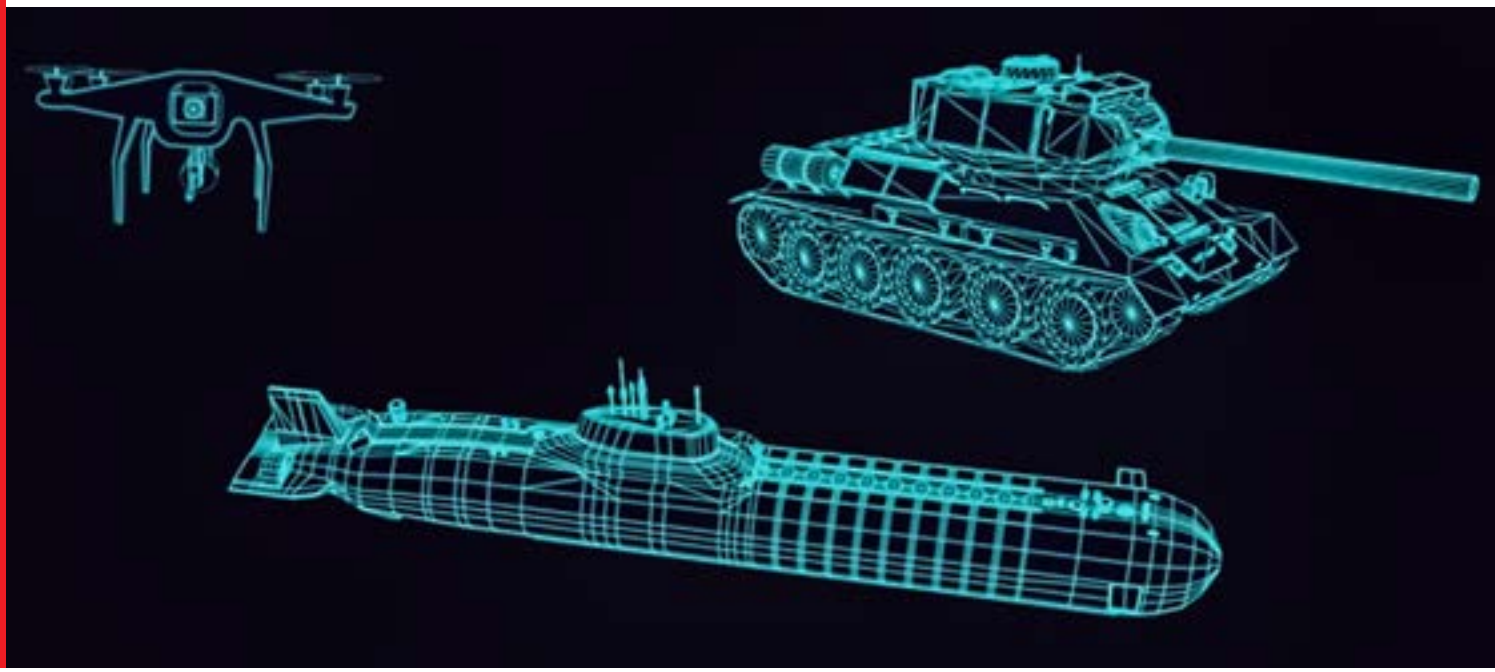
ICRC

POSITION PAPER

AUTONOMOUS WEAPON SYSTEMS AND INTERNATIONAL HUMANITARIAN LAW: SELECTED ISSUES



EXECUTIVE SUMMARY

The deployment of weapon systems with increasingly autonomous modes or functions is a fact of contemporary conflicts. Autonomous weapon systems (AWS) are already used against *military objectives by nature*, such as missiles, radars and warships, generally in environments where civilians are absent or excluded, and often under human supervision.

However, recent trends towards loosening the constraints on where, and against what, such weapons may be used, together with the development of swarm technologies and the integration of artificial intelligence (AI) into targeting processes, risk eroding existing limitations provided by international humanitarian law (IHL). These developments heighten concerns about the potential loss of human control over the use of force, which can have significant humanitarian and legal implications in armed conflict. Driven by the need to uphold and strengthen protections for people affected by armed conflict, the ICRC has called on States to urgently establish new legally binding rules, with clear prohibitions and restrictions on AWS.

In the view of the ICRC and many States and other actors, AWS are weapon systems that, once activated, can select and engage one or more targets without further human intervention. After activation, strikes are triggered in response to environmental inputs and generalized target profiles, which means the human user cannot determine in advance who or what will be struck, or exactly when or where the force will be applied. The loss of human control and judgement in decisions over life and death raises profound humanitarian, legal and ethical concerns, including the risk of harm to civilians and combatants, the danger of conflict escalation and challenges to compliance with IHL.

Under IHL, AWS are weapons and are therefore subject to existing prohibitions and restrictions, including those applicable to biological, chemical, incendiary and indiscriminate weapons. Weapons that cannot be reliably directed at specific military objectives or whose effects cannot be limited as required by IHL are prohibited. This is particularly relevant for AWS that function in opaque ways, for example where machine learning allows their functioning to change after being activated, which prevent the user from understanding, predicting or explaining the system's output. This results in a lack of control over their effects and potentially renders them indiscriminate by nature.

Moreover, IHL rules on the conduct of hostilities, namely distinction, proportionality and precautions in attack, presuppose context-specific human judgement. In particular, it is humans who must determine the lawfulness of attacks that they plan, decide upon or execute, and they remain accountable for these determinations, which cannot be delegated to machine processes. Because AWS do not allow the user to choose, or even know, the specific target or the precise timing and location of a strike, ensuring compliance with IHL requires commanders to anticipate and assess in advance the lawfulness of all possible strikes by an AWS. In particular, commanders and other users of AWS must ensure that any and all possible strikes using an AWS will comply with IHL, whoever or whatever may trigger such strikes, throughout the entire area and during the whole period in which the AWS is active. When doing so, they need to account for all reasonably foreseeable and relevant changes in circumstances between the AWS activation and the strike. Such an assessment is only possible where strict constraints are imposed on the variables applying to the AWS and its operating environments in order to limit the number of potential outcomes.

Challenges to ensuring compliance with IHL arise especially when AWS are used against *objects that are military objectives by purpose or use*, as their characterization depends on nuanced, context-dependent cues ill-suited to generalized target profiles. Ensuring compliance is especially difficult when attacks are directed against persons, whose status and protection can change rapidly before a strike. Without clarity on limits, future practice risks eroding existing protections under IHL and undermining the principles of humanity and the dictates of public conscience.

Past practice shows that States have repeatedly adopted prohibitions and restrictions to address the humanitarian and legal risks posed by new weapons. In response to the indiscriminate effects of victim- or contact-detonated mines, States moved from restrictions and precautionary measures to a full prohibition on anti-personnel landmines. Similar approaches were taken for other weapons with indiscriminate effects. Notably, some of the most effective humanitarian disarmament efforts were those that prohibited weapons before they were widely developed and used on the battlefield, thereby preventing suffering rather than responding only after unacceptable consequences had already become apparent. This was the case for the 1868 St. Petersburg Declaration on certain exploding bullets and Protocol IV's prohibition on blinding laser weapons.

Building on this time-tested approach, to uphold and reinforce IHL and provide clarity to developers and users, it is essential to articulate specific conditions and limits on AWS, including those that cannot be employed in compliance with IHL. A new international instrument must prohibit unpredictable AWS, whose functioning and effects cannot be sufficiently understood, predicted or explained, and AWS designed or used to target humans directly (anti-personnel AWS). It must also restrict all other AWS through limits on the type of targets; limits on the duration, scale and geographic scope of operation; limits on the situations of use, namely constraining them to situations where civilians or civilian objects are not present; and requirements for effective human supervision, timely intervention and deactivation, or, where necessary, self-destruction or self-neutralization.

Without timely agreement on a legally binding instrument that prohibits unacceptable AWS and regulates all others, emerging patterns in the research, development and use of AWS may become difficult to reverse, with lasting humanitarian and legal consequences. Establishing these prohibitions and restrictions in clear and binding international law is therefore an urgent humanitarian and legal priority.

TABLE OF CONTENTS

I. INTRODUCTION

The International Committee of the Red Cross (ICRC) welcomes the expert and intergovernmental discussions on autonomous weapon systems (AWS) that have taken place over the past decade, including in particular in the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems (GGE) within the framework of the Convention on Certain Conventional Weapons (CCW), as well as in the United Nations General Assembly (UNGA) First Committee and in the Human Rights Council.

In 2019, the High Contracting Parties to the CCW reaffirmed the applicability of international humanitarian law (IHL) to the development and use of AWS,[1] which was echoed by UNGA resolutions in 2023 and 2024.[2] Although IHL does not explicitly refer to AWS,[3] existing IHL principles and rules, which are technology neutral, impose strict constraints on AWS and prohibit the development and use of certain types of AWS altogether.

To uphold and strengthen the existing framework and to provide developers and users with the clarity needed to comply with legal obligations, it is essential to clearly articulate the specific conditions and limits on the development and use of AWS that may be derived from these principles and rules, including the types of AWS that cannot be employed in compliance with IHL.[4]

This paper lays out the ICRC's views on selected IHL issues relating to the development and use of AWS. It first proposes a definition of AWS (Part II), and then offers a snapshot of existing military practices in the use of AWS (Part III). Turning to IHL, the paper underscores the IHL requirement for context-specific assessments and the challenges raised in this regard by the functioning of AWS (Part IV). The paper then analyses how the IHL rules prohibiting and restricting weapons (Part V) and the IHL rules on the conduct of hostilities (Part VI) apply to AWS. In Part VII, it explores avenues to address the challenges raised by AWS to compliance with IHL.

