

Pakistan Journal of Criminal Justice
ISSN: 2958-9363 ISSN (E): 2958-9371
Volume 6, No.8, 2026, 142-162
DOI: https://doi.org/10.59075/pjcj.v6i8.308
Journal Homepage: https://journals.centeriir.org/index.php/pjcl

Artificial Intelligence and Autonomous Weapons in Armed Conflict: Rethinking Human Control under International Humanitarian Law

Dr. Jibran Jamshed

Associate Professor, Department of Law, The Islamia University of Bahawalpur. (Currently working as GM Legal PAFDA, Home Department), E-Mail: Jibran_jamshed@yahoo.com

Abstract

Autonomous weapon systems (AWS) or weapons that can identify and engage targets without any further human intervention once activated, are being moved from the fringes of military planning to actual implementation. There is broad consensus that international humanitarian law (IHL) applies to these systems, but it is contested how to apply the core IHL rules of distinction, proportionality and precautions in attack to autonomous decision-making. This article examines whether existing IHL rules are sufficient to regulate AWS or whether a new international legal instrument is required. The article argues that the existing IHL is applicable and already excludes some forms of autonomy, but it is not enough. This argument is made through doctrinal analysis of the Geneva Conventions and Additional Protocol I, customary IHL, the Article 36 weapons review obligation and a comparative analysis of state positions in the Convention on Certain Conventional Weapons and the United Nations General Assembly. Law only implicitly requires human involvement, weapons reviews are uneven and ill-suited to learning systems, and there is a real accountability gap when harm results from unpredictable system behaviour. The article proposes a five-part evaluative framework: predictability, targets, bounding, intervention and traceability and concludes that a legally binding instrument based on a two-tier approach of prohibitions and regulations is needed to make the human element explicit in the use of force.

Keywords: autonomous weapon systems, artificial intelligence, international humanitarian law, meaningful human control, Article 36 weapons review, accountability

OPEN ACCESS
© 2026 The Authors, Published by Center of Innovation in Interdisciplinary Research (CIIR).
This is an Open Access Article under the Creative Common Attribution Non-Commercial 4.0

Article History: Received: 13-08-2026 Accepted: 28-09-2026 Published: 13-10-2026

Corresponding Author's Email: Jibran_jamshed@yahoo.com

Introduction

In March 2021 a United Nations Panel of Experts on Libya reported that retreating forces affiliated with Khalifa Haftar had been "hunted down and remotely engaged" by loitering munitions such as the STM

Kargu-2 which, according to the Panel, "were programd to attack targets without requiring data connectivity between the operator and the munition" (United Nations Security Council, 2021, para. 63). The paragraph was short and, as several commentators later pointed out, it did not actually confirm anyone had been killed by a machine acting on its own. But for many who have been following the debate, it was a moment when the question of "killer robots" stopped being a thought experiment and became something that might already be happening on real battlefields. Since then, the wars in Ukraine and Gaza have demonstrated just how rapidly artificial intelligence (AI) is being employed in targeting, from drones that can fly their final approach without a radio link to decision support systems that generate long lists of suggested targets for human approval (Abraham, 2024; United Nations Secretary-General, 2024).

These developments pose a difficult legal issue which I want to address in this article. International humanitarian law (IHL), or the law of armed conflict, was drafted by people for people. Its core rules on the conduct of hostilities – distinction, proportionality and precautions in attack – assume there is a judgement being made: a commander who judges a building to be a military objective, or an officer who weighs the expected civilian harm against the anticipated military advantage. But when some of that judgement is handed over to software that selects and engages targets without further human intervention, it is no longer clear who is making the judgement, how it should be evaluated, or who should be accountable when something goes wrong.

States and scholars broadly agree that IHL applies to autonomous weapon systems (AWS). This was confirmed in the first of the eleven Guiding Principles adopted by the Group of Governmental Experts (GGE) under the Convention on Certain Conventional Weapons (CCW) in 2019 (CCW GGE, 2019, Annex IV). It's the difference in other places. Some states and writers argue that existing law, applied in good faith and supported by national weapons reviews, is sufficient (Schmitt, 2013; United States Department of Defence, 2023). Others contend that the current rules, while applicable, were never meant to govern machines that learn, adapt or operate in unpredictable ways, and that a new legally binding instrument is needed to ensure "meaningful human control" over the use of force (Human Rights Watch, 2012; International Committee of the Red Cross [ICRC], 2021). In October 2023, the United Nations Secretary-General and the President of the ICRC went so far as to call on states to conclude such an instrument by 2026 (United Nations Secretary-General & ICRC, 2023).

Research Problem and Research Gap

The research problem of the paper can be simply formulated as follows. The existing IHL rules clearly apply to new weapons technologies. However, how distinction, proportionality and precautions should apply to autonomous decision-making is contested. The literature thus far is largely of two kinds. The first group, which mostly wrote between 2012 and 2016, raised the question of whether AWS should be banned in principle (Asaro, 2012; Sharkey, 2012; Schmitt & Thurnher, 2013). The second group, which has been growing since around 2018, is focused on the idea of human control and on practical elements that might give it content (Boulanin et al., 2020; Santoni de Sio & van den Hoven, 2018). What's still missing, at least in my reading of the literature, is a structured way of connecting these two strands: a framework that takes each core IHL obligation, asks what kind of human involvement it actually requires, and then identifies where the existing tools (especially Article 36 weapons reviews and the rules on individual and state responsibility) leave gaps. This is the void the present article is trying to fill.

Research Question and Aims

The main research question is, are the existing rules of IHL adequate for regulating autonomous weapon

systems or should we develop a new international legal instrument?To answer this question, the article aims to fulfilll three objectives. It first considers the basic concepts of IHL of distinction, proportionality and precautions and considers their application to systems that select and engage targets with some degree of autonomy. Secondly, it assesses the notion of human control and the role of legal reviews of new weapons under Article 36 of Additional Protocol I. Thirdly, it identifies the accountability gaps that occur for individuals and States when AWS are used.

Organization of the Article

The article goes on to say: After a brief methodology section, the next part defines AI-enabled weapons and meaningful human control and offers a simple spectrum of human control that will be used throughout the analysis. The section below considers the relevant legal framework under the Geneva Conventions, Additional Protocol I and customary law, including the Article 36 review obligation and the Martens Clause. I then apply the rules of discrimination, proportionality and precautions to autonomous targeting." Part five deals with individual criminal responsibility and state responsibility. The sixth part compares the main regulatory proposals and the positions of the states at the CCW and the UN General Assembly. The last part presents the findings, proposes an evaluative framework and concludes.

Methodology

This is a doctrinal legal research. Doctrinal research, as it is generally understood in law schools, involves finding the relevant legal sources, interpreting them and structuring them into a coherent account of what the law requires (Hutchinson & Duncan, 2012). The principal sources for this article are the four Geneva Conventions of 1949, Additional Protocol I of 1977 (AP I), the CCW and its protocols, the Rome Statute of the International Criminal Court and the International Law Commission's Articles on Responsibility of States for Internationally Wrongful Acts (ARSIWA). I will interpret these treaties according to the ordinary meaning of their terms in context and in the light of their object and purpose as required by Article 31 of the Vienna Convention on the Law of Treaties (1969).

