ICRC POSITION ON AUTONOMOUS WEAPON SYSTEMS

ICRC POSITION ON AUTONOMOUS WEAPON SYSTEMS

The ICRC's concerns about autonomous weapon systems

Autonomous weapon systems select and apply force to targets without human intervention. After initial activation or launch by a person, an autonomous weapon system self-initiates or triggers a strike in response to information from the environment received through sensors and on the basis of a generalized "target profile". This means that the user does not choose, or even know, the specific target(s) and the precise timing and/or location of the resulting application(s) of force.

The use of autonomous weapon systems entails risks due to the difficulties in anticipating and limiting their effects. This loss of human control and judgement in the use of force and weapons raises serious concerns from humanitarian, legal and ethical perspectives.

The process by which autonomous weapon systems function:

- brings **risks of harm for those affected by armed conflict, both civilians and combatants**, as well as dangers of conflict escalation
- raises **challenges for compliance with international law, including international humanitarian law**, notably, the rules on the conduct of hostilities for the protection of civilians
- raises **fundamental ethical concerns for humanity**, in effect substituting human decisions about life and death with sensor, software and machine processes.

The ICRC's recommendations to States for the regulation of autonomous weapon systems

The International Committee of the Red Cross (ICRC) has, since 2015, urged States to establish internationally agreed limits on autonomous weapon systems to ensure civilian protection, compliance with international humanitarian law, and ethical acceptability.

With a view to supporting current efforts to establish international limits on autonomous weapon systems that address the risks they raise, the **ICRC recommends that States adopt new legally binding rules**. In particular:

1 **Unpredictable autonomous weapon systems should be expressly ruled out**, notably because of their indiscriminate effects. This would best be achieved with a prohibition on autonomous weapon systems that are designed or used in a manner such that their effects cannot be sufficiently understood, predicted and explained.

2 In light of ethical considerations to safeguard humanity, and to uphold international humanitarian law rules for the protection of civilians and combatants *hors de combat,* **use of autonomous weapon systems to target human beings should be ruled out**. This would best be achieved through a prohibition on autonomous weapon systems that are designed or used to apply force against persons.

3 In order to protect civilians and civilian objects, uphold the rules of international humanitarian law and safeguard humanity, **the design and use of autonomous weapon systems that would not be prohibited should be regulated**, including through a combination of:

- **limits on the types of target**, such as constraining them to objects that are military objectives by nature
- **limits on the duration, geographical scope and scale of use**, including to enable human judgement and control in relation to a specific attack
- **limits on situations of use**, such as constraining them to situations where civilians or civilian objects are not present
- **requirements for human–machine interaction**, notably to ensure effective human supervision, and timely intervention and deactivation.

The ICRC supports initiatives by States aimed at establishing international limits on autonomous weapon systems that aim at effectively addressing concerns raised by these weapons, such as efforts pursued in the Convention on Certain Conventional Weapons to agree on aspects of a normative and operational framework. Considering the speed of development in autonomous weapon systems' technology and use, it is critical that internationally agreed limits be established in a timely manner. Beyond new legal rules, these limits may also include common policy standards and good practice guidance, which can be complementary and mutually reinforcing. To this end, and within the scope of its mandate and expertise, the ICRC stands ready to work in collaboration with relevant stakeholders at international and national levels, including representatives of governments, armed forces, the scientific and technical community, and industry.

Geneva, 12 May 2021

BACKGROUND PAPER

1. INTERNATIONAL DISCUSSIONS ON AUTONOMOUS WEAPON SYSTEMS

International discussions on the humanitarian, legal and ethical concerns raised by autonomous weapon systems (AWS) have spanned the past decade. These include the work of High Contracting Parties to the Convention on Certain Conventional Weapons (CCW), which have discussed AWS since 2014, in a formal Group of Governmental Experts (GGE) on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems since 2016.

In 2019, the High Contracting Parties to the CCW agreed to work towards consensus recommendations on "aspects of the normative and operational framework" on AWS while adopting 11 Guiding Principles reflecting agreement to date.[1] During 2020 many States elaborated on their understanding of these principles in national commentaries submitted to the GGE and during deliberations at the GGE's September 2020 meeting. This demonstrated increasing convergence of views among States, as noted by consecutive GGE chairpersons during, and following, the 2020 meeting.[2] The GGE is due to hold further sessions in 2021 in advance of the Sixth Review Conference of the CCW – a key moment in States Parties' response to concerns raised by AWS.

