THE LAWWAY WITH LAWYERS JOURNAL

Website: www.the lawway with lawyers.com

VOLUME:-38 ISSUE NO:- 38 , SEPTEMBER 13, 2026

ISSN (ONLINE):- 2584-1106

Email: thelawwaywithelawyers@gmail.com

Digital Number : 2025-23534643

CC BY-NC-SA

Authored By:- Mahashree Gautom
B.A. LL.B. (Hons.), 3rd Semester, 2nd Year

WHEN WAR BECOMES ALGORITHMIC: AUTONOMOUS WEAPONS AND THE FUTURE OF INTERNATIONAL CRIMINAL RESPONSIBILITY

Abstract

War is an unavoidable alternative in today's national security. Although various international organisations and bodies are working to put a stop to the already brewing, more than 50, wars across the globe weapons acquisition and building of indigenous weapons is crucial at the current hour.

The increasing integration of autonomy and artificial intelligence into weapon systems is transforming the relationship between human decision-making and the use of lethal force. Autonomous weapon systems (AWS), capable of selecting and engaging targets with varying degrees of human intervention, raise significant questions concerning compliance with international humanitarian law (IHL) and the attribution of individual criminal responsibility. Although existing IHL principles, including distinction, proportionality, and precautions in attack, continue to govern the conduct of hostilities, the growing distance between human decisions and machine-generated actions creates difficulties in establishing attribution, causation, and the requisite mens rea.

This chapter critically examines whether the existing framework of international criminal law is capable of addressing unlawful conduct involving autonomous weapons. It analyses the principles of individual criminal responsibility under Article 25 and command responsibility under Article 28 of the Rome Statute, alongside the mental element required under Article 30, to assess whether commanders, operators, programmers, or other human actors may be held accountable when autonomous systems contribute to international crimes.

The chapter argues that autonomous weapons do not create an entirely new category of legal responsibility; rather, they expose an emerging "accountability gap" where increasing machine autonomy may weaken the connection between human agency and unlawful outcomes. It therefore proposes a human-centred accountability framework based on meaningful human control, traceable chains of responsibility, auditable decision-making, and lifecycle accountability.

At the end, the chapter concludes that technological autonomy cannot become a shield against individual criminal responsibility and that future international regulation must preserve a legally traceable human chain of accountability.

Keywords: Autonomous Weapons; International Criminal Law; International Humanitarian Law; Human Responsibility; Artificial Intelligence

Introduction

When one imagines war, the first image that comes is that of a human. A commander studying a map, a soldier making a split-second decision under pressure, or an operator pulling a trigger. Warriors have been praised and sung ballads about for their action and prowess during war. May it be in history or even mythology, a human has always been the center of a war.

The archer stood within sight of the battlefield. The artillery commander could be separated from the target by kilometres. The drone operator may be physically located thousands of kilometres away.

That is where the Autonomous Weapons System or AWS challenges the paradigm and the imagination of poets.

Autonomous weapons introduce a profound separation in war: the human may decide whether to activate a system, while the system subsequently determines which target to engage, when to engage it and, in some circumstances, where the resulting force will be applied.

The International Committee of the Red Cross (ICRC) defines autonomous weapons as systems that, once activated, can select and engage one or more targets without further human intervention. The significance of autonomy therefore lies not merely in automation, precision or technological sophistication. Despite its technological prowess, the AWS, therefore, largely depends on human intent and capability.

This produces a unique problem in the field of law.

The machine may determine the immediate target, but it cannot presently be treated as a subject of international criminal responsibility.

Evidently, Article 25 of the Rome Statute expressly places the jurisdiction of the International Criminal Court over “natural persons.” Responsibility must therefore ultimately return to human actors.

Yet identifying the relevant human actor is not a cakewalk.

Consider a hypothetical autonomous weapon deployed against military vehicles within a defined area.

