an overreliance on algorithmic outputs without engaging in the critical deliberations and assessments that are needed to exercise human agency, especially in critical targeting decisions that might inflict death, destruction, and severe harm. Such biases are typically exacerbated by the increased speed of AI-assisted military decision-making, especially in contexts where there are high levels of pressure to act rapidly. They can also be exacerbated by AI DSS that are used for prescription or recommendations, because such systems restrict the options or courses of action available to human decision makers.

Moreover, given that AI DSS are likely to be employed not individually but rather as part of a network of systems, the increased complexity of interactions can result in situations where humans act upon some outputs suggested by AI DSS, but do not overall exercise a high quality of agency. Due to these and many other concerns related to interactions between humans and AI DSS, **there is a need to further investigate challenges of human-machine interaction that result in AI DSS not positively 'supporting' humans but rather undermining humans' ability to exercise agency**.[5]

The risks of not addressing challenges of distributed agency are substantial. First, situations where humans are restricted in their exercise of agency **raise questions about compliance with international humanitarian law**, which requires that humans be held accountable and legally responsible for violations of legal principles. Although humans remain officially in control of the selection and engagement of targets, there are concerns about the exact role played by humans in context of using AI DSS in practice.

Second, these concerns also extend to the risk of **negatively affecting moral agency and responsibility in warfare**. Challenges of human-machine interaction that result in distributed agency would allow humans to feel less morally responsible

[3] Anna Nadibaidze, Ingvild Bode, and Qiaochu Zhang, *AI in Military Decision Support Systems: A Review of Developments and Debates* (Odense: Center for War Studies, 2024), https://www.autonorms.eu/ai-in-military-decision-support-systems-a-review-of-developments-and-debates/.

[4] Ingvild Bode, *Human-Machine Interaction and Human Agency in the Military Domain*, Policy Brief No. 193 (Waterloo, ON: Centre for International Governance Innovation, 2025), https://www.cigionline.org/publications/human-machine-interaction-and-human-agency-in-the-military-domain/.

[5] Anna Nadibaidze, "Do AI Decision Support Systems 'Support' Humans in Military Decision-Making on the Use of Force?" *Opinio Juris*, November 29, 2024, https://opiniojuris.org/2024/11/29/do-ai-decision-support-systems-support-humans-in-military-decision-making-on-the-use-of-force/.

for decisions that could affect other people's lives. They also risk making the human role a nominal, 'box-checking' exercise which can *de facto* be compared with AI DSS playing an 'autonomous' role because the human role is substantially reduced.

Third, there are **security and operational risks related to distributed agency dynamics**, especially when they give too prominent roles to AI DSS and algorithmic outputs. AI systems often malfunction, are trained on biased sets of data which do not apply beyond the training context or specific contexts of use, as well as integrate assumptions that might not be strategically or operationally beneficial.

Various types of biases, issues of trust, uncertainties, targeting and military doctrines, political and societal contexts in which AI DSS are used – all these aspects **can lead to dynamics of distributed agency which limit the exercise of human agency and prioritize algorithmic outputs**. It is important to investigate these dynamics and ensure that distributed agency provides more opportunities than limitations to human decision makers in warfare.

Relative under-development of the international debate on AI DSS

Despite increasing reports about the use of AI DSS in recent and ongoing armed conflicts, and the significant challenges and risks they pose to the effective exercise of human agency, **the international debate on human-machine interaction in the use of AI DSS remains insufficiently developed**, particularly within intergovernmental UN settings. Current discussions on AI in the military domain, including those within the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems (GGE on LAWS), have focused on the use of AI at the tail-end of the targeting process, specifically autonomy and AI in weapon systems. This narrow focus risks overlooking or failing to address critical normative, legal, ethical, security, and operational risks that can proliferate and compound throughout the entire targeting decision-making process.

An increasing, albeit still limited, number of stakeholders are raising this issue at international multistakeholder forums, such as the Summits on Responsible Artificial Intelligence in the Military Domain (REAIM). Some international non-governmental organisations and research institutes – such as the International Committee of the Red Cross (ICRC), the Stockholm International Peace Research Institute (SIPRI), the UN Institute for Disarmament Affairs (UNIDIR), and the Asser Institute – have initiated discussions on challenges posed by AI in the military domain, beyond the issue of autonomy in weapon systems. Despite this progress, there remains **a clear need to develop a more comprehensive and inclusive international multistakeholder debate to guide the responsible development and deployment of AI DSS in military contexts.**

Way forward

In closing, we sketch three ways intended to move the international debate about applications of AI in the military domain forward:

1. **Increase awareness for the implications of practices of designing, developing, and using AI DSS**. States and other stakeholders across industry, civil society, and academia engaged in the governance, development, and use of AI DSS for military targeting must consider the implications of their practices. These practices influence what counts as 'appropriate' ways of considering and employing AI DSS and thereby shape what becomes the accepted, requisite quality of human oversight and agency exercised over the whole process of use-of-force decision-making. To increase such awareness, the debate pursuant to A/RES/79/239 at the UNGA First Committee should centrally focus on the issue of AI DSS in the military domain.

2. **Consistently map both 'best' practices and 'problematic' practices associated with the design, development, and use of AI DSS**. To get a better sense of the direction that the design, development, and use of AI DSS take, states and other stakeholders need to closely map their own (and others') practices. While there have been some limited efforts to exchange potential best practices, we also need to be attentive to practices with potentially problematic effects. This should encompass practices exercised across the full life cycle of AI systems from development to use and post-use review. Mapping such practices would offer stakeholders a better overview of which practices may be beneficial, i.e., provide opportunities for the exercise of human agency, and which practices may be problematic, i.e., limit the exercise of human agency, and therefore assess the desirability of particular practices.

3. **Pursue the debate on AI DSS within a multistakeholder format**. States should work with diverse stakeholders – including academics across social sciences and technical disciplines, civil society representatives, and international organizations – to develop normative guidance and regulation, especially regarding the human role in military decision-making. Moreover, top-down processes towards governing AI DSS should be accompanied by a bottom-up, standard-setting process focused on establishing operational standards. Such an inclusive approach could strike a balance between national security and humanitarian concerns, while reinforcing the need to ensure that humans can exercise agency in use-of-force decisions.

Global Commission on Responsible Artificial Intelligence (AI) in the Military Domain

[11 April 2025]

1. Introduction

The Global Commission on Responsible Artificial Intelligence (AI) in the Military Domain (GC REAIM) welcomes the opportunity to contribute to the United Nations Secretary-General's report pursuant to resolution A/RES/79/239.

GC REAIM recognises that military applications of AI present both opportunities and challenges for global peace and security. Accordingly, the establishment of responsible and ethical governance – consistent with States' obligations under applicable international law – is essential. The global community must take proactive steps to ensure that military AI is developed and deployed in a manner that de-escalates rather than escalates conflicts; respects and enhances, rather than compromises, the sovereignty and territorial integrity of states; promotes rather than threatens the security and safety of civilians; constrains and supports rather than erodes the existing rules-based international order.

In line with GC REAIM's resolute commitment to advancing international governance efforts, this note outlines some of the – non-exhaustive – views expressed by GC REAIM Commissioners and Experts on the implications of AI in the military domain to peace and security. The views presented are general in nature and will be further elaborated in the forthcoming GC REAIM report. While the Commission plans to present substantive and actionable recommendations for stakeholders in September 2025, this note does not yet include concrete proposals. As discussions among Commissioners and Experts are still ongoing, it instead highlights some of the key opportunities, challenges, benefits and risks posed by AI in the military domain to peace and security.

2. Technological Foundations

GC REAIM holds that meaningful policy deliberations on AI in the military domain must be grounded in a shared, foundational understanding of the underlying technologies and their potential trajectories. The complexity of AI technologies often gives rise to misunderstandings, inflated expectations, or misguided applications. Consequently, it is imperative to demystify AI through formal and well-defined frameworks that distinguish between current capabilities and speculative future developments. To support this objective, GC REAIM is developing a taxonomy which seeks to map the full spectrum of AI applications across military and broader peace and security contexts. The taxonomy differentiates between the implications of AI in operational activities – such as warfighting and intelligence – and administrative activities – such as logistics and personnel training and helps identify the specific applications of AI that should be prioritised in governance deliberations.

In its approach to the creation of a taxonomy, GC REAIM highlights the need for and contributes to a concerted effort to clarify, standardise, and encourage the accurate use of technical language with different layers of abstraction for policymakers, experts, and the public, thereby enhancing transparency, mutual understanding, and public trust. GC REAIM also cautions against the uncritical multiplication or adoption of new terminologies in AI governance discourse, unless these are clearly defined; and to ensure such terms are not used to circumvent or obscure existing legal obligations. Precision and consistency in language are the basis of responsible AI governance.

3. Implications for Peace, Security, and Stability

GC REAIM recognises that the integration of AI into the military domain presents benefits as well as both foreseeable and unforeseeable risks to international peace and security. A balanced approach to the range of opportunities and challenges emerging throughout the AI life cycle lies at the core of GC REAIM's method and is essential for responsible AI governance.

AI in the military domain may contribute to international peace and security in several important ways. At the developmental stage, the advancement of military AI capabilities may act as a deterrent to violence, as the mere development and presence of advanced technologies by responsible actors can encourage restraint by aggressors. Military AI may enhance early warning systems, strengthening conflict prevention strategies, and supporting arms control verification through AI-driven tools that foster transparency, trust, and cooperation among states – fundamental elements in conflict prevention. AI can also bolster national security and defence by improving the precision, accuracy, and efficiency of intelligence analysis and situational awareness, enabling real-time threat detection, and facilitating more efficient counterterrorism operations through predictive analytics and autonomous systems. AI-powered systems can rapidly process vast amounts of complex data, enabling military forces to make timely, informed decisions that may prevent escalation and support conflict de-escalation efforts. These traits can also help improve targeting accuracy and precision, potentially reducing the risk of collateral damage or fratricide – attacks on one's own forces – and aiding compliance with International Humanitarian Law (IHL) to protect the security of protected persons, such as civilians and non-combatants, during armed conflict. Military AI may also reduce certain forms of human bias and enhance accountability by providing precise data, surveillance, and real-time monitoring, enabling clear attribution of actions to specific actors. In these ways, AI offers meaningful opportunities to reinforce adherence to international law and ethical standards, strengthening the normative foundation of the rules-based international order underpinning global peace and security.

AI in the military domain also presents a range of risks. As with the development of other general-purpose technologies, the development of AI in the military domain may accelerate arms races. AI technologies driven by the commercial market may be repurposed by militaries or soldiers in need or increase the access of violent non-state actors to AI-enabled military capabilities, which may intensify ongoing conflicts and contribute to broader instability. There are also concerns that states could employ AI technologies to suppress human rights, entrench internal repression, and destabilise both regional and global peace.

Concurrently, as with AI more broadly, the environmental consequences of military AI – such as the energy-intensive demands of AI systems, resource extraction, and ecological damage from AI-enabled military systems – could aggravate resource scarcity and environmental degradation, fuelling tensions and undermining long-term peace. However, given the impact militaries have on civilian technology development, efforts to reduce the environmental impact of AI in defence settings could have far-reaching beneficial consequences for all uses of AI. As such, considerations of environmental impacts should be a component of responsible AI governance in the military domain.

The large-scale data extraction required for AI development could intensify geopolitical rivalries, facilitate intrusive surveillance, and create distrust through opaque and exploitative data practices. Such deployment of military AI may perpetuate discrimination and exacerbate social divisions, undermining stability and ultimately international peace and security.

There are simultaneously significant concerns regarding the potential of integration of AI within the command, control, and communication (C3) structures of nuclear weapons. A number of Commissioners and Experts have emphasised that this is a red line that must not be crossed. The commitment of several nuclear-armed states to human decision-making surrounding the employment of nuclear weapons is therefore applauded. Further, the development of large-scale lethal autonomous weapon systems – such as swarms of anti-personnel devices – risks creating a new category of weapons of mass destruction, posing serious threats to global peace and security. Relatedly, AI may lower the barriers to creation and use of nuclear, chemical, or biological weapons by state or non-state actors, thus generating new challenges for arms control and non-proliferation regimes.

Beyond these strategic risks, AI may affect the character of war and lower the thresholds for armed conflict. By increasing the speed of armed escalation and driving changes in the capabilities of weapons systems, AI in the military domain may reduce states' confidence in their deterrent capabilities – particularly in the face of cyber infiltration risk – thus influencing how decision makers receive, process, and act on information. AI in the military domain could also exacerbate asymmetric warfare and violence by widening technological disparities that could increase the likelihood of force being used prematurely or disproportionately.

Operationally, inaccurate AI systems used for targeting can undermine the security of protected persons under IHL by increasing the risk of indiscriminate attacks, violations of proportionality, and failure to distinguish between combatants and civilians. Closely related to this is the risk of fratricide due to potential errors in target identification or decision-making, which can undermine operational effectiveness, escalate conflict, and erode trust within militaries and alliances. Finally, there are views that the use of certain AI systems in the military domain can create accountability gaps absent clear rules. By complicating the attribution of responsibility for unlawful actions, the deployment of AI in the military domain could undermine key principles of international law and state responsibility for internationally wrongful acts. This may complicate efforts to hold individuals or

states responsible for violations, leading to a reduced deterrent effect against unlawful conduct. Without avenues to hold actors legally responsible, the enforcement of international law weakens, potentially destabilising peace, encouraging impunity, and exacerbating global insecurity.

4. Decision-Making and Responsibility

GC REAIM acknowledges the ethical and legal challenges that arise from integrating AI into military decision-making which may have a direct impact on preservation of peace and security. The relationship between human judgment and machine outputs is complex and without measures to ensure lawful, responsible and effective development and deployment, there can be an erosion of accountability and increased risks of unintended harm. As AI systems become more sophisticated and integrated within military capabilities, it is plausible that algorithmic decisions may become more commonplace across global battlefields, introducing moral and legal challenges regarding human control, oversight and judgment in diverse contexts.

To address these risks, GC REAIM promotes the need for context-appropriate human judgement over specific uses, capabilities and decisions of AI in military applications. The GC REAIM report will list considerations and conditions that underpin and support human responsibility, judgment and means of adequately evaluating relevant actions and decisions. This could include the introduction of technical standards for explainability, as well as maintaining appropriate human oversight in targeting decisions, assessments of precautions, proportionality and distinction, and other critical operational choices. However, given that the very definition of autonomy in machines suggests the minimisation or removal of the human, ensuring human responsibility and accountability may require focusing on human decision-making at earlier stages of a system's life cycle, as the systems structure the behaviour of all who work with it. Human oversight is essential to uphold state obligations under applicable international law, in particular, IHL.

Military AI systems must be designed not only to support all individual and collective agents in the military domain to be effective in safely carrying out their lawful tasks, but also to do so responsibly and without compromising or undermining their status as moral human agents. GC REAIM suggests that military AI based socio-technical systems need to be explicitly and demonstrably designed to adequately attribute and apportion responsibilities and is determined to contribute to this process. For the security of protected persons, parties to armed conflicts should at all times be able to demonstrate that everything possible has been undertaken to create the conditions under which military personnel can effectively apply extant and widely shared principles and laws of armed conflict to their own situation, when using or relying upon AI components in the execution of their tasks.