The ICRC submits this paper to support this aspect of States' deliberations on AWS, in line with its mission and mandate to prevent suffering during armed conflicts by promoting and strengthening IHL and universal humanitarian principles. The paper builds on positions expressed by States and others in the CCW GGE, the UNGA and elsewhere; on insights from expert meetings convened and reports published by the ICRC on the humanitarian, ethical and legal concerns raised by AWS; and on the ICRC's recommendation to adopt new legally binding rules to address these concerns.[5]

This text was originally published in October 2025 as a position paper. This version includes an executive summary and incorporates various editorial updates.

1 UN, *Annex III – Guiding Principles affirmed by the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems, in the Final Report of the Meeting of 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*, CCW/MSP/2019/9, December 2019 (hereafter GGE Guiding Principles), letter (a), available at: https://docs.un.org/en/CCW/MSP/2019/9.

2 UN Doc. A/RES/78/241, PP1; UN Doc. A/RES/79/62, PP2; Group of 70 States, *Joint Statement on Lethal Autonomous Weapons Systems*, First Committee of the United Nations General Assembly, October 2022 (hereafter Group of 70 Joint Statement), available at: https://reachingcriticalwill.org/images/documents/Disarmament-fora/1com/1com22/statements/21Oct_LAWS.pdf.

3 For the specific case of anti-personnel landmines, see below Part VII, B.

4 ICRC, *Submission On Autonomous Weapon Systems to the United Nations Secretary-General*, March 2024 (hereafter ICRC, *Submission on AWS to the UNSG*), available at: https://www.icrc.org/sites/default/files/wysiwyg/war-and-law/icrc_submission_on_autonomous_weapons_to_unsg.pdf; 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.

5 ICRC, *Submission on AWS to the UNSG*; 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*, September 2024 (hereafter *2024 Challenges Report*), Part V(2), available at: https://shop.icrc.org/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-pdf-en.html; ICRC, *Position on Autonomous Weapon Systems and Background Paper*, May 2021 (hereafter ICRC, *Position Paper on AWS*), available at: https://www.icrc.org/sites/default/files/document_new/file_list/icrc_position_on_aws_and_background_paper.pdf; ICRC, *Autonomy, Artificial Intelligence and Robotics: Technical Aspects of Human Control*, August 2019, available at: https://www.icrc.org/sites/default/files/document/file_list/autonomy_artificial_intelligence_and_robotics.pdf; ICRC, *Ethics and Autonomous Weapon Systems: An Ethical Basis for Human Control?*, April 2018 (hereafter ICRC, *Ethics and AWS*), available at: https://www.icrc.org/sites/default/files/document/file_list/icrc_ethics_and_autonomous_weapon_systems_report_3_april_2018.pdf.

This paper does not provide a definitive analysis of all IHL issues arising from increasing autonomy in weapon systems. Instead, it focuses on key IHL principles and rules that govern the use of AWS in warfare, in particular the principles of distinction, proportionality and precautions. The conditions and limits identified below are illustrative, and further efforts to clarify the implementation of existing IHL obligations may yield additional conditions and limits derived from these or other rules.

Several issues are beyond the scope of this paper, including:

- challenges that AWS raise for accountability and, in particular, the ability to hold AWS users to account for harm done and violations of IHL[6]
- the ethical concerns raised by the development and use of AWS and that must be addressed in addition to satisfying legal requirements
- the Martens Clause, which confirms the continued protection afforded by the principles of humanity and the dictates of public conscience[7]
- issues raised by the use of AWS in law enforcement operations, whether during or outside an armed conflict, and limits arising from international human rights law[8] or other rules of international law
- the specific proliferation risks that AWS pose, and the potential human cost of their widespread availability and misuse during or outside armed conflicts.[9]

6 GGE Guiding Principles, letters (b) and (d); Marta Bo, *et al., Retaining Human Responsibility in the Development and Use of Autonomous Weapon Systems: On Accountability for Violations of International Humanitarian Law Involving AWS*, SIPRI, October 2022, available at: https://www.sipri.org/publications/2022/policy-reports/retaining-human-responsibility-development-and-use-autonomous-weapon-systems-accountability.

7 Geneva Conventions of 1949 (I-IV), Art. 63, 62, 142, 158 respectively. For the ethical challenges raised by AWS, see ICRC, *Ethics and AWS*, note 5 above.

8 See e.g. UN Human Rights Council, *Report of the Special Rapporteur on Extrajudicial, Summary or Arbitrary Executions*, Christof Heyns, A/HRC/23/47, para. 94, available at: https://docs.un.org/en/A/HRC/23/47; Human Rights Watch, *A Hazard to Human Rights: Autonomous Weapons Systems and Digital Decision-Making*, 2025, available at: https://www.hrw.org/report/2025/04/28/hazard-human-rights/autonomous-weapons-systems-and-digital-decision-making.

9 See e.g. CARICOM Conference, *CARICOM Declaration on Autonomous Weapons Systems*, 6 September 2023, Section I: Fundamental Principles and Concerns, available at: https://www.caricom-aws2023.com/__files/ugd/b69acc_4d08748208734b3ba849a4cb257ae189.pdf.

II. DEFINITION OF AUTONOMOUS WEAPON SYSTEMS

For the purposes of this paper, and consistent with the ICRC's broader understanding and that of several States and other actors,[10] AWS are weapon systems that select and engage (that is, apply force to) to one or more targets without human intervention. "Without human intervention" means 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.[11]

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 (AI) – including machine learning – technology. The use in armed conflict of weapons that function in this manner is already a reality. Mines and sensor-fused munitions, certain "loitering" munitions, close-in weapon systems (CIWS) and active protection systems (APS) that are already in use match this characterization of AWS or have modes that function in this manner.[12]

By contrast, autonomy in aspects of a weapon system other than the selection of and application of force to targets – for instance in navigation or in "locking on" to pre-defined targets – does not of itself render that system an AWS under the definition used in this paper, and the analysis presented here does not extend to weapon systems that rely on autonomy solely in such other functions. Computerized tools that are designed to aid human decision-making, sometimes called "decision support systems" (DSS) and some of which rely on AI, are also outside the scope of this paper.[13]

III. EXISTING MILITARY PRACTICES IN THE USE OF AUTONOMOUS WEAPON SYSTEMS

In present-day military practice, AWS are typically used only to attack military objectives such as enemy missiles or other projectiles, military radars or warships – namely, objects that under IHL are considered military objectives *by nature* (see below Part VI(A)(2)). The majority of AWS are used only in situations where civilians and civilian objects are not present, or measures are taken (e.g. barriers) to exclude civilians from the areas in which the AWS is active. Many AWS are fixed in place (not mobile), and almost all AWS are supervised in real time by a human operator who can intervene. For several years, militaries have employed defensive systems such as human-supervised CIWS and APS, which allow the human user to intervene and cancel or suspend an attack when necessary.[14] These regulations and conditions of use are not uniformly effective to address all humanitarian, ethical and legal concerns, but they demonstrate the ways in which militaries are already imposing parameters on AWS use to facilitate compliance with IHL obligations. Mines, however, which as explained above are 'unsophisticated' autonomous weapons,

10 Implementation Support Unit of the Convention on Certain Conventional Weapons (CCW), *Non-exhaustive compilation of definitions and characterizations*, March 2023, available at: https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-Group_of_Governmental_Experts_on_Lethal_Autonomous_Weapons_Systems_(2023)/CCW_GGE1_2023_CRP.1.pdf.

11 ICRC, *Submission on AWS to the UNSG*, note 4 above.

12 ICRC, *Autonomous Weapon Systems: Implications of Increasing Autonomy in the Critical Functions of Weapons, Expert Meeting*, 2016, pp. 10-14, available at: https://shop.icrc.org/autonomous-weapon-systems-implications-of-increasing-autonomy-in-the-critical-functions-of-weapons-pdf-en.html; Geneva Academy, *Autonomous Weapon Systems under International Law*, Academic Briefing No. 8, P.3-4, available at: https://www.geneva-academy.ch/joomlatools-files/docman-files/Publications/Academy%20Briefings/Autonomous%20Weapon%20Systems%20under%20International%20Law_Academy%20Briefing%20No%208.pdf; Autonomous Weapons Watch, *Weapons*, available at: https://autonomousweaponswatch.org/weapons.

13 For this, see ICRC, *Submission to the United Nations Secretary General on Artificial Intelligence in the Military Domain*, 2025, available at: https://www.icrc.org/sites/default/files/2025-04/ICRC_Report_Submission_to_UNSG_on_AI_in_military_domain.pdf.

14 Paul Scharre and Michael Horowitz, *An introduction to autonomy in weapons systems: Working Paper*, Center for a New American Security, February 2015, p. 12, available at: https://s3.us-east-1.amazonaws.com/files.cnas.org/documents/Ethical-Autonomy-Working-Paper_021015_v02.pdf?mtime=20160906082257&focal=none.

have been used in a much less restrictive manner, with unacceptable humanitarian consequences. This led to the development of specific instruments to prohibit or restrict their use (see below Part VII(B)(1)).

Even within these limitations, the possibility of an AWS strike being triggered by a person or object that is not the intended target of attack endangers protected persons and objects. This risk depends on the reliability of the AWS in the specific circumstances of use and the predictability of its effects (see below Part V(B) for a more detailed discussion) and may increase with the complexity of the operational environment in which the AWS is used. Furthermore, the use of AWS in armed conflicts poses dangers to civilians and civilian objects due to hardware and software vulnerabilities, which may be exploited by hacking, spoofing, unforeseen interactions with other weapon systems (including those of the enemy) and the operational environment, as well as deceptive or other countermeasures designed to impede AWS operation. These factors are not present to the same degree in peacetime contexts, making the operation of AWS in armed conflict inherently less reliable.