The commander authorises its use, the operator activates it, the system processes sensor information and the algorithm subsequently identifies a civilian vehicle as a military target. The vehicle is destroyed. The commander did not select that vehicle. The operator did not know that the system would identify it as a target. The programmer did not intend the system to attack civilians.

The legal question is not simply whether a violation occurred. It is:

Whose conduct constituted the crime, and whose mental state can legally be connected to it?

This question has produced what is commonly described in the world of academia as the “responsibility gap” debate.

Some scholars argue that autonomous weapons may expose a structural gap between harmful outcomes and the human mental states required for international criminal responsibility. Others caution to deny the assumptions of such gaps and instead emphasize the continued applicability of ordinary principles of individual responsibility.

The disagreement is significant because international criminal law cannot respond to technological novelty by simply abandoning its foundational requirements.

Criminal responsibility remains individualised. The prosecution must establish the relevant act, causal relationship, mode of liability and mental element. At the same time, technology cannot become a convenient means of making human responsibility disappear.

This paper puts forth three principal questions:

1. To what extent can existing IHL and international criminal law govern unlawful conduct involving autonomous weapons?
2. Can existing modes of individual and superior responsibility adequately connect unlawful autonomous outcomes to human actors?
3. What additional accountability mechanisms are necessary to prevent technological autonomy from producing practical impunity without imposing impermissible strict liability?

Autonomous Weapons And The Transformation Of The Kill-Chain

A. From Automation to Autonomy

The distinction between automation and autonomy is imperative to the legal analysis of the problem.

An automated system performs a function according to predetermined conditions. Its response may be highly sophisticated, but the relationship between input and output is established in advance. An autonomous system operates with a greater degree of independence in responding to information obtained from its environment.

Autonomy should therefore be understood as a continuum rather than a binary condition.

At one end, a human selects the target, authorises the attack and directly controls the weapon. At another, a human activates a system within broad parameters while the system independently identifies and engages targets. Between these points are systems in which humans supervise operations and retain varying capacities to intervene.

The 2019 Guiding Principles of the Convention on Certain Conventional Weapons (CCW) Group of Governmental Experts recognise that human-machine interaction may take different forms and occur at different stages of a weapon's lifecycle. They further affirm that human

responsibility for decisions concerning the use of weapons must be retained because accountability cannot be transferred to machines.

The importance of this position is immediately apparent. The question is not whether a human technically interacted with the weapon. The question is whether the human exercised legally meaningful control.

B. The Changing Kill-Chain

The conventional kill-chain operates in the following manner:

A human identifies a target. A human decides to hit the said target. The human orders or executes the attack. The weapon of choice produces the effect.

An autonomous regime system however, changes this in the following manner:

A human demarcates the parameters. The machine makes an observation of the environment. The machine identifies the target. The machine selects the target. The machine engages the target.

The case of international laws of engagement and conventions occurs between identification and engagement.

That is where a machine may replace a human judgment.

This creates a distinction between responsibility for deploying a system and responsibility for a particular engagement. A person may have responsibility for the previous one without necessarily possessing the intent required for the latter.

Therefore, any serious analysis of AWS must resist the temptation to ask only: *Who used the weapon?*

The more precise questions to be asked, however, are: *Who designed the system? Who assessed its reliability? Who legally reviewed it? Who authorised its deployment? Who established its operational parameters? Who supervised its operation? Who had the ability to intervene? Who knew of its limitations?*

The answer may involve different people at different stages.

The Ihl Foundation: Autonomy Can Co-Exist With The Law

The emergence of AWS does not necessarily create a legal vacuum. The CCW Guiding Principles expressly reaffirm that IHL applies fully to all weapons systems, including lethal autonomous weapons systems, bringing it under an umbrella of laws from the very conception.

The ICRC in a similar manner maintains that existing IHL applies to AWS while recognising that existing law may not resolve every humanitarian, legal and ethical difficulty created by autonomy.

Three principles are particularly important from this framework.

A. Distinction

The principle of distinction requires parties to distinguish between civilians and combatants and between civilian objects and military objectives.