I also rely on customary IHL, as many major military powers such as the United States, India, Israel, Pakistan, Iran and Türkiye are not parties to AP I. The main reference for this is the ICRC study on customary law (Henckaerts & Doswald-Beck, 2005) combined with the more critical literature on whether particular rules have actually reached customary status (Jevglevskaja, 2018).

The third component of the methodology is the examination of Article 36 weapons reviews. Here I rely on the ICRC's guide to legal reviews (ICRC, 2006), publicly available national practice such as United States Department of Defence Directive 3000.09 (2023), and the academic literature on whether reviews can cope with machine learning systems (Boothby, 2016; Copeland et al., 2023).

The article ends with a comparative analysis of the positions of the states. The main sources are the reports of the CCW GGE, working papers submitted by states, United Nations General Assembly resolutions 78/241 and 79/62, the 2024 report of the Secretary-General compiling state views, and secondary literature tracking state positions (Human Rights Watch, 2020; Rosert & Sauer, 2021). I group the states into broad clusters instead of looking at each state individually; this is a simplification, but it makes the comparison feasible in a single article.

There are some limitations I must be honest about. Many of the relevant state practice is classified and states do not often publish their weapons reviews. I also lack technical training in computer science, so

my discussion of how artificial intelligence systems work is based on secondary sources written for non-specialists (Boulanin & Verbruggen, 2017; Scharre, 2018). Finally, I have deliberately limited the sources to those published up to the end of 2025, meaning that some very recent developments in the diplomatic process are not covered.

Meaningful Human Control & Weapons with AI

What are Autonomous Weapons Systems?

This alone is part of the problem, as there is no agreed legal definition for an autonomous weapon system (Crootof, 2015; McFarland, 2020). The most commonly cited working definition is that used by the ICRC, which defines AWS as weapons that, once activated by a person, "select and apply force to targets without human intervention" (ICRC, 2021, p. 2). A similar formula is used in the United States: "A weapon system that, once activated, is able to select and engage targets without further intervention by an operator" (United States Department of Defence, 2023, p. 21). Both definitions hinge on the point that autonomy concerns the critical functions of selecting and engaging targets, not navigation, take-off or other functions that have been automated for decades.

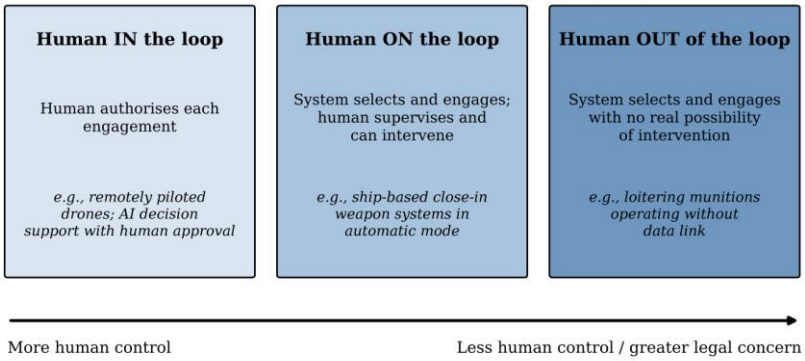
It is important to point out that autonomy and AI are not identical. Several weapons that already fit the ICRC definition do not use AI in any modern sense. Close-in weapon systems like Phalanx, which protect ships from incoming missiles, have been operating in an automatic mode since the 1980s with fairly simple rules (Scharre, 2018). In a way, anti-personnel landmines, which are banned for states party to the Ottawa Convention, are the most primitive autonomous weapon of all, because once laid they "decide" to explode when someone steps on them. What AI, and especially machine learning, brings is the ability to recognise targets in complex, cluttered environments, to adapt to new information, and to operate over larger areas and longer timescales (Boulanin & Verbruggen, 2017). This is precisely what appeals to militaries about AI, but it also takes away the predictability of the system, as the designers themselves may not be able to explain why a certain output was produced (Davison, 2017).

Separating AWS from AI decision support systems is also useful. Tools like the reported "Lavender" and "Gospel" systems in Gaza do not shoot weapons themselves. Instead, they recommend targets to human operators, who then authorise strikes (Abraham, 2024). Formally, a human stays in the loop. But if the human has to approve each recommendation in a few seconds, one can reasonably question whether the control is meaningful in any real sense. The problem of 'automation bias' illustrates that the distinction between autonomous and non-autonomous systems is less clear in reality than it looks on paper.

The Spectrum of Human Control

The literature often refers to human control in terms of "loops". When a human has to approve every engagement, that human is said to be in the loop. Human on the loop or off the loop, a human being is always standing by to supervise and intervene to stop the system if it can fire itself. Where the system operates with no real possibility of human intervention after activation, the human is out of the loop (Human Rights Watch, 2012; Scharre, 2018). These categories are illustrated in Figure 1, along a spectrum with examples and an overall indication of how legal concern correlates with decreasing human control.

Fig. 1: The Spectrum of Human Control over Autonomous Weapons Systems



Remark. Adapted from the categories "in/on/out of the loop," by Human Rights Watch (2012), and Scharre (2018), by the author. The examples are illustrative. Depending on how it is configured and used, the same system may operate at different points on the spectrum.

The loop language is a useful starting point but I think it has two weaknesses. First, it privileges the moment of engagement and ignores what happens before and after: decisions about where and when the system is deployed, what target profiles it is given, and how it is tested. Second, it presumes any human involvement is good enough. A person "on the loop" supervising a swarm of one hundred drones cannot realistically intervene in all engagements. These reasons have pushed the more recent literature to shift toward the notion of meaningful human control.

Robust human control

The United Kingdom-based non-governmental organization Article 36 brought the phrase "meaningful human control" to prominence as the central concept in the international debate around 2013 (Roff & Moyes, 2016). The idea is that it is not enough for a human to be somewhere along the chain; the human involvement needs to be of a quality that allows the person to make informed legal and moral judgements and to be held responsible for them. Roff and Moyes (2016) argued that meaningful control requires predictable, reliable and transparent technology, accurate information for the user, timely human action and possibility of intervention, and accountability to a certain standard.

Santoni de Sio and van den Hoven (2018) provide a more philosophical account, arguing that meaningful human control requires two conditions. The "tracking" condition requires the system to respond to the moral and legal reasons of the humans who deploy it, and the "tracing" condition requires that its outcomes be traceable to at least one human who understands the system's capabilities and their own moral responsibility. This account is particularly useful to me in that it directly links control to responsibility, which is the central legal concern.

In a joint study by the Stockholm International Peace Research Institute (SIPRI) and the ICRC, a more pragmatic approach was taken in identifying three types of measures that can collectively give effect to human control (Boulanin et al., 2020). (a) Controls on weapon parameters, such as limiting the type of targets, the duration of operation and the geographical area; (b) Controls on the environment of use, such as limiting use to areas where civilians are not present; and (c) Controls through human-machine interaction, such as the ability to supervise, intervene and deactivate. This framework has been very

influential and is reflected in the ICRC's 2021 position (discussed below).