Finally, current trends in military interest and investment include:

- the rapid integration of AI into targeting processes
- the development of swarming capabilities
- the increased autonomy in weapon systems
- the broadening of the types of targets and the situations of use, including the development of "offensive" AWS such as loitering munitions.[15]

These trends risk eroding existing limitations on the use of AWS and exacerbating related humanitarian, legal and ethical concerns, thus underscoring the urgency of reaching international agreement on new legally binding rules on AWS, especially on the conditions and limits necessary for AWS users to comply with IHL.

IV. IHL REQUIREMENT FOR CONTEXT-SPECIFIC HUMAN ASSESSMENTS AND THE CHALLENGES RAISED BY AUTONOMOUS WEAPON SYSTEMS

Parties to an armed conflict must respect and ensure respect for IHL in all circumstances.[16] This includes the obligation to determine whether their weapons, means and methods of warfare are capable of being used in accordance with IHL, and to ensure that they are in fact so used. Because States and non-state armed groups are abstract entities, the application of IHL relies on decisions and actions taken by human beings. Some IHL rules refer expressly to the requirement for human assessment and decision-making. For instance, some precautionary obligations are addressed specifically to "those who plan or decide upon an attack".[17]

In light of that, and despite the growing development of AWS and associated sensor, software and robotics technologies, it is worth emphasizing that IHL obligations regarding the conduct of hostilities must be fulfilled by human commanders and combatants. These humans must determine the lawfulness of the attacks that they

15 European Parliament Research Service, *Defence and Artificial Intelligence*, April 2025, available at: https://www.europarl.europa.eu/RegData/etudes/BRIE/2025/769580/EPRS_BRI(2025)769580_EN.pdf; Kurtis H. Simpson *et al.*, *Militarizing AI: How to Catch the Digital Dragon?*, Centre for International Governance Innovation, February 2025, available at: https://www.cigionline.org/articles/militarizing-ai-how-to-catch-the-digital-dragon/; Automated Research, *Increasing autonomy in weapons systems: 10 examples that can inform thinking*, December 2021, available at: Increasing-Autonomy-in-Weapons-Systems-Single-page-view.pdf; Ingvild Bode, *Loitering Munition and Unpredictability: Autonomy in Weapon Systems and Challenges to Human Control*, AutoNorms, June 2023, available at: https://www.autonorms.eu/loitering-munitions-and-unpredictability-autonomy-in-weapon-systems-and-challenges-to-human-control/; United Nations Security Council, *Final report of the Panel of Experts on Libya established pursuant to Security Council resolution 1973 (2011)*, 8 March 2021, para. 63, available at: https://docs.un.org/en/S/2021/229.

16 Geneva Conventions of 1949 (I-IV), common Article 1; ICRC, Customary IHL Study, Rule 139.

17 Additional Protocol I (hereafter AP I), Art. 57(2).

plan, decide upon or execute, and they remain accountable for these determinations.[18] While some proponents of AWS describe the systems as making a "decision", the decision to launch the weapon, and to carry out the attack, is always made by a human. As such, although certain technical tasks can be carried out by machine processes, determining the lawfulness of an attack – including for example whether an object is a military objective – and assessing whether those processes are sufficient for the attack to comply with IHL require context-specific human assessment and judgement, for which humans remain accountable. The need for such context-specific human judgement is particularly evident when determining the status of persons and objects, which can change over time and depending on the circumstances, as discussed below (Part VI(A)).

Because of the way AWS function, however, their users do not choose, and likely will not know, which specific person or object will trigger a strike, precisely when and where the strike will occur, or who and what will be affected, and how, by the strike. This creates challenges to IHL compliance that do not exist, or do not exist to the same degree, with the use of weapons or weapon systems in which the user chooses to apply force to a specific target at a specific time and location.

These challenges make it all the more important to ensure that context-specific human assessment and judgement are retained when AWS are used. This has repeatedly been emphasized by the GGE, which has affirmed that "[h]uman judgement is essential in order to ensure that the potential use of [AWS] is in compliance with international law, and in particular IHL".[19] The IHL requirement of "context-specific judgement by humans" has emerged as a common feature in national commentaries on the GGE's Guiding Principles,[20] in statements by States at multilateral forums (UNGA and GGE sessions),[21] as well as in the Group of 70 Joint Statement at the UNGA, where they "emphasise[d] the necessity for human beings to exert appropriate control, judgment and involvement in relation to the use of weapon systems in order to ensure [...] compliance with" IHL.[22]

18 ICRC, *Commentary on the 'Guiding Principles' of the CCW GGE on 'Lethal Autonomous Weapons Systems'*, July 2020, p. 3, available at: https://documents.unoda.org/wp-content/uploads/2020/07/20200716-ICRC.pdf; ICRC, 2019 Challenges Report, p. 30, available at: https://shop.icrc.org/international-humanitarian-law-and-the-challenges-of-contemporary-armed-conflicts-recommitting-to-protection-in-armed-conflict-on-the-70th-anniversary-of-the-geneva-conventions-pdf-en.html; ICRC, Customary IHL Study, Rule 12.

19 UN, Group of Government Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems, *Report of the 2019 Session*, CCW/GGE.1/2019/3, September 2019, para 17(e).

20 UN, Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems, *Commonalities in National Commentaries on Guiding Principles*, CCW/GGE.1/2020/WP.1, October 2020, para 11; *Switzerland, Switzerland's food for thought as requested by the Chair of the GGE*, June 2021, available at: https://documents.unoda.org/wp-content/uploads/2021/06/Switzerland.pdf.

21 Reaching Critical Will, *AWS Diplomacy Report*, Vol. 2 No. 2, May 2025, P. 9-10, 16, available at: https://reachingcriticalwill.org/images/documents/Disarmament-fora/aws/consultations/2025/reports/AWSR2.2.pdf.

22 Group of 70 Joint Statement.

V. APPLYING IHL RULES PROHIBITING AND RESTRICTING WEAPONS

AWS can be considered as constituting "weapons" for the purposes of IHL[23] and are therefore subject to IHL rules and principles governing the choice of weapons, as well as to the prohibitions and restrictions applicable to certain types of weapons.

A. PROHIBITIONS OF AND RESTRICTIONS ON SPECIFIC CATEGORIES OF WEAPONS

As with any other weapon, AWS are subject to all the relevant specific prohibitions and restrictions found in treaty and customary law. For instance, to the extent that AWS would constitute biological or chemical weapons, they are prohibited under the Biological and Chemical Weapons Conventions and corresponding customary rules.[24] AWS of a nature to cause superfluous injury or unnecessary suffering are likewise prohibited,[25] as is the use of AWS specifically designed to cause permanent blindness to unenhanced vision.[26] Moreover, an AWS "primarily designed to set fire to objects or cause burn injury to persons through the action of flame, heat, or combination thereof, produced by a chemical reaction of a substance delivered on the target" is considered an incendiary weapon and regulated as such under the CCW Protocol III on Incendiary Weapons and corresponding customary rules.[27] Lastly, the trade in AWS is subject to the Arms Trade Treaty and other applicable international instruments.

B. PROHIBITION OF WEAPONS THAT ARE BY NATURE INDISCRIMINATE

Customary IHL prohibits weapons that are indiscriminate by nature, that is, weapons that cannot be directed at a specific military objective or the effects of which cannot be limited as required by IHL.[28] As opposed to the prohibition on indiscriminate attacks,[29] which requires compliance with the rule in the specific circumstances of the attack (see below Part VI(A)(3)), the prohibition on indiscriminate weapons requires the AWS to be capable of being used in compliance with IHL in all of the normal or expected circumstances of their use.

Reliability, accuracy and predictability of the AWS, in particular system reliability in the critical function of selecting a target, and its accuracy in striking the target it identified, are critical to ensuring compliance with the prohibition of weapons that are by nature indiscriminate. The system's performance will depend on the system's design in terms of ensuring the target profile fulfils the definition of military objective; the accuracy of its sensor in detecting the information required to select a target that matches the target profile; the performance and reliability of the software, robotic, fuse or other technology used to trigger the strike based on the information received; and the accuracy of the strike itself. An AWS that precludes a human user from being able to understand, predict or explain its output cannot be used in compliance with IHL because its design, performance or operating features would render its use tantamount to blind firing.

This prohibition is particularly relevant for an AWS that functions in opaque ways (the "black-box" challenge),[30] which may prevent the human – at the time of the weapon's legal review, at the time of its use, or both – from being able to understand, predict or explain the system's output. This impossibility prevents the user from being able to direct it against a specific military objective and effectively results in a lack of control over the weapon's effects, making it indiscriminate by nature. This concern is heightened in AWS incorporating machine

23 Additional Protocol I, Art. 36.
24 Convention on the Prohibition of the Development, Production and Stockpiling of Bacteriological (Biological) and Toxin Weapons and on Their Destruction, 1972; Convention on the Prohibition of the Development, Production, Stockpiling and Use of Chemical Weapons and on Their Destruction, 1993; ICRC, Customary IHL Study, Rules 73 and 74.
25 AP I, Art. 35; ICRC, Customary IHL Study, Rule 70.
26 ICRC, Customary IHL Study, Rule 86; Protocol (IV) to the 1980 CCW, on Blinding Laser Weapons, 1995, Art. 1.
27 ICRC, Customary IHL Study, Rules 84 and 85; Protocol (III) to the 1980 CCW, on Prohibitions or Restrictions on the Use of Incendiary Weapons, 1980, Art. 1.
28 ICRC, Customary IHL Study, Rule 71.
29 AP I, Art. 54(1).
30 ICRC, Autonomy, *Artificial Intelligence and Robotics: Technical Aspects of Human Control*, August 2019, note 6 above, p. 27; ICRC, *Artificial Intelligence and Machine Learning in Armed Conflict: A Human-Centred Approach*, June 2019, note 6 above, p. 10.

learning, where its functioning can change after the commencement of an attack such that force will be applied in circumstances and in a manner unforeseen to the human user.[31]

Complex swarm technologies may also exhibit emergent behaviours. The predictability of the performance of such systems cannot be absolutely known and instead will need to be based upon probability distributions tied to anticipated environmental inputs. In such cases, it is difficult to accept that a commander could reasonably anticipate the effects that the AWS would have, in order to direct the AWS or limit its effects as required by IHL. This leads to the conclusion that such machine learning–based AWS would likely be indiscriminate by nature.

