

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

13 March 2023

English only

Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons System
Geneva, 6-10 March, and 15-19 May 2023

Agenda item 5

Intensify the consideration of proposals and elaborate, by consensus, possible measures, including taking into account the example of existing protocols within the Convention, and other options related to the normative and operational framework on emerging technologies in the area of lethal autonomous weapon systems, building upon the recommendations and conclusions of the Group of Governmental Experts related to emerging technologies in the area of lethal autonomous weapon systems, and bringing in expertise on legal, military, and technological aspects

Draft articles on autonomous weapon systems – prohibitions and other regulatory measures on the basis of international humanitarian law (“IHL”)

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

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”):

Recognizing that the research and development of new technologies in the field of artificial intelligence is progressing at a rapid pace, potentially enabling novel and more sophisticated weapons with autonomous functions, including those weapon systems that, once activated, can identify, select, and engage targets with lethal force without further intervention by an operator (“autonomous weapon systems” for the purposes of these draft articles and without prejudice to any other understandings of this or similar terms for other purposes);

Mindful that these technologies could be used to improve the protection of civilians, while also mindful of the risks from their misuse in armed conflict, including risks of unintended engagements;

Following the Guiding Principles affirmed by the Group of Governmental Experts on emerging technologies in the area of lethal autonomous weapons systems (“Group”), endorsed by High Contracting Parties to the CCW, and annexed hereto;

Building on the consensus conclusions of the Group, which have been compiled thematically and annexed hereto;

Recalling that the International Court of Justice has opined that cardinal principles of IHL, including those annexed hereto, apply to all forms of warfare and to all kinds of weapons, those of the past, those of the present, and those of the future;

Desiring to clarify how existing cardinal principles of IHL, which are derived from established custom, from the principles of humanity, and from the dictates of public

conscience, remain applicable and impose requirements that protect the civilian population and combatants;

Affirm the following draft articles on the development, deployment, and use of autonomous weapon systems, which specify prohibitions and other regulatory measures for States to implement to effectively satisfy these IHL requirements in armed conflict.

Article 1: Preventing Autonomous Weapon Systems That, By Their Nature, Are Incapable of Use in Accordance With IHL

IHL prohibits the use of an autonomous weapon system 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 IHL. To prevent the development and use of such systems that cannot, under any circumstances, be used in compliance with IHL:

1. Autonomous weapon systems must not be designed to:
 - (a) Target civilians or civilian objects, or to spread terror among the civilian population;
 - (b) Conduct engagements that would invariably result in incidental loss of civilian life, injury to civilians, and damage to civilian objects excessive in relation to the concrete and direct military advantage anticipated; or
 - (c) Conduct engagements that would not be the responsibility of the commanders and operators using the system.
2. Autonomous weapon systems may only be developed such that their effects in attacks are capable of being anticipated and controlled as required in the circumstances of their use, by the principles of distinction and proportionality. Measures during development to this end should include the following:
 - (a) Rigorous testing and evaluation to inform an assessment of how the weapon system will perform in the anticipated circumstances of its use;
 - (b) Legal reviews that consider, *inter alia*, whether the autonomous weapon system is capable of use in accordance with the principles of distinction and proportionality;
 - (c) Consideration of potential precautions or features to be implemented in the design and use of the system to mitigate the risk of harm to civilians and civilian objects, such as measures to:
 - i. control, limit, or otherwise affect the types of targets that the system can engage;
 - ii. control, limit, or otherwise affect the duration, geographical scope, and scale of the operation of the weapon system, such as the incorporation of self-destruct, self-deactivation, or self-neutralization mechanisms into munitions or the system;
 - iii. reduce automation bias in system operators;
 - iv. reduce unintended bias in artificial intelligence capabilities relied upon in connection with the use of the weapon system; and
 - v. otherwise enhance control or improve decision-making over the use of force, including relating to timing, precision, and accuracy.