The International Committee of the Red Cross (ICRC) first publicly drew attention to its concerns about AWS in 2011. **Since 2015, the ICRC has been calling on States to urgently establish internationally agreed limits on AWS** to respond to the rapid developments toward expanding the use of AWS, and the humanitarian, legal and ethical concerns they raise. The ICRC has subsequently made proposals to States on the general types of limit on AWS needed – in particular in terms of predictability, types of target, duration and scope of use, situations of use, and human supervision – most recently in the ICRC's commentary on the CCW GGE's Guiding Principles.[3] Thus far, the ICRC has left open the question of whether these limits should take the form of new legally binding rules, policy standards or shared practices.

The ICRC's position and its recommendations to States are based on its analyses of associated humanitarian, legal, ethical, technical and military implications of AWS, and insights published in a series of reports, such as the June 2020 report *Limits on Autonomy in Weapon Systems: Identifying Practical Elements of Human Control*, jointly published with the Stockholm International Peace Research Institute (SIPRI), and regular engagement with States and experts at the CCW and bilaterally.[4]

---

1   UN, 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, Geneva, 13-15 November 2019, *Final report*, CCW/MSP/2019/9, 13 December 2019.

2   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, 26 October 2020; UN, *Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons System: Chairperson's Summary*, CCW/GGE.1/2020/WP.7 (Advance copy), 19 April 2021.

3   ICRC, *ICRC Commentary on the "Guiding Principles" of the CCW GGE on "Lethal Autonomous Weapons Systems"*, July 2020.

4   ICRC, Statement of the ICRC to the UN CCW GGE on Lethal Autonomous Weapons Systems, 21−25 September 2020, Geneva; ICRC, *ICRC Commentary on the "Guiding Principles" of the CCW GGE on "Lethal Autonomous Weapons Systems"*, July 2020; V. Boulanin, N. Davison, N. Goussac, and M. Peldán Carlsson, *Limits on Autonomy in Weapon Systems: Identifying Practical Elements of Human Control*, ICRC & SIPRI, June 2020; ICRC, *International Humanitarian Law and the Challenges of Contemporary Armed Conflicts*, 33rd International Conference of the Red Cross and Red Crescent, Geneva, October 2019, pp. 22−24; ICRC, *Autonomy, Artificial Intelligence and Robotics: Technical Aspects of Human Control*, August 2019; ICRC, Statements of the ICRC to the UN CCW GGE on Lethal Autonomous Weapons Systems, 25−29 March 2019, Geneva; ICRC, *The Element of Human Control*, working paper submitted at the Meeting of High Contracting Parties to the CCW, Geneva, 21−23 November 2018, CCW/MSP/2018/WP.3, 20 November 2018; ICRC, *Ethics and Autonomous Weapon Systems: An Ethical Basis for Human Control?*, 3 April 2018; ICRC, *Views of the ICRC on Autonomous Weapon Systems*, 11 April 2016; ICRC, *Autonomous Weapon Systems: Implications of Increasing Autonomy in the Critical Functions of Weapons*, March 2016; ICRC, *Autonomous Weapon Systems: Technical, Military, Legal and Humanitarian Aspects*, March 2014.

On that basis, the ICRC can now **provide more detailed recommendations on what specific limits on AWS are needed** to ensure civilian protection, compliance with international humanitarian law (IHL), and ethical acceptability. Furthermore, **the ICRC is convinced that these limits should take the form of new legally binding rules that specifically regulate AWS**. These rules should clarify how existing rules of international law, including IHL, constrain the design and use of AWS, and supplement the legal framework where needed, including to address wider humanitarian risks and fundamental ethical concerns raised by AWS.

The negotiation of new legally binding rules on AWS and other efforts to develop aspects of an operational and normative framework under consideration in the CCW GGE[5] can be complementary and mutually reinforcing. Such efforts may include initiatives aimed at effectively addressing concerns raised by AWS by way of international commitments agreed among States in a political declaration, the elaboration of international technical standards on testing, validation or verification, as well as national moratoria on the development or procurement of AWS, and measures to support domestic implementation of internationally agreed limits, including in military doctrine and other guidance.