5. Governance and Regulation

In light of both the opportunities and risks associated with military AI, GC REAIM supports a comprehensive governance framework that implements authentic international law. GC REAIM reiterates that existing legal regimes provide a solid foundation for regulating AI technologies. Governance must incorporate and account for procedural safeguards (due diligence and legal reviews, transparency of testing, evaluation, and validation, accreditation, and verification), substantive obligations drawn from various branches of international law, and soft law tools (military doctrines, national policies and strategies, norms and standards). In principle, all relevant international legal frameworks must be considered and applied. These include, but are not limited to, the following: (1) international law (*jus ad bellum*) which regulates when and how states use force, codifying a general prohibition on the use of force and exceptions such as in the case of self-defence, (2) international

humanitarian law (*jus in bello*) which governs conduct during armed conflict and ensures the security of protected persons, (3) international human rights law.

GC REAIM further emphasises the critical role of international, regional, and domestic institutions in implementing and enforcing these legal norms. Effective governance requires collaboration across these levels and the inclusion of both binding (hard law) and non-binding (soft law) instruments. Soft law mechanisms, such as codes of conduct and ethical principles, can complement existing treaties and facilitate rapid, flexible responses to technological developments.

To address the diverse range of challenges surrounding the integration of AI into the military domain, GC REAIM supports proactive risk-mitigation and confidence-building measures. While binding regimes are challenging for general purpose technologies, there may be opportunities for rigorous monitoring, verification, and enforcement mechanisms inspired by successful global arms control regimes. For example, Commissioners and Experts have discussed ideas such as an Autonomous Incidents Agreement to reduce the risks of miscalculation among AI-enabled autonomous systems, or a committee or consortium that could set guidelines and recommendations surrounding the testing and evaluation of AI systems, including generative AI. GC REAIM also suggests that states and industries should consider adopting human-centred safety-by-design principles, implement red-teaming practices throughout AI system life cycles, and maintain clear chains of accountability for all actors. Only through robust multilateral dialogue and inclusive multi-stakeholder cooperation can AI be effectively governed to enhance peace and security rather than exacerbate global instability.

GC REAIM acknowledges that the development of a comprehensive governance framework for military AI faces several key challenges. First, there is the challenge of diverse interests and perspectives, with states, private companies, and civil society holding varying and sometimes conflicting views on the regulation of military AI. Second, the sensitivity surrounding national security and defence poses a significant barrier, as many states are reluctant to subject their military technologies to international scrutiny or regulation due to legitimate security interests. Third, achieving meaningful and substantive inclusivity in discussions is often difficult, as key stakeholders may be excluded or marginalised in decision-making processes. Fourth, a trust deficit between states, international organisations, and the private sector complicates efforts to establish cooperative governance. Fifth, the presence of crosstalk, incommensurability, and discursive dissonance arises due to the diverse backgrounds and expertise of stakeholders, making consensus-building challenging. Finally, these obstacles are compounded by the lack of clear frameworks that address the complex ethical, legal, and technical issues at the nexus of AI and the military domain. In light of these challenges, the final GC REAIM report will offer strategies to navigate and overcome these barriers in developing a robust governance framework.

6. Conclusion

GC REAIM observes that the rapid advancement and deployment of AI technologies in military contexts poses opportunities, challenges, benefits and risks for global peace and security. Balancing these considerations must be met with a technologically sound, inclusive, principled, and legally grounded approach to governance.

A clear understanding of AI's technological foundations is necessary to properly address its role in modern warfare. Ethical and legal responsibility should remain human-centred, and governance frameworks must rely on the robust application of international law, supplemented by cooperative multilateral efforts and soft law

instruments when appropriate. In its formation and deliberations, GC REAIM has had the opportunity to reflect upon the conversations happening in broader governance processes, finding ways to effectively bridge gaps between disciplines and regional perspectives.

GC REAIM urges the United Nations and all State Parties to place these principles at the heart of global discussions on the implications of AI in the military and broader peace and security, for the present and future generations. Only through concerted international cooperation, guided by a shared commitment to human dignity, peace, and justice, can we ensure that the future of AI in the military domain is one that strengthens our common security.

InterAgency Institute

[11 April 2025]

The InterAgency Institute was established in December 2020 as a digital think tank, founded by expatriate and Global South women as a collective of researchers. It is in this condition that we address this submission on "opportunities and challenges posed to international peace and security by the application of artificial intelligence in the military domain, with specific focus on areas other than lethal autonomous weapons systems," following A/RES/79/239. With this, we seek to craft a complementary set of suggestions to develop the policy discussion in points where understanding that AI encompasses a wide array of data-processing techniques, and may be integrated into different types of warfare, in multiple parts of the organization, and at different levels.

The InterAgency Institute would like to point to overarching trends that fall within our areas of expertise, namely: (1) a focus on the global south, specially in how to prevent furthering the security gap; and (2) in how interagency cooperation in a time of greater mistrust may be leveraged to ensure the integration of AI in the military does not. Additionally, we make the point that Decision Support Systems (DSS) create analogous problems when compared to Autonomous Weapon Systems (AWS).

1. Addressing the security gap between the Global South and the Global North

The increasing technological intensity and digitalization of the battlefield are likely to increase the capacity gap between countries in the Global North & South. The "optimization of war" entails furthering this discrepancy, augmenting threats, and deteriorating the global security landscape. The wide range of AI-enabled solutions represents discrepant utility levels across tools.

While some tools require a low threshold (thus providing usually an equally low ceiling), the systems that pose the biggest military advantage require a high knowledge threshold to be implemented, therefore, will likely not be open source, and will only be available to entities with sufficient means to develop or acquire them. Given the experience in past decades on multilateral forums, it is important to recognize that interest in access to these technologies will play a role in the negotiations.

In the long term, the current trend of "technological sovereignty" (or more specifically of restricted technological access due to global inequalities) may be transformed to undermine such technological control, creating far-reaching implications of this new revolution in warfare, involving stakeholders that may be reluctant to shape modern discussions due to a lack of current development of these technologies in their ecosystem.

2. InterAgency cooperation in times of distrust

These issues call for interagency cooperation at both the strategic and operational levels. The lack of interagency cooperation might lead to threat escalation and the eroding foundations for peace and security. Interagency cooperation should focus on formalizing specific channels for communication between different States, developing strategies for AI implementation that will not damage diplomatic relations, and generating more transparency in the interactions between agencies and contractors. The participation of different branches of government at the UN-level discussions is pivotal for a whole-of-government perspective in the deliberations. Beyond interagency cooperation at the governmental level, the wide array of applications of military AI calls for different sets of Confidence Building Measures (CBMs).

Since AI may be integrated in different warfare types and at different levels, its applications for different contexts have different ethical implications and consequences. Therefore, a monolithic understanding of risks posed by AI in the military context and consequently a unique set of CBMs would be inadvisable. CBMs for AI use in the strategic level of cyberspace will not be the same as CBMs for AI use in the tactical level of aerial warfare. Therefore, thinking about CBMs for military AI as a monolith will lead to inaccurate and in some cases inapplicable measures, undermining its effectiveness.

There is a necessity for sharing best practices in the introduction of AI into military procedures. In this sense, a trade-off should be made, prioritizing best practices that contribute to strengthening the aforementioned points of interagency cooperation and CBMs, and other practices that fall within the larger umbrella of strengthening international peace and security. Sharing of best practices relating to cybersecurity and reliability of the technology could also take place, but they should give priority to CBMs that focus on integration of AI at the strategic level and in manners that avoid the escalation of threats.

3. Decision Support Systems

Target identification or recognition via AI-enabled Decision Support Systems (DSS) entail analogue problems to Autonomous Weapon Systems (AWS). Digital dehumanization, lowering the threshold of violence, and automation bias are byproducts of that process that may only be avoided by the creation of red lines prohibiting such systems that replicate those concerns.

This problem stems not only from AI, but from a wider trend. Other data processing techniques that involve deterministic sorting of data that is not adequately processed by human operators also generate these problems. This caveat should be made to understand that not only systems with AI-enabled technology in DSS pose these kinds of threats, but a wider array of data gathering/processing techniques.

Conclusions and recommendations

- Formal interagency bodies to interface with multilateral AI/military tech negotiations
- Funding and support for academic research in the Global South focused on military AI implications;
- Regular technical-diplomatic summits focused on transparency, shared definitions, and threat perception;
- Prioritize capacity-building initiatives for Global South actors;

- Red lines and confidence building measures could be tailored to the specific technology and operational context;
- The discussions on Autonomous Weapon Systems encapsulate worries around AI-enabled Decision Support Systems. The creation of red-lines for these systems could benefit from building upon recommendations of the GGE on LAWS;

International Committee for Robot Arms Control

[11 April 2025]

The International Committee for Robot Arms Control (ICRAC) values the opportunity to submit our views to the United Nations Secretary-General in response to Resolution A/RES/79/239 "Artificial intelligence in the military domain and its implications for international peace and security."

Founded in 2009, ICRAC is a civil society organization of experts in artificial intelligence, robotics, philosophy, international relations, human security, arms control, and international law. We are deeply concerned about the pressing dangers posed by AI in the military domain. As members of the Stop Killer Robots Campaign, ICRAC fully endorses their submission to this report, and wishes to provide further detail regarding the concerns raised by AI-enabled targeting.

Increasing investments in AI-based systems for military applications, specifically AI-enabled targeting, present new threats to peace and security and underscore the urgent need for effective governance. ICRAC identifies the following concerns in the case of AI-enabled targeting:

1. AI-enabled targeting systems are only as valid as the data and models that inform them. 'Training' data for targeting requires the classification of persons and associated objects (buildings, vehicles) or 'patterns of life' (activities) based on digital traces coded according to vaguely specified categories of threat, e.g. 'operatives' or 'affiliates' of groups designated as combatants. Often the boundary of the target group is itself poorly defined. Although this casts into question the validity of input data and associated models, there is little accountability and no transparency regarding the bases for target nominations or for target identification. AI-enabled systems thus threaten to undermine the Principle of Distinction, even as they claim to provide greater accuracy.

2. Human Rights Watch research indicates that in the case of IDF operations in Gaza, AI-enabled targeting tools rely on ongoing and systematic Israeli surveillance of all Palestinian residents of Gaza, including with data collected prior to the current hostilities in a manner that is incompatible with international human rights law.

3. The increasing reliance on profiling required by AI-enabled targeting furthers a shift from the recognition of persons and objects identified as legitimate targets by their observable disposition as an imminent military threat, to the 'discovery' of threats through mass surveillance, based on statistical speculation, suspicion and guilt by association.

4. The questionable reliability of prediction based on historical data when applied to dynamically unfolding situations in conflict raises further questions regarding the validity and legality of AI-enabled targeting.

5. The use of AI-enabled targeting to accelerate the scale and speed of target generation further undermines processes for validation of the output of targeting systems by humans, while greatly amplifying the potential for direct and collateral

civil harm, as well as diminishing the possibilities for de-escalation of conflict through means other than military action.

Justification for the adoption of AI-enabled targeting is based on the premise that acceleration of target generation is necessary for 'decision-advantage', but the relation between speed of targeting and effectiveness in overall military success, or longer-term political outcomes, is questionable at best. The 'need' for speed that justifies AI-enabled targeting is based on a circular logic, which perpetuates what has become an arms race to accelerate the automation of warfighting. *Accelerating the speed and scale of target generation* effectively *renders human judgment impossible or, de facto, meaningless.* The risks to peace and security – especially to human life and dignity – are greatest for operations outside of conventional or clearly defined battlespaces. Insofar as the use of AI-enabled targeting is shown to be contrary to international law, the mandate must be to *not* use AI in targeting.

In this regard, ICRAC notes that the above systems present challenges to compliance with various branches of international law such as international humanitarian law (IHL), *jus ad bellum* (UN law on prohibition of use of force), international human rights law (IHRL) and international environmental law. In the context of military AI's implications for peace and security, *jus ad bellum*, a framework that prohibits aggressive military actions and regulates the conditions under which states may lawfully resort to the use of force, is the most relevant. In the same manner IHRL is important in this context because it is designed to uphold human dignity, equality, and justice – values that form the foundation of peaceful and secure societies.

International Humanitarian Law and Youth Initiative

[11 April 2025]

Artificial intelligence (AI) has gained a universal recognition during the 1950s'. Technological emergence has assisted humans in almost all facets of their lives thereby making work easier and faster. Moreso, the rapid growth of Artificial intelligence in technological field enthralling commercial investors, law makers, defense intellectuals and international competitors can be evidential in theoretical premises of international security. The use of Artificial intelligence (AI) in modern warfare particularly in the In the Middle East and North Africa, Ukraine/Russian armed conflict which has resulted in the killings of thousands of innocent civilians with women and children being the most vulnerable. The emergence of AI is expected to be utilized in improving all sectors in our daily lives However, its Negative application in the military domain continues to create Humanitarian crisis between warring parties making it of regional and international concern. The war in Gaza is one of the deadliest and most destructive war in history with technology playing a central role in enabling mass slaughter and destruction ranging from supplying the dystopian systems used to automate the killings and bombing. [1] Following the October 7 2023, there have been extensive reports evidencing the Israeli occupation forces use of surveillance technology, artificial intelligence, and other digital tool to determine who, what and when to attack in Gaza trip. Thus, this violates the principles of international humanitarian law which emphasize the necessity of distinguishing those in active combat and not[2] and to take necessary precautions when conducting an attack to minimize civilian harm.

[1] Accessnow. (October 2024) Big Tech and the risk of genocide in Gaza: what are companies doing? Available at https://www.accessnow.org/gaza-genocide-big-tech/.

[2] Article 48 of Additional protocol I of the Geneva convention.

IHLYI in this paper, responding to the request of the UN Secretary-General pursuant to a resolution A/RES/79/239, adopted by the General assembly on 24 December 2024 on Artificial intelligence in the military domain and its implication to international peace and security therefore, it analyzes AI In modern warfare, its implication to international peace and security and the role of technological companies in armed conflict.

Artificial Intelligence in Modern Warfare: A Legal and Humanitarian Perspective

The rules of international humanitarian law do not explicitly address the use of modern technological tools and artificial intelligence (AI) during armed conflicts. However, its core principles – such as distinction, proportionality, and precaution – remain applicable and binding on all parties. These principles require the differentiation between military objectives and civilians, and oblige parties to take all feasible measures to avoid or minimize harm to civilian populations. In recent years, militaries have contracted private companies to develop autonomous weapons systems. However, the armed conflict in Gaza stands out as one of the most prominent cases where commercially developed AI models – originally created in countries like the United States – have been employed in actual combat operations, despite the fact that these systems were not initially designed to make life-or-death decisions.

This shift highlights a troubling rise in the militarization of technology without clear legal or ethical oversight. While some of these tools may enhance operational efficiency, their unregulated use poses serious risks of human rights violations, especially amid a lack of transparency about how these tools function, the origin of the data they rely on, and the accuracy of their outcomes[3].

One of the most pressing concerns recently raised is the deployment of digital military tools based on unreliable data or flawed algorithms. Some of these systems depend on mass surveillance of Gaza's[4] population, including the collection of personal data prior to the outbreak of hostilities. Such practices raise legal and ethical questions regarding their compatibility with international obligations to safeguard privacy and prevent the misuse of personal information for the purpose of direct targeting.