Even where an AWS is not indiscriminate by nature, the specific circumstances and manner of use may make its employment fall under the prohibition of indiscriminate attacks (see Part VI(A)(3)).

Based notably on these considerations, the ICRC recommends enshrining in an international treaty a prohibition of AWS that are designed in such a manner that their effects cannot be sufficiently understood, predicted and explained.

31 ICRC, *Autonomy, Artificial Intelligence and Robotics: Technical Aspects of Human Control*, August 2019, note 6 above, p. 16.

VI. APPLYING IHL RULES ON THE CONDUCT OF HOSTILITIES

This part will now examine how key IHL rules on the conduct of hostilities – in particular the principles of distinction, proportionality, and precautions in attack – govern the use of AWS that are not prohibited in and of themselves, and the specific legal and practical challenges that arise.

Many of the IHL rules on the conduct of hostilities are formulated as regulating "attacks". This is the case, in particular, for the prohibition of attacks against civilians and civilian objects, the prohibitions of indiscriminate and disproportionate attacks, and many of the precautionary obligations.[32] Article 49(1) AP I defines "attacks" as "acts of violence against the adversary, whether in offence or in defence". Strikes by AWS are "acts of violence", and therefore every one of them must comply with all the IHL rules governing the conduct of attacks. However, the way AWS function raises specific questions with regard to the notion of attack. For example, for an AWS capable of multiple strikes, does each strike constitute an attack or is the entire operation of the AWS after its activation a single attack? Does an "attack" using an AWS begin at the time of activation or only at the time when a specific person or object triggers the AWS to select it for a strike and engage it?

Divergent views exist – both in general and specifically with regard to AWS – on the constitutive elements of an "attack" under IHL and its temporal limitations. It is submitted that each instance where an AWS selects and engages a specific person or object – whether with one or multiple strikes at once – is an attack that commences from the moment of activation by a human, i.e. the point at which the system no longer requires human intervention to autonomously select and engage targets. This is consistent with the requirement that IHL assessments must be made by humans, and such assessments can only be made before the activation of the AWS since, by definition, AWS select and engage targets without human intervention after activation. This is also consistent with Amended Protocol II to the CCW, which identifies the emplacement of a mine as the trigger for the application of several rules usually applied to attacks.[33]

Commanders and other users of AWS need to anticipate and assess in advance (i.e. at the time of activation) the lawfulness of all possible strikes by the AWS, and ensure that any and all possible strikes using an AWS will comply with IHL, whoever or whatever may trigger such strikes, throughout the entire area and during the whole period in which the AWS is active. When doing so, they need to account for all reasonably foreseeable and relevant changes in circumstances between the AWS activation and the strike.[34] Such an exercise is possible only if strict constraints are imposed on the variables applying to the AWS and its operating environment in order to limit the number of potential outcomes, including constraints on the type of targets as will be discussed below.

A. THE PRINCIPLE OF DISTINCTION

The principle of distinction provides that parties to armed conflicts must at all times distinguish between civilians and combatants and between civilian objects and military objectives.[35] Attacks may only be directed against combatants or military objectives. The characterization of an object as a "military objective" or of a person as a "combatant" is based on IHL definitions which call for context-specific and values-based assessments, and which resist being entirely reduced to a quantitative formula.

Applying the principle's prohibitions on directing attacks against civilians, combatants *hors de combat* and civilian objects, as well as on employing a method or means of combat that cannot be directed at a specific military objective, requires limiting the types of targets against which AWS can be used to only those whose legal classification will likely remain stable between the point of the AWS activation by a human and the eventual strike(s). Given that AWS rely on pre-programmed, generalized target profiles to trigger a strike, their use against humans is difficult if not impossible to reconcile with this requirement because of the infinite number of constellations of human movement and behaviour in armed conflict situations. Similarly, it will be difficult, if not impossible, to ensure that targets are military objectives unless they are such by nature, since all others are civilian objects that become military objectives only in specific circumstances and under specific conditions – ones that can vary infinitely.

32 AP I, Arts 51, 52 and 57; ICRC, Customary IHL Study, Rules 1, 7, and 11 to 21.

33 See, in particular, CCW Amended Protocol II, Arts 3(8) and 3(11); Vincent Boulanin, *et al.*, *Autonomous Weapon Systems and International Humanitarian Law: Identifying Limits and the Required Type and Degree of Human-Machine Interaction*, SIPRI, June 2021, pp. 23-24, available at: https://www.sipri.org/sites/default/files/2021-06/2106_aws_and_ihl_0.pdf.

34 ICRC, *2024 Challenges Report*, note 5 above, p. 61.

35 ICRC, Customary IHL Study, Rules 1 and 7; AP I, Art. 48.

1. Concerns raised by using autonomous weapon systems against humans

Under the principle of distinction, attacks must not be directed against civilians, unless and for such time as they directly participate in hostilities.[36] Attacking persons – including combatants or enemy fighters – who are *hors de combat*, which includes those wounded, sick or surrendering, is equally prohibited.[37] These protections reflect fundamental humanitarian considerations and legal obligations, and the use of AWS raises particularly acute challenges for their application when attacks are directed against persons.

Determining whether a person in an armed conflict is protected under one or more of these categories calls for highly contextual legal assessments relating to status, conduct, intent and causality. For instance, a civilian stealing weapons from a military base in an armed conflict may do so for personal gain or to aid the adverse belligerent – the former would not constitute direct participation in hostilities, and the civilian would remain protected against attack, but the latter would constitute direct participation in hostilities, and for the duration of such participation the civilian would be targetable under IHL. In situations of doubt as to whether a person is a civilian, an assessment must be made of the conditions and restraints governing a particular situation, as to whether there are sufficient indicators to warrant an attack, and Additional Protocol I (AP I) establishes a presumption of civilian status.[38] Such determinations are already complex for human decision makers. They are contextual, qualitative assessments, as required by IHL, which cannot be reduced to mathematical formulae or numerical values that could be coded into the AWS target profile.

With regard to combatants, there are a number of ways in which a soldier may express an intention to surrender, and the meaning of such gestures may only be comprehensible in their specific context. There is no single, legally agreed method of expressing an intent to surrender; for example, while some interpret the waving of a white flag as a clear expression of surrender, others regard it merely as an intent to negotiate or communicate.[39] The context in which an act takes place is therefore critical and can only be reliably assessed by humans. Indeed, an AWS, operating on generalized target profiles and sensor inputs, is unable to reliably interpret context or resolve ambiguous gestures, increasing the risk of misidentifying a sign of surrender.[40] For example, an AWS might be unable to distinguish between a soldier laying down their weapon and raising their hands to surrender, and one preparing to shoulder a weapon and launch an attack – movements that, when reduced to sensor inputs, could appear similar to the system despite their very different legal implications.

Similarly, there are many ways in which a soldier may react after having been wounded in battle or be otherwise defenceless, including lying down unconsciously, gesturing incoherently because of unbearable pain, or moving or otherwise gesturing towards the enemy line to surrender and receive medical care, among others. The status of a person may change within seconds from being a lawful target to being protected,[41] and in all these situations, the combatant must not be attacked.

Since AWS users do not select a specific person to attack at a precise time and place, there is a challenge to the user's ability to assess, at the time of activating the AWS, the legal status of the as-yet undetermined person(s) who will trigger an AWS strike at the as-yet undetermined time and place of the strike(s). In the ICRC's view, it is therefore difficult to envisage realistic combat situations where using AWS against persons would not pose a significant risk of IHL violations, since target selection is performed by the system, based on a generalized target profile that is unlikely to be able to account for the non-exhaustive range of contextual signals that a person is protected from attack.

---

36 ICRC, Customary IHL Study, Rule 6; AP I, Art. 51(3); Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of Non-International Armed Conflicts (Protocol II), 8 June 1977 (AP II), Art. 13(3).

37 ICRC, Customary IHL Study, Rule 47; AP I, Arts 41(1) and 85(3)(e); Common Article 3 to the Geneva Conventions.

38 ICRC, Customary IHL Study, Rule 6; AP I, Art. 50(1).

39 States of the opinion that a white flag is an intent to negotiate include Canada, Côte d'Ivoire, and the US, among others, while states of the opinion that a white flag is an intent to surrender include Belgium, Cameroon, France, and the Dominican Republic, see: Russell Buchan, *The rule of surrender in International Humanitarian Law*, Israel Law Review, Vol. 51, No. 1, 2018, p. 21, available at: https://www.cambridge.org/core/services/aop-cambridge-core/content/view/714B1EAB954811EB2907A046EA069504/S0021223717000279a.pdf/the-rule-of-surrender-in-international-humanitarian-law.pdf; Laura Brunn, *et al.*, *Compliance with international humanitarian law in the development and use of autonomous weapons systems: what does IHL permit, prohibit and require?*, SIPRI, March 2023, available at: https://www.sipri.org/sites/default/files/2023-03/ihl_and_aws.pdf.