---

[5] UN, 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, Geneva, 13–15 November 2019, *Final report*, CCW/MSP/2019/9, 13 December 2019.

2. CURRENT AND EMERGING AUTONOMOUS WEAPON SYSTEMS

The ICRC understands AWS to be weapons that select and apply force to targets without human intervention. After initial activation or launch by a person, an AWS self-initiates or triggers a strike in response to information from the environment received through sensors and on the basis of a generalized "target profile" (technical indicators function as a generalized proxy for a target).

In simple terms, **AWS are weapons that fire themselves when triggered by an object or person, at a time and place that is not specifically known, nor chosen, by the user**. Indeed, the distinction between a non-AWS and an AWS can be understood by asking whether a person chooses the specific target(s) to strike or not.[6] This process of applying force is a feature that could be implemented with a wide variety of weapon systems, platforms and munitions, especially unmanned systems that are presently remote-controlled.

Some AWS are already in use for specific tasks in narrowly defined circumstances, for example: **air defence systems** used on board warships or at military bases to strike incoming missiles, rockets or mortars; **"active protection" weapons** used on tanks to strike similar types of incoming munitions; **loitering weapons with autonomous modes** used against radars and possibly vehicles; and **certain missiles and sensor-fused munitions** used for example against warships and tanks. Mines have also been described as crude AWS.[7] According to proponents, AWS offer several potential military benefits over directly controlled and remote-controlled weapon systems, including:

- *Increased speed in targeting*: accelerating the process of detecting, tracking and applying force to targets. This provides a military advantage but risks loss of control over the use of force, and escalation.

- *Automated area denial*: AWS can deny adversaries access to or passage through areas without requiring the presence of soldiers or constant monitoring. This is a similar military rationale to laying minefields.

- *Continuing an attack when communications are denied*: Remote-controlled armed drones (air/land/sea) rely on communication links for the operator to trigger a strike but are vulnerable to communications being jammed, cut or hacked. AWS could operate without communications.

- *Operating in greater numbers, including swarms*: Since AWS remove operator involvement in individual strikes, they facilitate greater numbers of unmanned armed systems being deployed with fewer human resources than required for remote-controlled systems.

Some proponents also claim they are pursuing AWS to enable greater precision and/or accuracy in targeting compared to using directly controlled or remote-controlled weapons (non-AWS). AWS actually weaken precision and accuracy because of the shift to a more generalized decision-making in targeting, with *less* knowledge about the eventual target(s), and the precise timing and/or location of the resulting application(s) of force. Constraining AWS, however, does not prevent militaries from using new technologies to ensure greater precision and accuracy in targeting.

Another common argument put forward by proponents is that the use of AWS will be "better than humans" for compliance with IHL. However, to evaluate the risks posed by AWS, we need not compare humans and AWS. Rather, we need to compare (a) the consequences of humans using non-AWS against targets they choose with (b) the consequences of humans using AWS against targets they do not choose specifically. Whatever challenges human decision makers face today in anticipating and constraining the effects of their attacks in accordance with IHL, these are exacerbated, not reduced, by AWS due to the process by which AWS function.

Existing military practice in the use of AWS is characterized by strict limits that can help avoid risks for civilians and "friendly forces" and facilitate compliance with IHL, and that are likely influenced by ethical considerations. These include limits on:

- *Targets*: AWS are generally used to target military objects such as projectiles, aircraft, naval vessels, military radars, tanks or other military vehicles. To the ICRC's knowledge, there are no anti-personnel AWS in use (except anti-personnel landmines whose use is prohibited by the Anti-Personnel Mine Ban Convention and regulated by the CCW Amended Protocol II).

---

6   ICRC, *International Humanitarian Law and the Challenges of Contemporary Armed Conflicts,* 33rd International Conference of the Red Cross and Red Crescent, Geneva, October 2019, pp. 22–24.

7   V. Boulanin, N. Davison, N. Goussac, and M. Peldán Carlsson, *Limits on Autonomy in Weapon Systems: Identifying Practical Elements of Human Control,* ICRC & SIPRI, June 2020, p. 18. See also, ICRC, *Autonomous Weapon Systems: Implications of Increasing Autonomy in the Critical Functions of Weapons*, March 2016, pp. 13–14.