Among the tools reportedly in use is a system that tracks population movement through mobile phone data to monitor evacuations from certain areas. Another generates lists of structural targets to be hit militarily. A third tool classifies individuals based on levels of suspicion regarding their affiliation with armed groups, while a fourth seeks to determine the precise location of a target in order to carry out a strike at the opportune moment. These tools largely rely on data extracted from mobile devices – whether through cell tower location information or GPS[5]. However, from a technical perspective, such data is insufficiently precise to confirm an individual's presence at a specific location at a given time, particularly in conflict zones where individuals frequently change phones or numbers. Over-reliance on this technology may lead to fatal mistakes, especially when a mobile phone is used as a substitute for verifying a person's actual presence in a targeted area. Legally, the use of such systems without taking all feasible precautions to protect civilians constitutes

[3] Human Rights Watch, "Israel: AI-Powered Targeting Systems May Be Committing War Crimes in Gaza", 2024.

[4] Associated Press, "Documents Reveal Israel's Use of AI Tools in Targeting Gaza", Investigative Report, 2024.

[5] Human Rights Watch (2024). Questions and Answers: Israeli Military's Use of Digital Tools in Gaza Available at Questions and Answers: Israeli Military's Use of Digital Tools in Gaza | Human Rights Watch.

a clear violation of international humanitarian law – particularly Article 57[6] of Additional Protocol I to the Geneva Conventions, which obliges parties to take constant care to spare civilian lives during military operations.

Given this reality, urgent questions must be raised about the future of AI in warfare and the legislative and legal mechanisms needed to regulate it. Without proper oversight, these tools risk becoming instruments of systematic human rights abuses rather than technologies aimed at ensuring greater protection for those affected by war.

Implications of Artificial Intelligence on International Peace and Security

Armed conflicts in various regions around the world, such as Gaza, Lebanon, Syria, Ukraine, and Libya, have had catastrophic humanitarian and security consequences. These conflicts have led to the mass displacement of civilian populations, depriving thousands of people of their basic rights such as food, water, shelter, and healthcare. These individuals live in dire humanitarian conditions, with a significant increase in deaths due to famine, thirst, and diseases caused by contaminated water, in addition to exposure to harsh weather conditions without protection.

In this context, the increasing use of artificial intelligence and drones as weapons in conflicts, particularly by Israel in the Gaza Strip[7], stands out. Since October 2023, there has been a notable escalation in the use of "quadcopters" to carry out precise and targeted strikes against civilians. These drones are equipped with data analysis algorithms and offensive capabilities, enabling them to target individuals based on tracking their movements or mobile phone signals.

According to documented reports, this technology has led to the death of more than 1,000 Palestinians by May 2024, including a significant number of women and children. This constitutes a grave violation of international humanitarian law, particularly Articles 51 and 57 of Additional Protocol I to the Geneva Conventions, which prohibit attacks on civilians and obligate parties to the conflict to take all necessary precautions to avoid harming them.

The concerns are not limited to the use of artificial intelligence against individuals but extend to the misuse of data. Relying on mobile phone tracking technologies (either through GPS data or cell tower signals) to pinpoint individuals' locations presents serious risks. Recent studies have shown that these systems do not provide enough accuracy to reliably determine someone's location, especially in conflict zones where phones may be swapped or disconnected frequently. This means that relying on these methods without field verification can lead to erroneous decisions, resulting in unlawful killings.

In a well-known case, a Palestinian woman named "Silah" was killed while carrying a white flag and leading her family to safety. After stepping onto a main street, she was targeted by a small drone that shot her in the head. This incident, witnessed by those around her, serves as a stark example of the disastrous outcomes of unregulated use of technology on the battlefield[8].

In Libya, drones played a decisive role in the battles between conflicting parties, particularly as many of these drones, including Turkish and Chinese models, were operated using data analysis systems to target objectives. Some of these systems are

[6] Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts (Protocol I), Article 57.

[7] TRTWORLD (2024) Quadcopter strikes: 1000 Palestinians killed by Israeli drones in one year. Available at Quadcopter strikes: 1000 Palestinians killed with drones in a year.

[8] Gaza grandmother gunned down by Israeli sniper as child waved white flag," *Times Kuwait*, November 2024, https://timeskuwait.com/news/gaza-grandmother-gunned-down-by-israeli-sniper-as-child-waved-white-flag.

believed to rely on artificial intelligence techniques for targeting, without legal oversight. The use of these tools in urban areas like Tripoli and Sirte has led to the deaths of civilians and extensive damage to infrastructure[9].

All of these events indicate that integrating artificial intelligence into managing and directing armed conflicts without an internationally binding legal framework to regulate its use could open the door to widespread violations, especially if these systems are not subject to independent and transparent oversight to ensure compliance with international humanitarian law and human rights.

Roles of Companies Developing AI in Armed Conflicts

Through a rapid increase in artificial intelligence and computer services, U.S. tech corporations have discreetly given Israel the ability to monitor and kill many more militants in Gaza and Lebanon more quickly. However, the death toll among civilians has also skyrocketed, raising concerns that these instruments may be causing the deaths of innocent people. Israel's recent wars are a leading example of commercial AI models developed in the United States being used in active warfare, despite concerns that they were not originally designed to help decide who lives and who dies.

For years, militaries have hired private companies to create customized autonomous weapons. Numerous American software companies have backed Israel's battles in recent years, including Microsoft and the San Francisco-based startup OpenAI. Under "Project Nimbus," a $1.2 billion contract signed in 2021[10] when Israel first tried out its in-house AI-powered targeting systems, Google and Amazon offer cloud computing and artificial intelligence services to the Israeli military. The military has made use of Dell and Cisco data centers and server farms. Palantir Technologies, a Microsoft partner in U.S. defense contracts, has a "strategic partnership" that provides AI systems to support Israel's war efforts, while Red Hat, an independent IBM company, has also supplied cloud computing technologies to the Israeli military.

Furthermore, through a number of programs, Microsoft also supplies Israel's government with services that have allegedly been used to help the Israeli military, police, Israeli Prison Service (IPS), and illegal settlement operations. Over 10,000 Palestinians are being held by the IPS as of October 2024; half of them have been detained without being charged or having a trial date scheduled. At least 310 medical professionals, UN employees, women, and children are among the Palestinian prisoners from Gaza who are presently detained in prolonged, secret, and incommunicado detention, where they are subjected to torture, mistreatment, and sexual violence and abuse, according to the UN Human Rights Office.

Companies are under obligation to respect human rights within their scope of operations. Companies that directly aid the offender – for example, by offering financial, logistical, military, or intelligence support – may be held criminally responsible for a crime committed during an armed conflict. Companies and their managers or executives may be held accountable in certain situations even if they had no direct involvement in the crime or no intention of supporting it. As the Office of the High Commissioner on Human Rights (OHCHR) noted, companies "should treat

[9] France 24. (2021). "Have Killer Drones Been Deployed in Libya?". France 24. Retrieved from https://rb.gy/1m6k43.

[10] APNEWS (2025). As Israel uses US-made AI models in war, concerns arise about tech's role in who lives and who dies. Available at How US tech giants' AI is changing the face of warfare in Gaza and Lebanon | AP News.

this risk in the same manner as the risk of involvement in a serious crime, whether or not it is clear that they would be held legally liable[11]."

In light of the concerns raised in this submission and their implications for international peace and security, IHLYI urges states to:

1. **Refrain from the use of AI in military applications**: States should immediately halt the use of artificial intelligence in military activities and establish national regulations and laws to prevent its deployment in warfare.

2. **Work towards a global ban on the military use of AI**: States should actively pursue international agreements and frameworks to ban the use of AI in military contexts, ensuring that no country utilizes AI for warfare.

3. **Avoid the development of autonomous and AI-enabled weapon systems**: States should refrain from developing autonomous weapon systems or AI-powered weaponry that could be used to target humans, ensuring human oversight and decision-making in military actions.

4. **Ensure the protection of personal data**: States must guarantee that personal data is protected from misuse by military forces, law enforcement agencies, border control, and private contractors collaborating with these entities.

5. **Promote accountability in AI development**: Technology companies, researchers, engineers, and financial institutions should commit to not supporting the development or funding of AI technologies designed for military applications, advocating for responsible innovation in line with humanitarian principles.

Peace Movement Aotearoa and Stop Killer Robots Aotearoa New Zealand

[21 May 2024]

Peace Movement Aotearoa and Stop Killer Robots Aotearoa New Zealand welcome the opportunity to contribute our views to the UN Secretary-General's report on artificial intelligence (AI) in the military domain and its implications for international peace and security. Our submission briefly outlines our involvement in this issue, and has three sections summarising our position on: a) A new international instrument on military use of AI and autonomy in weapon systems is urgently needed; b) Key focuses of a new international instrument; and c) Scope of a new international instrument. The points below are based on discussions with our member and supporting groups about the content of this submission.

Introduction

Peace Movement Aotearoa is the national networking peace organisation in Aotearoa New Zealand, established in 1981 and registered as an Incorporated Society in 1982. Our purpose is networking and providing information and resources on peace, humanitarian disarmament, human rights and social issues; and we have extensive national networks of member and supporting groups and individuals. We are a founding member of the Stop Killer Robots campaign and coordinate the national Stop Killer Robots Aotearoa New Zealand (SKRANZ) campaign.

SKRANZ was launched in April 2013 to support the global campaign, with a specific national focus on urging New Zealand to take national action to prohibit the development, production and use of autonomous weapon systems; and to take

[11] Accessnow (2024) Big Tech and the risk of genocide in Gaza: what are companies doing? https://www.accessnow.org/gaza-genocide-big-tech/.

international action to support negotiations on a new treaty to prohibit autonomy in weapon systems. Since 2023 we have widened our focus to include military use of AI as its perils became increasingly obvious.

(a) A new international instrument on military use of AI and autonomy in weapon systems is urgently needed

As outlined in our submission for the UN Secretary-General's report on autonomous weapon systems (A/RES/78/241) last year, it has been clear for some years now that rapidly developing technological advances in the use of force and increasing autonomy in weapon systems pose an unprecedented threat both to humanity and to the foundations of international human rights and humanitarian law, which are based on respect for human life and dignity, protection of humanity in times of oppression and armed conflict, and human responsibility and accountability for harm.

The serious ethical, humanitarian, legal, and security concerns posed by these developments have been discussed for more than a decade within United Nations bodies – including the Human Rights Council, meetings related to the Convention on Certain Conventional Weapons and in the UN General Assembly – as well as in regional and national governmental and non-governmental forums.

Even as these discussions have taken place, some states have increasingly incorporated autonomy into military use of force in ways that have already resulted in gross violations of international law with disastrous consequences for civilian populations. It is apparent that the absence of specific international law on autonomy in weapon systems, and with differing interpretation by some states as to how existing law applies to new technological developments, the risk of proliferation of ever more dangerous and uncontrollable weapon systems is increasing rapidly.

The need for urgency for international action on this has been highlighted over the past eighteen months by, for example, Israel's use of AI-powered target suggestion systems in Gaza to make high explosive strikes on numerous targets possible in a short time frame, resulting in indiscriminate slaughter of civilians and systematic destruction of life-sustaining infrastructure. The reality of digital dehumanisation with catastrophic consequences is now very evident, as is the increasing tendency towards the development and use of autonomous weapon systems that will remove any remaining vestige of humanity from war.

We have noted with concern that states who brought forward A/RES/79/239 include states that have armed and supported Israel's genocidal attacks on Gaza, and where big data tech companies contributing data storage and AI capabilities to Israel's military systems are based.

Similarly, 'responsible AI in the military domain' (surely an oxymoron) is being promoted by states already developing their own AI targeting and autonomous weapon systems, as a way of undermining the push towards a binding instrument to prohibit these critical threats to international peace and security.

The US 'Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy' has highlighted for us the risk of horizontal proliferation of both military use of AI and autonomous weapon systems as states that do not have their own capability in this regard move from interoperability to integration with the states of armed forces that do or that are developing it. In the case of New Zealand, for example – as it seeks to be a 'combat capable force multiplier with enhanced lethality' [12] – this involves closer military integration with the armed forces of

[12] See, for example, the 2025 Defence Capability Plan released this week.

Australia and the US: New Zealand endorsed the US 'Political Declaration' early last year specifically to be compliant with US military doctrine.

These unfortunate developments underscore the urgent need for a new international instrument on military AI and autonomy in weapon systems to clarify and strengthen existing law. The instrument must include both prohibitions and regulations, as outlined below, and must include military use of AI in combat.

As emphasised in the UN Secretary-General's 2024 report on autonomous weapon systems[13], negotiations on a new instrument must begin without any further delay, in a multilateral forum where states can come together to work constructively, where the voices of those whose lives have already been impacted by military use of AI and increasing autonomy in weapon systems can be heard, and where UN agencies, the International Committee of the Red Cross (ICRC), and NGOs are active participants.

(b) Key focuses of a new international instrument

While much of the work around military use of AI and autonomous weapon systems has focused on the issue of meaningful human control over the use of force, it is our view that the key underlying ethical imperative is preventing human beings from being targeted or attacked by any system utilising digital code and/or sensors. A prohibition on military use of AI and autonomy in weapons systems that are designed or used to target human beings must be the starting point.

Meaningful human control over the use of force clearly has an ethical component, but it is also a practical and legal means to ensure accountability for any autonomy in weapon systems that breach the key dictates of humanitarian law.

(c) Scope of a new international instrument

It is our view that a new international instrument should include overarching rules to establish a framework for evaluating current and future technological developments, while promoting increased compliance with international human rights and humanitarian law.

Such overarching rules would prohibit autonomous weapon systems that are designed or used to target humans, and lay out specific obligations to ensure meaningful human control over other systems: for example, that the human operator/s understand the capabilities and limitations of the system, are able to fully evaluate the context in which the system will be used, and are making mindful firing decisions rather than assuming the technology is accurate – this would act to regulate autonomy in weapon systems. It would be useful to specify that decisions made by states on their assessment of new or altered weapon systems that incorporate autonomous features or functions must be transparent.

Furthermore, in the context of the UN Secretary-General's forthcoming report on AI in the military domain and in the light of the awful consequences of military use of AI in Gaza, the scope of a new international instrument must go beyond autonomous weapon systems. It is very clear that there is a spectrum of harmful military use of autonomy, ranging from target decision support systems (as some have described systems such as Lavender), data-based targeting systems, generation of target lists by algorithm or AI, sensor-based targeting systems, through to weapon systems that combine these elements and incorporate varying degrees of machine learning to make target selection decisions and attack autonomously.

We note the 2023 Joint Call by the UN Secretary-General and ICRC President stated *"The autonomous targeting of humans by machines is a moral line that we must*

[13] Lethal autonomous weapons systems: Report of the Secretary-General (A/79/88), 1 July 2024.

not cross"[14], yet that has already happened – a point reiterated in the UN Secretary-General's 2024 report[15].

It is therefore our view that a new instrument must cover military use of AI – including systems that automate significant decision-making in the use of force, such as target generation, force deployment, and engagement – as well as autonomous weapon systems.

Finally, although we have referred in this submission to military use of AI and autonomy in weapon systems, prohibitions and regulations in a new international instrument must also apply to all coercive agencies of the state, including those used for policing and internal security, for border control, in corrections facilities and in places of detention.