40 Robert Sparrow, *Twenty Seconds to Comply: Autonomous Weapon Systems and the Recognition of Surrender*, International Law Studies, U.S. Naval War College, Vol. 91, 2015, p. 727, available at: https://digital-commons.usnwc.edu/cgi/viewcontent.cgi?article=1413&context=ils.

41 ICRC, 2016 Commentary on GC I, commentary on common Article 3, para. 746.

Furthermore, employing anti-personnel AWS that, at the moment of selecting a human as a target, are not capable of recognizing an enemy's clear expression of intent to surrender or that an enemy is defenceless because of unconsciousness, wounds or sickness, and refrain from engaging them, may – depending on the circumstances and manner of use – violate the prohibition against conducting hostilities on the basis that there shall be no survivors (denial of quarter).[42] This would in particular be the case if such anti-personnel AWS were used in swarms or otherwise at scale, because it could be expected to lead to multiple strikes on combatants already *hors de combat* at the time the AWS selects them, leaving no survivors. Using AWS to target persons directly also raises serious ethical concerns, although these are beyond the scope of this paper.[43]

As discussed below (Part VII(B)(2)), there are precedents within IHL for distinguishing between weapons designed to target humans directly and those designed for anti-materiel use, with regulations on the anti-personnel ones, including banning anti-personnel mines whose explosion is triggered by the presence, proximity or contact of a person, based on experience demonstrating their indiscriminate effects.[44]

Drawing on this experience, and based notably on these considerations, the ICRC recommends enshrining in an international treaty a prohibition on AWS designed or used to attack humans directly (anti-personnel AWS). This prohibition is grounded in present practice, in which AWS are used against objects rather than persons, and it would be most effective if agreed upon before anti-personnel AWS are developed.

2. Challenges raised by attacking objects with autonomous weapon systems

Insofar as objects are concerned, military objectives are limited to those objects which, by their nature, location, purpose or use, make an effective contribution to military action and whose total or partial destruction, capture or neutralization, in the circumstances ruling at the time, offer a definite military advantage.[45] Civilian objects are all objects that are not military objectives.[46] Whether an object qualifies as a military objective is context-dependent and time-bound: if the destruction, capture or neutralization of an object does not yet offer or no longer offers a definite military advantage, it does not constitute a military objective and attacking it is prohibited.

To comply with the obligation to direct attacks only against military objectives, when activating the system, users of the AWS must ensure that whatever will trigger an AWS strike throughout the area and duration of AWS operation will satisfy the two-pronged criteria of the definition of a military objective at the time of the strike, while accounting for all reasonably foreseeable and relevant changes in circumstances.

However, since AWS users do not select a specific object to attack at a precise time and place, there is a challenge to the user's ability to assess, at the time of activating the AWS, the contribution that the as-yet undetermined object that will trigger an AWS strike makes to the enemy's military action, and the military advantage to be gained from the destruction of this as-yet undetermined object at the as-yet undetermined time and place of the strike.

In light of this challenge, the risk of an unlawful attack on a civilian object is highest when an AWS is used against objects that may qualify as military objectives by location, purpose or use,[47] as opposed to military objectives by nature. This is because the effectiveness of the contribution that the former types of objects make to the adversary's

---

42 ICRC, Customary IHL Study, Rule 45; AP I, Art. 40; Hitoshi Nasu, *LAWS Debate at the United Nations: Moving Beyond Deadlock*, Articles of War, Sept 21, ('Deploying lethal autonomous weapons systems on the battlefield without appropriate capabilities to identify and spare victims of war would run contrary to the prohibition of no quarter orders—showing no mercy or clemency to spare life in return for surrender'), available at: https://lieber.westpoint.edu/kargu-2-autonomous-attack-drone-legal-ethical/ For an ethical perspective on the issue, see e.g. Robert Sparrow, *Twenty Seconds to Comply: Autonomous Weapon Systems and the Recognition of Surrender*, note 40 above.

43 ICRC, *Position Paper on AWS*, note 5 above, pp. 8-9; ICRC, *Ethics and Autonomous Weapon Systems: An Ethical Basis for Human Control?*, note 5 above. As to the significance of ethical concerns in the regulation of AWS, see the working papers, e.g., Argentina *et al.*, *Roadmap*, 2022, Section II, para. 4; Australia *et al.*, *Principles and Good Practices*, 2022, paras 20(e) and 28; Non-Aligned Movement Working Paper, paras 7, 12, 18(d); China Working Paper, 2022, Section II. Working papers of the 2022 session are available here: https://meetings.unoda.org/ccw/convention-certain-conventional-weapons-group-governmental-experts-2022.

44 Convention on the Prohibition of the Use, Stockpiling, Production and Transfer of Anti-Personnel Mines and on their Destruction, 1997 (Anti-Personnel Mine Ban Convention), Art. 1.

45 ICRC, Customary IHL Study, Rule 8; AP I, Art. 52(2).

46 ICRC, Customary IHL Study, Rule 9; AP I, Art. 52(1).

47 E.g. a hill, a hotel temporarily used to accommodate troops, or a bridge about to be crossed by enemy forces.

military action, and the definite military advantage offered by their destruction, capture or neutralization, may vary and change significantly and rapidly. For example, a civilian vehicle that is temporarily requisitioned to transport soldiers to the front line only makes an effective contribution to the enemy's military action – and thus could only be characterized as a military objective by use – for the duration of this use. Similarly, its destruction at any other time is unlikely to offer any military advantage. Further, characterizing an object as a military objective by purpose requires ascertaining enemy intentions, the cues for which are nuanced, context-dependent and non-exhaustive, making them ill-suited to standardization in the kind of generalized target profile used by AWS. Consequently, and except to some extent for objects that make an effective contribution to military action by nature, sweeping or anticipatory classification of objects as military objectives is not permissible, since that "would negate the obligation to continually validate the nature of a proposed target".[48]

In case of doubt as to whether an object is protected, an assessment must be made of the conditions and restraints governing a particular situation as to whether there are sufficient indicators to warrant an attack. This is a particularly demanding requirement for objects normally dedicated to civilian purposes – which may be military objectives by use, purpose or location, but aren't military objectives by nature – and for which AP I establishes a presumption of civilian status.[49] In contrast, the effective contribution made by objects that qualify as military objectives by nature is less variable and less context-dependent. These objects generally include enemy weapons and weapon systems, military radars and typical military vehicles such as tanks, armoured personnel carriers, fighter jets and warships.

Furthermore, military objectives by nature are typically far more easily distinguishable from civilian objects than other types of military objectives. For instance, an AWS designed to be triggered by the signature (sonar, shape, direction of travel, etc.) of a type of enemy military submarine is less likely to be triggered by a merchant ship. On the other hand, an AWS intended to strike repurposed utility trucks used by enemy soldiers carries a high risk of being triggered by similarly shaped trucks being used by fleeing civilians.

Using AWS against objects that are military objectives by nature thus presents a lower risk than with other types of objects that the object eventually struck by the AWS will in fact be protected as a civilian object at the time and place of the strike.

Based notably on these considerations, the ICRC recommends enshrining in an international treaty a restriction on the design and use of AWS to only target objects that are military objectives by nature. This restriction is grounded in present practice, in which AWS are used against typically military objectives, and it would be most effective if agreed upon before AWS are developed to be used against other types of targets.

3. Prohibition of indiscriminate attacks

The prohibition against indiscriminate attacks, which stems from the principle of distinction, defines three types of attacks, each of which is of a nature to strike military objectives and civilians or civilian objects without distinction because:[50]

a) they are not directed at a specific military objective;
b) they employ a method or means of combat which cannot be directed at a specific military objective; or
c) they employ a method or means of combat the effects of which cannot be limited as required by IHL.

The first two types of indiscriminate attacks are an application of the obligation to direct attacks only against military objectives. The third one is based on the logical argument that employing methods or means of combat whose effects cannot be limited as required by IHL is prohibited, and includes the use of weapons, means and methods of warfare whose effects cannot be controlled in time or space.[51]

48 International Law Association (ILA) Study Group on the Conduct of Hostilities in the 21st Century, *The Conduct of Hostilities and International Humanitarian Law: Challenges of 21st Century Warfare*, International Law Studies, Vol. 93, 2017, p. 328 (hereafter ILA, *2017 Conduct of Hostilities Challenges*), available at: https://digital-commons.usnwc.edu/cgi/viewcontent.cgi?article=1709&context=ils.

49 ICRC, Customary IHL Study, Rule 10; AP I Art. 52(3).

50 ICRC, Customary IHL Study, Rules 11-13; AP I, Art. 51(4).

51 Sandoz *et al.* (eds), Commentary on the Additional Protocols, 1987, p. 623, para. 1963; Henckaerts and Doswald-Beck (eds), ICRC Study on Customary International Humanitarian Law, Vol. I, commentary on Rule 12, p. 43, available at: https://library.icrc.org/library/docs/DOC/icrc-1105-002-01.pdf#page=104; Michael Schmitt, "War, technology and the law of armed conflict", in A.M. Helm (ed.), *The Law of War in the 21st Century: Weaponry and the Use of Force*, International Law Studies, Vol. 82, 2006, p. 140.