- *Duration and geographical scope of use:* The majority of AWS are in autonomous mode for short periods only, and many are not mobile but rather fixed in place.
- *Situations of use*: The majority of AWS are used only in situations where civilians and civilian objects are not present, or measures are taken (e.g. barriers, warning signs, exclusion zones) to exclude the presence of civilians in the area where the AWS operates.
- *Human–machine interaction:* Almost all AWS are supervised in real time by a human operator that can intervene to authorize, override, veto or deactivate the weapon as needed.

However, the expanding infrastructure of weapon systems that could become future AWS is vast, ranging from hand-held armed quadcopters with facial recognition to autonomous combat aircraft, from "sentry guns" to autonomous tanks, and from armed speedboats to autonomous ship-hunting underwater drones. It includes networks of connected systems, where software for target identification and selection may trigger separate weapons, and autonomous cyber weapons.

Many remote-controlled systems can already *identify, track or select* targets autonomously and it is only a small investment – a software upgrade or even just a change of doctrine – for these systems to *apply force* autonomously. This could also occur due to a malfunction or deliberate hacking of the weapon. For example, remote-controlled "sentry guns" deployed at certain borders and military bases are used to autonomously *select* human targets. To the ICRC's knowledge, users must still specifically authorize the application of force by remote control, although commercial developers have already offered AWS versions.

Current trends in military interest and investments indicate that, without internationally agreed limits, future AWS may be:

- increasingly reliant on artificial intelligence and machine learning software, raising concerns about unpredictability by design
- used to target people and a greater variety of objects
- increasingly mobile and used over wider areas for longer periods, carrying out multiple strikes
- used in cities and towns where civilians would be most at risk
- used without effective human supervision, timely intervention or deactivation.

These trends are not limited to well-resourced States but are a feature of current rapid military technology and doctrinal developments, and proliferation among States and non-State armed groups. All these trends dramatically exacerbate the humanitarian, legal and ethical concerns outlined in the next section. They highlight the urgency of reaching international agreement on new legally binding rules on AWS as well as other aspects of a normative and operational framework on AWS under consideration in the CCW GGE.

3. LIMITS NEEDED ON AUTONOMOUS WEAPON SYSTEMS

The process by which AWS function leads to a loss of human control and judgement over the use of force and weapons, raising serious concerns from humanitarian, legal and ethical perspectives. Generally, the use of AWS introduces a significant increase in risk to those affected by armed conflict, by undermining civilian protection, challenging the rule of law and raising concerns under the principles of humanity.

AWS, as a means of warfare, must be capable of being used and must be used in accordance with IHL. The requirements under the IHL rules on the conduct of hostilities must be fulfilled by the users of an AWS, not by the weapon itself. It is parties to armed conflict – ultimately, human beings – who are responsible for applying IHL and who can be held accountable for violations.[8] However, **the process by which AWS function poses a challenge for compliance with these IHL rules.**

3.1 Addressing concerns about unpredictability in autonomous weapon systems

Humanitarian concerns

A degree of unpredictability is inherent in the effects of using all AWS due to the fact that the user does not choose, or know, the specific target(s), and the precise timing and/or location of the resulting application(s) of force. This brings risks of harm for those affected by armed conflict, serious challenges in applying IHL, and dangers of conflict escalation.

The **trends identified in section 2** (specifically, the use of AWS against a wider range of targets; over longer durations and wider areas; in more dynamic, congested and complex environments; and with reduced human involvement) **will increase the unpredictability of AWS effects, and therefore the risks for civilians**.

In addition, the **development of AWS controlled by artificial intelligence, and especially machine learning software, introduces an additional dimension of unpredictability at the design level**. Machine learning techniques make it extremely difficult for humans to understand and, therefore, to predict and explain the process by which an AWS functions (the "black-box" challenge), irrespective of its environment of use.[9]

International humanitarian law concerns

Unpredictability in AWS poses a fundamental challenge to IHL. Customary IHL prohibits weapons that are by nature indiscriminate, that is, weapons that in their normal or expected circumstances of use, cannot be directed at a specific military objective or whose effects cannot be limited as required by IHL.[10]

Certain AWS would be inherently indiscriminate and, thus, prohibited under existing IHL. These would include, notably, AWS whose effects, in their normal or expected circumstances of use, could not be sufficiently understood, predicted and explained. For instance, if humans responsible for the use of an AWS could not reasonably anticipate what would trigger an AWS strike, they could not control and limit its effects as required by IHL, nor could they explain why a particular person or object was struck in a manner that would allow holding perpetrators of IHL violations to account.