Ploughshares

[11 April 2024]

Project Ploughshares, a Canadian peace research institute, has for over a decade focused its advocacy and research on the military applications of emerging technologies, including artificial intelligence (AI) and autonomous weapons. As AI systems are rapidly advancing and being tested in contemporary conflict zones, international governance frameworks have struggled to keep pace. Meanwhile, intensifying geopolitical competition increases the likelihood that AI technologies will be deployed in complex, dynamic environments for which they are not suited – raising significant risks for civilians.

The wide-ranging use of AI in military applications demands urgent and coordinated international attention. We encourage the Secretary-General and member states to focus on three particularly pressing areas: the use of AI in decision-support systems related to the use of force, the dual-use nature of AI technologies, and the widening capacity gap among states engaging in multilateral discussions.

AI decision-support systems

One area that remains insufficiently addressed in current international discussions is the use of AI in military decision-making, especially decisions about the use of force. Of particular concern are AI-enabled targeting tools such as "Lavender" and "Gospel," reportedly used in in Gaza. These systems are classified as "decision support" because a human is technically required to approve target selections. However, there is little transparency regarding how these decisions are made, how frequently AI-generated recommendations are rejected, or whether human operators fully understand how the AI systems reach their conclusions.

In practice, these systems raise the risk of "rubber-stamping," in which human oversight becomes superficial, thereby undermining the principle of meaningful human control and increasing the likelihood of harm to civilians. The potential use of such AI systems in early-warning, surveillance, reconnaissance, and nuclear command-and-control systems further amplifies these concerns.

To mitigate these risks, states must work toward clear norms, regulations, and training requirements that enhance operator understanding, counter automation bias, and ensure genuine human engagement in decision-making processes.

[14] Joint call by the United Nations Secretary-General and the President of the International Committee of the Red Cross for States to establish new prohibitions and restrictions on Autonomous Weapon Systems, 5 October 2023.

[15] As at note 3.

Dual-use challenges

AI's dual-use nature – its applicability to both civilian and military domains – creates further governance complexity. Civilian-developed technologies can be repurposed for military use without appropriate testing or safeguards, increasing the risk of conflict escalation, misuse, and error. Additionally, the accessibility of certain AI tools means that nonstate armed groups may also gain access, potentially using them to target civilians and infrastructure.

We urge states to develop policy mechanisms, including export controls, technology impact assessments, and multistakeholder engagement, to account for dual-use risks and promote responsible innovation.

Capacity- and knowledge-building

Current multilateral discussions reveal stark capacity disparities among states, many of which do not have the resources or technical expertise to participate meaningfully in governance efforts. To ensure inclusive and equitable global engagement, we recommend that states collaborate with the UN Office for Disarmament Affairs to strengthen capacity-building initiatives.

The scientific and academic communities also have a role to play in supporting the development of accessible resources and training materials. International forums, such as the upcoming REAIM Summit in Spain, should include dedicated sessions for knowledge-sharing, especially to support representatives from under-resourced states.

Final thoughts

The international community is at a crossroads. The accelerating militarization of AI demands robust diplomatic responses. We can – and must – move from aspirational principles to concrete, enforceable frameworks, by employing political will, inclusive dialogue, and cross-sector collaboration.

AI-powered warfare is no longer a theoretical risk; it is a present reality. Whether this new era enhances global security or undermines it will depend on the steps states take now to strengthen governance, manage technological competition, and uphold international humanitarian norms.

Without timely, coordinated action, the risks of accidental escalation and unintended conflict will only increase.

Soka Gakkai International

[10 April 2025]

The Soka Gakkai International (SGI) welcomes the opportunity to share our views on the important issue of artificial intelligence (AI) in the military domain. As an NGO whose work is guided by Buddhist principles, we urge that the United Nations, its Member States and other stakeholders take into careful consideration the impact of AI in the military domain from a standpoint of upholding and respecting human dignity.

Introduction

AI in the military domain is rapidly evolving and transforming modern warfare and international peace and security. These systems are being used for various purposes, including surveillance, autonomous weapons, decision-making support, and logistics. With such wide-ranging applications, the integration of AI technologies in military systems poses significant challenges. To better ensure compliance with

international humanitarian law (IHL) and uphold protection for civilians and combatants alike there are several issues that we may consider.

Lack of transparency and accountability

- If an AI system were to make an error – such as identifying a target incorrectly – it could be difficult to pinpoint the cause of the error, "the black box problem". Was it a flaw in the data used to train the AI, an issue with the algorithm, or a problem in the operational context? Without transparency within these systems, assigning responsibility is difficult.
- International laws and treaties, such as the Geneva Conventions, were created before AI systems became commonplace in warfare. Without global norms and legal frameworks, there is no consistent approach to ensuring accountability for AI decisions made in warfare.
- With inadequate accountability mechanisms in place, AI could be used for military strategies that violate human rights, suppress civil liberties, or engage in unethical operations.

Speed of decision-making and risk of escalation

- The ability of a military force to make decisions and execute actions faster than its opponent is increasingly viewed as having a strategic advantage. However, the drive for speed can lead to unintended and costly consequences.
- Decisions made too quickly without proper analysis or consideration can lead to poor outcomes, including tactical blunders, strategic missteps, or ethical violations.
- Instead of diffusing a tense situation or negotiating, if combatants react too quickly it could provoke an even greater confrontation, further escalation and prolonged conflict resulting in more human suffering including amongst civilians.
- The acceleration of decision-making processes closes down the possibility of meaningful human control, the growing trend to automate decision-making threatens the ability to achieve human oversight which is essential to facilitate compliance with IHL.

Bias in AI in the military domain

- AI bias refers to the presence of systematic and unfair discrimination in AI systems, such as historical bias, where systems may reinforce harmful stereotypes, bias in data processing and algorithm development which can lead to making biased decisions and bias in how the systems are used.
- AI bias in the military domain is a significant concern, particularly as AI systems are increasingly being integrated into defense and security operations. The potential for AI bias to emerge in these areas can result in human rights implications, exacerbating existing inequalities and lead to deadly consequences for certain groups.
- AI heavily relies on vast amounts of high-quality and reliable data for decision-making. There are several potential violations when it comes to obtaining this data including issues around privacy and surveillance, challenges of bias also arise when dealing with incomplete and inaccurate data.
- When AI systems are biased, they not only perpetuate inequalities but also contribute to the digital dehumanization[16] of marginalized groups.

[16] Digital dehumanization is a process where humans are reduced to data, which is then used to make decisions and/or take actions that negatively affects their lives.

Proliferation

- Nations may rush to develop AI-based military technologies to outpace their adversaries, which could lead to a destabilizing arms race and increased global tensions.
- Without regulation autonomous weapons systems in particular, could proliferate globally, including amongst non-state actors which could increase crime nationally and regionally, exacerbating social inequalities, overwhelm resources and infrastructures of countries, as well as undermine social and national security.

Conclusion

The issue of AI within the military contexts is complex, and without regulation, it could lead to serious consequences for global peace and security. The desire to speed up decision-making processes within this context has yet to be proven as an effective way of resolving conflicts and achieving peace and security. Furthermore, you cannot divorce AI in the military and AI in civil uses, a failure to address AI in a military context could have widespread repercussions in all spheres of civil life including law enforcement, border control, education, housing and health care. Fundamentally, AI is here to stay, how we utilize it in the military and in our lives will shape the course of humanity. We have the possibility and the responsibility to decide how we want to use technology, knowledge, and the world's resources. To use it in a way that uplifts humanity or degrades it? This is an urgent question that requires moral, ethical and courageous leadership.

Stop Killer Robots

[11 April 2025]

The Stop Killer Robots campaign welcomes the opportunity to submit our views to the United Nations Secretary-General in response to Resolution A/RES/79/239.

Established in 2012, we are a coalition of more than 270 non-governmental organisations working across 70 countries.[1] We seek to counter threats to humanity and human dignity through the adoption of a new international treaty to prohibit and regulate autonomous weapons systems.[2] We support the development of legal and other norms that ensure meaningful human control over the use of force, counter digital dehumanisation, and reduce automated harm.[3]

Building an effective international response to emerging technologies

Autonomous weapons systems, 'AI in the military domain,' and trends and developments in increasingly automated decision-making and action in the use of force – as well as in our lives and societies more broadly – are all part of the same concerning picture:

The growing influence of computer processing and algorithmic thinking increasingly shapes our interactions in the world and the outcomes available to us. There are clear threats to peace, justice, dignity, human rights, equality, responsibility and accountability, and respect for law. We are getting closer to machine processes determining whom to kill.

[1] See www.stopkillerrobots.org/about-us and www.stopkillererobts.org/a-global-push/member-organisations.
[2] See https://www.stopkillerrobots.org/our-policies/.
[3] See www.stopkillerrobots.org/vision-and-values/.

To address these challenges effectively, a comprehensive and holistic response is needed from the international community.

Adopting a legally binding instrument on autonomous weapons systems will be one critical component: we must draw basic red lines for humanity against the automation of killing, which brings under jeopardy both international humanitarian law and international human rights law, in particular the presumption of innocence, the right to equality and non-discrimination, dignity, and wipes away contextual circumstances of the target(s) in question. The UN Secretary-General's comprehensive report last year reiterated his urgent call on states to negotiate a legally binding instrument to prohibit and regulate these systems by 2026.

But, a new international treaty on autonomous weapons systems alone may not be enough. States must also reach agreement on preventing and addressing grave harm from other uses of emerging technologies. A whole set of strong international rules are needed that stop the erosion of meaningful human control and the slide towards greater digital dehumanisation and automated harm, across international and domestic practice, in armed conflict and in civilian life.

'Military applications of AI' are already contributing to civilian harm

The risks of integrating AI into the use of force in armed conflict reach far beyond those to peace and security between states: a holistic consideration of peace and security that considers dimensions such as ethical, legal, and humanitarian issues must be taken into account in the UN Secretary-General's report under resolution 79/239.

We are already seeing grave threats to civilian protection and human rights and huge harm being caused by AI and automation in the use of force. This is arising from the quest for speed in warfare, the reduction of people to objects, and issues such as automation bias that Stop Killer Robots has raised the alarm about for years.

We have been horrified by reports of the use of AI-powered 'decision support systems' by Israel in Gaza, which suggest human targets to strike.[4] According to reports, human approval of these suggestions in vast volumes at high speed has been minimal – entailing digital dehumanisation, the erosion of meaningful human decision-making and control (including through automation bias), and directly contributing to massive and devastating harm to civilians in Gaza, alongside other tools.[5]

Many other states are developing and using such 'decision support systems', which raise concerns around international humanitarian law, human rights law, and ethics. So far there are few reports on how these are being deployed, with what constraints and with what impacts. Nevertheless, the push by many states to develop and integrate AI and autonomy into decision-making and the use of force is a huge concern. The further use in hostilities of these kinds of tools by any state in the unacceptable ways that we have seen in Gaza must be prevented. Stop Killer Robots struggles to see how such uses could meet the definition of the responsible application of AI in the military domain given in resolution 79/239.

Further risks to peace and security, rights, and human dignity

The quest for greater speed through AI and automation – towards the goal of increasing the tempo of conflict to a point beyond human cognition in the pursuit of a military and strategic edge – is an extremely dangerous one for international peace

[4] 'Lavender': The AI machine directing Israel's bombing spree in Gaza, +972 Magazine https://www.972mag.com/lavender-ai-israeli-army-gaza/.

[5] Questions and Answers: Israeli Military's Use of Digital Tools in Gaza, Human Rights Watch, https://www.hrw.org/news/2024/09/10/questions-and-answers-israeli-militarys-use-digital-tools-gaza.

and security. These risks are further to the impact 'AI in the military domain' is already having on civilian protection. Risks include unwanted escalation, lowered political thresholds to the use of force, and arms race dynamics.

Technologies that can contribute to target selection (such as threat detection tools) and remote biometric surveillance (such as facial recognition) have already had documented negative impacts on human rights such as the rights to privacy, equality and non-discrimination, freedom of expression and peaceful assembly, and the freedom of movement. In the case of facial recognition for identification (1:n), the technology is considered by many legal experts as wholly incompatible with international human rights law.

That AI systems inevitably encode and reproduce the biases of our societies – including racism, sexism and ableism – and that such bias cannot be eliminated, is also well established. The use of such systems to process people in the use of force will inevitably lead to disproportionate – and multiplied – impacts on already marginalised and minoritised people. Integrating automation and AI into decisions and actions in the use of force against people contributes to digital dehumanisation – the process where humans are reduced to data, which is then used to make decisions and/or take actions that negatively affects their lives.

The relationship with autonomous weapons systems

Stop Killer Robots notes that the UN Secretary-General's report will be on the "application of artificial intelligence in the military domain, with specific focus on areas other than lethal autonomous weapons systems." It is important nevertheless to highlight that various applications beyond the boundary of autonomous weapons systems are closely linked to them.

Firstly, such tools could be integrated as components of autonomous weapons systems now or in the future. For example, a 'decision support system' could be used as an autonomous targeting system, connected to a platform tasked to strike targets on the list generated, based on processing sensor data. Secondly, these tools are linked not only practically, but raise and are part of the same picture of concern. Strikes undertaken based on the nominal human approval of targets generated by a decision support system do not sit far from strikes undertaken with an autonomous weapons system.

It is therefore important that states consider these issues in dialogue: many of the rules and principles developed for autonomous weapons systems on keeping control and rejecting automated killing will need to be extended (with adaptations) to other tools; and, how the development of AI in the military domain more broadly will impact the direction and challenges posed by autonomous weapons systems will need consideration.

Recommendations

Technologies incorporating AI and automation into the use of force in armed conflict are currently being deployed without specific agreed rules; the principles various states have proposed and committed to so far have been too weak and vague to prevent civilian harm and risks to peace and security.

All developments in autonomy and AI in the use of force which threaten our safety, security, and humanity must be urgently and adequately addressed through strong regulation by the international community, with unacceptable uses prevented.

States must:

- Move with urgency to negotiate and adopt a new international treaty to prohibit and regulate autonomous weapons systems;

- In International discussions, critically and meaningfully engage with the implications and real-world consequences of current practice in the use of tools that fall under ‘AI in the military domain,’ including acknowledging and examining humanitarian harm;
- Fully consider the legal, ethical, humanitarian, and peace and security risks of further development and use of such systems, whatever the perceived ‘benefits’ may be
- Work urgently to prevent unacceptable uses of technology and trends in development, through committing to develop strong norms for meaningful human control and against digital dehumanisation:
  - This should take place domestically, regionally, and internationally.
  - It must involve a comprehensive and holistic international response, including a legally binding instrument prohibiting and regulating autonomous weapons systems alongside other measures.
  - It should include consideration and development of the other legal instruments necessary to preserve meaningful human control and to protect human dignity against AI in the use of force.

Stop Killer Robots Youth Network

[10 April 2025]

The Stop Killer Robots Youth Network welcomes the opportunity to submit recommendations for consideration by the United Nations Secretary-General in response 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. As a global network of young people under age 30 in over 50 countries working to secure a future free of automated killing, we have consistently advocated for the creation of a new treaty on autonomous weapons systems (AWS) – in particular, we insist on a total prohibition of anti-personnel autonomous weapons as we wish to build a world without such dehumanising weapons. While youth will inevitably face the risks of new weapons technologies, we remain underrepresented in the decision-making process and are often sidelined in forums that shape our interests. As youth who have grown up in an increasingly digital world, we wish to create a future where technology is used to promote peace, justice, equality, and human rights, not perpetuate violence.