Not every state is buying the language of meaningful human control. For example, the United States prefers the formulation “appropriate levels of human judgement over the use of force” arguing that it is human judgement that is important not moment-by-moment control (United States Department of Defence, 2023, p. 3). The United Kingdom uses the term “context-appropriate human involvement” (UK Ministry of Defence, 2022). Some scholars have also argued that the focus on human control is misguided because humans themselves make many mistakes in war and the law should focus on outcomes rather than insisting on human error-prone judgement (Jensen, 2020). This dispute about language is not just semantic. As I discuss further below, the type of instrument states will be willing to agree to is directly affected by whether the standard is "control" or "judgement."

Legal Framework Applicable under the Geneva Conventions and Additional Protocol I.

Additional Protocol I and the Geneva Conventions

The four Geneva Conventions of 1949 have been universally ratified but most of their provisions concern the protection of persons in the power of a party, like the wounded, prisoners of war, and civilians in occupied territory. The most important AWS rules are about how it can conduct hostilities. Those rules are mostly found in AP I, adopted in 1977. AP I has 174 states parties, a large number, but as noted above several major military powers are not among them.

The most relevant articles of AP I for this article are Article 48 (the basic rule of distinction), Article 51 (protection of the civilian population, including the prohibition of indiscriminate attacks and the proportionality rule in Article 51(5)(b)), Article 52 (definition of military objectives), Article 57 (precautions in attack), and Article 36 (legal review of new weapons). The right to choose methods or means of warfare is not unlimited; weapons causing superfluous injury or unnecessary suffering are prohibited – basic rules are set out in Article 35.

Customary International Humanitarian Law

Most of the rules on the conduct of hostilities in AP I are considered to be reflective of customary international law, and therefore are binding on all states, including those that have not ratified the Protocol. In its study on customary law, the ICRC found that, among other rules, Rule 1 (distinction between civilians and combatants), Rule 7 (distinction between civilian objects and military objectives), Rule 11 and 12 (indiscriminate attacks), Rule 14 (proportionality), and Rules 15 to 21 (precautions in attack) are customary in both international and non-international armed conflicts (Henckaerts & Doswald-Beck, 2005). While some States, notably the United States, have questioned aspects of the methodology of the study, there is little real dispute that the core rules of distinction, proportionality and precautions are customary. This is important because the legal analysis in the next section applies to all states that may develop or use AWS, not just parties to AP I.S

The Obligation to Review Weapons Under Article 36

AP I Article 36 reads:

> A High Contracting Party in the study, development, acquisition or adoption of a new weapon, means or method of warfare is under an obligation to determine whether its employment would,

> in some or all circumstances, be prohibited by this Protocol or by any other rule of international law applicable to the High Contracting Party.

This is probably the most important existing instrument for the regulation of AWS at the national level, and is often cited by states opposed to a new treaty as evidence that existing law is sufficient. But I want to emphasise some limitations.

First, it remains uncertain whether the obligation under Article 36 constitutes customary law. According to the ICRC's guide, all states have an interest in conducting reviews, and the obligation stems from the general duty to respect and ensure respect for IHL (ICRC, 2006). Having considered state practice, Jevglevskaja (2018) concluded that a customary obligation to hold formal weapons reviews has not developed clearly, considering that states are bound by the substantive rules that a review is meant to verify. Second, formal review procedures are known to exist in few states, with estimates often putting the figure at less than twenty (Boulanin & Verbruggen, 2017). Third, reviews are done by the states themselves and are seldom published, meaning there is no international oversight or common standard.

Fourth, and most important for AWS, Article 36 reviews were designed for weapons whose effects could be tested and predicted prior to use. This model presents a challenge for a machine learning system that learns after deployment or that behaves differently in environments that are different from its training data. Is the system behaviour changed? If so, is it still the same weapon which was reviewed? Copeland et al (2023) argue that reviews can still have a useful role, but only if they become continuous rather than one-off, and if they look not just at the weapon itself but also at the intended context and the training and procedures of the people who use it. I think this is correct, but it also shows how much the review process would have to change and Article 36 itself says nothing about how this should happen.

The Martens' Clause

The Martens Clause - in its modern form in Article 1(2) of AP I states that in cases not covered by treaties, civilians and combatants remain under the protection of the principles of international law derived from established custom, from the principles of humanity and from the dictates of public conscience. The International Court of Justice also confirmed in the Nuclear Weapons advisory opinion that the clause "has proved to be an effective means of addressing the rapid evolution of military technology" (Legality of the Threat or Use of Nuclear Weapons, 1996, para. 78).

Campaigners have heavily relied on the Martens Clause, claiming that the transfer of life and death choices to machines violates the standards of humanity and the dictates of public conscience (Human Rights Watch, 2012; Asaro, 2012). The ICRC has also mentioned it as a reason for ethical considerations to guide the regulation of AWS (ICRC, 2021). But it is unclear whether the clause has any legal force. Most scholars interpret it as an interpretive guide, or reminder, that the absence of a specific prohibition does not mean that something is permitted rather than an independent source of prohibitions (Sassòli, 2014). I thus consider the Martens Clause in this article to be relevant to the policy question of whether a new instrument is needed, but not an existing rule that alone precludes AWS.

International Law of Human Rights

While this article focuses on IHL, it is essential to keep in mind that international human rights law continues to apply in armed conflict, and fully applies when the use of force outside of armed conflict, such as in law enforcement. In 2013, the former Special Rapporteur on extrajudicial, summary or arbitrary

executions, Christof Heyns, argued that robots should not be granted the power of life and death over human beings and called for national moratoria (Heyns, 2013). Specific concerns arise for the right to life and the right to dignity when AWS are used in policing or border control. These issues are beyond the scope of this article, but are part of the reason why many states want any new instrument to include AWS "in all circumstances", not just in armed conflict.

Distinction, Proportionality and Cautions

This part makes the essence of the doctrinal analysis. For each principle, I ask three questions: what does the rule require; can an autonomous system comply with it, at least in some circumstances; and what kind of human involvement is needed to ensure compliance? There is one introductory point to be made. IHL obligations are owed to States and to persons, not to weapons. A machine cannot 'comply' with IHL in the legal sense, only the humans who design, deploy and use it can (Davison, 2017; ICRC, 2021). So when I ask whether an AWS can comply with distinction, what I really mean is whether a commander can lawfully use that system in a particular context and have reasonable assurance that the attack will comply with the rule.

The Rule of Division

The principle of distinction, as enshrined in Article 48 AP I and Rules 1 and 7 of the customary study, obliges parties to a conflict to distinguish at all times between civilians and combatants and between civilian objects and military objectives, and to direct their operations only against military objectives. Article 51(4) prohibits indiscriminate attacks, including attacks which employ a method or means of combat the effects of which cannot be limited to a specific military objective.

For certain autonomous systems, distinction is not especially problematic. A system to detect and destroy incoming anti-ship missiles, working at sea where there are few civilians, aims at objects that are military in nature and have very distinct signatures, such as speed and radar profile. And the same applies to systems designed to attack tanks or military radars in remote areas. As Schmitt and Thurnher (2013) correctly pointed out, there is no basis in law or technology for a blanket assumption that AWS cannot make the distinction. The legality of a weapon is determined by whether it can be used in a discriminate way in some circumstances and not in all circumstances.

