SUBMISSION ON AUTONOMOUS WEAPON SYSTEMS TO THE UNITED NATIONS SECRETARY-GENERAL

RE: ODA-2024-00019/LAWS

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 78/241 "Lethal autonomous weapon systems", adopted by the General Assembly on 22 December 2023, which requested the Secretary-General to seek views on "ways to address the related challenges and concerns [that autonomous weapon systems] raise from humanitarian, legal, security, technological and ethical perspectives and on the role of humans in the use of force."

The ICRC is a neutral, impartial and independent humanitarian organization. Through the Geneva Conventions of 1949 (Geneva Conventions) and other international legal instruments, the ICRC is mandated by States to protect and assist persons affected by armed conflict. The ICRC also endeavours to prevent suffering by promoting and strengthening (including, where necessary, through contributing to the development of) international humanitarian law (IHL) and universal humanitarian principles.

We make this submission based on our 160 years' experience of humanitarian action, during which we have witnessed the significant humanitarian consequences of armed conflict, whether that be direct or indirect harm to people, objects, communities and societies.

Over the course of our history, the ICRC has played a significant role in the development of many of the IHL rules limiting or prohibiting the use of weapons of concern. By drawing the attention of States and the public to the unacceptable effects of certain weapons on combatants and civilians, the ICRC has helped to create the conditions for the development of the law in this area, including with respect to nuclear, chemical and biological weapons, blinding laser weapons, anti-personnel mines, explosive remnants of war and cluster munitions.

The ICRC's work related to the regulation of weapons is always driven by an "effects-based" approach. This means that we observe the actual, or, in the case of new weapons not yet deployed, the foreseeable effects of the use of weapons – both on civilians and combatants. We then raise our concerns regarding particular weapons that pose legal or ethical challenges, or present other risks of harm to those affected by armed conflict. Our assessment is that the unconstrained development and use of autonomous weapon systems (AWS) raises these concerns.[1]

1. NEED FOR NEW, BINDING INTERNATIONAL LAW

Existing IHL, in particular, the rules on means and methods of warfare and those governing the conduct of hostilities for the protection of civilians, already regulate and constrain the use of AWS. They also prohibit certain types of AWS, such as inherently indiscriminate AWS. However, the ICRC considers that existing IHL rules do not hold all the answers to the humanitarian, legal and ethical questions raised by AWS, and States hold different views about what limits and requirements for the design and use of AWS derive from existing rules of IHL.

The ICRC is convinced that new rules are urgently needed to clarify and specify how IHL applies to AWS, as well to address wider humanitarian risks and fundamental ethical concerns. New, legally binding rules would offer the benefits of legal certainty and stability. We are concerned that, without such rules, further developments in the design and use of AWS may give rise to practices that erode the

[1] ICRC position paper on autonomous weapon systems, May 2021, available at: https://www.icrc.org/en/document/icrc-position-autonomous-weapon-systems; Joint Call by the United Nations Secretary-General and the President of the ICRC, October 2023, available at: https://www.icrc.org/en/document/joint-call-un-and-icrc-establish-prohibitions-and-restrictions-autonomous-weapons-systems; ICRC commentary on the guiding principles of the CCW GGE, 2020, available at: https://documents.unoda.org/wp-content/uploads/2020/07/20200716-ICRC.pdf

protections presently afforded to the victims of war under IHL and the principles of humanity and the dictates of public conscience.

The recommendations that the ICRC has made are in line with our long-standing mandate and practice to prepare the development of IHL, including specific prohibitions and restrictions on weapons.[2] The ICRC makes this submission to the United Nations Secretary-General in order to assist States in conceiving of how a legally binding instrument could be drafted to address the specific concerns raised by AWS.

The groundwork has been done, with over ten years of UN discussions including in the Human Rights Council, under the Convention on Certain Conventional Weapons, and the UN General Assembly. Now, faced with the realities of current weapon technologies and their use in today's conflicts, States must shift their focus from the horizon and negotiate new international rules to respond to the concrete threats before us.