Article 2: Cardinal Principles of IHL Regulating the Use of an Autonomous Weapon System

1. IHL requires that the use of an autonomous weapon system to conduct attacks be consistent with, *inter alia*, distinction, proportionality, and precautions in attack.
2. A combatant's reliance on autonomous functions to identify, select, or engage targets:

- (a) Must be in good faith and in light of the information available at the time.
- (b) Must be consistent with due diligence in the implementation of the requirements and principles of distinction, proportionality, and precautions in attack, under which the lawfulness of such reliance may depend on, *inter alia*: (i) the expected performance of the autonomous function; (ii) the alternatives, consistent with customary military practice, to relying on the autonomous function; and (iii) the urgency of the situation.

Article 3: Regulatory Measures to Ensure Distinction in Conducting Attacks

To ensure effective implementation of the principle of distinction in attacks involving the use of autonomous weapon systems:

1. Commanders and operators must have the intention of striking specific or potential targets that constitute military objectives or of operating the system within specific locations constituting military objectives to deny enemy forces access to such locations.
2. The autonomous weapon system needs to perform with adequate reliability to enable, in the circumstances of its use, force to be directed against such targets or to remain within such locations.

Article 4: Regulatory Measures to Ensure Proportionality in Conducting Attacks

To ensure effective implementation of the principle of proportionality in attacks involving the use of autonomous weapon systems:

1. A commander must not direct or authorize subordinates to use the weapon system when the commander has assessed that the expected loss of civilian life, injury to civilians, and damage to civilian objects incidental to the use of the weapon system will be excessive in relation to the concrete and direct military advantage anticipated.
2. In making such assessments, commanders should consider all information that is relevant and available at the time, which may include information about, *inter alia*:
 - (a) The presence of civilians or civilian objects within the area where and during the time when the weapon system is expected to be operating;
 - (b) The incidence of targets constituting military objectives that could be engaged by the weapon system in the operational area and time period;
 - (c) The performance of the weapon system's autonomous functions in identifying, selecting, and engaging targets that constitute military objectives;
 - (d) The danger posed to civilians and civilian objects when the weapon system engages such targets;
 - (e) The effectiveness of any precautions taken to reduce the danger to civilians and civilian objects; and
 - (f) How the use of the weapon system is expected to protect friendly forces or civilians, neutralize enemy forces, or divert enemy forces' resources, attention, and movement.

Article 5: Regulatory Measures to Ensure Precautions in Attack

To ensure effective implementation of the requirement to take feasible precautions in planning and conducting attacks involving the use of autonomous weapon systems:

1. Commanders and planners need to assess what precautions are feasible in the circumstances of using autonomous weapon systems, which may include:

- (a) Adjusting the location where or times when the system is operating to reduce the likelihood of civilians being present;
 - (b) Giving warnings that enable steps to reduce the danger to civilians, such as avoiding locations where the system is operating;
 - (c) Monitoring the operation of the weapon system; and
 - (d) Other choices in the size and type of munitions and in the operation of the weapon system that reduce the danger to civilians and civilian objects, while offering the same or superior military advantage.
2. Commanders and planners may need to consider the use of an autonomous weapon system, among other weapons or alternative courses of action, when such use could constitute a feasible precaution in an attack.

Article 6: Regulatory Measures to Ensure Accountability

To ensure comprehensive accountability for the use of autonomous weapon systems:

1. A State should only deploy such weapon systems within the State's general framework for the implementation of IHL, including:
 - (a) dissemination of, and training on, IHL;
 - (b) domestic law under which the State can hold its personnel accountable;
 - (c) operation of such systems within a responsible chain of human command and control;
 - (d) internal mechanisms for the reporting of incidents that may involve violations of IHL;
 - (e) investigations or other reviews of such incidents; and
 - (f) appropriate actions in response to such investigations or reviews of incidents, including accountability for personnel.
2. States should provide with respect to their autonomous weapon systems:
 - (a) Readily understandable human-machine interfaces and controls;
 - (b) Guidance (*e.g.*, policies, doctrine, and procedures), consistent with applicable IHL, for personnel regarding the proper use of the weapon system;
 - (c) Training of personnel to understand such guidance and the capabilities and limitations of the weapon system's autonomous functions in the anticipated circumstances of its use; and
 - (d) Appropriate rules of engagement or other directives or orders circumscribing the use of the weapon system in military operations, consistent with applicable IHL.