The problem is when AWS are used against people or in populated, dynamic environments. It's not as easy as looking for a uniform or a weapon to tell a combatant from a civilian. Civilians are only deprived of protection under Article 51(3) of AP I "for such time as they take a direct part in hostilities", and whether someone is directly participating in hostilities is a matter of assessing their behaviour and intentions in context (Sharkey, 2012). A man with a rifle can be a policeman, a rancher guarding his cows, or a soldier. A wounded soldier, unconscious or trying to surrender, is hors de combat under Article 41 and must not be attacked. There are subtle cues to recognising surrender, in particular some that even trained humans miss. I think it is quite unlikely that present machine learning systems can make these judgements reliably. This is further complicated by the fact that adversaries will deliberately try to fool them, for example by spoofing sensors.

Another problem is doubt. AP I, Article 50(1) provides that in case of doubt whether a person is a civilian, that person shall be considered a civilian, and AP I, Article 52(3) contains a similar presumption for objects normally dedicated to civilian purposes. Humans, at least in theory, can tell when they are uncertain. Machine learning classifiers generate confidence scores, but these can be poorly calibrated, and

a system can be very “confident” and yet completely wrong, especially when faced with inputs outside of its training data. This means that the legal presumption in favour of civilian status is not easily translatable into code. I think this is one of the strongest arguments for insisting that systems used against people be kept under tight human control.

I conclude that the principle of distinction does not prohibit AWS as such, but severely limits the conditions under which they can be lawfully used. In practice compliance is a function of the limitation of target types (e.g. to objects that are military by nature), the environment (e.g. to areas free of civilians) and the duration and geographical scope of operation. These are just the types of controls identified by Boulanin et al. (2020).

Principle of Proportionality

Article 51(5)(b) AP I sets out the proportionality rule prohibiting 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”. It is repeated as a precautionary obligation in article 57(2)(a)(iii) and (b), and is recognised as custom in Rule 14 of the ICRC study (Henckaerts & Doswald-Beck, 2005).

Proportionality is generally considered to be the most difficult rule for AWS. The assessment is in two parts. The first part, estimating expected civilian harm, is in some ways amenable to computation. Militaries already have collateral damage estimation methodologies that use software to estimate blast effects and the number of persons likely to be in a building at a given time (Schmitt, 2013). One could imagine that an AWS could do this part of the analysis at least as well as the human.

The second part is far more difficult. The rule requires a balance to be struck between civilian harm and military advantage, two completely different values which have no common unit of measurement. The military advantage of destroying a particular target is not fixed, but is dependent on the overall plan of operations, the current situation on the battlefield, and the commander's intent, all of which are subject to change very quickly. The standard that has emerged in practice is that of the “reasonable military commander”, as shown by the statements made by several States upon ratification of AP I and the jurisprudence of the International Criminal Tribunal for the former Yugoslavia (Sassòli, 2014). So proportionality is a good faith value judgement by a human being in context.

Some authors argue that an AWS could be programd with pre-set thresholds, for example that it may not attack a target if more than a certain number of civilians are detected nearby, and that these thresholds would be set by commanders who make the proportionality judgement in advance (Schmitt & Thurnher, 2013). This is a reasonable argument and I concede it may work for short narrowly bounded missions. But it suffers from a defect that increases with the length and breadth of operation. A proportionality judgement at the time of activation is valid only as long as circumstances stay roughly the same. If a system can loiter over an area for hours or days the military advantage of attacking a particular target can change, civilians can move into the area and the original judgement may no longer be reasonable. Article 57(2)(b) explicitly requires that an attack be aborted or suspended if it appears it would be disproportionate, which presupposes that someone is able to notice and act. In my view this is the strongest legal basis for arguing that human control must continue throughout the operation, not just at the moment of activation.

Precautions against attacks

AP I, Article 57, sets out a series of obligations to take precautions in attack. Those who plan or decide upon an attack must: • Take all feasible measures to verify that targets are military objectives; • Take all feasible precautions in the choice of means and methods of attack with a view to avoiding, and in any event to minimising, incidental civilian harm; and • Refrain from attacks which would be disproportionate. As indicated above, an attack must also be cancelled or suspended when it becomes apparent that the objective is not a military one or that the attack would be disproportionate. Where possible, advance notice should be given and should be effective. Article 57 (1) provides a general obligation to take constant care to spare the civilian population.

These precautions cut both ways for AWS and this is a point I think is sometimes missed in the campaigning literature. Some writers have argued that the obligation to select means and methods that minimise civilian harm could, in some circumstances, require the use of more precise autonomous systems if they are genuinely better at avoiding civilian casualties than the alternatives (Schmitt, 2013; Jensen, 2020). The duty to cancel or suspend attacks and the duty of continuous care, in contrast, both presuppose an ongoing human capacity to supervise and intervene. (Henckaerts & Doswald-Beck, 2005, Rule 15) The term ‘feasible’ is understood to mean ‘what is practicable or practically possible, having regard to all the circumstances at the time including humanitarian and military considerations’. It could be argued in defence of a state's choice of a weapon that cannot be cancelled that the state has deliberately made precautions impossible. In my opinion, this is not a persuasive reading of the law. A state cannot choose its own technology to escape its obligations.

Summary of the theological study

Combining these three principles, I come up with a conclusion that is somewhere between the two main camps. No general ban on autonomy in the critical functions of weapons exists in the existing IHL. There are some situations where AWS could be lawfully used under current rules, for example defensive systems against missiles or attacks on purely military objects in remote areas. However the three principles combined imply requirements for human involvement. The user must know the system well enough to forecast its effects, must restrict its operation in time, space and target type so that the judgements made at activation remain valid, and must retain the ability to intervene when circumstances change. If the system is unpredictable it cannot meet these requirements and is not lawful to use. The problem is that these requirements are implicit, scattered through different provisions, and subject to very different interpretations by different states. Here's the first gap.

Individual and State Responsibility

A second important question is who is responsible when things go wrong, even if the primary rules can be satisfactorily applied to AWS. The first to express the concern about a “responsibility gap” or “accountability gap” in the philosophical literature was Sparrow (2007). According to Sparrow, it would be unfair to hold either the programmer, the commander or the machine responsible for an unlawful killing by a truly autonomous system and this made the use of such systems unethical. Human Rights Watch later developed this argument in a legal version (Human Rights Watch, 2015). In this section I consider whether such a gap really exists under the rules of individual criminal responsibility and state responsibility.

Criminal Responsibility of Individuals

The Rome Statute criminalises the deliberate targeting of attacks against the civilian population, or against individual civilians not taking direct part in hostilities, in an international armed conflict as a war crime (Article 8(2)(b)(i)). It further criminalises the deliberate launching of an attack with knowledge that it will cause clearly excessive incidental civilian harm as a war crime (Article 8(2)(b)(iv)). The general mental element is set out in Article 30: unless otherwise provided, a person is criminally responsible only if the material elements are committed with intent and knowledge.