For autonomous weapons, the difficulty is not simply whether an algorithm can identify an object. A legally permissible targeting decision may depend upon circumstances, context and changing information.

An individual may be a combatant at one moment and hors de combat at another. A civilian object may become a military objective depending upon its use. A vehicle that appears to satisfy a general military profile may in fact contain civilians.

An autonomous system may therefore possess extraordinary object-recognition capabilities while remaining unsuitable for an operational environment in which lawful targeting requires contextual human judgement.

The legal question consequently becomes one of predictability and appropriateness of deployment.

B. Proportionality

Proportionality prohibits an attack expected to cause incidental civilian harm excessive in relation to the anticipated concrete and direct military advantage.

This principle presents an even greater difficulty for autonomy because proportionality is not reducible to target classification.

The assessment requires a contextual judgment concerning anticipated civilian harm and military advantage. A machine may calculate blast radius, estimate collateral effects or process vast amounts of data, but computational capacity does not automatically equate to legal judgement.

Accordingly, an autonomous system's ability to calculate consequences should not be confused with its ability to make the normative determination required by IHL.

C. Precautions in Attack

The obligation to take feasible precautions requires parties to take reasonable measures to verify targets and minimise civilian harm.

For AWS, this obligation extends beyond the instant at which the weapon engages a target. It affects whether the system should be deployed in the first place.

A commander who knows that an autonomous weapon becomes unreliable in a particular environment cannot necessarily avoid responsibility by arguing that the machine, rather than the commander, made the final selection.

The ICRC's 2026 position paper emphasises that predictability, human supervision and the circumstances of use are central to assessing compliance with IHL.

Thus, the legality of an autonomous attack may depend substantially upon decisions made before the weapon becomes autonomous.

Article 36 And The Preventive Architecture Of Accountability

Article 36 of Additional Protocol I requires States or Parties, when studying, developing, acquiring or adopting a new weapon, means or method of warfare, to determine whether its employment would be prohibited by the Protocol or another applicable rule of international law.

The provision is particularly important for systems like an AWS because it introduces a preventive dimension to legal accountability.

International criminal law is most of the time retrospective in nature. A crime occurs and the legal system subsequently determines responsibility.

Article 36 operates as per prospective or probability.

It asks: the following question*: Should this weapon be used at all, and under what circumstances can it lawfully be used?*

The problem is particularly acute for autonomous systems which use the technology of machine-learning functions. If the system behaviour is difficult to predict, a State may struggle to establish whether the weapon can reliably comply with distinction, proportionality and precaution requirements.

The ICRC's 2026 analysis emphasises that an AWS must be sufficiently understandable and predictable for an effective legal review and notes that modifications affecting system functionality may require renewed review.

This has a major implication for individual criminal responsibility.

Weapons review should not be treated as merely a technical procurement exercise. It can become part of the evidentiary chain of future responsibility.

If a system was known to produce unreliable target classifications, whether that information was documented, disclosed and acted upon may become relevant to determining what later decision-makers knew.

Accordingly, the accountability chain should begin: *before deployment, not after destruction.*

When Failure To Act Becomes Criminal: Omission And Command Responsibility

A. Commission by Omission

In traditional armed conflict, criminal liability under international humanitarian law (IHL) is predominantly active and needs to be committed by the perpetrator actively by their hand.

In the case of an autonomous weapon system (AWS), however, executing a lethal strike independently, the immediate physical act stems from automated code and machine mechanics rather than direct human muscle contact.

This technological reality shifts the doctrinal inquiry from active commission to commission by omission. The central legal challenge is determining under what precise conditions a human actor's failure to intervene or deactivate an AWS crosses the threshold into criminal liability.