2. GUIDING PRINCIPLE: NEED FOR HUMAN CONTROL OVER THE USE OF FORCE AND EFFECTS

IHL requires weapon users in armed conflict to be able to anticipate, control and limit the effects of those weapons. For example, the prohibition against indiscriminate attacks means that, to be lawful, a weapon must be able to be "directed" at a specific military objective (which necessitates the ability to control the weapon's effects), and its effects must be capable of being "limited".[3] In terms of anticipation, the rule on proportionality calls for weapon users to form reasonable "expectations" about potential incidental loss of civilian life, injury to civilians, damage to civilian objects, or a combination thereof, and to reasonably "anticipate" the concrete and direct military advantage from an attack.[4]

While these obligations do not necessarily demand direct human control over the weapon system itself at all stages of its deployment and use and therefore do not prohibit all AWS in all circumstances, they do require human control over the weapon system's *effects* in the circumstances of a specific attack. This calls for a combination of prohibitions and restrictions both on the design of the weapon system itself, and its operating parameters.

Many States, civil society and others, including in the context of the Convention on Certain Conventional Weapons' Group of Governmental Experts (GGE),[5] have also consistently emphasized the importance of human control with regard to autonomous weapons.[6] Indeed, the GGE, in its 2023 report, confirmed that "control [with regard to AWS] is needed to uphold compliance with international law, in particular IHL, including the principles and requirements of distinction, proportionality and precautions in attack."[7]

This principle should underpin and guide the drafting and interpretation of a legally binding instrument on AWS, even if "human control" is not explicitly included as a requirement.

[2] ICRC position paper on autonomous weapon systems, May 2021, note 1 above.
[3] Customary international humanitarian law (CIHL), Rule 12.
[4] Additional Protocol I, Art. 51(5)(b); CIHL, Rule 14.
[5] For example, working paper submitted by Austria (CCW/GGE.1/2023/WP.1) para. 2; Draft Protocol on AWS submitted by Argentina and others (CCW/GGE/1/2023/WP.6), Art. 2(2); working paper of the Russian Federation, 2023, p. 2.
[6] Joint Statement on Lethal Autonomous Weapons System, UN General Assembly First Committee, 21 October 2022, delivered by Austria on behalf of a group of 70 States, available at: https://estatements.unmeetings.org/estatements/11.0010/20221021/A1jJ8bNfWGlL/KLw9WYcSnnAm_en.pdf
[7] GGE.1/2023/2, para. 21.

3. PREAMBLE

The new legally binding instrument could include preambular paragraphs that, *inter alia*:

Reaffirm the existing framework, and the need for States to comply at all times with applicable law, including international humanitarian law and international human rights law.

> *Rationale:* The United Nations General Assembly has affirmed that international law, in particular the Charter of the United Nations, international humanitarian law and international human rights law, applies to AWS.[8] Further, the CCW's Group of Governmental Experts on lethal autonomous weapon systems (LAWS) has repeatedly emphasized that the rules and principles of IHL, including distinction, proportionality and precautions in attack, must be adhered to in the development, deployment and use of autonomous weapon systems.[9]

Acknowledge the range of concerns that have been expressed about the risks posed by autonomous weapons.

> *Rationale:* The United Nations General Assembly has stressed the serious challenges that AWS raise from humanitarian, legal, security, technology and ethical perspectives.[10] The UN Secretary-General, in the New Agenda for Peace, also recognized that, in the absence of specific multilateral regulations, the design, development and use of AWS raise humanitarian, legal, security and ethical concerns.[11]

Reaffirm the need to continue the codification and progressive development of the rules of international law applicable in armed conflict.