The issue for AWS is the mental side of things. If a commander uses an AWS in good faith believing that the system will only attack military vehicles, and the system attacks a bus full of civilians because of a malfunction or misclassification, the commander did not intend to attack civilians. Under the Rome Statute, recklessness and negligence generally do not suffice to constitute the war crime of attacking civilians, although the jurisprudence of the International Criminal Tribunal for the former Yugoslavia accepted a form of recklessness (dolus eventualis) for some crimes (Bo, 2021). The result is that in many realistic situations no one will meet the mental element. Bo (2019) has argued that the gap is smaller than is sometimes claimed, because a commander who knowingly uses an unpredictable system in a populated area may well have the requisite awareness that civilian deaths will occur "in the ordinary course of events". This is persuasive in the obvious cases, but it does not help in the hard cases where the commander had reason to trust the system.

The problem is also not solved by command responsibility under article 28 of the Rome Statute. The doctrine encompasses the failure of a superior to prevent or punish crimes of subordinates. A machine is not a 'subordinate' and it cannot commit a crime as it has no mental state. The extension of command responsibility by analogy has been proposed by some authors, but this is contrary to the principle of legality in criminal law (Chengeta, 2016).

And the manufacturers and programmers? Theoretically, they could be liable as accessories if they knowingly designed a system to commit war crimes. In practice, programming is usually done well in advance of any armed conflict, often by large teams in private companies, and it would be very difficult to show that any individual developer had the required intent and knowledge in relation to a specific attack (Crootof, 2016). There is also the question of whether the IHL war crimes framework applies to conduct that occurred before the armed conflict.

What I conclude is that there is a real, but sometimes over-wrought, gap in individual criminal responsibility. It's not a problem when a human uses an AWS on purpose to commit a crime, because then the AWS is just a tool. It occurs in the much more frequent scenario of an unintentional illegal outcome as a consequence of the unpredictable action of the system, where no human had the necessary mental state. This is a direct result of reduced human control: The more the human is removed from the particular decision, the more difficult it is to assign criminal responsibility to that human.

State Responsibilities:

The better way is the road of state responsibility. An internationally wrongful act under article 2 of ARSIWA is an act attributable to a state that constitutes a breach of an international obligation of that state (International Law Commission, 2001). The use of an AWS by a state's armed forces is clearly attributable to the state under Article 4 of ARSIWA. This is because the armed forces are organs of the state. Further, Article 91 of AP I provides that a party which violates the Conventions or the Protocol shall

be liable to pay compensation and shall be responsible for all acts committed by persons forming part of its armed forces. Unlike criminal responsibility, state responsibility does not usually require fault or intent, unless the primary rule itself requires it (Crootof, 2016).

But there are complications. Some of the major rules, like the prohibition on targeting civilians, probably have a mental component to them, since you have to "target" an attack at somebody. And if an AWS malfunctions and kills civilians, it is not clear that the attack was ‘directed’ at civilians by the state. The state may, however, still be liable for failing to take feasible precautions under Article 57 or for failing to conduct an adequate weapons review under Article 36. In my view, the most useful hook for state responsibility in AWS cases is the precautions obligation, which focuses on what the state did or did not do before and during the attack, rather than the intent behind the specific engagement.

The more serious practical problem is enforcement. Courts rarely hold states responsible for violations of IHL, there is no established mechanism for individual victims to seek damages from a foreign state and the facts about how an AWS was programd and tested are likely to be classified. A new paradigm of "war torts" has been proposed by Crootof (2016) which would hold states strictly liable for injuries caused by AWS regardless of whether IHL was violated. This is an interesting suggestion and I come back to it in the findings section.

Table 1. Responsibility for Damage caused by Autonomous Weapon Systems

Form of accountability	Legal basis	Main obstacle in AWS cases	Possible response
Individual criminal responsibility (direct perpetrator)	Rome Statute arts. 8 and 30; customary war crimes law.	Intent and knowledge often lacking where harm results from unpredictable system behaviour	Clarify that knowingly deploying unpredictable AWS in populated areas satisfies the mental element
Command responsibility	Rome Statute art. 28; AP I arts. 86–87	Applies to crimes of subordinates A machine is not a subordinate No mental state	Duties of commanders to understand, test and supervise AWS made explicit
2. Liability of developers or manufacturers	2.1. Liability for accessories; domestic liability for products	2.2. Remoteness in time and place; large teams; proof of intent; conduct may precede the conflict	2.3. Design and testing obligations; domestic civil liability
ARSIWA arts. Responsibility of States	2 and 4; AP I art. 91	Rules with mental element; weak enforcement; classified evidence	Strict liability for harm caused by AWS (“war torts”); duties of record keeping and investigation
Weapons review	AP I art. 36	Not clearly customary; few states review; no transparency; one-off model	Ongoing review for learning systems; publication of procedures

Note. AWS = autonomous weapon systems . AP I = Additional Protocol I . ARSIWA = Articles on Responsibility of States for Internationally Wrongful Acts . Author’s compilation from Bo (2021), Chengeta (2016), Crootof (2016), Human Rights Watch (2015) and Jevglevskaja (2018).

The main forms of accountability are summarised in Table 1, together with the obstacles when applied to AWS and possible responses. Overall, the picture is that the existing frameworks do not leave a complete

vacuum, but they leave significant gaps precisely in the cases that are most likely to be encountered in practice: unintended harm caused by unpredictable system behaviour.

Proposals for international regulation

The Group of Governmental Experts on the CCW

The CCW has used AWS as the main forum for intergovernmental discussion. Informal expert meetings began in 2014 and in 2016 the Fifth Review Conference established a Group of Governmental Experts on emerging technologies in the area of lethal autonomous weapons systems (LAWS), which has been meeting annually since 2017. The CCW was selected in part because it is supposed to be flexible: it is a framework convention to which protocols on particular weapons can be attached, as was the case with Protocol IV on blinding laser weapons in 1995, a rare example of a weapon being banned before it was widely deployed.

The most evident concrete result achieved so far is the set of eleven Guiding Principles adopted by consensus in 2019 (CCW GGE, 2019, Annex IV). These principles are: that IHL continues to apply fully to all weapon systems including LAWS (principle a); that human responsibility for decisions on the use of weapon systems must be retained as accountability cannot be transferred to machines (principle b); that human-machine interaction should ensure that the potential use of weapon systems is in compliance with international law (principle c); that accountability must be ensured in accordance with applicable international law including through a responsible chain of human command and control (principle d); and that legal reviews under Article 36 are an obligation for parties to AP I (principle e). The principles are useful as a consensus-building tool, but they are not legally binding and are written in broad terms. In particular, they do not even use the term “meaningful human control”.

In November 2023, the CCW High Contracting Parties mandated the GGE to “further consider and formulate, by consensus, a set of elements of an instrument, without prejudging its nature” with a view to finalising its work before the Seventh Review Conference in 2026. In 2024 and 2025, the Chair developed a “rolling text” of draft provisions on characterisation, prohibitions and regulations, and accountability (CCW GGE, 2024). The rolling text reflects what is often called the two-tier approach, i.e. a prohibition on AWS that cannot be used in compliance with IHL, combined with regulation (limits and requirements) for other AWS. The phrase “without prejudging its nature” is a diplomatic way of saying that states still disagree on whether the outcome should be a binding protocol or a political declaration.

The problem of CCW process is the consensus rule. Any one state can block an outcome, and in practice the Russian Federation, and sometimes other states, have blocked the GGE from adopting stronger language. Hence, many observers have concluded that the CCW will probably not be able to produce a binding instrument by itself (Rosert & Sauer, 2021).

