

Principles and Good Practices on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems

Proposed by: Australia, Canada, Japan, the Republic of Korea, the United Kingdom, and the United States

Preamble and Introduction

The High Contracting Parties to 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 (“the Convention” or “CCW”):

Recalling the objectives and purposes of the Convention;

Reaffirming that international law, in particular the United Nations Charter and International Humanitarian Law (IHL) as well as relevant ethical perspectives, should guide continued consideration and elaboration, by consensus, of possible measures and options related to the normative and operational framework on emerging technologies in the area of lethal autonomous weapons systems (LAWS);

Noting the potential challenges posed by emerging technologies in the area of lethal autonomous weapons systems to IHL;

Reaffirming without prejudice to the result of future discussions, the following guiding principles, which were affirmed by the High Contracting Parties:

- (a) International humanitarian law continues to apply fully to all weapons systems, including the potential development and use of lethal autonomous weapons systems;
- (b) Human responsibility for decisions on the use of weapons systems must be retained since accountability cannot be transferred to machines. This should be considered across the entire life cycle of the weapons system;
- (c) Human-machine interaction, which may take various forms and be implemented at various stages of the life cycle of a weapon, should ensure that the potential use of weapons systems based on emerging technologies in the area of lethal autonomous weapons systems is in compliance with applicable international law, in particular IHL. In determining the quality and extent of human-machine interaction, a range of factors should be considered including the operational context, and the characteristics and capabilities of the weapons system as a whole;
- (d) Accountability for developing, deploying and using any emerging weapons system in the framework of the CCW must be ensured in accordance with applicable international law, including through the operation of such systems within a responsible chain of human command and control;

- (e) In accordance with States' obligations under international law, in the study, development, acquisition, or adoption of a new weapon, means or method of warfare, determination must be made whether its employment would, in some or all circumstances, be prohibited by international law;
- (f) When developing or acquiring new weapons systems based on emerging technologies in the area of lethal autonomous weapons systems, physical security, appropriate non-physical safeguards (including cyber-security against hacking or data spoofing), the risk of acquisition by terrorist groups and the risk of proliferation should be considered;
- (g) Risk assessments and mitigation measures should be part of the design, development, testing and deployment cycle of emerging technologies in any weapons systems;
- (h) Consideration should be given to the use of emerging technologies in the area of lethal autonomous weapons systems in upholding compliance with IHL and other applicable international legal obligations;
- (i) In crafting potential policy measures, emerging technologies in the area of lethal autonomous weapons systems should not be anthropomorphized;
- (j) Discussions and any potential policy measures taken within the context of the CCW should not hamper progress in or access to peaceful uses of intelligent autonomous technologies;
- (k) The CCW offers an appropriate framework for dealing with the issue of emerging technologies in the area of lethal autonomous weapons systems within the context of the objectives and purposes of the Convention, which seeks to strike a balance between military necessity and humanitarian considerations.

Acknowledging the consensus achievements to date of the Group of Governmental Experts while also recognizing the importance of continued work, including in light of the evolving and dynamic nature of emerging technologies in the area of LAWS;

1. Affirm the following principles and good practices, without prejudice to the result of future discussions.

Purpose and Scope

2. The following principles and good practices are intended to:
 - a. strengthen the implementation of international humanitarian law and to promote responsible behavior with regard to emerging technologies in the area of LAWS;
 - b. be considered by States throughout the life-cycle of weapons systems based on emerging technologies in the area of LAWS, such as when designing, developing, deploying, and using such systems; and

- c. provide a basis for further international discussion and work, including the further elaboration of these principles and good practices.

