
**AUTONOMOUS WEAPONS SYSTEMS AND
ACCOUNTABILITY UNDER INTERNATIONAL
HUMANITARIAN LAW: BRIDGING GAPS IN LIABILITY
ATTRIBUTION AND ETHICAL GOVERNANCE**

Muskan Sangwan, Queen Mary University of London

ABSTRACT

This paper analyzes the ethical, legal, and accountability challenges posed by autonomous weapon systems (AWS) under international humanitarian law (IHL) and human rights frameworks. It evaluates AWS compatibility with core IHL principles—distinction, proportionality, and precautions—arguing that current AI lacks the nuanced human judgment needed for reliable target selection and force application. Proposing strict liability regimes and a binding treaty mandating meaningful human control, the study urges prohibitions on fully autonomous lethal systems to prevent dehumanization, discrimination, and unremedied harm.

INTRODUCTION

The ethical and legal implications of autonomous weapons systems (AWS) have been the subject of debate since the early 2000s, with some people advocating for the continued use of these systems and others advocating for their prohibition. Although the debate regarding the ethical and legal implications of AWS's use remains deeply polarised, there is a consensus regarding the ethical and legal aspects that should be taken into account when making this decision: International Humanitarian Law (IHL), respect for human dignity, and international stability. There is a general agreement that AWS can only be deployed following the principles of necessity, proportionality, and distinction of IHL, which are central to this debate¹. These principles are universally acknowledged; however, the issue is determining whether and to what extent autonomous artificial agents that facilitate AWS can adhere to them.

Targets are selected and engaged by AWS through sensor processing, rather than human inputs. Upon initial user activation, they depend on software, which frequently employs algorithms, input from sensors such as cameras, radar signatures, and heat structures, as well as other data, to identify a target². The payload is fired or released by the systems without the approval or evaluation of a human operator once they have identified a target. In other words, the application of force is determined by a machine rather than a human, and the location, timing, and force are all predetermined.

For years, certain weapons systems have existed with varying degrees of autonomy; however, their categories of targets, duration of operation, geographical scope, and operating environment have been restricted. They consist of missile defence systems, including the US Phalanx Close-In Weapon System and Israel's Iron Dome. Other examples include armed drones and loitering munitions, or so-called one-way attack drones that remain in the air in search of a target before launching an attack.³

The rapid development of AWS that operates without substantive human control is currently being driven by technological advancements and military investments. Individuals may also be

¹ Docherty, B. (2025). A hazard to human rights. Human Rights Watch. <https://www.hrw.org/report/2025/04/28/hazard-human-rights/autonomous-weapons-systems-and-digital-decision-making>.

² Cummings, M.L. and International Security Department and US and the Americas Programme (2017) 'Artificial intelligence and the future of warfare,' Chatham House [Preprint].

³ Connolly, C. (2022). Weapons systems with autonomous functions used in Ukraine. <https://automatedresearch.org/news/weapons-systems-with-autonomous-functions-used-in-ukraine/#:%7E:text=Ukraine%20has%20also%20been%20using,feature%20and%20object%20recognition%20capabilities>.

targeted by AWS, which have been granted the authority to make life-and-death decisions. These AWS pose a variety of ethical, moral, legal, accountability, and security concerns, including those that are governed by international human rights law.

INTERNATIONAL HUMANITARIAN LAW

IHL treaties do not specifically regulate autonomous weapon systems as they are defined. Nevertheless, it is undeniable that any autonomous weapon system must be capable of being used and must be used in accordance with IHL. Each state that is engaged in the development, deployment, and utilisation of armaments bears the primary responsibility for ensuring this. While the parties to an armed conflict are the primary subjects of IHL, the regulations governing the conduct of hostilities—particularly the rules of proportionality, distinction, and precautions in attack—are directed at those who plan, deliberate, and execute an attack⁴.