Under established international criminal justice system thoughts, establishing individual criminal responsibility through omission requires the act must meet four foundational criteria:

- A Legal Duty to Act: The individual (whether commander, operator, or programmer) must hold a recognized legal obligation under IHL—such as the duty to protect civilian populations or prevent indiscriminate attacks.
- Capacity to Act: The individual must possess the physical and structural ability to alter the operational outcome.
- Knowledge of the Developing Unlawful Attack: The actor must possess actual knowledge, or be in possession of sufficient indicators to know, that the AWS is exhibiting aberrant targeting behavior or drifting toward an IHL violation.
- Practical Ability to Intervene: Functionally accessible overrides, safety protocols, or kill-switches must be within the actor's immediate grasp in real time.

As Bo observes in his doctrinal analysis of commission by omission, criminal law cannot artificially manufacture duties where structural control is absent. If an AWS operates at algorithmic processing speeds that effectively lock human operators out of the engagement loop, the "capacity to act" is legally extinguished. Consequently, omission liability cannot legitimately

punish a human for failing to govern a system that technology structurally prevented them from managing. Instead, this doctrine forces accountability backward, penalizing the antecedent negligence or recklessness involved in deploying an uncontrollable autonomous architecture into a dynamic operational environment.

B. Human-on-the-Loop Responsibility

The concept of a *human-on-the-loop* presumes that human oversight remains preserved because a monitor is stationed to observe system functions and intercede if irregularities arise.

- The Absence of a Physical Trigger Pull: The lack of a direct, physical trigger pull by a human operator doesn't resolve the liability. If an operator willfully ignores the clear and continued remote sensing indicating imminent civilian harm while there still being an option to abort, their passive negligence can satisfy the actus reus of participation.
- The Threshold of Actual Intervention Capacity: However, legal doctrine must resist becoming a vehicle for legal fictions. If a weapon system operates at millisecond speeds that far exceed human cognitive reaction times, the notion of "supervision" becomes illusory. Criminal responsibility must hinge strictly upon actual intervention capacity rather than theoretical control. Where system opacity and processing speed deny the operator meaningful time to evaluate a target, treating the human as an effective failsafe collapses under legal scrutiny.

C. Article 28 and Command Responsibility

When evaluating military leadership in the context of advanced weapon systems, Article 28 of the Rome Statute provides the primary framework for establishing superior responsibility. To anchor criminal liability upon a commander for the unlawful acts of forces under their command, three cumulative requirements must be satisfied:

- Effective Command and Control: The superior must possess the material ability to prevent or repress crimes, underpinned by a genuine hierarchical authority.
- The Knowledge Standard: The commander either knew, or owing to the circumstances at the time, should have known that subordinates (or systems deployed under their authority) were committing or about to commit crimes—such as observing a documented pattern of algorithmic targeting failures.
- Failure to Take Necessary and Reasonable Measures: The commander failed to take all feasible steps within their power to prevent, halt, or punish the unlawful conduct.

D. The Limits of Treating AWS as Subordinates

A persistent hazard in contemporary legal scholarship is the temptation to fit artificial intelligence into the traditional superior-subordinate architecture contemplated by international criminal law. Machines cannot be prosecuted, detained, morally censured, or imprisoned by the International Criminal Court (ICC), nor do algorithms possess the psychological agency, cognitive capacity, and hierarchical comprehension required of subordinates under Article 28.

As Spadaro persuasively argues, treating an autonomous algorithm as a "subordinate" strains international criminal law beyond its conceptual breaking point. Subordinates under IHL are moral agents capable of receiving lawful or unlawful orders, formulating independent criminal intent (*mens rea*), and internalizing humanitarian prohibitions. An AWS possesses none of these attributes.

This doctrinal boundary is reinforced by landmark ICC jurisprudence, notably *Prosecutor v. Bemba* and *Prosecutor v. Ntaganda*. These decisions establish that effective control requires a genuine, dynamic causal and hierarchical relationship between superior and subordinate. A commander cannot exercise "effective control" over a machine-learning neural network whose proprietary, emergent outputs generate unpredictable operational pathways.