Characteristics and Concepts

3. The role and impacts of autonomous functions in the identification, selection, or engagement of a target are among the essential characteristics of weapons systems based on emerging technologies in the area of LAWS. (based on 2019 GGE Report ¶19a).
4. Emerging technologies in the area of LAWS can include novel advancements in the field of Artificial Intelligence.
5. These principles and good practices may be of particular relevance when considering uses of weapons systems based on emerging technologies in the area of LAWS in which the system operator relies on autonomous functions to select and engage targets with lethal force and, before activation, the system operator does not identify a specific target or targets for intended engagement.
6. The following considerations may continue to aid the identification of characteristics and concepts relevant to emerging technologies in the area of LAWS and to the application of these principles and good practices:
 - a. Characterization, or working definitions, should neither predetermine nor prejudice policy choices; they should be universally understood by stakeholders. (2018 GGE Report ¶22a).
 - b. Purely technical characteristics such as physical performance, endurance, or sophistication in target acquisition and engagement may alone not be sufficient to characterize LAWS, especially in view of rapid evolution in technology. (2018 GGE Report ¶22b).
 - c. Attempting to define a general threshold level of autonomy based on technical criteria alone could pose difficulty because autonomy exists on a spectrum, understandings of autonomy change with shifts in the technology frontier, and different functions of a weapons system could have different degrees of autonomy. (based on 2018 GGE Report ¶22c).
 - d. A focus on characteristics related to the human element in the use of force and its interface with machines is necessary in addressing accountability and responsibility. (based on 2018 GGE Report ¶22f).

Application of International Humanitarian Law

7. International humanitarian law continues to apply fully with respect to weapons systems based on emerging technologies in the area of LAWS.

8. The right of parties to an armed conflict to choose methods or means of warfare, **including weapons systems based on emerging technologies in the area of LAWS**, is not unlimited. (CCW preamble with insertion in bold).
9. In cases involving weapons systems based on emerging technologies in the area of LAWS not covered by the Convention and its annexed Protocols or by other international agreements, the civilian population and the combatants shall at all times remain under the protection and authority of the principles of international law derived from established custom, from the principles of humanity, and from the dictates of public conscience. (2019 GGE Report ¶17g).

Weapons Prohibited from Use in All Circumstances

10. A weapons system based on emerging technologies in the area of LAWS must not be used if it is of a nature to cause superfluous injury or unnecessary suffering, if it is inherently indiscriminate, or if it is otherwise incapable of being used in accordance with international humanitarian law. (Sixth RevCon Declaration ¶19).
11. To prevent the development of such weapons systems based on emerging technologies in the area of LAWS that could not, under any circumstances, be used in compliance with international humanitarian law:
 - a. weapons systems must not be designed to be used to conduct attacks against the civilian population, including attacks to terrorize the civilian population;
 - b. weapons systems must not be designed to cause incidental loss of civilian life, injury to civilians, and damage to civilian objects that would invariably be excessive in relation to the concrete and direct military advantage expected to be gained;
 - c. the autonomous functions in weapons systems must not be designed to be used to conduct attacks that would not be the responsibility of the human command under which the weapon system would be used; and
 - d. weapons systems are to be developed such that their effects in attacks can be anticipated and controlled, as may be required, in the circumstances of their use, by the principles of distinction and proportionality and such that attacks conducted with reliance upon their autonomous functions will be the responsibility of the human command under which the system was used.

Other Prohibitions or Restrictions on the Use of Weapons Systems Based on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems

12. The potential use of weapons systems based on emerging technologies in the area of LAWS must be conducted in accordance with applicable international law, in particular international humanitarian law and its requirements and principles, including, *inter alia*, distinction, proportionality, and precautions in attack. (2019 GGE Report ¶17a).

- a. These international humanitarian law requirements and principles must be applied through a chain of responsible command and control by the human operators and commanders who use weapons systems based on emerging technologies in the area of LAWS. (based on 2019 GGE Report ¶17d).
 - b. Compliance with these international humanitarian law requirements and principles in the potential use of weapons systems based on emerging technologies in the area of LAWS requires, *inter alia*, that human beings make certain judgements in good faith based on their assessment of the information available to them at the time. (based on 2019 GGE Report ¶17f).
13. *Distinction*. Civilians and civilian objects must not be made the object of attacks involving the use of weapons systems based on emerging technologies in the area of LAWS. Attacks involving the use of weapons systems based on emerging technologies in the area of LAWS may only be directed against military objectives.
14. *Proportionality*. The expected loss of civilian life, injury to civilians, and damage to civilian objects incidental to attacks involving the use of weapons systems based on emerging technologies in the area of LAWS must not be excessive in relation to the concrete and direct military advantage expected to be gained.
15. *Precautions in attack*. Feasible precautions must be taken in planning and conducting attacks involving the use of weapons systems based on emerging technologies in the area of LAWS to spare, as far as possible, civilians and civilian objects from the loss of life, injury, and damage or destruction. Feasible precautions are those that are practicable or practically possible, taking into account all circumstances ruling at the time, including humanitarian and military considerations.