The following are the fundamental legal responsibilities of a commander or operator in the use of weapon systems: to ensure that there is a clear distinction between military objectives and civilian objects, combatants and civilians, and active combatants and those hors de combat; to ascertain whether the attack is likely to result in incidental civilian casualties and damage to civilian objects, or a combination of these, that would be excessive in relation to the concrete and direct military advantage anticipated, as required by the rule of proportionality; and to cancel or suspend an attack if it becomes apparent that the target is not a military objective, is subject to special protection, or that the attack may be expected to violate the rule of proportionality, as required by the rules on precautions in attack⁵. These regulations establish obligations for human combatants in the use of all armaments to guarantee compliance with IHL.

The development and use of AWS implicates at least 4 core obligations and principles of international human rights law:

Right to Life: The right not to be arbitrarily deprived of life necessitates that the use of force

⁴ Davison, N. and International Committee of the Red Cross (2018) Autonomous weapon systems under international humanitarian law, UNODA Occasional Papers, pp. 5–

8. https://www.icrc.org/sites/default/files/document/file_list/autonomous_weapon_systems_under_international_humanitarian_law.pdf.

⁵ Davison, N. and International Committee of the Red Cross (2018) Autonomous weapon systems under international humanitarian law, UNODA Occasional Papers, pp. 5–

8. https://www.icrc.org/sites/default/files/document/file_list/autonomous_weapon_systems_under_international_humanitarian_law.pdf.

be proportionate and necessary to achieve a legitimate objective⁶. Lethal force, in particular, must be employed as a last resort to safeguard human life, as mandated by the right. This three-part exercise would present significant challenges for AWS. The current technology and other technical limitations are among the obstacles that imply automated systems would never be able to approximate or surpass the distinctly human abilities required to complete specific types of tasks⁷. AWS would be unable to interpret the necessity of an attack by identifying subtle cues of human behaviour, would lack the human judgement to assess proportionality, and would be unable to communicate effectively with an individual to defuse a situation and ensure that lethal force is a last resort. The use of force by them would be arbitrary and unlawful as a consequence.

Human Dignity: The principle of human dignity is the foundation of all human rights, including the right to life, and it posits that individuals possess an inherent value that is both universal and inviolable. The process of rendering life-and-death determinations by AWS would violate that foundational principle. Despite their lack of the unique human capacity to comprehend or appreciate the true value of a human life, these machines would murder, as they are not living beings. Additionally, they would dehumanise and instrumentalise their targets by employing algorithms that reduce individuals to data points⁸.

Non-discrimination: The principle of non-discrimination mandates the safeguarding and advancement of human rights for all individuals, regardless of their race, gender, ability, or other legal status⁹. There are numerous reasons why AWS would be discriminatory. For instance, the design of a system and subsequent decision-making processes may be influenced by the prejudices of developers, such as their programming or selection of training data. After the deployment of an autonomous weapon system that employs artificial intelligence (AI), a human operator may be unable to inspect recommended targets and intervene to rectify errors prior to the application of force due to a lack of comprehension regarding the system's decision-making process. Algorithmic bias can have a disproportionately negative impact on already

⁶ Docherty, B. (2025). A hazard to human rights. Human Rights Watch. <https://www.hrw.org/report/2025/04/28/hazard-human-rights/autonomous-weapons-systems-and-digital-decision-making>.

⁷ Docherty, B. (2025). A hazard to human rights. Human Rights Watch. <https://www.hrw.org/report/2025/04/28/hazard-human-rights/autonomous-weapons-systems-and-digital-decision-making>.

⁸ Geiß, R. et al. (no date) Lethal Autonomous Weapons Systems Technology, Definition, Ethics, Law & Security, Lethal Autonomous Weapons Systems Technology, Definition, Ethics, Law & Security. <https://www.auswaertiges-amt.de/blob/204830/5f26c2e0826db0d000072441fdeaa8ba/abruestung-laws-data.pdf>.

⁹ Docherty, B. (2025). A hazard to human rights. Human Rights Watch. <https://www.hrw.org/report/2025/04/28/hazard-human-rights/autonomous-weapons-systems-and-digital-decision-making>.

marginalised groups, as evidenced by other AI technologies.

