



ICRC

**Rio Seminar on Autonomous Weapon Systems
Naval War College, Rio de Janeiro
19-20 February 2020**

AUTONOMOUS WEAPON SYSTEMS¹

The ICRC understands autonomous weapon systems as: *Any weapon system with autonomy in its critical functions. That is, a weapon system that can select and attack targets without human intervention.* Autonomy in critical functions – already found in some existing weapons to a limited extent, such as air defence systems, active protection systems, and some loitering weapons – is a feature that could be incorporated in any weapon system.

The most important aspect of autonomy in weapon systems – from a humanitarian, legal and ethical perspective – is that the weapon system self-initiates, or triggers, an attack in response to its environment, based on a generalized target profile. To varying degrees, the user of the weapon will know neither the specific target nor the exact timing and location of the attack that will result. Autonomous weapon systems are, therefore, clearly distinguishable from other weapon systems, where the specific timing, location and target are chosen by the user at the point of launch or activation.

The ICRC's primary concern is loss of human control over the use of force as a result of autonomy in the critical functions of weapon systems. Depending on the constraints under which a system operates, the user's uncertainty about the exact timing, location and circumstances of the attack(s) may put civilians at risk from the unpredictable consequences of the attack(s). It also raises legal questions, since combatants must make context specific judgements to comply with IHL. And it raises ethical concerns as well, because human agency in decisions to use force is necessary in order to uphold moral responsibility and human dignity.

Fuller understanding of the legal,² military,³ ethical,⁴ and technical⁵ aspects of autonomous weapon systems has enabled the ICRC to refine its views.⁶ It continues to espouse a human-centred approach, based on its reading of the law and ethical considerations for humans in armed conflict.⁷

Human control under IHL

The ICRC holds that legal obligations under IHL rules on the conduct of hostilities must be fulfilled by those persons who plan, decide on, and carry out military operations. It is humans, not machines, that comply with and implement these rules, and it is humans who can be held accountable for violations. Whatever the

¹ Extract from ICRC, *International Humanitarian Law and the Challenges of Contemporary Armed Conflicts*, 2019 pp 29-31; available at <https://www.icrc.org/en/document/icrc-report-ihl-and-challenges-contemporary-armed-conflicts>

² Neil Davison, "A legal perspective: Autonomous weapon systems under international humanitarian law", in UNODA Occasional Papers, No. 30, November 2017; available at <https://www.icrc.org/en/document/autonomous-weaponsystems-under-international-humanitarian-law>; ICRC, *Autonomous Weapon Systems: Technical, Military, Legal and Humanitarian Aspects*, 2014; available at <https://www.icrc.org/en/document/report-icrc-meeting-autonomousweapon-systems-26-28-march-2014>.

³ See ICRC, *Autonomous Weapon Systems: Implications of Increasing Autonomy in the Critical Functions of Weapons*, 2016; available at <https://www.icrc.org/en/publication/4283-autonomous-weapons-systems>.

⁴ See ICRC, *Ethics and Autonomous Weapon Systems: An Ethical Basis for Human Control?*, 2018; available at <https://www.icrc.org/en/document/ethics-and-autonomous-weapon-systems-ethical-basis-human-control>.

⁵ See ICRC, *Autonomy, Artificial Intelligence and Robotics: Technical Aspects of Human Control*, 2019; available at <https://www.icrc.org/en/document/autonomy-artificial-intelligence-and-robotics-technical-aspects-human-control>.

⁶ See ICRC, *IHL Challenges Report 2011*, pp. 39–40. On definitions in particular, see ICRC, *IHL Challenges Report 2015*, p. 45.

⁷ See ICRC, *Statements to the Group of Governmental Experts on Lethal Autonomous Weapons Systems*, March 2019; available at [https://www.unog.ch/80256EE600585943/\(httpPages\)/5535B644C2AE8F28C1258433002BBF14?OpenDocument](https://www.unog.ch/80256EE600585943/(httpPages)/5535B644C2AE8F28C1258433002BBF14?OpenDocument).

machine, computer program, or weapon system used, individuals and parties to conflicts remain responsible for their effects.

Certain limits on autonomy in weapon systems can be deduced from existing rules on the conduct of hostilities – notably the rules of distinction, proportionality and precautions in attack – which require complex assessments based on the circumstances prevailing at the time of the decision to attack, but also during an attack. Combatants must make these assessments reasonably proximate in time to the attack. Where these assessments form part of planning assumptions, they must have continuing validity until the execution of the attack. Hence, commanders or operators must retain a level of human control over weapon systems sufficient to allow them to make context-specific judgments to apply the law in carrying out attacks.

Human control can take various forms during the development and testing of a weapon system (“development stage”); the taking of the decision to activate the weapon system (“activation stage”); and the operation of the weapon system as it selects and attacks targets (“operation stage”). Human control at the activation and operation stages is the most important factor for ensuring compliance with the rules on the conduct of hostilities. Human control during the development stage provides a means to set and test control measures that will ensure human control in use. However, control measures at the development stage alone – meaning control in design – will not be sufficient.

Importantly, however, existing IHL rules do not provide all the answers. Although States agree on the importance of human control – or “human responsibility”⁸ – for legal compliance, opinion varies on what this means in practice. Further, purely legal interpretations do not accommodate the ethical concerns raised by the loss of human control over the use of force in armed conflict.

Towards limits on autonomy in weapon systems

In the ICRC’s view, the unique characteristics of autonomous weapon systems, and the associated risks of loss of control over the use of force in armed conflict, mean that internationally agreed limits are needed to ensure compliance with IHL and to protect humanity.

Insofar as the sufficiency of existing law – particularly IHL – is concerned, it is clear, as shown above, that existing IHL rules – in particular distinction, proportionality, and precautions in attack – already provide limits to autonomy in weapon systems. A weapon with autonomy in its critical functions that is unsupervised, unpredictable and unconstrained in time and space would be unlawful, because humans must make the context-specific judgments that take into account complex and not easily quantifiable rules and principles.

However, it is also clear that existing IHL rules do not provide all the answers. What level of human supervision, intervention and ability to deactivate is needed? What is the minimum level of predictability and reliability of the weapon system in its environment of use? What constraints are needed for tasks, targets, operational environments, time of operation, and geographical scope of operation?

Moreover, the limits dictated by ethical concerns may go beyond those found in existing law. Anxieties about the loss of human agency in decisions to use force, diffusion of moral responsibility, and loss of human dignity are most acute with autonomous weapon systems that present risks for human life, and especially with the notion of anti-personnel systems designed to target humans directly. The principles of humanity may demand limits on or prohibitions against particular types of autonomous weapon and/or their use in certain environments.

At a minimum, there remains an urgent need for agreement on the type and degree of human control necessary in practice to ensure both compliance with IHL and ethical acceptability.

⁸ United Nations, Report of the 2018 Session of the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems, CCW/GGE.1/2018/3, 23 October 2018.