With escalating conflicts and the rapid deployment of new weapons technologies around the world, there is an urgent need to reinvest in international law as a measure to build trust and achieve sustainable peace and security. The application of artificial intelligence (AI) in the military domain presents numerous challenges that concern us as youth, including digital dehumanisation, the gamification of violence, and the further erosion of human control and involvement over the use of force.

Military AI & AI systems already in use

Artificial intelligence has been progressively implemented in the military domain over the past decade, however, due to the opacity of military activities and development, the wide public has not been aware of this issue until recently when the active uses of AI systems have been mediatized. We have seen and monitored the use of AI systems to support the targeting of both objects and people. Unfortunately, the use of such systems have not been able to alleviate civilian suffering, for example, in Gaza where one third of victims are children and where too many civilian infrastructures, including critical infrastructures such as humanitarian camps,

hospitals[1], and schools[2], have been either directly targeted or indirectly impacted by the hostilities.

There have been other concerning uses[3] of AI systems outside of the military which need to be considered as they might be implemented in the military domain, mainly predictive AI and facial recognition. Predictive AI technologies have been used in the police and judicial systems since the early 2010s and have been shown to be ineffective, incorrect, and subject to reinforcing discriminatory behavior.[4] If predictive AI were to be implemented in the military domain, it could lead to the increasing risk of civilians being targeted as they could be labeled as possible fighters or being indirect victims of military activities due to the multiplications of targets with predicted military advantages. Facial recognition technologies (FRTs) are also of concern as they are also unreliable especially when it comes to the identification of non-white males. Facial recognition-enabled targeting in military operations must be prohibited as those systems cannot comprehensively analyse every factor that makes military personnel or civilians a target or not.

Digital dehumanisation

One of the main concerns we have about the use of AI systems in the military domain is the proliferation and banalisation of **"Digital dehumanisation"**. We define digital dehumanisation as the process whereby humans are reduced to data, which is then used to make decisions and/or take actions that negatively affect their lives. This process deprives people of dignity, demeans individuals' humanity, and removes or replaces human involvement or responsibility through the use of automated decision-making in technology.[5] Additionally, the increased speed and scale of target production through military AI erodes moral restraints in war and lowers the impact and capacity of decisions from human operators[6], thus enabling the AI systems to make decisions without meaningful human control, which further dehumanises the decision-making process.

Relying on (Big) data leads to problems

We also believe that the use of (big) data in the military leads to multiple issues which need to be considered.

One of the primary issues is the challenge of data labeling – the process of categorizing and tagging data to train algorithms. Inaccurate or biased labeling can have far-reaching consequences, particularly in the context of distinguishing between

[1] World Health Organization (2025), 'oPt Emergency Situation Update'. https://www.emro.who.int/images/stories/Sitrep_57.pdf.

[2] Save the Children (2025), 'Education Under Attack In Gaza, With Nearly 90% Of School Buildings Damaged Or Destroyed'. https://www.savethechildren.net/blog/education-under-attack-gaza-nearly-90-school-buildings-damaged-or-destroyed.

[3] Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ProPublica (2016), 'Machine Bias'. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

[4] Will Douglas Heaven, MIT Technology Review (2020), 'Predictive policing algorithms are racist. They need to be dismantled'. https://www.technologyreview.com/2020/07/17/1005396/predictive-policing-algorithms-racist-dismantled-machine-learning-bias-criminal-justice/.

[5] Automated Decision Research (2022), 'Autonomous weapons and digital dehumanisation'. https://automatedresearch.org/news/report/autonomous-weapons-and-digital-dehumanisation-a-short-explainer-paper/.

[6] Marta Bo and Jessica Dorsey, OpinioJuris (2024), 'Symposium on Military AI and the Law of Armed Conflict: The 'Need' for Speed – The Cost of Unregulated AI Decision-Support Systems to Civilians'. http://opiniojuris.org/2024/04/04/symposium-on-military-ai-and-the-law-of-armed-conflict-the-need-for-speed-the-cost-of-unregulated-ai-decision-support-systems-to-civilians/.

combatants and non-combatants in conflict zones. If the data used to train military AI systems is flawed or biased, it can lead to disastrous mistakes, such as the targeting of innocent civilians or misidentification of threats.

A critical issue when relying on big data is that the nature data itself is often broken and is incomplete. This means that the data used to train AI models can be incomplete, outdated, or unrepresentative of real-world situations. Such flaws in data can lead to systems that fail to generalize properly, resulting in inaccurate or incorrect predictions and decisions. For example, in combat situations, a lack of diversity in the data used to identify individuals could lead to inaccurate targeting, with devastating consequences. Important data might be missing or poorly represented, such as the exact location of civilians or combatants, which can lead to AI failing to make informed and balanced decisions. In a war scenario, a system trained with data from a specific past conflict may not be capable of handling a new, unpredictable situation. For instance, an AI system that has been fed data from one particular type of conflict might struggle to apply that data to a war with entirely different characteristics, resulting in errors in target identification or incorrect decision-making.

Another significant problem is that many AI systems operate as black boxes. This means that while these systems make decisions and predictions based on the data they process, the decision-making process is not transparent or easily understood. In military scenarios, where the consequences of decisions are extremely serious, the lack of transparency is particularly concerning. If an AI system makes an error, such as wrongly identifying a civilian as a combatant, the absence of clarity about how the system reached that conclusion makes it nearly impossible to understand the origin of the error. This makes accountability difficult, as we cannot determine why the system acted in a particular way. The lack of explanation regarding the decision-making processes of AI also makes it impossible to correct or adjust the system's behavior, potentially perpetuating errors without the ability to fix them effectively.

Linguistic and cultural bias embedded in data which is used to train AI systems can create security vulnerabilities and catastrophically misinterpret communications, behaviors, and intentions across diverse cultural contexts, potentially triggering lethal automated responses to misunderstood signals.[7] These systems risk automating and amplifying existing prejudices at unprecedented scale and speed with life or death consequences in conflict zones where cultural misunderstandings could rapidly escalate into devastating military actions causing dire consequences.

Accountability

The inclusion of AI systems in the command and decision-making chains will indubitably lead to a lack of accountability and liability for those relying on these systems to make decisions. It will create a sense of distance and lack of liability on the consequences of a decision which mean that decisions may be made without specific, consistent and thorough analysis of the lawfulness and humane characters of the decision. Then, if an action taken using AI systems violates IHL, the people involved in the implementation and those involved in the decision-making should be held accountable and the use of an AI system shall never exempt people from their responsibilities.

We recognize that military operations are bound by multiple bodies of law – national law, International Humanitarian Law (IHL) and International Human Rights

[7] Jimena Sofía Viveros Álvarez, Humanitarian Law & Policy (2024), 'The risks and inefficacies of AI systems in military targeting support'. https://blogs.icrc.org/law-and-policy/2024/09/04/the-risks-and-inefficacies-of-ai-systems-in-military-targeting-support/.

Law (IHRL) – which need to be respected and implemented in order for operations to be lawful. Unfortunately, rules of engagement and of targeting – and all the exceptions – cannot be fully understood and implemented by AI systems. Concepts like doubt, proportionality, and the balance between humanity and necessity are inherently human judgments that cannot be captured by an algorithm. Machines cannot be trusted to uphold these standards on their own. Therefore, it is critical that AI systems never act in a vacuum and that humans retain oversight and decision-making power at all times.

What the future might look like

While AI theoretically has the potential to enhance precision and efficiency in military operations, its integration into warfare raises significant concerns about the future of global security. Autonomous weapons systems, capable of making life-or-death decisions without human control, introduce ethical dilemmas and risks of unintended consequences. The use of AI in military technology is likely to aggravate the existing arms race, as nations compete to develop increasingly sophisticated AI systems, widening the power gap between technologically advanced countries and those less developed, leaving them vulnerable in terms of military readiness. The deployment of autonomous weapon systems and AI-driven tools makes conflict more unpredictable, scalable, and asymmetric, granting certain nations the ability to unleash devastating technologies that smaller states or non-state actors may not be able to counter. The proliferation of AI in the military sphere also raises the threat of terrorism, as organized actors could easily access advanced AI-powered systems. Moreover, the fast-paced, constantly evolving nature of AI development turns military strategies into a “cat and mouse” game, where advancements are met with equally rapid countermeasures. In light of these challenges, the future of military AI must be handled with extreme caution, emphasizing robust ethical frameworks, international regulations, and stringent human oversight to prevent these technologies from destabilizing global peace.

What we need

We call for the establishment of a meaningful legally binding instrument for the use of AI-driven systems in the military requires comprehensive integration of the technical sector alongside state actors, addressing the urgent need for standardized verification protocols and trust-building mechanisms between nations. Such an instrument should define clear autonomy thresholds that specify permissible levels of independence in target selection and engagement, mandate extensive documentation of algorithmic decision processes and testing methodologies and establish explicit red lines that cannot be crossed including prohibited deployment scenarios, target categories, and operational environments. This framework should apply consistently across developing and developed nations, incorporate independent verification bodies with appropriate technical expertise to conduct regular compliance audits, and establish enforcement mechanisms with meaningful consequences for violations, all while facilitating technical data sharing and research that builds confidence between stakeholders in this domain.

These systems present an unprecedented threat to global security and human rights, and the risks they pose to non-combatants are immense. It is crucial that it implements a robust framework of monitoring, accountability and oversight. Firstly, the states need to be bound by positive obligations to ensure the responsible use of AI in the military domain. Accountability is a fundamental aspect of this framework. We call for comprehensive mechanisms that oversee every stage of the AI system life cycle, from development and updates to transfers and research. States must ensure that any uses of AI systems are monitored, with clear reporting structures in place to

address incidents promptly. Furthermore, it is vital that human operators using these systems receive thorough training and guidance to make ethical decisions in the field. The principle of meaningful human control must remain central when it comes to the use of AI in the military domain to ensure that ultimate responsibility for any actions remains with human decision makers.

Unione degli Scienziati Per Il Disarmo

[6 April 2025]

Introduction

USPID (***U****nione degli* ***S****cienziati* ***P****er* ***I****l* ***D****isarmo, Union of Scientists for Disarmament*) is an association of concerned scientists – founded in 1983 and based in Italy – which promotes arms control and disarmament initiatives based on scientific understanding of risks posed by military applications of science and technology. USPID submits to the United Nations Secretary-General its views on "Artificial intelligence in the military domain, with specific focus on areas other than lethal autonomous weapons systems, and its implications for international peace and security", in accordance with the invitation formulated in operative paragraphs 7 and 8 of Resolution 79/239 adopted by the UN General Assembly on 24 December 2024.

Hazards for peace and security arising from AI military applications

USPID expresses its deep concern about new hazards for peace, international security, and the respect of International Humanitarian Law (IHL) which arise on account of the ongoing and accelerating military efforts to incorporate Artificial Intelligence (AI) into multiple facets of warfare. Major sources of these hazards have been identified in current limitations of our capability to understand, predict precisely, and control the behavior of AI systems developed by machine learning methods and their interactions with other human or artificial agents. Initially identified in connection with the operation of AI-enabled Autonomous Weapons Systems (AWS), these hazards are now spreading to AI systems supporting intelligence collection, the achievement of situational awareness, and human decision-making in warfare.

Exceptionally grave concerns are raised by proposals to integrate AI in Nuclear Command, Control, and Communication (NC3) and in adjacent systems supporting nuclear decisions, and to let AI perform tasks that might directly or indirectly affect nuclear decision-making. A significant case in point is the proposal to use AI technologies in nuclear early warning and decision-support systems, which is being advanced with the expectation that AI accuracy will reduce potential errors, and its processing speed will buy more time for nuclear decision makers. However, on account of the probabilistic nature of AI information processing, one cannot exclude the risk of AI perception leading to false positives of a nuclear attack or producing perniciously unreliable recommendations given the impossibility of ensuring that the underlying models are aligned with human values and the UN overarching goal of preventing and removing threats to peace. If such mistakes occur, no matter how infrequent, large-scale and even existential implications for humanity might ensue. Accordingly, it would be imperative to proceed with time-consuming verifications of AI responses in nuclear early warning. But these verifications would be hindered by the black-box nature of much AI information processing and by the reliance on mostly simulated data, eventually thwarting the expectation of buying more time for human decision makers.

Additional concerns are raised by proposals to exploit the rapid pace at which AI operates to speed up battlefield decision-making and targeting cycles. These proposals are fueled by the goal of gaining military advantage over potential adversaries.

However, fighting at machine speed jeopardizes both the effectiveness of human oversight on AI-enabled decision support systems and the fulfilment of ethical and legal roles that are attached to human oversight of warfare action. Indeed, overly tight temporal windows for decision-making hinder effective human control over IHL threats raised by machine suggestions. Human interventions which aim at preventing inadvertent conflict escalations prompted by fighting at machine speed are similarly hampered. In addition to this, excessive speed in human-machine interactions has been identified as a factor inducing automation biases on the battlefield, and potentially skewing human decision-making even in the absence of AI failures.

Further hazards arise in connection with inherent vulnerabilities of AI learning methods and systems. Malicious manipulation of input data might be exploited to induce classification mistakes by AI systems. Moreover, poisoning attacks corrupting learning datasets may impair learning processes and the accuracy of resulting AI systems. These risks are compounded by our current inability to fully align AI systems with human goals and values, potentially causing them to deviate from strategic objectives.

Recommended actions

Mindful of these and other emerging hazards posed by the rapid adoption of AI technologies and systems in the military domain, USPID recommends

- to integrate discussion of AI in NC3 into the Non-Proliferation Treaty framework and in dedicated high-level dialogues and forums such as the Summit on Responsible Artificial Intelligence in the Military Domain (REAIM);
- to develop sustained international dialogue, good practices, and confidence-building measures concerning new and emerging risks for peace and IHL respect raised by AI warfare applications;
- to support a comprehensive and detailed inquiry aimed at identifying actual and potential AI applications in the military domain, jointly with situations of use that pose serious threats to peace, international stability, and the respect of IHL;
- to consider and investigate the need to introduce international regulations or prohibitions for those AI military applications that pose serious threats to peace, international stability, and the respect of IHL.

Women’s International League for Peace and Freedom

[24 May 2024]

The Women’s International League for Peace and Freedom (WILPF) has opposed war and the development of technologies of violence since its founding in 1915. WILPF has consistently condemned military spending and militarism as detrimental to human life and wellbeing. Our concerns with artificial intelligence (AI) in the military domain and its implications for international peace and security are grounded within our wider opposition to weapons, war, and violence, as well as in our opposition to patriarchal, racist, and colonial power relations that are embedded within AI technology.

While there are many perils of the military use of AI; WILPF’s submission is focused on the following issues:

1. The need for human emotion, analysis, and judgement in relation to the use of force;

2. The existence of gender, racial, and other bias in AI technology and the implications for digital dehumanisation;

3. The impacts of military use of AI on privacy and personal data;

4. The environmental harms exacerbated by the military use of AI; and

5. The dangers of war profiteering and arms racing.

Due to the concerns raised in this submission and in other spaces, WILPF opposes the military use of AI. This technology, rather than placing limits on violence or harm, expands both. Governance is insufficient in the face of the profits and power the developers of these technologies seek.