Right to Remedy: The right to remedy, which is activated after an autonomous weapon system's lifecycle, requires states to pursue prosecutions for severe violations of IHL and gross violations of international human rights law, as well as to provide various forms of reparation. In particular, AWS that rely on AI may make determinations through opaque, "black box" processes, which present challenges in terms of holding individual operators criminally liable for the unpredictable actions of a machine they cannot comprehend¹⁰. Additionally, the identification of programmers and developers who are accountable under civil law presents legal obstacles. This would result in an accountability breach if AWS were implemented.

The spectrum of issues that AWS raise in times of armed conflict and law enforcement operations is exemplified by the infringements on these four human rights obligations and principles. Human rights are a critical lens through which to examine this rapidly emerging technology, although these systems also present ethical, security, IHL, and other threats¹¹. International human rights law is not the exclusive method to frame the concerns with AWS.

LIABILITY/ACCOUNTABILITY

Governments may adopt autonomous weapons, such as military robots, to obtain an advantage in armed conflict as artificial intelligence advances. These weapons are capable of autonomously navigating environments and making decisions regarding who to attack and who to spare. Unintended injury is inevitable, even though the aforementioned example is unlikely to occur. The consideration of these scenarios assists in the formulation of critical questions that governments and researchers must jointly address, including: How can we ensure that human beings are held accountable for the actions of autonomous systems? If the perpetrator is fundamentally a computer, how is justice served? It has been discovered that there is no simple solution to this issue. A human soldier is held accountable for the killing of defenceless civilians when they commit an atrocity. However, when autonomous weapons are responsible for murdering, it is challenging to assign responsibility for their errors.

¹⁰ Geiß, R. et al. (no date) Lethal Autonomous Weapons Systems Technology, Definition, Ethics, Law & Security, Lethal Autonomous Weapons Systems Technology, Definition, Ethics, Law & Security. <https://www.auswaertiges-amt.de/blob/204830/5f26c2e0826db0d000072441fdeaa8ba/abruestung-laws-data.pdf>.

¹¹ Docherty, B. (2025). A hazard to human rights. Human Rights Watch. <https://www.hrw.org/report/2025/04/28/hazard-human-rights/autonomous-weapons-systems-and-digital-decision-making>.

The "decision" of an autonomous weapon to murder innocent civilians is analogous to a computer's "decision" to pause the screen and delete an unsaved project. People hardly believe that the computer is intended to complicate their lives, despite the frustration of a frozen computer. Autonomous weapons may exhibit outward indicators of decision-making and intention, but they still operate on a code that is as impersonal as the code that glitches and freezes a computer screen. Therefore, intention must be demonstrated to prosecute an individual for a war crime. Similar to computers, these systems are not legal or moral agents, and it is unclear whether or not they can be held culpable for their errors.

Therefore, who is accountable for the loss of innocent lives caused by autonomous weapons? Is it permissible for them to murder at all?

Due to their potential for being significantly more hazardous than semi-autonomous and non-lethal autonomous weapons, lethal autonomous weapon systems should be subject to a rigorous liability standard. In general, strict liability is administered with a greater degree of precision and under more stringent conditions than any other tort law liability theory¹². Although the utilisation of fully autonomous weapon systems, particularly those that are lethal, does not fall within any of the categories that were initially listed, it is possible that their utilisation could be classified as ultra-hazardous activities. In this situation, a defendant is held accountable for an injury, despite the absence of any negligence, as a result of the execution of an ultra-hazardous or dangerous activity.

The origin of this concept can be traced back to the original case that established strict liability: *Rylands v. Fletcher*. In this case, the English court distinguished between the "natural" and "non-natural" uses of land¹³. *Rylands* has been typically adopted by American courts; however, they are still hesitant to impose strict liability without further consideration. This implies that the courts will consider the activity, as well as the location and circumstances in which it is being conducted.

When discussing strict liability, the Restatement (Third) of Torts evaluates four primary variables. Strict liability will apply to an abnormally hazardous activity if the following

¹² Gregory C. Keating, *The Theory of Enterprise Liability and Common Law Strict Liability*, 54 Vand. L. Rev. 1285, 1289 (2001).