Article 7: Status of These Draft Articles and Their Relationship to Law

These draft articles seek to clarify the requirements imposed by existing IHL and specify measures to effectively satisfy these requirements, but these draft articles neither create, nor authorize derogation from, any legal obligation.

Annex I

Relevant Cardinal Principles and Requirements of International Humanitarian Law

Prohibited Weapons

Any weapon system 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.

Distinction

Civilians and civilian objects must not be made the object of attack. Attacks may only be directed against combatants and other military objectives. A military objective, insofar as objects are concerned, means any object which by its nature, location, purpose or use makes an effective contribution to military action and whose total or partial destruction, capture or neutralization, in the circumstances ruling at the time, offers a definite military advantage.

Proportionality

The expected loss of civilian life, injury to civilians, and damage to civilian objects incidental to attacks must not be excessive in relation to the concrete and direct military advantage anticipated.

Precautions in Attack

Feasible precautions must be taken in planning and conducting attacks 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.

Annex II

Guiding Principles affirmed by the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons System and endorsed by High Contracting Parties to the Convention on Certain Conventional Weapons

It was affirmed that international law, in particular the United Nations Charter and International Humanitarian Law (IHL) as well as relevant ethical perspectives, should guide the continued work of the Group. Noting the potential challenges posed by emerging technologies in the area of lethal autonomous weapons systems to IHL, the following were affirmed, without prejudice to the result of future discussions:

- (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.

Annex III

Thematic compilation of the consensus conclusions and recommendations of the Group of Governmental Experts on emerging technologies in the area of lethal autonomous weapons systems from 2017-2022

Concepts and Characteristics

1. Characterization, or working definitions, should neither predetermine nor prejudice policy choices; they should be universally understood by stakeholders. (2018 report ¶22a)
2. Purely technical characteristics such as physical performance, endurance or sophistication in targeting acquisition and engagement may alone not be sufficient to characterize lethal autonomous weapons systems, especially in view of rapid evolution in technology. (2018 report 22b)
3. Technical characteristics related to self-learning (without externally-fed training data) and self-evolution (without human design inputs) have to be further studied. Similarly, attempting to define a general threshold level of autonomy based on technical criteria alone could pose difficulty as autonomy is a spectrum, its understanding changes with shifts in the technology frontier, and different functions of a weapons system could have different degrees of autonomy. (2018 report 22c)
4. Lethality as made explicit in the mandate of the Group of Governmental Experts (GGE) does not prejudice the application of and respect for all rules relevant to the conduct of hostilities. (2018 report 22d)
5. Autonomy in the military targeting and engagement cycle has to be studied further keeping in view that autonomy can exist throughout or during parts of the targeting cycle and could start to be applied increasingly in other contexts such as close combat. (2018 report 22e)
6. In the context of the CCW, 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. (2018 report 22f)
7. 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 lethal autonomous weapons systems, which is of core interest to the Group. (2019 report ¶19a)
8. Identifying and reaching a common understanding among High Contracting Parties on the concepts and characteristics of lethal autonomous weapons systems could aid further consideration of the aspects related to emerging technologies in the area of LAWS. (2019 report 19b)

Application of International Humanitarian Law

1. International humanitarian law continues to apply fully to all weapons systems, including the potential development and use of lethal autonomous weapons systems. (Guiding Principle (a)).
2. 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. (Guiding Principle (h)).
3. IHL imposes obligations on States, parties to armed conflict and individuals, not machines. (2019 Report 17b). In addition, States must comply with international

humanitarian law. Humans responsible for the planning and conducting of attacks must comply with international humanitarian law. (2022 Report 19).