In light of the concerns raised in WILPF's submission and the implications for international peace and security, WILPF urges states:

- To refrain from using AI in the military domain and to develop national laws and regulations to this end;
- To pursue a global prohibition on the military use of AI;
- To not develop autonomous weapon systems or AI-enabled weapon systems, including those that can be used to target human beings;
- To ensure protection of personal data from use by militaries, police, border enforcement, and private companies and contractors collaborating with these institutions;
- To uphold human rights and dignity online and offline; and
- To address the environmental harms generated by data centres, cloud computing, and AI by reducing the number of these centres and energy consumption and water use, which will include reducing the overall use of AI.

WILPF also urges:

- Technology companies, tech workers, scientists, engineers, academics and others involved in developing AI or robotics to pledge to never contribute to the development of AI technologies for military use;
- Financial institutions such as banks and pension funds to pledge not to invest money in the development or manufacture of AI for military use; and
- Activists, academics, affected communities, and other concerned about privacy rights, digital dehumanisation, environmental and climate justice, gender-based violence, and other issues to collaborate and strategise to oppose the development and use of AI in the military and other violent domains.

D. Scientific Community

AI, Automated Systems, and Resort-to-Force Decision Making Research Project, The Australian National University

[11 April 2025]

Introduction

This executive summary highlights policy recommendations outlined in ***AI, Automated Systems, and Resort-to-Force Decision Making – Policy Recommendations: Submission to the UN Secretary General Pertaining to A/RES/79/239*** **(11 April 2025)**, available on the UNODA website. For a complete account of the underlying research and associated research papers, please refer to the full submission.

Underlying Research Project

This research has arisen from a **two-and-half-year research project (2022-2025),** entitled ***Anticipating the Future of War: AI, Automated Systems, and Resort-to-Force Decision Making***, led by Professor Toni Erskine (Australian National University) and funded by the Australian Government through a grant by the Department of Defence.

Its focus is **distinctive and critical**. While the attention of academics and policy makers has been overwhelmingly directed towards the use of AI-enabled systems in the *conduct of war* – including, prominently, on the emerging reality of 'lethal autonomous weapons systems' ('LAWS'), this project has addressed the **relatively neglected prospect of employing AI-enabled tools at various stages and levels of deliberation over the *resort to war.*** In other words, 'it takes us from AI on the battlefield to **AI in the war-room'**.[1]

This research project has brought together **leading scholars and practitioners** working on different aspects of international politics and security, strategic and defence studies, and artificial intelligence (AI) to contribute to a multi-disciplinary study and set of **policy recommendations on the risks and opportunities of introducing AI, machine learning (ML), and automated systems** into state-level decision making on the **initiation of war**. Our interventions are made from the perspectives of political science, international relations, law, computer science, philosophy, sociology, psychology, engineering, and mathematics.

Project participants presented and discussed their research at two workshops (June 2023 and July 2024) at the Australian National University (ANU), convened by Professor Toni Erskine and Professor Steven E. Miller (Harvard). Participants also received feedback on their initial research-based policy recommendations from senior Australian Government delegates from the federal civil service as part of a one-day policy roundtable (July 2024) at the ANU.

'Four Complications'

For all the potential **benefits** of AI-driven systems – which are able to analyse vast quantities of data, make recommendations and predictions by uncovering patterns in data that human decision makers cannot perceive, and respond to potential attacks with a speed and efficiency that we could not hope to match – challenges abound. Through this project, we have sought to address **four thematic 'complications'** that we propose will accompany the gradual infiltration of AI-enabled systems in **decisions to wage war:**[2]

- **Complication 1** relates to the displacement of human judgement in AI-driven resort-to-force decision making and possible implications for deterrence theory and the unintended escalation of conflict.
- **Complication 2** highlights detrimental consequences of automation bias, or the tendency to accept without question computer-generated outputs – a tendency that can make human decision makers less likely to use (and maintain) their own expertise and judgement.

[1] T. Erskine and S. E. Miller, 'AI and the Decision to Go to War: Future Risks and Opportunities', *Australian Journal of International Affairs*, Vol. 78: 2 (2024), pp. 135–147 (p. 138).

[2] For an account of these 'four complications', see T. Erskine and S. E. Miller, 'AI and the Decision to Go to War: Future Risks and Opportunities', *Australian Journal of International Affairs,* Vol. 78: 2 (2024), pp. 135–147 (pp. 139–40).

- **Complication 3** confronts algorithmic opacity and its potential effects on the democratic and international legitimacy of resort-to-force decisions.
- **Complication 4** addresses the likelihood of AI-enabled systems impacting organisational structures and chains of command, whether degrading or enhancing strategic and operational decision-making processes.

Contributors to this project have explored these proposed complications in the context of either **automated self-defence** or the use of **AI-driven decision-support systems (DDS)** that would inform human resort-to-force deliberations. We have identified risks and opportunities of using AI-enabled systems in these contexts and make recommendations on how risks can be mitigated and opportunities promoted.

Complication 1: Displacement of human judgement

AI in Nuclear Crisis Decision Making

One key area of research undertaken in response to this complication is the nuanced interplay between AI and human decision making in the high-stakes context of **nuclear crisis management**. **Risks (including the increased fragility of nuclear deterrence relationships, crisis signalling becoming more complex**, and **unintended escalation)** have been explored in two broad areas: i) automation in military deployments, or taking the human 'out of the loop' in the decision to use nuclear or strategic non-nuclear weapons (SNNW); and, ii) the integration of AI into human decision-making (particularly in early warning threat assessments). Although much of this research has focused on risks, **novel benefits** of introducing AI-driven decision-support systems (DSS) into human-led nuclear crisis management have also been proposed.

Policy Recommendations:

- **Always incorporate human-in-the-loop safeguards**: Ensure AI systems in nuclear command and control are always overseen by human operators and that human decision-making remains central to determining when and how nuclear-weapon states resort to the use of their arsenals.
- **Promote a holistic approach to AI-safety**: AI safety should account for both technical and socio-technical dimensions. Assess safety challenges in AI-enabled DSS comprehensively, including issues of security, trust, and liability.
- **Broaden the scope of risk assessments**: Apply risk assessments relating to the deployment of AI and ML not only to obvious areas such as nuclear launch orders, but also to less obvious areas such as early warning intelligence assessments (including by non-nuclear allies) and SNNW capabilities (including by non-nuclear allies).
- **Restrict the use of AI-assisted warning data**: The key to balancing the benefits of incorporating AI into early warning against the risks is limiting what AI-assisted warning data is used for. In AI research, prioritise tasks such as calculating effective evasive manoeuvres in the event of an attack and using pattern recognition and anomaly detection to improve arms control verification.
- **Pursue** informal arms control and confidence-building: Advance informal measures such as regular dialogue, red-line agreements, and information-sharing mechanisms. Expand unilateral initiatives like moratoriums where feasible.
- **Explore AI's potential to promote empathy and enhance decision making**: Decision makers must exercise 'security dilemma sensibility' (SDS) in times of crisis. Decision makers and diplomats exercise SDS when they are open to the

possibility that the other side is behaving the way they are because they are fearful and insecure, and crucially, recognize the role that their own actions may have played in this. Explore ways that the balanced integration of AI and human judgement could enhance SDS during nuclear crises by promoting empathy and trust.

AI Mistakes in the Resort to Force

Another area addressed in relation to this complication is **state responsibility** when **errors** occur in AI-driven or autonomous systems involved in resort-to-force decisions. Such errors may arise from poor system training, data poisoning by adversaries, or two AI-driven systems interacting in unintended ways. It is essential to develop legal standards and practices that reduce the risk of unintended conflict resulting from such failures.

Policy Recommendations:

- **Adopt robust security and cyber hygiene**: States should adopt robust protections against AI data poisoning and cyber attacks to meet jus ad bellum standards of good faith and reasonable conduct.
- **Clarify legal guidelines on delegating the use of force to autonomous systems**: Senior leadership within states should set clear domestic legal standards regarding when and how autonomous systems may be authorised to use force.
- **Commit to transparency in after-action reviews**: States should commit to being transparent and deliberate about after-action reviews of any AI errors that occur in the field, potentially drawing on civilian casualty review processes as a model.

Complication 2: Automation bias

Our research in response to the second complication focuses on the relationship between human actors and **AI-driven DSS** in resort-to-force decision making. It includes a detailed survey-based study of **military trust in AI** during strategic-level deliberations and a robust account of the importance of ensuring that there are human **'experts-in-the-loop'** when AI-driven systems contribute to decisions on war initiation. This body of work also addresses the **benefits** of employing DSS to **enhance our cognitive capacities** in strategic decision making and, conversely, uncovers the potential **dangers** of such reliance if these systems **dull our sensitivity to the tragic qualities of war** or contribute to the **erosion of restraint** by creating the illusion that they replace us as responsible actors.

Policy Recommendations:

- **Consider the multidimensionality of trust**: Recognize that soldiers' trust in AI is not a forgone conclusion. Rather, it is complex and multidimensional, and further complicated by biases, uncertainty, and lack of education.
- **Interrogate norm compliance**: In terms of governance, explain how policies on increasingly autonomous capabilities coincide or diverge from international norms and laws informing their use.
- **Embed experts in decision structures**: Enshrine an 'expert-in-the-loop' organisational structure – i.e., high-level experts as core decision makers.
- **Prohibit automation**: Prohibit automation of resort-to-force decisions.

- **Increase AI literacy of domain experts**: Provide and require basic technical training for high-level domain experts so they understand the logics of AI and can thus incorporate AI decision inputs from an informed position.
- **Provide on-going, substantive training for domain experts**: Sustain substantive training for, and assessment of, high-level experts to bolster and ensure substantive competencies.
- **Regulate non-autonomous AI**: While autonomous AI agents, e.g., lethal autonomous weapons systems (LAWS), need regulation, so do non-autonomous AI systems, which leave humans vulnerable to new forms of influence, moral and cognitive atrophy, and undermined responsibility.
- **Design AI-driven DSS to promote more accurate perceptions of their capacities**: Ensure AI-driven DSS are not easily mistaken for responsible agents in themselves by avoiding anthropomorphism, building in warnings about system limitations, and incorporating features that emphasise human agency and accountability.

Complication 3: Algorithmic opacity

Our research in response to the third complication addresses **how the lack of transparency of AI-driven decision making can threaten the legitimacy** of AI-informed decisions on the resort to force. This body of work includes original research on **large language models (LLMs)** and their potential to exacerbate existing **pathologies in intelligence analyses**. It also examines the role that the **'architecture of AI'** and its hidden vulnerabilities play in deliberations surrounding the resort to force. Moreover, research within this pillar conceives of military decision-making institutions as **'complex adaptive systems'** – a conceptual framework that yields a range of insights, including that human-machine teams possess a form of **'cognitive diversity' that could be leveraged for more efficient decision-making,** but also **exploited to poison information flows**, and that technical explanations for algorithmic opacity will not solve accountability concerns.

Policy recommendations:

- **Develop policy to limit epistemic pathologies of LLMs**: States should clearly determine defence and intelligence policy towards either a) procurement of LLMs, b) state development of LLMs, or c) a combination of both. They should use this guidance to develop policy which seeks to limit the epistemic pathologies of LLMs in autonomous decision-making.
- **Commit to sector-wide procurement guidelines and oversight of generative AI tools** used in decision-making chains.
- **Commit to regulating data markets** and access to those markets through alliance relationships.
- **Promote understanding of the tech ecosystem and its fragilities**: Increase understanding of the inherent interdependencies and vulnerabilities in the tech ecosystem, including by creating technology literacy training programs designed specifically for politicians and policy, intelligence, and military leaders.
- **Invest in research** to develop a comprehensive picture of the architecture – physical and digital – that underpins AI, including critical dependencies and vulnerabilities and how access and power are distributed.

- **Invest in research on social media** and its impact on functions of government, including its potential to disrupt democracies, facilitate foreign interference, and influence decision making on the use of force.
- **Recognize AI's current influence**: Significantly increase awareness of government reliance on the architecture of AI, especially for critical government functions, including resort-to-force decision making.
- **Invest in research and development** to maximize the benefits of human-machine cognitive diversity.
- **Implement responsible AI governance programs** that carefully balance accountability with operational efficiency.
- **Perform regular red-team exercises** to ensure that the integration of AI in decision-making institutions does not induce systemic blind spots and vulnerabilities in military decision-making.

Complication 4: Impact on organisational structures

Our research regarding the fourth complication explores both the **beneficial and damaging effects that AI-driven systems can have on institutional structures** in the context of resort-to-force decision making. Studies focus on how AI-driven **DSS can improve 'adaptive culture'** within military organisations, thereby improving wartime decisions, and how the urgent need to **upgrade AI literacy and educate human analysts** should lead us to **reform institutional structures and cultures**. The novel notion of **'proxy responsibility'** is proposed as an institutional response to ensure that responsibility can be meaningfully assigned to humans for resort-to-force decisions that are informed by AI systems. Moreover, original research highlights the significance of the **neglected category of AI 'integrators'** – sandwiched between the 'developers' and 'users' of AI within organisational structures – when it comes to strategic military applications of AI.

Policy recommendations

- **Set (and evolve) measures of effectiveness**. If AI-enabled adaptive capacity is to work effectively, measures of military effectiveness must guide which direction adaptation might take. Establish such measures at the tactical (battlefield) and strategic (war-room) levels to guide development and implementation of AI-enabled adaptation.
- **Know where adaptation relevant data is found, stored and shared**. An enhanced adaptive stance in military institutions must have enhanced data awareness as a foundation. Data awareness and management must become one of the basic disciplines taught to military personnel.
- **Scale AI support from individual to institution**. There is unlikely to be a one-size-fits-all algorithm or process that can enhance learning and adaptation at every level of military endeavours. Create a virtual 'arms room' of adaptation support algorithms as part of an institution-wide approach to adaptation.
- **Routinely question AI-enabled outputs**: Build mindsets, protocols, institutional cultures, and inter-agency structures in 'normal' pre-crisis times to routinely question AI-enabled output from human-machine teams.
- **Institute an advisory body**: In order to support the notion of 'proxy responsibility' as an institutional response to 'responsibility gaps' when decisions on war initiation are informed by AI-enabled systems, establish and/or strengthen state-level 'AI departments'. These departments would integrate

technical, political, and ethical competence and expertise and advise on resort-to-force decision-making processes.

- **Support research on AI integration**: Fund research on the integration of AI in strategic decision-making.
- **Provide standards**: Outline minimum standards for the responsibilities of AI developers and integrators.
- **Facilitate inter-group discussions** between developers, integrators and users during development, integration, and longer-term maintenance processes.
- **Create accountability guidelines**: Provide well-defined guidelines and rules indicating who is accountable if something goes 'wrong'.

Queen Mary University of London, T.M.C Asser Institute, University of Southern Denmark, University of Utrecht

[11 April 2025]

Views of members of the scientific community and civil society; specifically, we are a group of academics with expertise in ethical, legal and political dimensions of military Artificial Intelligence and herewith put forward our shared views 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.