Specifically, if an AWS' functioning is opaque, then humans responsible for the application of IHL rules – both persons entrusted with the legal review of an AWS and persons responsible for compliance with IHL during its use – could not reasonably determine its lawfulness under IHL. The functioning could be opaque notably due to reliance on artificial intelligence and machine learning techniques, or because it changes during use in a way that affects the use of force (e.g. machine learning enables changes to targeting parameters over time).

ICRC recommendation: Ruling out unpredictable autonomous weapon systems

In light of this analysis, **unpredictable AWS should be expressly ruled out, notably because of their indiscriminate effects**: the user cannot know whether they will target civilians or combatants, civilian or military objects or whether their effects will be limited as required by IHL. **This could best be achieved with a prohibition on AWS that are designed or used in a manner such that their effects cannot be sufficiently understood, predicted and explained.**

---

8 ICRC, *International Humanitarian Law and the Challenges of Contemporary Armed Conflicts*, 33rd International Conference of the Red Cross and Red Crescent, Geneva, October 2019, pp. 22–24.

9 ICRC, *Autonomy, Artificial Intelligence and Robotics: Technical Aspects of Human Control,* August 2019; ICRC, *Artificial Intelligence and Machine Learning in Armed Conflict: A Human-Centred Approach*, June 2019.

10 ICRC, *Customary IHL Study*, Rule 71, 2005.

This prohibition would build on the recognition by States of the need for sufficient predictability in the use of AWS for compliance with IHL and for practical military operational reasons. Such a prohibition would find support in the general agreement that inherently indiscriminate weapons are prohibited under existing IHL. A treaty-based prohibition on unpredictable AWS would also help clarify which AWS would be deemed indiscriminate.

3.2 Addressing concerns raised by the use of autonomous weapon systems against persons

Particular ethical concerns and legal challenges also arise with AWS that are designed or used to target persons, as highlighted previously by the ICRC[11] and others.

Ethical concerns

The process by which AWS function raises fundamental ethical concerns for humanity, in effect substituting human decisions about life and death with sensor, software and machine processes. **In sum, most agree that an algorithm – a machine process – should not determine who lives or dies**, even though it is not always explicit whether this concern should rule out: all AWS, AWS that endanger humans, or only AWS that target humans directly.

These concerns have been raised by many States,[12] the UN Secretary-General,[13] civil society[14] and leading figures in the technology industry and scientific community.[15]

These concerns centre on the interrelated loss of human agency, moral responsibility and human dignity in life-and-death decisions. Humans have moral agency and responsibilities that guide their decisions and actions, whereas inanimate objects (e.g. weapons, machines and software) do not. This remains the case regardless of the "sophistication" of an AWS.

Preserving human agency requires effective human deliberation. Without this it can be said that there has not been morally responsible decision-making, nor recognition of the human dignity of those targeted or affected. Removing human agency is a dehumanizing process that undermines a shared sense of humanity. In decisions about life and death, it also removes the possibility for restraint, a human quality that means people may decide *not* to use force even if it would be lawful.

In the view of the ICRC, these ethical concerns apply to AWS that endanger human beings and they are most acute with AWS designed or used to target persons directly (as opposed to AWS that target unmanned military objects such as missiles). The latter would facilitate death and injury based on a generalized target profile, where human life is reduced to sensor data and machine processing.[16] It would effectively amount to "death by algorithm" – the final frontier in the automation of killing.

---

11 ICRC, *Ethics and Autonomous Weapon Systems: An Ethical Basis for Human Control?*, 3 April 2018.

12 See, V. Boulanin, N. Davison, N. Goussac, and M. Peldán Carlsson, *Limits on Autonomy in Weapon Systems: Identifying Practical Elements of Human Control*, ICRC & SIPRI, June 2020, p. 10, note 22.

13 UN Secretary-General, "Machines Capable of Taking Lives without Human Involvement Are Unacceptable, Secretary-General Tells Experts on Autonomous Weapons Systems", SG/SM/19512-DC/3797, 25 March 2019.

14 e.g. Human Rights Watch, *Losing Humanity: The Case against Killer Robots,* 18 November 2012; Article 36, "Targeting People", *Policy Note*, November 2019.