Recommendations of the ICRC

The ICRC issued its position on AWS in May 2021, which has become a key reference point. The ICRC recommended three elements of new legally binding rules for States to adopt (ICRC, 2021). First, unpredictable autonomous weapon systems should be expressly ruled out, not least because of their indiscriminate effects; this would best be achieved by prohibiting systems that are designed or used in such a way that their effects cannot be sufficiently understood, predicted and explained. Secondly, the use of AWS against human beings must be ruled out due to ethical concerns for humanity protection. Third,

all other AWS should be designed and used subject to a combination of limits on the types of target, limits on the duration, geographical scope and scale of use, limits on situations of use, and requirements for human-machine interaction, especially to ensure effective human supervision and timely intervention and deactivation.

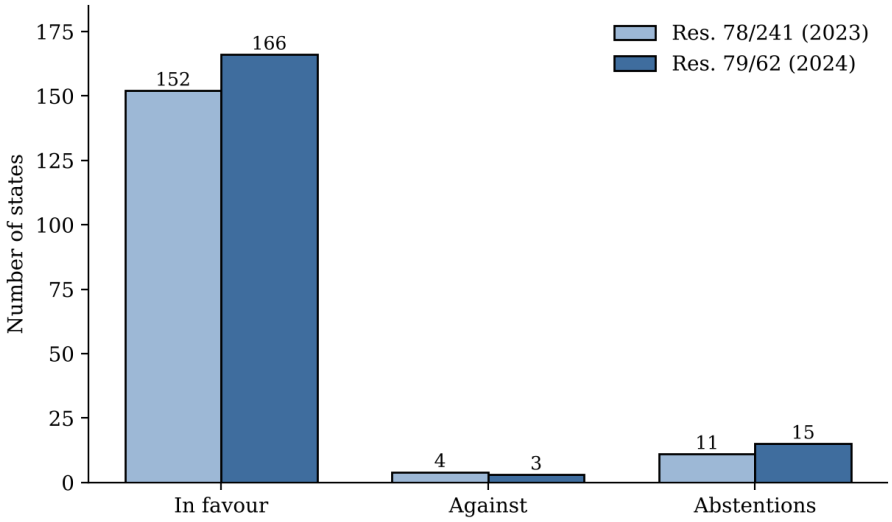
The noteworthy point about the ICRC position is that it does not assert that all AWS are already prohibited under existing IHL. Instead, it argues that the current rules are insufficient to ensure clarity and that new rules are required to address humanitarian, legal and ethical concerns. This is a powerful statement from the body that is generally perceived as the guardian of IHL and has shaped the position of many states.

UN General Assembly & Other Initiatives

Frustrated with the CCW, a group of states led by Austria took the issue to the UN General Assembly. The General Assembly adopted Resolution 78/241 in December 2023, the first resolution dedicated specifically to LAWS, in which it expressed concern about the potential negative consequences of AWS and requested the Secretary-General to seek the views of states and other stakeholders (United Nations General Assembly, 2023). The report of the Secretary-General, issued in July 2024, consolidated submissions from a large number of states and organisations and reiterated his call for a legally binding instrument by 2026 (United Nations Secretary-General, 2024). In December 2024 the General Assembly adopted Resolution 79/62, deciding to hold open informal consultations in New York in 2025 to consider the report, in addition to, rather than instead of, the CCW process (United Nations General Assembly, 2024). The consultations were held in May 2025 and were open to all states and civil society, which is a key difference from the CCW.

The voting results of these two resolutions are shown in Fig. 2. "Between 2023 and 2024, support increased and the number of states voting against was very small. "There is a wide political majority now to at least take the question of a new instrument seriously.

Figure 2. Voting on UN General Assembly Resolutions on Lethal Autonomous Weapons Systems, 2023-2024



Note. Information on the voting results of United Nations General Assembly Resolution 78/241 (22 December 2023) and Resolution 79/62 (2 December 2024) (2023, 2024).

In parallel with the UN processes, a number of regional conferences have taken place. The Belén Communiqué, adopted at the Latin American and Caribbean Conference on the Social and Humanitarian Impact of Autonomous Weapons, which took place in Costa Rica in February 2023, called for the urgent negotiation of a legally binding instrument with prohibitions and regulations. In April 2024, Austria hosted the conference "Humanity at the Crossroads" in Vienna, concluding with a Chair's summary calling for the same.

At the same time, the United States has promoted a different track. In November 2023, the United States published the Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy, which by the end of the year had been supported by about fifty states (United States Department of State, 2023). The Declaration is not binding, and deals with military AI generally, not just weapons. It includes commitments to legal reviews, to training personnel on the capabilities and limitations of standard artificial intelligence systems, and to ensuring systems have explicit, well-defined uses and are subject to appropriate testing. These commitments are not worthless, but the Declaration deliberately excludes any prohibitions, and does not use the concept of meaningful human control.

Comparative Assessment of State Stances

The positions of states as given in the sources above can be roughly classified into four clusters. These are presented in Table 2. It is important to stress that this is a simplification. States are not fixed and states within each group differ on important details. The clustering does, however, help to point to where the main lines of disagreement lie.

Table 2:Clusters of State Positions on the Regulation of Autonomous Weapons Systems

Cluster	Illustrative countries	Core stance	Desired outcome
Legally binding prohibition and regulation	Austria, Brazil, Chile, Costa Rica, Mexico, New Zealand, Pakistan, Philippines, Sierra Leone, many states of the Non-Aligned Movement and African Group	Existing IHL insufficient; meaningful human control should be a legal requirement; ethical concerns (Martens Clause) central	New treaty with prohibitions and regulations, negotiated in the CCW or other forum
Two-tier approach, open to binding or political outcome	France, Germany, Netherlands, Switzerland, Belgium, Ireland, Norway	Some AWS already unlawful; others require specific limits; human control or judgement required	Two-tier instrument; flexible as to binding nature; CCW as main forum
Existing law sufficient, and non-binding guidance	United States, United Kingdom, Australia, Japan, Republic of Korea, Israel	IHL sufficient if applied through weapons reviews and good practice; "appropriate human judgment"	Non-binding principles and good practices; Political Declaration (2023)
New instrument or narrow approach	Russian Federation, India; China (partial)	Too soon to regulate; definitions are unclear; CCW only legitimate forum. China supports ban on use of narrowly defined "unacceptable" LAWS, but not on development	Continued discussion in CCW by consensus; no new binding treaty at present

Note. 8. Positions have changed over time . Groupings are simplified. CCW: Convention on Certain

Conventional Weapons; LAWS: lethal autonomous weapons systems. The author compiled the data from the CCW GGE (2019, 2024), Human Rights Watch (2020), Rosert and Sauer (2021), United Nations General Assembly (2023, 2024), United Nations Secretary General (2024) and United States Department of State (2023).