¹³ Charles E. Cantu, *Distinguishing the Concept of Strict Liability for Ultra-Hazardous Activities from Strict Products Liability Under Section 402A of the Restatement (Second) of Torts: Two Parallel Lines of Reasoning That Should Never Meet*, 35 Akron L. Rev. 31, 33 (2001)

conditions are met¹⁴:

- (1) the activity generates a foreseeable risk of physical harm;
- (2) the risk is "highly significant"; and
- (3) the risk persists "even when reasonable care is exercised by all actors." and
- (4) "the activity is not a matter of common usage."

The most compelling case in this strict liability domain is when a defendant is aware of and comprehends the substantial risk that the activity presents, yet they choose to proceed. Strict liability will undoubtedly prevail in these circumstances.

Autonomous weapons are inherently associated with a highly significant risk. Strict liability would be the appropriate regulation for entirely autonomous weapons when analysed through both the Rylands approach and the Restatement (Third) of Torts approach. Rylands posits that the military is granted the privilege of employing autonomous weaponry; however, these weapons will not engage in significant human interaction¹⁵. This excludes them from the category of "ordinary" applications for comparable items, particularly because human control would not be guaranteed with entirely autonomous weapons.

Stronger justification for the regulation of lethal autonomous weapon systems under strict liability is provided by the Restatement. Regardless of their level of sophistication, autonomous weapons are inherently hazardous. There is no indication as to how the paths of killer robots will change from the point at which they depart to the point at which they strike a target, as they will have the capacity to select and target independently, rather than through human guidance¹⁶. This risk is further exacerbated by machine learning, as these killer robots may eventually develop the ability to reason and act independently. In this scenario, a negligence standard would also fail, as there is no way to avoid the danger when the machines have a mind

¹⁴ Restatement (Third) of Torts § 20 (Am. Law Inst. 2010).

¹⁵ Fuzaylova, E. (no date) 'War torts, autonomous weapon Systems, and liability: Why a limited strict liability tort regime should be implemented,' *Cardozo Law Review*, 40(3), pp. 1–15. <https://cardozolawreview.com/war-torts-autonomous-weapon-systems-and-liability/>.

¹⁶ Fuzaylova, E. (no date) 'War torts, autonomous weapon Systems, and liability: Why a limited strict liability tort regime should be implemented,' *Cardozo Law Review*, 40(3), pp. 1–15. <https://cardozolawreview.com/war-torts-autonomous-weapon-systems-and-liability/>.

of their own, unless these weapons are not used at all.

For instance, the "Stop Autonomous Weapons" video that was screened at the CCW convention features a miniature drone-like lethal autonomous weapon¹⁷. If the weapon was initially deployed with a specific target in mind, but it later reevaluated the data in its software and determined that a different individual should be the target, it would unquestionably follow the new route rather than the original. In the interim, the military, which was intended to supervise the operation, lacks any significant input or control due to the assassin robot's complete autonomy¹⁸. Consequently, there is no feasible method to circumvent the danger that is generated, as a human would be incapable of either altering the machine's trajectory or halting it. Despite the potential peril, the state may elect to deploy the weapon in this hypothetical scenario.

The world has yet to reach a consensus on a plan for the rise of autonomous weapons, despite the fact that numerous United Nations conventions have been convened to discuss the aforementioned ban on these weapons. The likelihood of catastrophic incidents occurring without remedies would be exceedingly high in the absence of an enforced strict liability tort regime for accidental injuries sustained from killer robots. A state would be compelled to reconsider the use or deployment of specific AI in the presence of such a liability system.