15 e.g. Future of Life Institute, *An Open Letter to the United Nations Convention on Certain Conventional Weapons,* 2017, and Future of Life Institute, *Autonomous Weapons: An Open Letter from AI & Robotics Researchers*, 2015 (4,502 artificial intelligence and robotics researchers, 26,215 other scientists and experts, and the founders and CEOs of 100 artificial intelligence and robotics companies in 26 countries signed open letters calling for prohibitions and regulations on AWS); Google, *AI Principles*, 2018.

16 V. Boulanin, N. Davison, N. Goussac, and M. Peldán Carlsson, *Limits on Autonomy in Weapon Systems: Identifying Practical Elements of Human Control*, ICRC & SIPRI, June 2020, p. 14: "Fundamental ethical concerns do appear to be heightened in situations where AWS are used to target humans, and in situations where there are incidental risks for civilians (though such concerns could also be raised in relation to inhabited military targets, such as military aircraft, vehicles and buildings)."; ICRC, *Ethics and Autonomous Weapon Systems: An Ethical Basis for Human Control?*, 3 April 2018, p. 22: "The combined and interconnected ethical concerns about loss of human agency in decisions to use force, diffusion of moral responsibility and loss of human dignity could have the most far-reaching consequences, perhaps precluding the development and use of anti-personnel autonomous weapon systems, and even limiting the applications of anti-materiel systems, depending on the risks that destroying materiel targets present for human life."

International humanitarian law concerns

From a legal perspective, AWS pose a real risk of harm to persons protected under IHL. In particular, the use of AWS to target human beings entails a significant risk that protected civilians and combatants *hors de combat* may trigger an AWS strike.

Effectively protecting combatants/fighters who are placed *hors de combat* and civilians who are not, or no longer, taking a direct part in hostilities calls for difficult and highly contextual, conduct-, intent- and causality-related legal assessments by humans in the context of a specific attack. Two interrelated challenges make it difficult to envisage how anti-personnel AWS could be used lawfully under IHL. First, the ways in which a civilian might take part in hostilities are extremely diverse, as are the ways in which a combatant, or a civilian taking part in hostilities, may surrender or react to being wounded; a determination of whether a person is protected against attack, or is a lawful target, is therefore highly contextual and does not lend itself to being standardized in a target profile. Second, these legal characterizations can change quickly, meaning that an assumption about the targetability of persons within an AWS' area of operation made by a commander upon launching an attack are subject to change before the AWS strikes. The legal protection of persons from attack varies more easily depending on the circumstances compared to objects that are military objectives by nature (see section 3.3 below).

In today's combat situations, increasingly involving fighting in the midst of urban areas – dynamic and congested places – compliance with the principle of distinction and rules protecting combatants *hors de combat* already presents formidable challenges. The introduction of AWS to target persons can only increase these challenges. In the view of the ICRC, **it is difficult to envisage realistic combat situations where AWS use against persons would not pose a significant risk of IHL violations.**

ICRC recommendation: Ruling out anti-personnel autonomous weapon systems

In light of ethical considerations to safeguard humanity, and to uphold IHL rules for the protection of civilians and combatants *hors de combat,* **use of AWS to target human beings should be ruled out**. This would best be achieved through a prohibition on AWS that are designed or used to apply force against persons.

Such a prohibition is grounded in present practice, where AWS are not yet used to target humans directly. It also finds support in concerns expressed by many States, scientists, philosophers, human rights specialists, civil society, and the public at large, that humans must not delegate life-and-death decisions to machines.

The prohibition of anti-personnel landmines in the Anti-Personnel Mine Ban Convention provides a precedent for excluding AWS that are triggered by persons. The recommended prohibition of anti-personnel AWS will draw an important normative line.

3.3   Addressing concerns raised by other autonomous weapon systems

The use of any AWS must comply with IHL rules aimed at protecting civilians and civilian objects during the conduct of hostilities, notably, the principle of distinction, the prohibitions of indiscriminate and disproportionate attacks and the obligation to take all feasible precautions in attack. Use of AWS raise humanitarian, legal and ethical concerns even in situations other than those discussed above and for which the ICRC recommends a prohibition.