> *Rationale:* Such reaffirmation would directly mirror the one found in the preamble to the CCW itself.[12] It also reflects the need to uphold protections for those affected by armed conflict and, when necessary, to strengthen these protections in the face of the development of new weapons, means and methods of warfare. A similarly progressive approach was taken with the St Petersburg Declaration, which pre-emptively prohibited the use in war of certain weapons that were foreshadowed by technological developments at the time.[13] It did so on the basis that their use would be contrary to the laws of humanity, and in order to prevent unacceptable harm before it occurred.

Reaffirm that, in cases not covered by the instrument or by other international agreements, civilians and combatants 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.

> *Rationale:* The "Martens clause", first found in the 1899 Hague Convention (II) with respect to the laws and customs of war on land,[14] provides a link between ethical considerations and IHL, which makes it particularly relevant to the assessment of autonomous weapon systems. Ethical considerations provide an important basis for the adoption of prohibitions and restrictions on the development and use of autonomous weapon systems. The Guiding Principles affirmed by the GGE on LAWS state that relevant ethical principles should guide the GGE's work.

---

[8] UNGA Resolution A/RES/78/241, PP 1.
[9] For example, CCW/GGE.1/2023/2 subparas 21(a) and 22; CCW/MSP/2019/9 (Guiding Principles Affirmed by the GGE), subpara. (a).
[10] UNGA Resolution A/RES/78/241, PP 3.
[11] "Our Common Agenda Policy Brief 9: A New Agenda for Peace", p. 25.
[12] 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 (CCW), Preamble.
[13] Declaration Renouncing the Use, in Time of War, of Explosive Projectiles Under 400 Grammes Weight, St Petersburg, 29 November/11 December 1868.
[14] First appeared in the preamble to the 1899 Hague Convention (II); today, versions of the Martens Clause have found entry in many international treaties. In addition to Additional Protocol I, see also the common article on denunciation of the Geneva Conventions, i.e. First Geneva Convention (GC I), Article 63; Second Convention (GC II), Article 62; Third Convention (GC III), Article 142; and Fourth Convention (GC IV), Article 158; as well as Additional Protocol II, Preamble, para. 4. See also the Convention on Certain Conventional Weapons (1980), Preamble, para. 5, and the Convention on Cluster Munitions (2008), Preamble, para. 11. Elements of the Martens Clause can also be found in other treaties; see the Geneva Gas Protocol (1925), Preamble, paras 1–3; Biological Weapons Convention (1972), Preamble, para. 9; Anti-Personnel Mine Ban Convention (1997), Preamble, para. 8; ICC Statute (1998), Preamble, para. 2; and Treaty on the Prohibition of Nuclear Weapons (2007), Preamble, para. 11.

Reaffirm the need for States, in the study, development, acquisition or adoption of AWS, to determine whether the weapon system's employment would, in some or all circumstances, be prohibited by international law.

Rationale: Effective weapons reviews are a clear procedural obligation for States Party to Protocol I of 8 June 1977 additional to the Geneva Conventions and, for all States, they are a crucial step in ensuring compliance with international law.

(4) SCOPE

The instrument should provide that it applies to all High Contracting Parties and that, in case of armed conflicts not of an international character occurring in the territory of one or more High Contracting Parties, each party to the conflict shall be bound to apply the instrument's prohibitions and restrictions.

Rationale: The instrument should make clear that its prohibitions and restrictions also bind non-state armed groups party to non-international armed conflicts occurring in the territory of one or more High Contracting Parties.

5. DEFINITIONS

The instrument should contain an unambiguous definition covering the general category of AWS, to which the whole instrument will apply. Within that category, certain types of AWS subject to specific prohibitions should be further defined (see discussion under section 6 below).

"Autonomous weapon system" means a weapon system that is designed to select and engage one or more targets without the need for human intervention after activation.

Rationale: In the understanding of the ICRC, as well as many States and other actors, AWS are weapon systems that operate in such a manner that the selection of, and application of force to, targets occur without human intervention.[15] This understanding of AWS is therefore not based on a specific technology, but rather on the role of humans in the process of target selection and application of force. Such a technology-neutral approach to characterization of AWS is necessary to ensure that new rules address the key concerns raised by these weapons, and remain relevant in the face of technological developments.