Responsibility and Accountability

16. *General Considerations*. The following principles related to accountability and responsibility, although not exhaustive, should be considered across the entire life cycle of weapons systems based on emerging technologies in the area of LAWS:
- a. Human responsibility for decisions on the use of weapons systems must be retained since accountability cannot be transferred to machines. (Guiding Principle (b)).
 - b. Humans must at all times remain accountable in accordance with applicable international law for decisions on the use of force. (Sixth RevCon Declaration ¶20 / 2018 GGE Report ¶23a).
17. *Responsibility and International Humanitarian Law*. International humanitarian law imposes obligations on States, parties to armed conflict, and individuals, not machines. (2019 GGE Report ¶17b).
- a. States, parties to armed conflict, and individuals remain at all times responsible for adhering to their obligations under applicable international law, including international humanitarian law.

- b. States must also ensure individual responsibility for the employment of means or methods of warfare involving the potential use of weapons systems based on emerging technologies in the area of LAWS in accordance with their obligations under international humanitarian law. (2019 GGE Report ¶17c).

18. *State Responsibility*. Under principles of State responsibility:

- a. Every internationally wrongful act of a State, including such conduct involving the use of a weapons system based on emerging technologies in the area of LAWS, entails the international responsibility of that State.
- b. The conduct of a State's organs such as its agents and all persons forming part of its armed forces, is attributable to the State. This includes any such acts and omissions involving the use of a weapons system based on emerging technologies in the area of LAWS, in accordance with applicable international law.

Good Practices Related to Human-Machine Interaction

- 19. Human-machine interaction, which may take various forms and be implemented at various stages of the life cycle of a weapon, should ensure that the potential use of weapons systems based on emerging technologies in the area of LAWS is in compliance with applicable international law, in particular international humanitarian law. In determining the quality and extent of human-machine interaction, a range of factors should be considered including the operational context, and the characteristics and capabilities of the weapons system as a whole. (Guiding Principle (c)).
- 20. At various stages of the life-cycle of a weapon, the following good practices related to human-machine interaction can strengthen compliance with international humanitarian law, strengthen accountability, and mitigate risks in the use of weapons systems based on emerging technologies in the area of LAWS:
 - a. Conducting legal reviews (2019 GGE Report ¶23b), including the practices described in paragraph 24;
 - b. Conducting rigorous testing and evaluation of systems (2019 GGE Report ¶23b), such as to ensure that they function as anticipated in realistic operational environments;
 - c. Providing for physical security and appropriate non-physical safeguards, including cyber security against hacking or data spoofing (Guiding Principle (f));

- d. Incorporating readily understandable human-machine interfaces and controls (2019 GGE Report ¶23b);
- e. Establishing policies, doctrine and procedures (based on 2019 GGE Report ¶23b), such as guidance on the ethical development and use of emerging technologies;
- f. Training personnel (2019 GGE Report ¶23b), such as training to enable system operators and commanders to understand the functioning, capabilities, and limitations of the system's autonomy in realistic operational conditions;
- g. Ensuring a domestic legal framework under which a State can hold its personnel accountable;
- h. Circumscribing weapons use through appropriate rules of engagement (2019 GGE Report ¶23b);
- i. Conducting operations under a responsible command;
- j. Reporting incidents that may involve violations;
- k. Conducting assessments, investigations, or other reviews of incidents that may involve violations; and
- l. Taking measures to mitigate the risk of unintended engagements, such as those described in paragraph 29.