This comparison allows us to see a number of things. First, there is very little or no opposition (except in a few states) to the application of IHL to AWS, and very little support for the view that AWS should be unregulated. The real debate is between those who say existing law plus national implementation and non-binding guidance is enough and those who want a binding instrument. Second, the states least supportive of a binding instrument are also largely those states investing heavily in military AI. Third, the two-tier approach has emerged as a possible middle ground, because it is compatible with the view that some AWS are already unlawful under the existing IHL (the prohibitions) and also creates new, more specific obligations (the regulations). Fourth, the disagreement over terminology, "meaningful human control" versus "appropriate levels of human judgment" or "context-appropriate human involvement", reflects a deeper disagreement over whether human involvement should be a free-standing legal requirement or merely a means to ensure compliance with existing rules.

It is also worth mentioning that many states in the Global South, including Pakistan, who was one of the first states to call for a ban on fully autonomous weapons in the CCW, are worried that AWS will lower the threshold for the use of force and will be used primarily by technologically advanced states against less advanced ones (Human Rights Watch, 2020). This concern is not a purely legal argument about IHL, but it does explain in part why the debate is so polarised, and why many states have been unsatisfied with the mere statement that 'existing law is sufficient'.

Results and Conclusions

Addressing the Research Question

The main research question of this article was whether existing IHL rules are sufficient to regulate autonomous weapon systems or whether a new international legal instrument is needed. Based on the aforementioned analysis, my answer is that the existing IHL is applicable, partly adequate, but not sufficient. Summary of main findings is given below.

Finding 1: IHL applies and already prohibits certain AWS. AWS are fully subject to the rules of distinction, proportionality and precautions. Read together, they already prohibit the use of autonomous systems whose effects cannot be predicted, and they prohibit the use of AWS in circumstances where the system cannot reliably discriminate between lawful and unlawful targets. In this sense are mistaken those who say there is a legal vacuum.

Finding 2: Implicit, not explicit, requirement of human involvement. The analysis showed that these three principles together implied the assumption of a human who understands the system, limits its use and can intervene when things change. But nowhere in the treaties is this requirement mentioned. It must be interpreted, and states interpret it in very different ways, as illustrated by the gap between "meaningful human control" and "appropriate levels of human judgment." Such a fundamental question left to unilateral interpretation creates legal uncertainty and a risk of a race to the bottom.

Finding 3: Article 36 reviews are necessary, but not sufficient. The primary existing mechanism for verifying the legality of AWS before use is a weapons review, but the obligation is only binding as treaty law for parties to AP I, its customary status is uncertain, few states conduct formal reviews, there is no common standard or transparency, and the one-off review model is ill-suited for systems that learn or are updated after deployment.

Finding 4: An accountability gap exists around unintended harm. Intent and knowledge are required for individual criminal responsibility, and these are often absent when the harm comes from the unpredictable behaviour of a system. There is no command responsibility for machines. State responsibility is in principle but weakly enforced. May have trouble with rules that require a mental element. The gap is less than some campaigners say, but it is real and precisely where AWS is most likely to do harm.

Finding 5: Political support for a new instrument is growing but the forum problem remains. The voting trends in the General Assembly , the ICRC recommendations and the emergence of the two-tier approach in the CCW rolling text point to a broad convergence on the content of possible rules . The main obstacle is political, not legal: the consensus rule in the CCW and the opposition of a number of major military powers.

A Framework for Assessing Autonomous Weapons

The expected contribution of this article is a legal framework to evaluate AWS and to identify where additional safeguards might be needed. The results lead me to propose that all AWS should be evaluated by the following five questions, in the following order. I believe that this framework is useful for today's national reviewers and as a possible basis for the content of a future instrument.

1. Predictability test. Can the people that will use the system in the intended context understand, predict and explain the effects of the system sufficiently? Otherwise the system cannot be operated in accordance with differentiation and proportionality and should be prohibited.

2. Test target. Is the system designed or used to target humans? If so, the difficulties of establishing civilian status, direct participation in hostilities and surrender, together with the ethical concerns articulated in the Martens Clause, justify a prohibition or at least a requirement of direct human authorisation for each engagement.

3. Boundary tests. Are the target types, duration, geographical area and scale of operation sufficiently limited so that the proportionality and distinction judgements made at activation still hold good over the course of the operation? If not, the operational parameters should be narrowed.

4. intervention trial Is there a human in the loop who can supervise the system and intervene to cancel or suspend an attack in time, as required by Article 57(2)(b)? Otherwise, the system must be restricted to applications in situations where the circumstances are very unlikely to change, such as defensive use against incoming ordnance.

5. Test for traceability. Is each use of force attributable to an identifiable human decision, with records kept that can later be investigated, identifying state and individual responsibility? If not, additional record keeping and reporting requirements should be mandated.

The first two tests are linked to the "prohibition" level and the last three to the "regulation" level in the two-tier approach. Most of these tests can be derived from existing IHL by interpretation. This proves that a new instrument would not be creating new law. Its main value would be to put the implicit requirements on the table, to harmonise interpretation across states and to add the elements that are clearly missing in existing law, especially on transparency, continuous review and accountability.

Suggestions

Based on this analysis, I would make the following recommendations. First, a legally binding instrument based on the two-tier approach should be negotiated by states. The best way forward would be within the CCW, which includes the principal military powers, but if the consensus rule keeps progress blocked, states should be ready to negotiate in the General Assembly or in a separate process, as was done for the Ottawa Convention and the Convention on Cluster Munitions. Secondly, states should strengthen their Article 36 reviews, regardless of what happens at the international level, making them continuous in learning systems, sharing best practice, and publishing at least the procedures they follow, if not the results. Third, the accountability gap should be addressed by establishing a clear rule that states are strictly liable for harm caused by their AWS, akin to Crootof's (2016) war torts proposal, and by coupling this with obligations to keep records and investigate incidents. Fourth, more attention should be paid to AI decision support systems in the debate, which may not fit the definition of AWS but raise very similar concerns around the quality of human control.

Conclusion

When I began this research I thought that the existing law was either clearly enough adequate or clearly enough in need of a ban. The answer was more complicated. The existing IHL is not silent on autonomous weapons. Its fundamental values already require a level of human understanding, judgement and supervision that prevents the most dangerous forms of autonomy. But the law speaks only indirectly, in rules written decades before anyone could imagine a weapon that might decide for itself whom to attack. This indirectness allows for very different interpretations and the weapons review system that is supposed to fill the gap is opaque, uneven and poorly suited to systems that learn. The law is weakest in accountability, especially on the cases that matter most.

For these reasons I conclude that a new international legal instrument is necessary, not because existing IHL is irrelevant, but because it needs to be clarified and supplemented. The human element in the use of force should not be left to inference, it should be put down on paper. As the 2019 Guiding Principles state, accountability "cannot be transferred to machines" (CCW GGE 2019, Annex IV). Now it is up to the states to work out binding rules to translate that principle into, before the technology gets further ahead of the law.

This article could be built upon in several ways in future research. It would be useful to further investigate how specific states conduct their Article 36 reviews of AI-enabled systems, to revisit the use of AI decision support systems in recent conflicts as more reliable information becomes available, and to analyse how a future instrument could be verified ., which is a difficult question that I have not addressed here.