As already noted above, since the specific person or object that will trigger a particular strike, and the precise location, time and number of strikes may be undetermined at the time of activating the AWS, the user will not be able to assess the effects of a particular strike in concrete terms. Consequently, the AWS user must ensure that its use would not amount to an indiscriminate attack. This requires considering the effects of any and all strikes that may occur after the activation of the AWS – whoever or whatever might trigger a strike – across the entire area and duration of AWS operation, while accounting for all reasonably foreseeable and relevant changes in circumstances.

The assessment above concerning the application of the prohibition of weapons that are by nature indiscriminate to AWS (see Part V(B)) applies *mutatis mutandis* here. The difference, however, is that the prohibition of indiscriminate attacks focuses on whether an AWS, which is not indiscriminate by nature, can be used in a discriminate manner in the particular circumstances of the attack. For instance, activating an AWS to engage enemy military vehicles of a certain type, such as a troop carrier, when civilian vehicles within its area of operation could reasonably be expected to trigger AWS strikes, would amount to an indiscriminate attack. Similarly, activating an AWS without knowing whether objects other than enemy military vehicles, and if so what objects, could trigger strikes in the circumstances of use would constitute an indiscriminate attack.

More generally, even where an AWS is not indiscriminate by nature, commanders or other users who employ an AWS whose effects they cannot sufficiently understand, predict and explain *in the specific circumstances of use* would violate the prohibition of indiscriminate attack. This is because, in those circumstances of use, employing the system would be tantamount to blind firing, where the user, unable to anticipate its effects, would be unable to ensure that they are limited as required by IHL.

In this regard, the understanding of the functioning of the AWS must extend to the nature and functioning of its sensors, the definition of its target profile and the potential effects in the specific circumstances of use – including consideration of any risk of error, malfunction or other unintended consequences, whether inherent to the design of the AWS or its operating features, or dependent on the circumstances of use, or still due to tampering (e.g. adversarial attack), and the system's reaction to possible deception (such as "spoofing"), countermeasures or other defensive action by the enemy.

By way of analogy and looking at a system that does not constitute an AWS, understanding the nature of the guidance mechanism of precision-guided munitions (e.g. GPS or laser) is necessary to assess the lawfulness (and the military utility) of using the munition when faced with enemy countermeasures (e.g. GPS spoofing, smoke screens or other laser countermeasures). Similarly, understanding the manner in which physical phenomena are measured and processed by the AWS sensors and algorithms, and whether their performance might be affected by, among others, the time of use (e.g. day or night), the weather or enemy electronic or other countermeasures or deceptions, is necessary to ensure that an attack employing AWS complies with IHL.

Based notably on these considerations, the ICRC recommends enshrining in an international treaty a prohibition on the use of AWS whose effects cannot be sufficiently understood, predicted and explained so that a user can be reasonably satisfied that it will not amount to an indiscriminate attack. Specific restrictions recommended on the use of AWS that are not prohibited[52] *will also support compliance with the prohibition of indiscriminate attacks.*

B. PROHIBITION OF DISPROPORTIONATE ATTACKS

IHL prohibits launching an attack which may be expected to cause incidental loss of civilian life, injury to civilians, damage to civilian objects or a combination thereof, which would be excessive in relation to the concrete and direct military advantage anticipated.[53] The prohibition on disproportionate attacks calls for a prospective (*ex-ante*) assessment in light of the circumstances of a particular attack. Those planning, deciding upon or otherwise launching an attack must determine the incidental harm to civilians and civilian objects that is expected, and compare it to the anticipated, concrete and direct military advantage, to determine if the former may be excessive in relation to the latter.

52 See below Part VI *in fine* and Part VII(C).

53 ICRC, Customary IHL Study, Rule 14; AP I, Arts 51(5)(b) and 57(2)(a)(iii).

All direct and indirect harm to civilians and civilian objects that is reasonably foreseeable at the time of planning, deciding upon the attack and activating the AWS must be considered. This includes harm to civilians and civilian objects that may be expected to be in or near a target at the time of any AWS strike, as well as indirect or "reverberating" effects. While there must be a causal link between the attack and the reverberating effects, there are no temporal or geographic requirements for the effects to be considered, provided they are reasonably foreseeable.[54] Civilian harm can vary significantly, depending notably on the presence of civilians or civilian objects in and around the object being struck, which itself depends on its location and the timing of the strike. The fact that one or more of these variables may be undetermined at the time of AWS activation significantly constrains the user's ability to form the assessment required, under the limb of the rule, of the incidental civilian harm that may be expected.

The military advantage is also highly context-dependent and varies over time. The requirement that the military advantage be concrete and direct limits consideration to military advantage that is substantial and relatively close and excludes speculative or hypothetical advantage. The evolution of the wider military context affects the military advantage that may reasonably be anticipated from the destruction of identical objects.[55]

The context-specificity of this assessment is complicated by a long-standing controversy about the frame of reference for assessing proportionality in attack, with certain States emphasizing that the military advantage must be assessed in relation to an attack considered "as a whole and not only from isolated or particular parts" thereof.[56] Even if the military advantage offered by specific attacks must be assessed in light of the attack as a whole when such attacks are carried out in a concerted manner, an attack as a whole must nonetheless be a finite event with defined limits.[57] If the user expects that several strikes will be triggered during the period of activation of the AWS, whether some or all of them can be considered together (possibly together with the use of non-autonomous systems) as an attack as a whole to identify the anticipated concrete and direct military advantage when assessing the proportionality of these attacks depends on the relationship between the strikes – for example whether they are intended to achieve the destruction of a single, composite military objective. For instance, a belligerent that launches loitering AWS against enemy radar systems as part of a complex aerial attack directed against an enemy strategic military objective (e.g. its air-defence system) may, when assessing the proportionality of each of the expected AWS strikes, take into account the anticipated military advantage that each strike offers, assessed in light of the attack as a whole.[58] In the absence of such a relationship, the military advantage anticipated from any strike that may be expected from the AWS must be considered independently and separately for the purposes of applying the prohibition of disproportionate attacks.

Since, as already noted, AWS users do not select the particular person or object that could trigger an AWS strike or the time and location of the strike, they cannot assess in specific terms the expected incidental civilian harm and anticipated concrete and direct military advantage of the attack which they are responsible for launching. They must therefore ensure that any and all possible strikes will comply with the prohibition of disproportionate attacks, whoever or whatever may trigger such strikes throughout the entire area and during the whole period in which the AWS is active.

This is particularly challenging in urban warfare, because military and civilian people and objects are often intermingled in cities, and because of the wide array of indirect or reverberating effects that may be caused when damaging critical civilian infrastructure.[59] To reduce the risk of civilians and civilian objects being present in the AWS strikes' area effects, as well as the risk that the proportionality assessment is invalidated by changes in circumstances between AWS activation and the strike, it is essential to impose strict limits on their use – including

54 ICRC, *International Expert Meeting Report: The principle of proportionality*, 2018 (hereafter ICRC, *Principle of Proportionality Expert Meeting Report*), available at: https://www.icrc.org/sites/default/files/document/file_list/4358_002_expert_meeting_report_web_1_0.pdf; see also, ICRC, *2024 Challenges Report*, note 5 above, p. 41.

55 ICRC, *Principle of Proportionality Expert Meeting Report*, note 54 above, p. 12; ICRC, Commentary on the Additional Protocols, 1987, paras 2209 on AP I, Art. 57.

56 Australia, Belgium, Canada, France, Germany, Italy, the Netherlands, New Zealand, Spain and the United Kingdom, available at: https://www.icrc.org/applic/ihl/ihl.nsf/States.xsp?xp_viewStates=XPages_NORMStatesParties&xp_treatySelected=470; US, DoD Law of War Manual, 2023, § 5.12.2.1.

57 ILA, *2017 Conduct of Hostilities Challenges*, note 48 above, p. 343.

58 ICRC, *Principle of Proportionality Expert Meeting Report*, note 54 above, p. 17.

59 ICRC, *2019 Challenges Report*, note 18 above, Chapter 2.

restricting them to situations where civilians and civilian objects are absent, limiting the number of strikes in a single operation and constraining the duration and geographic scope of their deployment.

C. PRECAUTIONS IN ATTACKS

IHL requires parties to a conflict to take constant care in the conduct of their military operations to spare the civilian population, civilians and civilian objects. They must take all feasible precautions to avoid, and in any event to minimize, incidental loss of civilian life, injury to civilians and damage to civilian objects.[60] This entails doing everything feasible to verify that targets are military objectives;[61] taking all feasible precautions in the choice of means and methods of warfare to avoid, and in any event minimize, incidental civilian harm;[62] doing everything feasible to assess whether the attack may be expected to cause incidental civilian harm which would be excessive in relation to the concrete and direct military advantage anticipated;[63] doing everything feasible to cancel or suspend an attack if it becomes apparent that the target is not a military objective or that the attack may be expected to be disproportionate;[64] giving effective advance warning of attacks which may affect the civilian population, unless circumstances do not permit;[65] and, whenever a choice is possible between several military objectives for obtaining a similar military advantage, selecting the military objective the attack on which may be expected to cause the least danger to civilian lives and to civilian objects.[66] Feasibility, in the context of precautions, refers to that which is practicable or practically possible, taking into account all circumstances ruling at the time, including humanitarian and military considerations.[67]

Decisions about precautions must be reached by commanders and others responsible for planning, deciding upon and executing attacks on the basis of information from all sources reasonably available to them at the relevant time.[68] They must do everything feasible to obtain information that will allow for a meaningful assessment of the foreseeable consequences of the attack. The information must be timely and up to date, collected and analysed up to the launch of the attack, and, to the extent feasible, during the conduct of the attack. Instructions issued in advance of an attack will not, on their own, constitute sufficient precautions, as evidenced by the continuing obligation to do everything feasible to cancel or suspend an attack if it becomes apparent that it would violate the principles of distinction or proportionality.[69]

When applying the obligation of precautions to the use of an AWS, a question arises as to the extent to which the user may rely on the technical features of an AWS such as a self-deactivation mechanism, a sensor-based fuse or other technical features that may be developed in the future, in fulfilment of the user's obligation to take all feasible precautions. In the ICRC's view, the existence of such technical features does not in and of itself fulfil the legal requirement to take all feasible precautions in attack. Furthermore, the use of an AWS cannot itself justify a failure to take precautionary measures that would reduce the risk of harm to civilians and that are feasible in the circumstances of a particular attack. This includes measures that would have been feasible had the attacker chosen another means or method of attack. This consideration could arise in the context of an AWS which selects and engages targets at such a speed that it precludes real-time intervention by the human user to prevent civilian harm, or when using an AWS would make it impossible to give effective advance warning of an attack which may affect the civilian population. If it were feasible in the circumstances to use another weapon, including a non-autonomous weapon, or another means or method that would make it possible to minimize civilian harm by allowing human intervention before the strike or by enabling such a warning, the rules on precautions may require the user to choose that other weapon, means or method over an AWS.