Legal Reviews

- 21. In accordance with States' obligations under international law, in the study, development, acquisition, or adoption of a new weapon, means, or method of warfare, including such potential weapons systems based on emerging technologies in the area of LAWS, determination must be made whether its employment would, in some or all circumstances, be prohibited by international law. (based on Guiding Principle (e)).
- 22. Legal reviews, at the national level, in the study, development, acquisition, or adoption of a new weapon, means, or method of warfare are a useful tool to assess nationally whether potential weapons systems based on emerging technologies in the area of LAWS would be prohibited by any rule of international law applicable to that State in all or some circumstances. States are free to independently determine the means to conduct legal reviews, although the voluntary exchange of best practices could be beneficial, bearing in mind national security considerations or commercial restrictions on proprietary information. (2019 GGE Report ¶17i).
- 23. Weapons systems based on emerging technologies in the area of LAWS under development, or modification that significantly changes the use of existing weapons systems, must be reviewed as applicable to ensure compliance with international humanitarian law. (based on 2018 GGE Report ¶23(c)).

24. Legal reviews of weapons systems based on emerging technologies in the area of lethal autonomous weapons systems can include the following good practices:
- a. The legal review considers whether the weapon is of a nature to cause superfluous injury or unnecessary suffering, or if it is inherently indiscriminate, or is otherwise incapable of being used in accordance with international humanitarian law. (Building on and implementing paragraph 10 above).
 - b. If the use of the weapon is not prohibited, the legal review considers whether the use of the weapon is subject to the rules in any CCW Protocols or other rules applicable to certain types of weapons, applicable to the State in question.
 - c. The legal review is conducted with an appropriate understanding of the weapons' capabilities and limitations, its planned uses, and its anticipated effects in those circumstances.
 - d. The legal review advises on potential practical measures that would assist in ensuring compliance with international humanitarian law, such as the practices described in paragraph 20.

Risk Assessments and Mitigation Measures

25. Risk assessments and mitigation measures should be part of the design, development, testing, and deployment cycle of weapons systems based on emerging technologies in the area of LAWS. (based on Guiding Principle (g)).
26. During the design, development, testing, and deployment of weapons systems based on emerging technologies in the area of LAWS, the risks, *inter alia*, of civilian casualties, as well as precautions to help minimize the risk of incidental loss of life, injuries to civilians, and damage to civilian objects must be considered. Other types of risks should be considered, as appropriate, including but not limited to the risk of unintended engagements, risk of loss of control of the system, risk of proliferation, and risk of acquisition by terrorist groups. (2019 GGE Report ¶23a).
27. Where feasible and appropriate, verifiability and certification procedures covering all likely or intended use scenarios must be developed. The experience of applying such procedures should be shared, bearing in mind national security considerations and restrictions on commercial proprietary information. (based on 2018 GGE Report ¶23d).
28. Where feasible and appropriate, interdisciplinary perspectives must be integrated in research and development, including through independent ethics reviews, bearing in mind national security considerations and restrictions on commercial proprietary information. (2018 GGE Report ¶23b).
29. Measures to mitigate the risk of unintended engagements (*e.g.*, engagements against civilians, civilian objects, or unintended military targets) involving weapons systems based on emerging technologies in the area of LAWS, can include measures across the life-cycle of the weapons system to:

- a. control, limit, or otherwise affect the types of targets that the system can engage;
- b. control, limit, or otherwise affect the duration, geographical scope, and scale of the operation of the weapons system, such as the incorporation of self-destruct, self-deactivation, or self-neutralization mechanisms into munitions and weapons systems;
- c. reduce automation bias in system operators as well as unintended bias in artificial intelligence capabilities relied upon in connection with the use of the weapon system; and
- d. otherwise enhance control or improve decision-making over the use of force, including relating to timing, precision, and accuracy.

Implementation

30. The High Contracting Parties to the Convention intend to:

- a. Implement, as appropriate, these principles and good practices within each Party's respective national system;
- b. Share, on a voluntary basis, their national policies and experiences relevant to the implementation of these principles and good practices, bearing in mind national security considerations and restrictions on commercial proprietary information; and
- c. Keep these principles and good practices under review, and elaborate them by consensus as appropriate, while also continuing to consider and elaborate by consensus other possible measures and options related to the normative and operational framework on emerging technologies in the area of LAWS.