References

1. Abraham, Y. (2024, April 3). 'Lavender': The AI machine directing Israel's bombing spree in Gaza. *+972 Magazine*. https://www.972mag.com/lavender-ai-israeli-army-gaza/
2. Asaro, P. (2012). On banning autonomous weapon systems: Human rights, automation, and the dehumanization of lethal decision-making. *International Review of the Red Cross, 94*(886), 687–709.
3. Bo, M. (2021). Autonomous weapons and the responsibility gap in light of the mens rea of the war crime of attacking civilians in the ICC Statute. *Journal of International Criminal Justice, 19*(2), 275–299.
4. Boothby, W. H. (2016). *Weapons and the law of armed conflict* (2nd ed.). Oxford University Press.
5. Boulanin, V., Davison, N., Goussac, N., & Peldán Carlsson, M. (2020). *Limits on autonomy in weapon systems: Identifying practical elements of human control*. Stockholm International Peace Research Institute & International Committee of the Red Cross.
6. Boulanin, V., & Verbruggen, M. (2017). *Mapping the development of autonomy in weapon systems*. Stockholm International Peace Research Institute.
7. CCW Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems. (2019). *Report of the 2019 session of the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems* (CCW/GGE.1/2019/3). United Nations.
8. CCW Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems. (2024). *Rolling text, status date: 6 September 2024* [Chair's working document]. United Nations Office for Disarmament Affairs.
9. Chengeta, T. (2016). Accountability gap: Autonomous weapon systems and modes of responsibility in international law. *Denver Journal of International Law and Policy, 45*(1), 1–50.
10. 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, October 10, 1980, 1342 U.N.T.S. 137.
11. Copeland, D., Liivoja, R., & Sanders, L. (2023). The utility of weapons reviews in addressing concerns raised by autonomous weapon systems. *Journal of Conflict and Security Law, 28*(2), 285–316.
12. Crootof, R. (2015). The killer robots are here: Legal and policy implications. *Cardozo Law Review, 36*(5), 1837–1915.
13. Crootof, R. (2016). War torts: Accountability for autonomous weapons. *University of Pennsylvania Law Review, 164*(6), 1347–1402.
14. Davison, N. (2017). A legal perspective: Autonomous weapon systems under international humanitarian law. In *Perspectives on lethal autonomous weapon systems* (UNODA Occasional Papers No. 30, pp. 5–18). United Nations.
15. Geneva Convention Relative to the Protection of Civilian Persons in Time of War, August 12, 1949, 75 U.N.T.S. 287.
16. Henckaerts, J.-M., & Doswald-Beck, L. (2005). *Customary international humanitarian law: Volume I, Rules*. Cambridge University Press.
17. Heyns, C. (2013). *Report of the Special Rapporteur on extrajudicial, summary or arbitrary executions* (A/HRC/23/47). United Nations Human Rights Council.
18. Human Rights Watch. (2012). *Losing humanity: The case against killer robots*. Human Rights Watch & International Human Rights Clinic, Harvard Law School.
19. Human Rights Watch. (2015). *Mind the gap: The lack of accountability for killer robots*. Human

Rights Watch & International Human Rights Clinic, Harvard Law School.
20. Human Rights Watch. (2020). *Stopping killer robots: Country positions on banning fully autonomous weapons and retaining human control*.
21. Hutchinson, T., & Duncan, N. (2012). Defining and describing what we do: Doctrinal legal research. *Deakin Law Review, 17*(1), 83–119.
22. International Committee of the Red Cross. (2006). A guide to the legal review of new weapons, means and methods of warfare: Measures to implement Article 36 of Additional Protocol I of 1977. *International Review of the Red Cross, 88*(864), 931–956.
23. International Committee of the Red Cross. (2021). *ICRC position on autonomous weapon systems*.
24. International Law Commission. (2001). *Responsibility of states for internationally wrongful acts* (Yearbook of the International Law Commission, Vol. II, Part Two). United Nations.
25. Jensen, E. T. (2020). The (erroneous) requirement for human judgment (and error) in the law of armed conflict. *International Law Studies, 96*, 26–57.
26. Jevglevskaja, N. (2018). Weapons review obligation under customary international law. *International Law Studies, 94*, 186–221.
27. *Legality of the Threat or Use of Nuclear Weapons*, Advisory Opinion, 1996 I.C.J. 226 (July 8).
28. McFarland, T. (2020). *Autonomous weapon systems and the law of armed conflict: Compatibility with international humanitarian law*. Cambridge University Press.
29. Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of International Armed Conflicts (Protocol I), June 8, 1977, 1125 U.N.T.S. 3.
30. Roff, H. M., & Moyes, R. (2016). *Meaningful human control, artificial intelligence and autonomous weapons* [Briefing paper]. Article 36.
31. Rome Statute of the International Criminal Court, July 17, 1998, 2187 U.N.T.S. 90.
32. Rosert, E., & Sauer, F. (2021). How (not) to stop the killer robots: A comparative analysis of humanitarian disarmament campaign strategies. *Contemporary Security Policy, 42*(1), 4–29.
33. Santoni de Sio, F., & van den Hoven, J. (2018). Meaningful human control over autonomous systems: A philosophical account. *Frontiers in Robotics and AI, 5*, Article 15. https://doi.org/10.3389/frobt.2018.00015
34. Sassòli, M. (2014). Autonomous weapons and international humanitarian law: Advantages, open technical questions and legal issues to be clarified. *International Law Studies, 90*, 308–340.
35. Scharre, P. (2018). *Army of none: Autonomous weapons and the future of war*. W. W. Norton.
36. Schmitt, M. N. (2013). Autonomous weapon systems and international humanitarian law: A reply to the critics. *Harvard National Security Journal Features*.
37. Schmitt, M. N., & Thurnher, J. S. (2013). "Out of the loop": Autonomous weapon systems and the law of armed conflict. *Harvard National Security Journal, 4*, 231–281.
38. Sharkey, N. E. (2012). The evitability of autonomous robot warfare. *International Review of the Red Cross, 94*(886), 787–799.
39. Sparrow, R. (2007). Killer robots. *Journal of Applied Philosophy, 24*(1), 62–77. https://doi.org/10.1111/j.1468-5930.2007.00346.x
40. UK Ministry of Defence. (2022). *Ambitious, safe, responsible: Our approach to the delivery of AI-enabled capability in Defence*.
41. United Nations General Assembly. (2023). *Lethal autonomous weapons systems* (A/RES/78/241).
42. United Nations General Assembly. (2024). *Lethal autonomous weapons systems* (A/RES/79/62).
43. United Nations Secretary-General. (2024). *Lethal autonomous weapons systems: Report of the Secretary-General* (A/79/88).
44. United Nations Secretary-General & International Committee of the Red Cross. (2023, October 5). *Joint call by the United Nations Secretary-General and the President of the International*

Committee of the Red Cross for states to establish new prohibitions and restrictions on autonomous weapon systems [Press release].
45. United Nations Security Council. (2021). *Final report of the Panel of Experts on Libya established pursuant to Security Council resolution 1973 (2011)* (S/2021/229).
46. United States Department of Defense. (2023). *DoD Directive 3000.09: Autonomy in weapon systems*.
47. United States Department of State. (2023). *Political declaration on responsible military use of artificial intelligence and autonomy*.
48. Vienna Convention on the Law of Treaties, May 23, 1969, 1155 U.N.T.S. 331.