Introduction:

The rapid advancement and integration of AI technologies into targeting operations have sparked ongoing debates surrounding their ethical, legal, and operational implications. Over the past decade, the discourse on AI in warfare has largely centered on autonomous weapon systems (AWS),[1] driven in part by the initiation of discussions in 2013 and the formalization of a regulatory process under the UN Convention on Certain Conventional Weapons (CCW) and the Group of Governmental Experts on Lethal Autonomous Weapons Systems (GGE LAWS), which exclusively focuses on lethal AWS.[2] However, the increasing integration of AI-based decision-support systems (AI-DSS) into targeting practices[3] introduces new layers of complexity that demand closer attention from a broad range of stakeholders. This submission responds to that need, structured around three key components: (1) a

[1] The latest definition of AWS from the CCW GGE LAWS Rolling Text (26 November 2024): "A lethal autonomous weapon system can be characterized as an integrated combination of one or more weapons and technological components that enable the system to identify and/or select, and engage a target, without intervention by a human user in the execution of these tasks." On file with authors.

[2] For a brief overview of some of the latest developments of the GGE LAWS see Jeroen van den Boogaard, *Warning! Obstacles Ahead! The Regulation of Autonomous Weapons Systems in the GGE LAWS*, *Opinio* Juris, 4 March 2024 found at: https://opiniojuris.org/2024/03/04/warning-obstacles-ahead-the-regulation-of-autonomous-weapons-systems-in-the-gge-laws/.

[3] There have been several reported uses of AI-DSS by Israel in Gaza and potentially in Lebanon, by both Ukraine and Russia in the ongoing conflict, and by the United States in its actions against Houthi rebels in the Red Sea and in Yemen, to name a few. For a comprehensive overview of literature in this space, see e.g., Anna Nadibaidze, Ingvild Bode, and Qiaochu Zhang, *AI in Military Decision Support Systems, A Review of Developments and Debates,* Centre for War Studies, University of Southern Denmark, November 2024. Found here: https://www.autonorms.eu/ai-in-military-decision-support-systems-a-review-of-developments-and-debates/.

brief overview of how AI-DSS are currently used in targeting decisions; (2) an analysis of key concerns, including how these systems shape the potential exercise of human judgement and control and underline fundamental gaps in global governance; and (3) a concluding set of recommendations.

1. Overview of AI-DSS and the joint targeting cycle

Defined as "the process of selecting and prioritizing targets and matching the appropriate response to them, considering operational requirements and capabilities,"[4] targeting is a core military function at the very heart of warfare. While the potential range of use cases for AI-DSS in military decision-making is broad, in targeting, AI-DSS can be understood to serve as **tools** that use AI techniques to collect and analyze data, provide information about the operational environment as well as actionable recommendations, with the aim of aiding military decision makers in evaluating factors relevant to legal compliance such as taking precautions and ensuring proportionality in attacks.

More specifically, AI-DSS are increasingly integrated across multiple phases of the joint targeting cycle (JTC), including within target development and prioritization, capabilities analysis, and mission execution. The JTC is a reflective example of a structured process used by military forces to identify, evaluate, and engage targets while ensuring compliance with operational, legal, and ethical standards,[5] generally consisting of six (non-linear) phases:

1. **End-State and Commander's Objectives**: Defining strategic military goals and desired outcomes.

2. **Target Development and Prioritization**: Identifying, verifying/validating, and prioritizing targets based on intelligence and mission goals.

3. **Capabilities Analysis**: Assessing the available strike options and their effectiveness.

4. **Force Assignment**: Allocating specific military assets (e.g., airstrikes, artillery, cyber operations) to engage the target.

5. **Mission Execution**: Carrying out the targeting operation while ensuring compliance with relevant laws and the rules of engagement.

6. **Assessment**: Evaluating the effectiveness of the operation and adjusting for future operations, if necessary.

Within this framework, AI DSS are assumed to serve primarily as informational and analytical tools which support human decision-making rather than supplant it. However, this assumption and framing obscures how AI-DSS influence human cognitive processes within the JTC. This impact on human decision-making is often

[4] United States Department of Defense, *Dictionary of Military and Associated Terms*, March 2017, found at: https://www.tradoc.army.mil/wp-content/uploads/2020/10/AD1029823-DOD-Dictionary-of-Military-and-Associated-Terms-2017.pdf.

[5] Michael Schmitt et al, *Joint and Combined Targeting: Structure and Process*, Chapter 13 in Jens David Ohlin (ed) *Weighing Lives in War* (Oxford, 2017). See also, Jessica Dorsey and Marta Bo, *AI-Enabled Decision-Support Systems in the Joint Targeting Cycle: Legal Challenges, Risks, and the Human(e) Dimension*, forthcoming 2025, *International Law Studies.* "Targeting generally involves four key steps: (1) objectives and guidance, (2) planning, (3) execution, and (4) assessment. Encapsulating these four key steps, the United States and NATO outline their targeting processes through similar six-phase cycles [addressed in this submission]. As the reader can discern, different states employ different doctrines for targeting. What is important … is not necessarily the specific labels for various steps followed by any given state, but rather how and when compliance with the principle[s of IHL are] incorporated into the targeting process."

underestimated and remains insufficiently examined, leaving critical discussions about the role of AI-DSS largely absent from current policy debates.

2. Analysis of Key Concerns

(a) (Meaningful) Human Judgement and Control

AI-DSS are often portrayed as enhancing human decision-making and the quality of decisions therein. The perception of AI-DSS as mere subsidiary tools has led to a narrative that the integration of AI-DSS poses fewer challenges than AWS, given that these systems do not directly "engage" targets (i.e., they do not have an inherent capability to directly carry out the use of force) and are tools that assist human commanders. The outputs are ostensibly ultimately reviewed through (several layers of) human oversight, such as processes of verifying and validating targets using additional intelligence sources.[6] As a result, errors or inaccuracies in AI-DSS outputs are often seen as non-critical, based on the assumption that robust human oversight and appropriate control will compensate for them. However, closer examination reveals that this control is frequently superficial, offering only the appearance of, rather than actual meaningful, or context-appropriate, human judgement and control.

This is because AI-DSS structure and condition the quality of human control and oversight and limit the ways control and oversight can be exercised. The use of AI-DSS creates a shared decision-making space between human military personnel and AI technologies. States appear to have recognized and focused on many of the advantages of this shared decision-making space for military personnel, i.e., how the use of AI-DSS advances human decision-making through offering data-driven insights. But using AI-DSS also delimits the capacity to exercise human oversight and control because of the technologies' complexity and the increased speed (and therefore scale) it can bring to decision-making processes. Rather than supporting human oversight, using AI-DSS may risk humans becoming little more than reactive cogs in socio-technical systems.[7] Moreover, this configuration risks amplifying adverse human biases, such as automation bias, anchoring bias, or cognitive action bias, to the detriment of exercising qualitatively high levels of human control.[8] Considering AI-DSS as a distinct form of technology therefore reveals significant challenges associated with military AI and human oversight, challenges that extend beyond those that arise when simply integrating the technology in weapon systems.

Recent conflicts have shown the risks associated with AI-DSS being employed in critical functions, such as target selection and even nomination, and their conditioning and constraining of human involvement, affecting the fulfilment of core legal obligations embedded within the JTC. The use of AI-DSS raises fundamental concerns about whether human decision makers can retain adequate cognitive autonomy over the JTC process or whether humans will become overly reliant on algorithmic outputs for critical judgements in the context of armed conflict.[9] Consequently, there are significant legal concerns regarding the effects of such systems on decision-making processes and use of force decisions and ability for users to comply with IHL obligations, especially with respect to the obligation to take all

[6] Alexander Blanchard and Laura Bruun, *Automating Military Targeting: A Comparison Between Autonomous Weapon Systems and AI-Enabled Decision Support Systems*, Stockholm International Peace Research Institution (SIPRI) forthcoming 2025 (draft on file with authors).

[7] Ingvild Bode, *Human-Machine Interaction and Human Agency in the Military Domain*, Policy Brief No. 193 (Waterloo, ON: Centre for International Governance Innovation, 2025).

[8] Dorsey and Bo, supra n. 5.

[9] *Ibid*; see also Anna Nadibaidze, Ingvild Bode, and Qiaochu Zhang, *AI in Military Decision Support Systems, A Review of Developments and Debates,* Centre for War Studies, University of Southern Denmark, November 2024. Found at: https://www.autonorms.eu/ai-in-military-decision-support-systems-a-review-of-developments-and-debates/.

feasible precautions to minimize civilian harm to the greatest extent possible in attack and comply with the principles of distinction and proportionality.[10]

Importantly, these concerns are not new. There is extensive debate around how to preserve meaningful human judgment and human agency when conducting IHL-evaluative legal assessments, in the context of AWS. These discussions – which include expert analysis on accountability, human-machine interaction, automation bias, and the effect of AI systems on legal and ethical reasoning[11] – provide valuable lessons that must inform discussions around military AI and specifically the use of AI-DSS in military contexts.

(b) AI-DSS: Understudied, Under-Addressed and Unregulated

Framing AI-DSS as mere tools, has led to an underestimation and lack of analysis on the way their use affects the cognitive decision-making process within the JTC. The relative lack of attention paid to AI-DSS so far can partly be attributed to the fact that such systems are seen to be used with a human *in* or *on* the loop framework, with their outputs ostensibly reviewed by one or more individuals during the targeting process. As a result, current understandings of AI-DSS use appear to align with widely supported principles of human control and oversight. However, this gap in the debate is also caused by a lack of transparency around how specific AI-DSS function, and a consistent failure to comprehensively examine how they are being used in practice.

Additionally, the persistent focus on AWS at the expense of AI-DSS obscures the growing reliance on AI in shaping operational and strategic outcomes. Unlike AWS, which have been debated in the framework of the CCW for the past decade, AI-DSS lack a comparable institutional platform. Attention to AI-DSS remains scattered across various initiatives but these efforts have yet to provide the dedicated regulatory focus or coordination needed.

3. Recommendations:

i. **Reassert** the central role of human cognitive and legal reasoning in military operations by implementing safeguards that ensure key legal assessments remain grounded in human(e) judgment. Leverage existing insights from debates on AWS and research on human-machine teaming and human-computer interaction to inform discussions on AI-DSS.

[10] Article 57 of the First Additional Protocol to the Geneva Conventions. See also Dorsey, Bo supra n. 5 (on AI-DSS and their effects on the principle of precautions); Jessica Dorsey, *Proportionality under Pressure: The Effects of AI-Enabled Decision Support Systems, the Reasonable Commander Standard and Human(e) Judgment in Targeting*, forthcoming *International Review of the Red Cross* (2025) (on AI-DSS and their effects in the context of IHL proportionality assessments).

[11] Marta Bo, *Autonomous Weapons and the Responsibility Gap in light of the Mens Rea of the War Crime of Attacking Civilians in the ICC Statute*, 19 Journal of International Criminal Justice 2021; Bo, M., Bruun, L. and Boulanin, V., *Retaining Human Responsibility in the Development and Use of Autonomous Weapon Systems: On Accountability for Violations of International Humanitarian Law Involving AWS* (SIPRI: Stockholm, Oct. 2022), p. 41; Boulanin, V., Bruun, L. and Goussac, N., *Autonomous Weapon Systems and International Humanitarian Law: Identifying Limits and the Required Type and Degree of Human-Machine Interaction* (SIPRI: Stockholm, June 2021), p. 54; and Bruun, L., Bo, M. and Goussac, N., *Compliance with International Humanitarian Law in the Development and Use of Autonomous Weapon Systems: What Does IHL Permit, Prohibit and Require?* (SIPRI: Stockholm, Mar. 2023), p. 24. Elke Schwarz, "The (im)possibility of meaningful human control for lethal autonomous weapons systems," *Humanitarian Law and Policy*, 29 August 2018, found at: https://blogs.icrc.org/law-and-policy/2018/08/29/im-possibility-meaningful-human-control-lethal-autonomous-weapon-systems/.

ii. **Recognize** and address the incremental effects of AI-DSS design and use on human cognitive reasoning and critical deliberation. Promote awareness and attentiveness as a crucial part of reasserting and strengthening the exercise of human agency in targeting decision-making.

iii. **Reinforce** calls for greater attention to the implications of AI-DSS in armed conflict. Utilize platforms such as the UN General Assembly's First Committee on Disarmament and International Security and the Global Commission on the Responsible Use of AI in the Military Domain to foster inclusive and complementary discussions on the associated risks and systemic changes AI-DSS introduce.

United Nations Institute for Disarmament Research

[11 April 2025]

Artificial intelligence (AI) is rapidly transforming the military domain and profoundly influencing international peace and security. Initiatives such as the summits on Responsible AI in the Military Domain (REAIM) and the Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy, while not being universal processes, have significantly elevated international attention on the military applications of AI. In particular, they have moved the debate beyond lethal autonomous weapon systems (LAWS) and have successfully highlighted the multifaceted impacts of AI, fostering broader international policy engagement. Building on the political momentum generated by these initiatives, resolution 79/239 adopted by the United Nations General Assembly in December 2024 represented a significant milestone as the first UN resolution on AI in the military context and has offered Member States, international and regional organizations and the multi-stakeholder community the opportunity to share their views on opportunities and risks.

For many years, the United Nations Institute for Disarmament Research (UNIDIR) has played an important role in shaping and informing discussions on the broader impact of AI in the military domain, both within and beyond applications of this technology in weapon systems. It has undertaken research, facilitated multilateral dialogues, and offered policy insights that underline AI's transformative potential for international peace and security. This policy note draws from all the work conducted to summarise opportunities and risks and to offer a potential roadmap for future policy action.

The international community can now shape how AI is used in the military domain, putting principles of responsible AI at the core. A central challenge is the complexity of defining the "military domain". States and regions interpret the scope of this domain differently based on their unique security landscapes, realities and operational practices. For some countries, military roles extend to internal security tasks such as policing, border control, combating organized crime, protection of critical infrastructure or humanitarian relief in response to natural disasters. Others maintain a stricter definition, limiting military functions to battlefield engagements. These variations, rather than serving as barriers, offer important context for multilateral discussions. International governance frameworks should remain flexible and inclusive, acknowledging and adapting to diverse national and regional security perspectives.

In the many operational contexts within the military domain, AI acts as a force multiplier across several military tasks, including command and control (C2), information management and intelligence, advanced autonomy, logistics, training and simulation, and organizational and support functions. In C2, AI enhances the speed and quality of decision-making, thereby helping commanders rapidly analyse battlefield scenarios. It has the potential to improve adherence to international humanitarian law (IHL), for example by integrating detailed proportionality and other legal assessments. AI-driven intelligence tools analyse large volumes of data at speed, and so improve

situational awareness and threat detection. In logistics, AI optimizes supply chains and predictive maintenance, enhancing operational readiness and improving the sustainability of military operations over time. AI further supports advanced autonomy in drones, cybersecurity, and operations in the information domain. Training and simulation benefit from AI by creating personalized, realistic synthetic environments and scenarios. In short, if developed, deployed and used responsibly, AI could increase operational effectiveness while offering new ways to mitigate risks and reduce harm.

However, integrating AI in military contexts also presents significant risks and challenges – technological, security, legal, policy and ethical.