"Without the need for human intervention" could be further defined as meaning that, after initial activation by a human, the application of force is triggered in response to information from the environment received through sensors measuring phenomena such as heat, light, movement, shape, velocity, weight or acoustic or electromagnetic signals; and on the basis of a generalized "target profile" such as the shape, infrared or radar "signature", speed and direction of a type of military vehicle, etc.[16] "Human intervention", for these purposes, should be understood as excluding human inputs or actions that do not materially affect the autonomous functions of target selection or engagement.

The process of applying force in this manner can be implemented with a wide variety of weapons, munitions and platforms, which can be technically rudimentary or complex and may, but need not, rely on artificial intelligence technologies, including machine learning. The use in armed conflict of weapons that function in this manner is already a reality. Mines and sensor-fused munitions, certain "loitering" munitions, missile and rocket defence systems and "active protection" weapons that are already in use match this description of AWS, or have modes that operate in this manner.

By contrast, autonomy in aspects of a weapon system other than the selection of, and application of force to, targets – for instance in navigation, intelligence collection or decision support – does not of itself render that system an AWS, and the analysis presented here does not extend to weapon systems that rely on autonomy solely in such other functions.

[15] ICRC position paper on autonomous weapon systems, May 2021, available at: https://www.icrc.org/en/document/icrc-position-autonomous-weapon-systems

[16] The concept of a "target profile" has been referred to in the context of the GGE, e.g. 2019 Report CCW/GGE.1/2019/3, and could be the subject of definition in an instrument.

The following definitions, relevant to later provisions, could be incorporated from existing international agreements:

"Military objective" means, so far as objects are concerned, 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.

Rationale: This definition can be found in the Protocol I additional to the Geneva Conventions, Article 52(2) and in Amended Protocol II to the CCW.

"Self-destruction mechanism" means an incorporated automatically functioning mechanism which is in addition to the primary initiating mechanism of the AWS and which secures the destruction of the AWS into which it is incorporated.

"Self-deactivating" means automatically rendering an AWS inoperable by means of the irreversible exhaustion of a component, for example a battery, that is essential to the operation of the AWS.

Rationale: Contained in the 2008 Convention on Cluster Munitions.

6. PROHIBITIONS

In accordance with the principles and objectives set out in the preamble, the instrument must contain specific prohibitions on types of AWS. These should be as clear and well defined as possible.

UNPREDICTABLE AWS

The instrument should provide that it is prohibited in all circumstances to develop, produce, otherwise acquire, stockpile or retain, or transfer, directly or indirectly to anyone, or to use any autonomous weapon system that is designed or of a nature, or used in such a manner that does not allow a human user to both (1) understand, predict and explain how the AWS will function in any normal or expected circumstances of use, in particular what circumstances or conditions will trigger the system to apply force, and (2) predict and limit the effects of the AWS in all such circumstances as required by IHL.

Rationale: 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 prohibition will be particularly relevant for AWS that function in opaque ways (the "black box" challenge), such as AWS relying on artificial intelligence 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.

This concern would arise in AWS that incorporate machine learning, the functioning of which can change after the commencement of an attack so that force may be applied in circumstances and in a manner unforeseen to the human user. Complex swarm technologies may also exhibit emergent behaviours.

ANTI-PERSONNEL AWS

The instrument should provide that it is prohibited in all circumstances to develop, produce, otherwise acquire, stockpile or retain, or transfer, directly or indirectly to anyone, or to use any autonomous weapon system:

- that is designed or used in such a manner to be triggered by the presence, proximity or contact of one or more persons, or
- the target profile of which otherwise represents one or more persons.

Rationale: The ICRC finds it hard to envisage realistic combat situations where the use of autonomous weapons against humans would not pose a significant risk of IHL violations. While the user or commander may have made a general assessment that one or more people in the area constitute a lawful target at the time of launching the AWS, those peoples' actions, intentions (such as surrendering) and physical state (such as being wounded) – and hence their qualification as a lawful target – can change rapidly.