CONCLUSION AND RECOMMENDATIONS

The matter of the compatibility of AWS with IHL has been left in ambiguity, possibly intentionally, due to the absence of a settled and definitive regulatory position. However, the regulatory debate remains stagnant, and technological advancements in the autonomous weaponry sector persist, as nations inch closer to employing offensive autonomous weapons in combat. We must ascertain whether the deployment of autonomous weapons is following IHL in its current state. It is contended that the most effective approach to resolving this matter is to unwind the layers of the law and conduct an analysis of the matter by evaluating the compatibility of autonomous weapons with the principles of IHL. Ultimately, the principles are fundamental to the regime and can serve as a dependable and recognisable guide to a subject

¹⁷ Stop Autonomous Weapons (2017) *Slaughterbots*. <https://www.youtube.com/watch?v=9CO6M2HsoIA>.

¹⁸ fuzaylova, E. (no date) 'War torts, autonomous weapon Systems, and liability: Why a limited strict liability tort regime should be implemented,' *Cardozo Law Review*, 40(3), pp. 1–15. <https://cardozolawreview.com/war-torts-autonomous-weapon-systems-and-liability/>.

that is both emerging and intricate.

To resolve these concerns, it is necessary to establish a legally binding instrument (treaty) that will prohibit and regulate AWS. Three primary obligations should be incorporated into the treaty. Initially, it should include a general obligation to preserve meaningful human control over the use of force. This obligation would establish a principle that would govern the interpretation of the remaining portions of the instrument, and its generality would seal unexpected loopholes in other provisions. Given the potential for new issues to arise as technology continues to develop, these factors are of particular significance¹⁹. In order to account for future, potentially unforeseen technologies, the general obligation should concentrate on the regulation of conduct (i.e., the use of force) rather than a specific system. The phrase "use of force" guarantees that the obligation would be applicable to armed conflict and law enforcement operations, both of which are governed by international human rights law.

Secondly, the legally binding instrument should prohibit the development, production, and use of AWS that, by their very nature, present fundamental moral or legal issues. It should prohibit AWS that operate without meaningful human control and those that target individuals²⁰. The absence of meaningful human control is the cause of many of the threats posed by AWS, including legal, accountability, and ethical ones. Furthermore, the utilisation of target attributes, such as weight, heat, or sound, by AWS to identify and apply force to individuals would result in "digital dehumanisation," violations of human dignity, and discrimination²¹. The dual prohibitions would aid in the prevention of damage to civilians and other protected persons.

Third, the treaty should include rules, or positive commitments, to make sure that people still have a say in how any AWS that aren't banned are used. Some of the parts of meaningful human control that the rules could be based on are²²:

¹⁹ Docherty, B. (2025). A hazard to human rights. Human Rights Watch. <https://www.hrw.org/report/2025/04/28/hazard-human-rights/autonomous-weapons-systems-and-digital-decision-making>.

²⁰ Docherty, B. (2025). A hazard to human rights. Human Rights Watch. <https://www.hrw.org/report/2025/04/28/hazard-human-rights/autonomous-weapons-systems-and-digital-decision-making>.

²¹ Docherty, B. (2025). A hazard to human rights. Human Rights Watch. <https://www.hrw.org/report/2025/04/28/hazard-human-rights/autonomous-weapons-systems-and-digital-decision-making>.

²² Docherty, B. (2025). A hazard to human rights. Human Rights Watch. <https://www.hrw.org/report/2025/04/28/hazard-human-rights/autonomous-weapons-systems-and-digital-decision-making>.

- Decision-making parts (for example, operators need to know how an autonomous system works);
- Technological parts (for example, predictability, reliability, and the operator's ability to intervene); and
- Operational parts (for example, limits on the types of targets and the area and time of operation).

There have been several findings that have come out of this research. The first is that humanity and military necessity are not actually principles of IHL since they don't have the basic features of principles. The second is that autonomous weapons would need to have a lot of artificial intelligence in order to follow the real criteria of IHL, which are distinction, proportionality, and prudence. The third reason is that, right now, robots don't have the appropriate level of sophistication since they can't make good, context-sensitive decisions based on outside information. The fourth reason is that technology has come a long way, and robots are almost ready to follow IHL. The last piece of the jigsaw is for robots to get "high-level" intelligence, which scientists say might happen in only a few decades.