As a result, Article 28 needs to attach responsibility to the human command structure surrounding an AWS, rather than conceptualizing the machine itself as a subordinate. Because the algorithm cannot function as a legal subject or a moral agent, the chain of accountability must loop upward and backward. The "subordinate" in this equation is a legal ghost; therefore, a commander's culpability does not lie in managing a rogue human soldier, but in introducing an autonomous weapon system whose structural autonomy renders genuine command and control impossible from the outset.

Beyond The Commander: Programmers, Developers, Procurement Authorities And The Distribution Of Responsibility

A. The Distributed Responsibility Problem

The lifecycle of an Autonomous Weapon System (AWS) breaks away from the linear, top-down methods of traditional military procurement and deployment of weapons and weapon systems.

Bringing a weapon which runs on an algorithm from a laboratory to the practical battlefield is not a simple task, and usually involves a vast network of actors that includes programmers and software engineers, private manufacturers and defense contractors, military procurement authorities and acquisition officials, legal and technical reviewers, commanders and operational planners, operators and real-time supervisors, and intelligence personnel feeding target parameters into the system.

This large ecosystem gives rise to a complex and distributed responsibility problem, challenging international criminal law's traditional focus on direct, isolated perpetrators.

B. Why Technical Proximity Is Not Criminal Responsibility

As the web of contributors widens, a dangerous legal temptation arises: to cast the net so broadly that anyone who touches the technology shares the blame. International criminal law must display restraint to this impulse.

For instance, a software engineer should not automatically face war crimes charges merely because a line of code they wrote ultimately contributed to a targeting error, nor should a procurement official be held criminally culpable simply for purchasing a legally certified system off a defense contractor's shelf.

Criminal responsibility must remain rigorously and majorly individualised. Attributing war crimes requires more than showing that an individual's professional duties touched the weapon system at some point in its lifecycle.

C. When Responsibility May Become Legally Relevant

To map out liability across the entire technological lifecycle, individual criminal responsibility must be tied to specific, personally committed conduct linked to the ultimate harm. Across this expansive web, potential legal relevance arises only when distinct behavioral patterns are crossed.

Commanders may incur liability by knowingly authorizing unlawful deployment, while operators may be held accountable for deliberately using the system unlawfully or failing to intervene despite having the actual capacity. Programmers cross the criminal threshold by intentionally or knowingly designing unlawful functionality or targeting parameters.

Moreover, developers and manufacturers may face responsibility if they conceal known safety defects or algorithmic vulnerabilities relevant to unlawful operation.

Procurement authorities can become liable by knowingly acquiring a system fundamentally incapable of operating lawfully within the parameters of International Humanitarian Law, and legal or technical reviewers may be held liable for deliberately concealing or misrepresenting known legal or technical deficiencies during mandatory weapons reviews.

D. The Governing Principle

To navigate this expanded network of actors without collapsing the boundaries of criminal justice, the application of liability must be anchored by a single rule: Responsibility should follow knowledge, contribution, and control, rather than mere proximity to technology.

Adhering to this principle shields the legal framework from two opposite, catastrophic errors. The first is technological impunity, which creates a vacuum where no one is held accountable because "the machine acted autonomously," allowing human actors to hide behind a digital shield.

The second is technological strict liability, which represents an overcorrection where everyone involved in the weapon's lifecycle—from coders to buyers—faces criminal sanction regardless of their intent or control, thereby violating the foundational requirement of individual culpability.

From Accountability Gap To Lifecycle Accountability: A Human-Centred Framework

A. Meaningful Human Control

To prevent autonomous weapon systems from developing and operating in a space of legal and moral vacuum, international law must demand the preservation of meaningful human control.

A legally meaningful human role cannot be reduced to a symbolic rubber-stamp or a mere monitoring assignment. Instead, it must involve a robust set of safeguards and active capabilities to interfere.

Specifically, human oversight requires a sufficient understanding of the system's operational parameters, predictable boundaries of behavior, strict target and geographical restrictions, and temporary limitations.