4. A weapons system based on emerging technologies in the area of lethal autonomous weapons systems must not be used if it 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 the requirements and principles of IHL. (2019 Report 17h).

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

6. The IHL requirements and principles including inter alia distinction, proportionality and precautions in attack 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 lethal autonomous weapons systems. (2019 Report 17d).

7. Human judgement is essential in order to ensure that the potential use of weapons systems based on emerging technologies in the area of lethal autonomous weapons systems is in compliance with international law, and in particular IHL. (2019 Report ¶17e).

8. Compliance with the IHL requirements and principles, including inter alia distinction, proportionality and precautions in attack, in the potential use of weapons systems based on emerging technologies in the area of lethal autonomous weapons systems 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. (2019 Report 17f).

9. In cases involving weapons systems based on emerging technologies in the area of lethal autonomous weapons systems not covered by the CCW 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 Report 17g).

Responsibility and Accountability

1. 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. (Guiding Principle (b)).

2. 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. (Guiding Principle (d)).

3. Humans must at all times remain accountable in accordance with applicable international law for decisions on the use of force. (2018 Report 23a).

4. Responsibility for the deployment of any weapons system in armed conflict remains with States. States must ensure accountability for lethal action by any weapon system used by the State's forces in armed conflict in accordance with applicable international law, in particular international humanitarian law. (2017 Report 16c).

5. States, parties to armed conflict and individuals remain at all times responsible for adhering to their obligations under applicable international law, including IHL. 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 lethal autonomous weapons systems in accordance with their obligations under IHL. (2019 Report ¶17c).

6. Accountability for the use of force in armed conflict must be ensured in accordance with applicable international law, including through the operation of any emerging weapons systems within a responsible chain of command and control. (2018 Report 23e).

7. Human responsibility for the use of weapons systems based on emerging technologies in the area of lethal autonomous weapons systems can be exercised in various ways across the life-cycle of these weapon systems and through human-machine interaction. (2019 Report 21).

8. For the purposes of its work, the Group recognized that every internationally wrongful act of a State, including those potentially involving weapons systems based on emerging technologies in the area of LAWS entails international responsibility of that State, in accordance with international law. (2022 Report 19).

Human-Machine Interaction

1. 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. (Guiding Principle (c)).

2. Touch points in the human-machine interface include: (0) political direction in the pre-development phase; (1) research and development; (2) testing, evaluation and certification; (3) deployment, training, command and control; (4) use and abort; (5) post-use assessment. (2018 Report 23).

3. Necessary investments in human resources and training should be made in order to comply with IHL and retain human accountability and responsibility throughout the development and deployment cycle of emerging technologies. (2018 Report 23g).

4. Human responsibility for the use of force must be retained. To the extent possible or feasible, this could extend to intervention in the operation of a weapon if necessary to ensure compliance with IHL. (2018 Report 23f).

Weapons Reviews

1. 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. (Guiding Principle (e)).

2. 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 lethal autonomous weapons systems 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 Report 17(i)).

3. Weapons systems under development, or modification which significantly changes the use of existing weapons systems, must be reviewed as applicable to ensure compliance with IHL. (2018 Report 23(c)).

Risk Assessments and Mitigation Measures

1. Risk assessments and mitigation measures should be part of the design, development, testing and deployment cycle of emerging technologies in any weapons systems. (Guiding Principle (g)).

2. 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. (Guiding Principle (f)).

3. During the design, development, testing and deployment of weapons systems based on emerging technologies in the area of lethal autonomous weapons systems, 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 Report 23a).

4. Risk mitigation measures can include: rigorous testing and evaluation of systems, legal reviews, readily understandable human-machine interfaces and controls, training personnel, establishing doctrine and procedures, and circumscribing weapons use through appropriate rules of engagement. (2019 Report 23b).

5. 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 or commercial restrictions on proprietary information. (2018 Report 23d)

6. Where feasible and appropriate, inter-disciplinary 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 Report 23b).