A prohibition against anti-personnel AWS is also required based on the unacceptability of such weapons from an ethical perspective, which is reflected in public conscience and in the principles of humanity.[17] The killing of humans based on a machine process would simultaneously undermine the human agency of the person using force and the human dignity of the person against whom force is used.

This is consistent with the enhanced protection that the international community afforded to human beings compared to objects, when agreeing on prohibitions and restrictions on weapons such as bullets which explode within the human body, blinding laser weapons and anti-personnel mines.

7. RESTRICTIONS

The instrument should provide that, in the use of AWS other than those which are prohibited by the other articles, measures shall be taken to protect civilians and civilian objects, and other protected persons, from the effects of AWS. Such measures should include, but not be limited to:

- restricting targets of the AWS to only those which are military objectives by nature
- limiting the location where, time that and situation in which the AWS is operating, including to avoid concentrations of civilians or civilian objects[18]
- limiting the number of engagements that the AWS can undertake
- ensuring, to the maximum extent feasible, the ability for a human user:
    - to effectively supervise, and
    - to, in a timely manner, intervene and, where appropriate, deactivate operation of the AWS
- AWS that do not allow a human user to, where appropriate and in a timely manner, intervene and deactivate operation of the AWS, must be equipped with an effective self-destruction or self-neutralization mechanism and have a back-up self-deactivation feature, which is designed so that the AWS will no longer function as an AWS when it no longer serves the military purpose for which it was launched.[19]

Rationale: Even in the use of an AWS that is sufficiently predictable and designed and used only against objects, the user's reduced ability to know all specifics of the attack, including the ultimate target and any incidental harm, will still create residual challenges for their context-specific application of IHL's rules on the conduct of hostilities. To reduce the risk of violations, and to preserve control over the effects of any attacks, stringent conditions must be adopted on the use of these AWS; on the types of objects to be targeted, the duration, geographical scope, scale and situations of use, and incorporation of possibilities for effective human supervision, intervention and deactivation, and/or self-destruction or self-neutralization.

The instrument could exclude from the scope of these restrictions the non-autonomous use of weapon systems that have the capacity to function as AWS but are not used as such in specific circumstances.

8. FURTHER PROVISIONS

Relevant existing IHL provisions could be restated in an instrument, in addition to the more specific provisions listed under sections 6 and 7, such as the prohibition of indiscriminate and disproportionate attacks, and the obligations to take all feasible precautions in attack.

National legislation, policy and practical measures are necessary to implement any international instrument at the domestic level. The instrument should require States Parties to take all appropriate legal, administrative and other measures, including the imposition of penal sanctions, to prevent or suppress any activity prohibited to States Parties under the instrument undertaken by persons or on

[17] GC I, Art. 63; GC II, Art. 62; GC III, Art. 142; GC IV, Art. 158.
[18] NB similar language found in Additional Protocol I, Art. 51(5)(a) and CCW Amended Protocol II and Protocol III.
[19] Language of CCW Amended Protocol II.

territory under their jurisdiction or control. The obligation to carry out legal reviews could be included in a substantive article, whether in addition to or replacing reaffirmation in the Preamble (see above).

Requirements for testing and training may be considered for incorporation as part of national implementation measures.

The instrument may also need to contain other elements such as:

- reporting requirements
- verification and compliance, and cooperation mechanisms
- procedures for amendments and entry into force.

However, the finalization of elements such as verification mechanisms should not stand in the way of establishing clear prohibitions and restrictions on development and use. Many norms of IHL prohibiting certain weapons have been enacted without such verification provisions.[20]

The ICRC is grateful for the opportunity to share the above views and recommendations on ways to address the challenges and concerns raised by AWS 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 AWS.

19 March 2024

[20] For instance, the CCW protocols and the Biological Weapons Convention.