Humanitarian, legal and ethical concerns

AWS use carries a risk that determinations made by the AWS user upon launching an attack are invalidated by a change of circumstances, including determinations about whether the objects the AWS will strike are military objectives and about the proportionality of attack. This risk is heightened, *inter alia*, when targeting objects whose legal characterization as military objectives is subject to rapid change, by a longer duration of an AWS attack, a larger area over which the AWS operates, a higher number of strikes it can conduct, and a more dynamic, congested or complex operating environment. Whereas existing AWS are generally designed and employed in a manner that tries to minimize these risks and facilitate compliance with IHL, the **trends of AWS development identified in section 2 all point in the direction of increased risk** in these respects.

These trends also increase the risk that AWS users would not be in a position to recognize changed circumstances that warrant the suspension of an attack, and that they would be unable to intervene in time to prevent adverse humanitarian consequences and violations of IHL.

Viewed against the backdrop of the evolution of contemporary armed conflict, including the increase of warfare in urban settings, **unfettered AWS design and use bring significant humanitarian risk and risk of violations of IHL**.

Types of measure used to attenuate risks in present practice

Mutually reinforcing humanitarian, legal, ethical and military operational rationales strictly limit AWS design and use in present practice and provide examples of the types of limit on AWS needed to allow the exercise of sufficient human control and judgement over the use of force, and to attenuate the risks highlighted above. **This is done through a combination of technical and doctrinal limits**:

- **Targets pursued with AWS are generally limited to objects whose legal qualification as a military objective is relatively stable, namely military objectives by nature**, such as projectiles, military radar, or military naval vessels. The legal determination of whether other objects are military objectives (e.g. buildings or vehicles can become military objectives if used for military action by the adversary[17]) is typically highly dependent on the circumstances, and can therefore differ between physically similar objects in the AWS' area of operation (e.g. identical vehicles being used by civilians and by military) and vary quickly between the launch of an attack and an AWS strike (e.g. the adversary having stopped using, at the time of the AWS strike, a civilian vehicle they had been using for military action at the time of the AWS launch).

- **The use of AWS is generally limited in space, time and scale of force.** Limits pertain to the area within which an AWS may apply force, the duration of operation, and the scale or number of strikes it may conduct. These limits aim to enable AWS users to have the necessary situational awareness to anticipate the effects of an attack and be reasonably certain upon launching the attack that it will comply with IHL. These limits also reduce the risk that circumstances may change during an attack and facilitate supervision during the operation of the AWS.

- **AWS are generally used in places where civilians and civilian objects are not present.** The higher the number of civilians and civilian objects within the area where an AWS can apply force, the higher the risk of harm to civilians. First, civilian objects such as cars or buses might trigger an AWS whose target profile is meant to capture military jeeps or personnel carriers. Second, civilians and civilian objects may also be harmed incidentally if they are in or near a military objective (such as a military jeep or personnel carrier).

  These risks can be more easily managed in a situation where civilians and civilian objects are not present, e.g. on the high seas far from shipping lanes or fishing areas, or an area from where they can effectively and legitimately be excluded (e.g. through fencing off a military compound or an air exclusion zone). By contrast, use of an AWS in a dynamic, congested or complex civilian environment, such as a city or town, can put civilians at a significant risk of harm. In such environments, concern about compliance with IHL rules for the protection of civilians is heightened. So are ethical concerns about loss of human life as a result of machine processes or calculations in the use of AWS that accidentally or incidentally endanger persons even if they are not directly targeted.

- **AWS are generally used under constant human supervision and with the option of deactivation.** Measures taken in the design and use of AWS (including the limits discussed above on targets, time and space, scale of force, and situations of use) serve to enable real-time situational awareness and to safeguard a practical possibility for AWS users to intervene and deactivate an AWS if need be.

There is a risk that the trends identified in section 2, especially increasing speed, scale, and reliance on artificial intelligence and machine learning to control the selection and application of force to targets, will reduce human operators' capacity to make sense of information received, meaningfully deliberate on their choices and take timely action in line with humanitarian, legal and ethical principles. This, in turn, would reduce the prospect of holding AWS operators to account for harm done and violations of IHL.

ICRC recommendation: Regulation of other autonomous weapon systems

In light of this analysis, **the design and use of AWS that would not be prohibited should be regulated** to avoid harm to civilians and civilian objects, uphold the rules of IHL and safeguard humanity, including through a combination of legally binding:

- **limits on the types of target**, such as constraining them to objects that are military objectives by nature

- **limits on the duration, geographical scope and scale of use**, including to enable human judgement and control in relation to a specific attack

- **limits on situations of use**, such as constraining them to situations where civilians or civilian objects are not present

- **requirements for human–machine interaction**, notably to ensure effective human supervision, and timely intervention and deactivation.