Technologically, military AI systems face issues related to the quality, availability and inherent biases of data. These may lead to unpredictable and potentially harmful outcomes, including violations of international law. The "black box" nature of AI systems, often coupled with their adaptiveness and highly context-dependent nature, complicates trustworthiness assessments and may, at times, challenge the conduct of effective investigations into alleged violations of IHL. Cybersecurity vulnerabilities also expose AI systems to adversarial attacks, requiring stringent security measures.

Security challenges include risks of miscalculation and unintended escalation, particularly through AI-enabled rapid decision-making processes and AI-enabled autonomy, which may result in escalatory responses. The potential for an AI arms race exacerbates international and regional tensions, possibly leading to destabilizing outcomes similar to historical arms competitions. The proliferation of AI technologies to non-state actors further complicates threat landscapes and necessitates robust life cycle management of military AI systems. Additionally, AI-generated disinformation threatens societal stability by undermining trust in information and can have a direct impact on military operations.

Legal challenges revolve around ensuring compliance with international law, particularly IHL and international human rights law. Key debates focus on, among other things, accountability and both state and individual responsibility for AI-driven actions, especially regarding lethal decisions. States diverge on whether existing legal frameworks are sufficient or if new, specialized regulations are required. Beyond international law, ethical considerations emphasize maintaining human judgment in critical decision-making and preventing societal biases from infiltrating AI systems. The latter requirement calls for greater diversity and inclusivity in AI development. Additionally, bridging gaps between government, academia and the private sector remains challenging yet crucial for effective governance.

Addressing these challenges requires a comprehensive road map with actions at the multilateral, regional and national levels.

Multilaterally, establishing a United Nations-led comprehensive platform that enables a regular institutional dialogue to address military AI's broader implications on international peace and security is key as it would provide an institutional framework to advance policy discussions. This platform could build on the existing internationally developed AI principles and frameworks, such as UNESCO's recommendations or the commitments made in the Global Digital Compact (e.g. safe, secure and trustworthy AI) and further refine them for application in the military domain. These principles could be further developed into voluntary norms of responsible behaviour in the development, deployment and use of AI in the military domain and provide a solid foundation for future multilateral instruments. In addition, such platform could be leveraged to develop practical confidence-building measures (CBMs), lead inclusive multi-stakeholder engagement, and deliver global capacity-building programmes that enhance global security via transparency, cooperation and predictability.

Regionally, existing organizational frameworks can be used to tailor CBMs and guidelines that reflect local security contexts. Cross-regional dialogues would

facilitate mutual learning, prevent information silos, and include diverse perspective which would encourage globally coherent responses.

Nationally, states should develop comprehensive AI strategies that detail vision, priorities and governance frameworks, ensuring compliance with international norms and ethical standards. Robust governance structures (e.g., dedicated AI steering committees and ethics boards), alongside iterative legal reviews, would enhance accountability and safety. Transparent communication and clearly defined accountability protocols would further support responsible AI implementation. High standards of data governance, life cycle management approaches, rigorous training programmes and updated military operational guidelines complete these proposed national measures, ensuring the responsible integration of AI in the military domain.

Table A below provides an overview of the proposed roadmap for policy action.

Table A: A roadmap for future policy action

Level	*Action*	*Rationale*
Multilateral	**Establish a multilateral process under United Nations auspices to provide a comprehensive platform for discussion on military applications of AI and their impact on international peace and security.** This process could be leveraged to: a. Develop a set of overarching, core principles of responsible AI in the military domain to help align national efforts and reduce risk. b. In the future, further develop these core principles into international voluntary norms or guidelines for responsible state behaviour in the development, deployment and use of AI in the military domain. These guidelines could take the form of a code of conduct or a political declaration supplemented by more technical instruments as required (e.g., on AI assurances, and robust protocols for testing and evaluation). c. Develop confidence-building measures (CBMs) for military AI. States could agree on and implement practical CBMs to increase transparency and trust regarding AI in the military domain.	Collectively, these multilateral actions aim to foster cooperation, set common rules and share knowledge on military AI at the international level with a view to increasing predictability. They aim to shape the global landscape so that all states move towards safer and more transparent integration of AI in the military domain, thereby reducing the risks. While clustered under a single umbrella recommendation, each of the actions above could be implemented on its own, although their mutually reinforcing nature would amplify the impact achieved if they are implemented in combination.

Level	*Action*	*Rationale*
	d. Promote multi-stakeholder engagement in support of multilateral policy action. e. Develop and implement a coherent capacity-building programme.	
Regional	**Leverage regional and subregional organizations and dialogues to discuss the issue of AI in the military domain.** Regional and sub-regional organizations could: a. Develop region-specific CBMs, norms or guidelines that reflect local contexts. b. Set up networks for information-sharing on AI-related best practices suited to their security landscape. c. Develop joint AI-development projects, aligning operational, legal and technical requirements. **Initiate cross-regional dialogues Initiate cross-regional dialogues on AI**, where two or more regional groups exchange lessons and possibly align their approaches.	Regional and subregional approaches allow tailoring to specific security realities and threat perceptions, which could lead to concrete results that are more aligned with specific needs. In addition, regional and subregional approaches could be leveraged to inform and shape global dialogues and strengthen context-specific capacity-building. Cross-regional dialogue can be a useful tool to enable mutual learning and avoid echo chambers.
National	**Implement a comprehensive approach to AI governance in the military domain** to include the following actions: a. Develop a comprehensive national strategy or policy on AI in security and defence. b. Establish robust governance structures and review processes. c. Implement transparency and accountability measures d. Implement robust data practices and governance frameworks for all military AI applications.	A national strategy clarifies roles and responsibilities, and provides a clear direction for the development, acquisition, integration and use of AI in the military domain. Dedicated structures provide focus and accountability. They create effective checkpoints that AI projects must pass and comply with consistently (e.g., ethical approval, legal clearance, safety certification), reducing chances of unsafe or unlawful deployment. Transparency builds public trust and international confidence that a state is using AI responsibly.

Level	*Action*	*Rationale*
	e. Manage AI capabilities throughout their entire life cycle – from design and development, through testing and deployment, to updates and decommissioning – with continuous risk assessments and mitigation at each stage.	Accountability ensures that the presence of AI does not create a vacuum of responsibility – maintaining the ethical and legal norm that humans are accountable for military actions.
	f. Invest in human capital and training by developing extensive training programmes for military personnel on AI and cultivating a new generation of AI-literate officers and specialists. This includes not only technical training but also training on the ethical and legal aspects of AI use in operations.	By prioritizing robust data governance and the provision of the necessary infrastructure to enable it, militaries can improve the performance and trustworthiness of their AI systems and reduce error rates. A life cycle view ensures that safety and compliance are ongoing commitments reducing chances of failure in the field and ensuring that accountability is maintained throughout the system's use. Human expertise and judgment remain critical. Training reduces misuse and enables more effective human-machine teaming.
	g. Review military operational guidelines to strengthen AI governance in military contexts, including military documentation (e.g. doctrines, standard operating procedures and others), and rules of engagement.	Existing military governance tools and instruments can be used to strengthen the governance of AI in the military domain at a more practical, tactical level, thereby offering an impactful complement to the highest levels of governance and the associated obligations emanating from international, regional and national laws and regulations.

E. Industry

Microsoft

[24 May 2024]

Microsoft welcomes the opportunity provided by the United Nations General Assembly resolution A/RES/79/239 on "Artificial Intelligence in the Military Domain and its Implications for International Peace and Security", and UNODA's invitation to share perspectives on the opportunities and challenges posed to international peace and security by the application of artificial intelligence (AI) in the military domain, with specific focus on areas other than lethal autonomous weapons systems.

Our perspectives reflect Microsoft's deep commitment to our Responsible AI Principles and our Secure Future Initiative, emphasizing cybersecurity, safeguarding international norms, and promoting trust in technology, and our active participation in multi-stakeholder initiatives including the UNIDIR-led Roundtable for AI, Security, and Ethics (RAISE).

I. Opportunities

Microsoft recognizes substantial opportunities in responsibly applied AI within the military domain, particularly:

- *Enhancing cybersecurity and defense capabilities:* AI significantly strengthens cybersecurity defenses by automating threat detection, enabling faster and more accurate responses to cyber threats. Technologies such as Microsoft Security Copilot illustrate the transformative potential of AI in defense, empowering cybersecurity professionals to identify and mitigate risks efficiently. Initiatives like Microsoft's Zero Day Quest and collaboration with MITRE ATT&CK demonstrate proactive industry efforts to enhance global cybersecurity preparedness and resilience.
- *Broad spectrum of military applications:* Beyond cybersecurity, responsibly designed AI can significantly enhance efficiency and effectiveness across logistics, command and control systems, intelligence processing, military training, peacekeeping, humanitarian assistance, and disaster relief operations. Diverse applications underscore AI's transformative potential beyond combat scenarios alone.
- *Improving compliance with international humanitarian law:* AI technologies should improve the accuracy and effectiveness of targeting processes, aiding militaries to better adhere to principles of distinction, proportionality, and necessity. AI should significantly enhance protections for civilians and civilian infrastructure, thereby reducing unintended collateral damage in conflict.
- *Capacity building and international cooperation:* The adoption of AI in the military domain presents opportunities for global knowledge-sharing and capacity-building initiatives. International partnerships should support developing nations by sharing security capabilities, knowledge, and best practices, thus bridging technological divides and fostering global stability.

II. Challenges

Microsoft equally acknowledges significant challenges and risks associated with AI applications in the military domain:

- *AI-enhanced cyber threats:* AI has escalated cyber threat capabilities, empowering state-sponsored and criminal actors to carry out increasingly sophisticated cyber operations. These AI-driven threats include advanced phishing campaigns, automated exploitation of vulnerabilities, and adaptive malware, significantly increasing global cybersecurity risks.
- *Risks of escalation and miscalculation:* Integrating AI into military decision-making risks unintended escalation and/or miscalculation. Rapid, automated decision-making processes may inadvertently lower conflict thresholds, amplifying risks of destabilization or accidental conflict.
- *Proliferation and uncontrolled diffusion:* Uncontrolled diffusion, especially through open-source models and decentralized development, heightens the risk of malicious use by both state and non-state actors, including terrorist groups and cyber mercenaries. Increasingly accessible dual-use and proprietary AI

systems enable actors even with limited resources can gain access to capabilities that previously required significant investment or expertise, posing additional threats to international security and stability.

- *Algorithmic bias and ethical implications:* Algorithmic biases embedded within AI systems pose ethical and humanitarian concerns. Biases related to gender, race, age, or socioeconomic factors in AI datasets can intentionally and unintentionally perpetuate inequality and discrimination, particularly within sensitive military and security applications.
- *Digital divides and inequality:* Without deliberate policy actions, disparities between developed and developing nations in AI capabilities could deepen, increasing geopolitical tensions and socio-economic inequalities, thus undermining long-term global stability.

III. Relevant normative proposals

Microsoft recognizes several existing and emerging normative frameworks relevant to AI governance in the military domain, including:

- UNIDIR's RAISE initiative, facilitating international multi-stakeholder dialogues and governance proposals.
- The Responsible AI in the Military Domain (REAIM) Summits, emphasizing transparency, accountability, and human oversight at the international level.
- The US Department of Defense Responsible AI Strategy, highlighting responsibility, equitability, traceability, reliability, and governability.
- NATO's Principles of Responsible Use for AI in Defence, emphasizing reliability, governability, and traceability among member nations.

IV. Microsoft recommendations

To maximize opportunities and mitigate the challenges, Microsoft proposes several key recommendations:

- *Establish clear international norms and standards:* Develop explicit international norms and industry standards governing responsible use and development of military AI. These norms should delineate acceptable and unacceptable behaviors, providing robust frameworks to deter misuse and foster transparency and accountability, supported where appropriate by monitoring or compliance mechanisms. AI governance frameworks should explicitly differentiate operational contexts, such as peacekeeping, humanitarian assistance, crisis management, and conflict scenarios, to appropriately address varied ethical, legal, and humanitarian considerations. To ensure continued relevance, such norms should be periodically reviewed and updated to reflect evolving technological developments and operational realities.
- *Ensure human-centric oversight and accountability:* Adopt policies ensuring meaningful human judgment, oversight, and accountability remain central to military decisions involving AI, particularly regarding the use of force. Clear oversight mechanisms and enforceable accountability structures, including rigorous human control and review processes, are necessary to maintain ethical standards, avoid automation bias, and mitigate unintended consequences.
- *Advance secure and transparent AI development practices:* Promote rigorous technical standards and comprehensive life cycle management protocols covering pre-design, development, testing, deployment, operation, acquisition, and decommissioning. Robust vulnerability management, security audits, and

transparent development and deployment processes should be integral components, alongside clear capacity-building measures, ensuring AI systems remain secure, responsible, and resilient throughout their operational life cycle.

- *Enhance responsible data governance practices:* Establish clear international guidelines on responsible data governance specifically tailored to military AI applications. Transparent and accountable data management practices addressing collection, sharing, storage, training, and operational usage are crucial for managing dual-use risks, preventing misuse, and maintaining strict compliance with international legal and ethical frameworks.
- *Address and reduce algorithmic bias:* Prioritize addressing algorithmic bias through rigorous testing, transparent data practices, and inclusive AI development processes. Developers and users should establish clear policies to proactively identify, mitigate, and remediate biases, especially when AI systems are deployed in sensitive military or security contexts.
- *Promote responsible innovation and risk-based regulation:* Support regulatory frameworks that are risk-based, outcome-oriented, and balanced, ensuring they encourage innovation while adequately addressing security and ethical risks associated with AI deployment. Industry should advocate for flexible, adaptive regulations that keep pace with technological change, without imposing overly prescriptive or impractical requirements. Industry-led initiatives, such as voluntary codes of conduct, vulnerability disclosure standards, and collaborative red-teaming exercises, should be actively supported and integrated into broader international normative frameworks.
- *Strengthen international governance and alignment:* Support and actively engage in international initiatives, including REAIM Summits and dialogues at the UN General Assembly and UN Security Council. Robust international governance frameworks, characterized by transparency, clear accountability measures, and trust-building mechanisms, are essential for coherent and inclusive approaches to AI governance. Member States and stakeholders should coordinate closely through these forums to reduce fragmentation and ensure global alignment.
- *Support knowledge-sharing and awareness-raising with the UN system:* Encourage and actively contribute to efforts by the UN Secretariat and relevant UN entities to convene meaningful multi-stakeholder expert dialogues, workshops, and knowledge-sharing on AI in the military domain. Exchanges through voluntary contributions, technical expertise, and collaborative initiatives should aim at enhancing global understanding of AI's implications for international peace and security.
- *Strengthen international cooperation and information sharing:* Encourage robust international cooperation, emphasizing real-time threat intelligence sharing and joint attribution mechanisms. Industry actors should actively participate in collective cybersecurity efforts, enhancing global cybersecurity preparedness and response.
- *Foster multi-stakeholder dialogue and collaboration:* Actively participate in and support forums such as RAISE, involving states, international organizations, academia, civil society, and industry. Such inclusive dialogues are essential for mutual understanding, shaping responsible AI practices, and developing collaborative governance structures.

V. Conclusion

Microsoft is deeply committed to proactive collaboration with Member States, the UN system, industry, and civil society to implement these recommendations swiftly and effectively. Through sustained collective efforts and ongoing engagement in multi-stakeholder initiatives, Microsoft will continue supporting responsible AI governance, innovation, and practices that meaningfully contribute to international peace and security.