60 ICRC, Customary IHL Study, Rule 15; AP I, Art. 57(1).
61 ICRC, Customary IHL Study, Rule 16; AP I, Art. 57(2)(a)(i).
62 ICRC, Customary IHL Study, Rule 17; AP I, Art. 57(2)(a)(ii).
63 ICRC, Customary IHL Study, Rule 18; AP I, Art. 57(2)(a)(iii).
64 ICRC, Customary IHL Study, Rule 19; AP I, Art. 57(2)(b).
65 ICRC, Customary IHL Study, Rule 20; AP I, Art. 57(2)(c).
66 ICRC, Customary IHL Study, Rule 21; AP I, Art. 57(3).
67 Protocol on Prohibitions or Restrictions on the Use of Mines, Booby-Traps and Other Devices (Protocol II), 1980, Art. 3(4); Protocol on Prohibitions or Restrictions on the Use of Incendiary Weapons (Protocol III), 1980, Art. 1(5); Protocol on Prohibitions or Restrictions on the Use of Mines, Booby-Traps and Other Devices as amended on 3 May 1996 (Protocol II to the 1980 CCW Convention), Art. 3(10). For State practice, see the ICRC Customary IHL Study, explanation and practice related to Rule 15, available at: https://ihl-databases.icrc.org/en/customary-ihl/v1/rule15#title-4.
68 ICRC, Customary IHL Study, explanation on Rule 15.
69 ICRC, Customary IHL Study, Rule 19; AP I, Art. 57(2)(b).

The technical features of an AWS system are, however, directly relevant when choosing between various AWS that would be available. For instance, the obligation to take all feasible precautions in the choice of means and methods of warfare to avoid, or at least minimize, incidental civilian harm may in some circumstances require the choice of an AWS that allows for rapid intervention and deactivation over an AWS whose features preclude such intervention.

In this regard, and to ensure that the user can cancel or suspend an attack when required,[70] AWS need to be designed to allow the human user to monitor and assess the circumstances and, where appropriate, intervene and deactivate the operation of an AWS in a timely manner. Such a feature, allowing for human–machine interaction, reduces the risk that the development and use of AWS will diminish existing humanitarian protections provided by IHL. AWS that do not allow the user to do so must be equipped with an effective self-destruction or self-neutralization mechanism and have a backup 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.

As discussed in the preceding parts on distinction and proportionality, users of AWS must take into account all reasonably foreseeable changes in circumstances over the entire area and duration of the AWS operation between activation and any potential strike. But the difficulty for an AWS user to know the circumstances ruling at the time and place where the system's sensors will be triggered to strike, and what will trigger it, makes it difficult to determine which precautions may be required to avoid or minimize incidental civilian harm. There is a significant risk that changes in the operating environment would invalidate legal determinations made and precautions taken at the time of the AWS activation, a risk heightened in dynamic environments where civilians or civilian objects are present, particularly in or near cities and other populated areas. The likelihood and scale of such changes increase with the size of the area, the duration of the operation and the dynamic nature of the situation.

Based notably on the considerations developed in Part VI, the ICRC recommends enshrining in a new treaty a number of restrictions on the use of AWS that are not prohibited. In particular limits on the duration, geographical scope, scale and situation of AWS use, in particular restricting their use to only situations in which civilians and civilian objects are not present, are therefore essential to enable the user to assess, when planning and deciding to activate an AWS, whether any AWS strike will comply with the principles of distinction (including the prohibition of indiscriminate attacks), proportionality and precautions in attack. Further, requirements for human–machine interaction in terms of effective supervision and the ability to intervene and deactivate as appropriate would enable a user's ability to cancel or suspend the attack in case of risk of violation of IHL.

70 *Ibid*.

VII. ADDRESSING THE CHALLENGES RAISED BY AUTONOMOUS WEAPON SYSTEMS TO COMPLIANCE WITH IHL

A. LEGAL REVIEW OF AUTONOMOUS WEAPONS

In the study, development, acquisition or adoption of a new AWS, States party to AP I are under an obligation to determine whether its employment would, in some or all circumstances, be prohibited by IHL or by any other applicable rule of international law.[71] The legal review of new weapons is a critical tool to ensure that a State's armed forces are capable of conducting hostilities in accordance with its international obligations, regardless of whether they are party to AP I, and their importance has also been emphasized by the GGE.[72]

A reviewer will need to be satisfied of the legality of an AWS in its normal or expected circumstances and manner of use. It follows that the AWS must be understandable, predictable and explainable to allow the reviewer to anticipate the operation of the weapon system and assess conformity with the prospective and post-attack requirements of IHL and other international law. This has also been recognized in the GGE, with particular focus on the potential for self-learning "that could introduce a risk of unpredictability", which makes it imperative that weapon reviews be "conducted with a full understanding of the weapons' capabilities and limitations, in light of its normal or expected uses and sufficient confidence about its effects in those circumstances".[73]

Further, States are obliged to conduct a new review of a weapon, means or method of warfare which has been modified in a manner that alters its function or impacts its effects. Thus, a new review of an AWS must be carried out if the system's functioning changes, for example as a result of machine learning software, in a way that affects its selection and/or engagement functions, to ensure the continued lawfulness of AWS employment. The very nature of AWS creates significant practical challenges to carrying out strong and effective legal review.

As underscored throughout this paper, the recommendations the ICRC submitted to States on the development and use of AWS are grounded in existing IHL principles and rules. A rigorous legal review process should therefore already lead to the imposition of many of these limits, preventing the development of unacceptable AWS and identifying restrictions for others, depending on the normal or expected circumstances of use.

However, it is important to emphasize that, while robust legal reviews remain essential, they are not a sufficient response on their own to all the concerns raised by AWS and are not a substitute for States agreeing on international limits on AWS. In this regard, the ICRC considers that while existing IHL imposes stringent constraints as described in this paper, it does not hold all the answers to the humanitarian, legal and ethical questions raised by AWS. Furthermore, States hold different views about what limits and requirements for the design and use of AWS derive from existing rules of IHL. Developing specific prohibitions and restrictions for the development and use of autonomous weapons to address these challenges has the advantage of clarity and legal certainty, as it entails developing and codifying a common understanding among States. The development of humanitarian disarmament treaties over the past decades has demonstrated the effectiveness of such an approach and provides a number of lessons for regulation of AWS, which are explored in the following section.

71 AP I, Art. 36.

72 GGE Guiding Principles, letter (e).

73 UN, Group of Government Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems, Revised Chair's paper, *Draft elements on possible consensus recommendations in relation to the clarification, consideration and development of aspects of the normative and operational framework on emerging technologies in the area of lethal autonomous weapon systems*, September 2021, para 4(4)(g).

B. LESSONS TO BE DRAWN FROM THE EXISTING INTERNATIONAL REGULATION OF WEAPONS

In articulating prohibitions and restrictions on the use of AWS, it is useful to draw upon existing international regulations of analogous or similar weapon systems.

1. Prohibitions and restrictions related to the indiscriminate effects of weapons

It is particularly relevant to look at how States addressed victim- or contact-detonated mines, as well as booby traps. Contrary to command-detonated mines, these weapons detect the presence, proximity or contact of a person or object, which triggers their explosion without human intervention. They can therefore be considered rudimentary autonomous weapons.

As early as 1907, States agreed on several prohibitions regarding the use of automatic submarine contact mines "to ensure, as far as possible, to peaceful navigation the security to which it is entitled, despite the existence of war".[74] Some of the prohibitions relate to the design of the mines, and require that they become harmless in situations where the user loses control over the mines and their effects, while another prohibition pertained to the area and purpose of use.[75] The Convention also requires precautionary measures.[76]

The prohibition and restrictions on landmines developed in an iterative manner over two decades. The CCW High Contracting Parties first adopted Protocol II on Mines, Booby-Traps and Other Devices in 1980. This was quickly deemed insufficient, and they adopted Amended Protocol II in 1996. The restrictions and special precautionary requirements set out in the 1996 Amended Protocol II include fencing, monitoring and other measures to protect civilians against the effects of mines, and measures to limit the duration of their operability.[77] The Protocol also applies the prohibition on area bombardment to the use of these weapons, reflecting the particular protection concerns that arise in locations containing a concentration of civilians or civilian objects.[78] Nevertheless, an overwhelming majority of States deemed that this approach to mitigating the indiscriminate effects of landmines was insufficient for anti-personnel landmines, and opted in 1997 to ban altogether landmines designed to be triggered by a person.[79]

Similarly, "in order to avoid indiscriminate area effects", the Convention on Cluster Munitions (CCM) limits the scope of a cluster munition attack and requires that a cluster munition must "have fewer than ten explosive submunitions", each "designed to detect and engage a single target object" and equipped with electronic self-destruction and self-deactivation features.[80]

Finally, Amended Protocol II lays out specific restrictions on the use of booby traps in populated areas, while the 1980 Protocol (III) to the CCW on Incendiary Weapons imposes a specific prohibition and restrictions on the use of incendiary weapons within a concentration of civilians, because of the specific challenge to comply with IHL when using booby traps and incendiary weapons in such environments.[81]

These existing prohibitions and restrictions can provide indicative baselines for a similar framework for AWS, although any such framework would need to account for the specific ways in which AWS function, as highlighted throughout this paper.