---

[17] Art. 52 of Protocol I additional to the Geneva Conventions; ICRC, *Customary IHL Study,* Rules 7–10, 2005.

4. CONCLUSIONS AND SUMMARY OF THE ICRC'S RECOMMENDATIONS TO STATES

In the view of the ICRC, new legally binding rules are urgently needed to address the humanitarian, legal and ethical concerns raised by AWS that have been highlighted by many States, civil society and the ICRC.

With a view to supporting current efforts to establish international limits on AWS that address the risks they raise, the **ICRC recommends that States adopt new legally binding rules**. In particular:

1 **Unpredictable AWS should be expressly ruled out**, notably because of their indiscriminate effects. This would best be achieved with a prohibition on AWS that are designed or used in a manner such that their effects cannot be sufficiently understood, predicted and explained.

2 In light of ethical considerations to safeguard humanity, and to uphold IHL rules for the protection of civilians and combatants *hors de combat,* **use of AWS to target human beings should be ruled out**. This would best be achieved through a prohibition on AWS that are designed or used to apply force against persons.

3 In order to protect civilians and civilian objects, uphold the rules of IHL and safeguard humanity, **the design and use of AWS that would not be prohibited should be regulated**, including through a combination of:

- **limits on the types of target**, such as constraining them to objects that are military objectives by nature
- **limits on the duration, geographical scope and scale of use**, including to enable human judgement and control in relation to a specific attack
- **limits on situations of use**, such as constraining them to situations where civilians or civilian objects are not present
- **requirements for human–machine interaction**, notably to ensure effective human supervision, and timely intervention and deactivation.

Consistent with the ICRC's long-standing role to prepare the development of IHL, including specific prohibitions and restrictions on weapons,[18] these recommendations aim to uphold humanitarian principles and strengthen IHL in response to challenges raised by the application of science and technology developments to AWS as means and methods of warfare.

In the view of the ICRC, existing IHL rules do not hold all the answers to the humanitarian, legal and ethical questions raised by AWS. New rules are needed to clarify and specify how IHL applies to AWS, as well as to address wider humanitarian risks and fundamental ethical concerns. New legally binding rules would offer the benefits of legal certainty and stability. The ICRC is concerned that without such rules, further developments in the design and use of AWS may give rise to practices that erode the protections presently afforded to the victims of war under IHL and the principles of humanity.

The ICRC offers its recommendation to all States with a view to supporting both national policy development and current international efforts to address the risks posed by AWS, including the work of the CCW GGE to agree aspects of the normative and operational framework on AWS.

The ICRC is encouraged that many States recognize the need for international limits on AWS, with many having already called for new legally binding rules, and others more generally for internationally agreed limits along similar lines to those proposed by the ICRC. The ICRC also acknowledges that diverse views remain on where, and in what form, limits on AWS should be drawn, and that some States consider that national measures are sufficient to address AWS.

Against this backdrop, the ICRC intends with these recommendations to contribute to building shared understandings and fostering progress towards the establishment of effective internationally agreed limits on AWS. The ICRC is looking forward to further discussion with States on these recommendations, including to elaborate what exactly would fall under the purview of the proposed prohibitions and regulations.

---

18 K. Lawand and I. Robinson, "Development of Treaties Limiting or Prohibiting the Use of Certain Weapons: The Role of the International Committee of the Red Cross" in R. Geiß, A. Zimmermann and S. Haumer (eds), *Humanizing the Laws of War: The Red Cross and the Development of International Humanitarian Law*, Cambridge University Press, June 2017, pp. 141–184.

Within the scope of its mandate and expertise, the ICRC will continue to engage with all interested stakeholders and to support initiatives that aim to contribute to limits on AWS that effectively and in a timely manner address the concerns it has raised, including efforts within the framework of the CCW to agree on aspects of the normative and operational framework, such as a political declaration, common policy standards or good practice guidance. To this end, the ICRC stands ready to work in collaboration with relevant stakeholders at international and national levels, including representatives of governments, armed forces, the scientific and technical community, and industry.