It also demands the continuous capacity to intervene, reliable communication channels, and an unrestrained ability to deactivate or terminate the system's operation before a violation occurs.

B. Continuous Article 36 Review

Under Article 36 of Additional Protocol I to the Geneva Conventions, States are required to determine whether the employment of a new weapon would be prohibited by international law. In the context of autonomous systems, however, a single, static pre-deployment review is largely inadequate and insufficient.

Article 36 reviews must transition from a one-time clearance into a continuous, lifecycle-based obligation.

This requires an initial legal review prior to fielding, subsequent reviews following substantial software updates or machine-learning capability modifications, and detailed evaluations grounded in real-world operational experience and previously identified targeting failures or systemic anomalies.

C. Algorithmic Auditability

Post-conflict investigations and international criminal tribunals cannot adjudicate responsibility without transparent evidentiary trails. Because neural networks and proprietary algorithms often function as "black boxes," or systems which are barely transparent and legal systems must mandate strict algorithmic auditability and credibility.

Manufacturers and military authorities should be legally required to preserve comprehensive logs and operational data concerning system activation, target classification metrics, internal system warnings, confidence and error indicators, operator interventions and overrides, engagement decisions, and unprompted system failures or deviations. Without this rigid data trail, prosecuting unlawful conduct becomes practically impossible.

D. Traceable Responsibility Chains

The fragmentation of human agency across the lifecycle of the algorithm requires a structured method and procedure for legal reconstruction. Every single deployment of an autonomous weapon system must be logged and documented in a manner that allows investigators to effortlessly trace the chain of causation and intent.

An effective accountability framework must be able to systematically answer a precise sequence of investigative questions: *Who authorized the deployment? Who configured the system's parameters? Who reviewed its legal compliance? Who monitored its operation in real time? What did each actor know at their respective stages? What capacity did they possess to alter the outcome? And precisely what physical and algorithmic events occurred during the engagement?*

E. Non-Delegation of Responsibility

At the end of the day, the integration of artificial intelligence into warfare must not be allowed to rewrite the foundational rules of international criminal law. The use of a complex algorithm can never be deployed as a legal escape mechanism to absorb blame or shield human decision-makers from culpability.

At the same time, the regulatory framework must avoid the opposite extreme: converting every technological malfunction or unforeseen algorithmic drift into an automatic charge of individual criminal liability. Navigating this balance requires anchoring the legal architecture to a definitive core principle: Criminal responsibility must remain rigorously individualised, while accountability mechanisms must become comprehensively lifecycle-based.

Conclusion

It has been said by many scholars that one must learn to change with the times.

The birth and integration of autonomous weapon systems in modern day warfare does not, typically, make international criminal law irrelevant, nor do these technologies invalidate the foundational principles of International Humanitarian Law (IHL) and International Criminal Law (ICL).

Instead, the genuinely difficult cases emerge at the where middle ground of human agency is distributed across a massive technological supply chain which leads to system behavior becoming unpredictable due to machine-learning loops, evidence is technically less clearer, intervention capacity is mostly restricted, and the required mental state (*mens rea*) cannot easily be established under traditional systems.

In this situation legal pathways such as Articles 25, 28, and 30 of the Rome Statute remain indispensable, but they cannot work alone. They must be strengthened and reinforced by active, preventive, and evidence based mechanisms.

Measures which include legislations such as Article 36 legal reviews, enforceable standards of meaningful human control, and rigorous algorithmic auditability ensure that criminal investigations do not begin only after the digital evidentiary trail has vanished.

At the end, the law must reject both extremes: it must refuse technological impunity, which allows human actors to hide behind algorithmic shields, while equally rejecting automatic human strict liability for unforeseen machine malfunctions.

The machine cannot stand trial. But neither should technological autonomy allow the human beings who design, authorize, deploy, supervise, or knowingly permit unlawful systems to disappear behind the machine.

War may become increasingly autonomous, but the chain of human accountability must not.