---

74 Convention (VIII) relative to the Laying of Automatic Submarine Contact Mines. The Hague, 18 October 1907, Preamble.
75 *Ibid.*, Arts 1 and 2.
76 *Ibid.*, Art. 3.
77 CCW Amended Protocol II, Arts 3, 6, 9, 12 and Technical Annex.
78 *Ibid.*, Art. 3(9).
79 Anti-Personnel Mine Ban Treaty, Art. 1.
80 Convention on Cluster Munitions, 2008, Art. 2(2)(c).
81 CCW Amended Protocol II, Art. 7(3); Protocol III to the 1980 CCW, on Prohibitions or Restrictions on the Use of Incendiary Weapons (CCW Protocol III), 1980, Art. 2.

2. Prohibition and restrictions to provide enhanced protection for humans

Existing international law also provides many examples of weapon prohibitions and restrictions that afford stronger protections to humans than to objects, due to concern that these weapons will have indiscriminate effects, cause superfluous injury or unnecessary suffering, or both.

Landmines are among those weapons regulated (through prohibitions and restrictions) based on concerns that they would be indiscriminate when used against persons. This is particularly evident with the prohibition of anti-personnel landmines, while such a prohibition does not exist for anti-vehicle or other anti-materiel landmines. Similarly, to avoid being prohibited as cluster munitions, submunitions must meet certain conditions, including being designed to detect and engage a single target "*object*"[82] — which, as legal commentators note, provides the sense that "what is being referred to are anti-vehicle or anti-materiel weapons, rather than anti-personnel weapons".[83]

Other anti-personnel weapons, or their anti-personnel use, have been prohibited in line with the prohibition against causing superfluous injury or unnecessary suffering, including "laser weapons specifically designed [...] to cause permanent blindness",[84] and "the anti-personnel use of bullets which explode within the human body".[85] Further international law examples of the differentiated regulation of weapon technologies based on their effects on humans include "weapons the primary effect of which is to injure by fragments which are not detectable by X-rays in the human body"[86] and "bullets which expand or flatten easily in the human body".[87]

It is worth noting that these prohibitions have not hindered the development of similar technologies for anti-materiel purposes, which remain permissible under IHL. Examples include anti-vehicle mines, anti-materiel lasers and anti-materiel explosive weapons, from heavy explosive weapons to high explosive armour-piercing ammunitions and explosive weapons with delayed fuses.

A prohibition of AWS designed or used to apply force against persons (anti-personnel AWS) aligns with this human-centred approach, addresses specific concerns raised by AWS, advances ethical considerations to safeguard humanity and upholds IHL rules for the protection of civilians and combatants *hors de combat*.

3. Preventing the development of unacceptable weapons

History contains many examples of disarmament treaties that significantly reduced the harm caused by unacceptable weapons, from the success in curbing the use of biological and chemical weapons to the vast decrease in mine victims since the adoption of the Anti-Personnel Mine Ban Treaty.[88] However, anti-personnel mines have not completely disappeared, and their total elimination is hampered by the reluctance of military powers to renounce their use before potential adversaries do. The same phenomenon underlies the lack of progress in nuclear disarmament.

Even if they are less known, the most successful examples of humanitarian disarmament were those that prohibited weapons before their development or use on the battlefield. This was the case with certain exploding bullets, renounced by signatories to the 1868 St. Petersburg Declaration, and with blinding laser weapons, prohibited by Protocol IV of the 1980 Convention on Certain Conventional Weapons. These instruments proved successful by preventing suffering rather than waiting for unacceptable consequences to become apparent before acting. Full compliance by all States and parties to armed conflicts was also possible, because these weapon technologies had not been fielded.

---

82 Convention on Cluster Munitions, 2008, Art. 2(2)(c)(iii).

83 Bonnie Docherty, *et al.*, "Definitions", in Gro Nystuen and Stuart Casey-Maslen (eds), *The Convention on Cluster Munitions: A Commentary*, Oxford Commentaries on International Law, 2010, paras 2.123 and 2.132, available at: https://academic.oup.com/oxford-law-pro/book/57496/chapter/473376109.

84 ICRC, Customary IHL Study, Rule 86; Protocol IV to the 1980 CCW, on Blinding Laser Weapons, Art. 1.

85 ICRC, Customary IHL Study, Rule 78; St Petersburg Declaration Relating to Explosive Projectiles, 1868.

86 ICRC, Customary IHL Study, Rule 79; Protocol I to the 1980 CCW, on Non-Detectable Fragments.

87 ICRC, Customary IHL Study, Rule 77; Hague Declaration (IV, 3) Concerning Expanding Bullets, 1899.

88 See, for example, International Campaign to Ban Landmines, *Landmine Monitor 2024*, p. 41, available at: https://backend.icblcmc.org/assets/reports/Landmine-Monitors/LMM2024/Downloads/Landmine-Monitor-2024-Final-Web.pdf.

Humanitarian disarmament is not only about prohibiting weapons that can never be used in compliance with IHL. On the contrary, even in situations where the use of a weapon would not be absolutely prohibited by the general rules of IHL, States endeavoured, through the development of weapons treaties, to strike a balance between the military utility of that weapon and the various risks it entails from a humanitarian, legal, ethical or other perspective. In this regard, developing new means or methods of warfare that challenge the users' ability to understand, predict and limit their effects and ensure that attacks will comply with IHL, runs counter to the object and purpose of IHL and to the balance of military necessity and humanity considerations that sustains it. Without timely agreement on a legally binding instrument that prohibits unacceptable AWS and regulates all others, emerging patterns in the development and use of AWS may become difficult to reverse, with lasting humanitarian and legal consequences.

C. WAY FORWARD: AN INTERNATIONAL INSTRUMENT PROHIBITING UNACCEPTABLE AUTONOMOUS WEAPON SYSTEMS AND RESTRICTING OTHERS

In light of the serious risks of harm for those affected by armed conflict, the challenges to compliance with IHL detailed in this position paper, and ethical concerns raised by AWS, the ICRC has, since 2021, been calling for new, binding international law rules on the development and use of AWS.[89] These rules should clarify and formalize specific prohibitions and restrictions concerning the development and use of AWS and safeguard the protections afforded by IHL against being undermined by increasing autonomy in weapon systems. Any such limits would be additional and complementary to existing IHL rules, including weapons treaties, and would not displace them. They would clarify, strengthen and build on existing legal protections in order to respond to the specific risks and ethical concerns raised by AWS.

In particular, new rules must:

- prohibit unpredictable AWS, namely those that, due to their design or the circumstances and manner of use, do not allow a human user to understand, predict and explain the system's functioning and effects. Users of AWS must be able, with a reasonable degree of certainty, to 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. Examples of autonomous weapons which are likely to exhibit such unpredictable effects include those which incorporate machine learning, along with certain swarm technologies.
- prohibit AWS designed or used to target humans directly (anti-personnel AWS). This is required because of the significant risk of IHL violations and the unacceptability of anti-personnel autonomous weapons from an ethical perspective.

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 rules on the conduct of hostilities. To reduce the risk of violations, new rules must also strictly constrain the development and use of AWS, including through a combination of:

- restricting targets of the AWS to only those that are military objectives by nature
- limiting the duration and geographic scope of AWS operation
- limiting the scale of use, including the number of engagements that the AWS can undertake
- limiting the situations of use, namely constraining them to situations where civilians or civilian objects are not present
- ensuring, to the maximum extent feasible, the ability for a human user:
  - to effectively supervise, and
  - in a timely manner, to intervene and, where appropriate, deactivate operation of the AWS.

  Where this is not feasible, the AWS must be equipped with an effective self-destruction or self-neutralization mechanism.

89 ICRC, *Position Paper on AWS*, note 5 above.

Against the backdrop of the rapid and expanding development and use of AWS, the establishment of these prohibitions and restrictions on AWS in clear and binding international law is an urgent humanitarian priority. The ICRC president and the United Nations secretary-general called on States to take bold and principled political action to conclude negotiations on such rules by 2026.[90] The ICRC has submitted its views, for consideration by States and the United Nations secretary-generall, as to how these rules could be drafted in a legally binding instrument.[91]

90 ICRC, *Joint call by the United Nations Secretary-General and the President of the ICRC*, 2023, available at: https://www.icrc.org/en/document/joint-call-un-and-icrc-establish-prohibitions-and-restrictions-autonomous-weapons-systems.

91 ICRC, *Submission on AWS to the UNSG*, note 4 above.

ICRC

International Committee of the Red Cross
19, avenue de la Paix
1202 Geneva, Switzerland
T +41 22 734 60 01

www.icrc.org
facebook.com/icrc
x.com/icrc
instagram.com/icrc

4896/002 12.2025