Page 3199 - 3211

INTERNATIONAL JOURNAL OF LAW MANAGEMENT & HUMANITIES

[ISSN 2581-5369]

Volume 9 | Issue 3

2026

Algorithmic Warfare and Target Selection: Epistemic Uncertainty in AI-Enabled Military Operations

AKANCHHA SINGH*

ABSTRACT

The integration of AI-enabled targeting systems into military operations raises pressing questions about the application of international humanitarian law (IHL). This paper considers how AI-based decision-support tools influence the implementation of the fundamental principles of IHL, namely distinction, proportionality, and precautions in attack. It argues that the greatest challenge posed by AI is not a deficiency in the legal rules, but the emergence of epistemic uncertainty, which changes the nature of what military decision-makers are able to know, predict, verify, and justify at the point of attack. Drawing on doctrinal analysis of IHL principles, examination of state practice and military doctrine, and selected case studies from recent conflicts, including the Israel-Hamas conflict and the war in Ukraine, the paper evaluates the implications of AI-enabled targeting for legal compliance and accountability. It identifies three features of AI systems, namely opacity, probabilistic reasoning, and operational scale, as core conditions that alter the factual basis on which IHL is applied: they make target verification more difficult, distort proportionality assessments, and compress meaningful deliberation under the precautionary obligations, while fragmenting traditional structures of legal responsibility. Consequently, existing accountability mechanisms and weapons-review procedures are significantly constrained when applied to algorithmic decision-making environments. The paper concludes that IHL remains robust and technologically neutral, but that it must be reinterpreted and supplemented with stronger safeguards to govern AI-enabled weaponry effectively. Compliance is best secured by ensuring meaningful human control, strengthening Article 36 weapons-review processes, increasing transparency and auditability, and reinforcing accountability. The paper thereby contributes to ongoing debate by reframing AI-enabled targeting as a problem of epistemic uncertainty rather than a deficiency in the existing legal system.

Keywords: *Algorithmic warfare, AI-enabled targeting, international humanitarian law, epistemic uncertainty, meaningful human control, Article 36 weapons review, accountability.*

* Author is an LL.M. student at National Law University, Delhi, India.

 [ISSN 2581-5369]

I. INTRODUCTION

The integration of artificial intelligence (AI) into the military sector has given rise to a new concept known as "algorithmic warfare", which refers to the use of machine-learning systems to carry out or assist military functions, especially in intelligence processing, threat assessment, and operational decision-making. One aspect of this concept is target selection, that is, the process of identifying, classifying, and prioritising persons or objects as lawful military targets for attack. Increasingly, such selection is performed by algorithmic systems that base their operations on data patterns and predictive analytics rather than solely on human judgment, which introduces new legal and ethical issues. Machine-learning systems in this context do not merely support humans in target identification and engagement decisions; in some cases, they decide on their own. Although IHL has not formally changed, the use of AI-enabled decision-support systems (AI-DSS) drastically alters the factual context in which its main principles, distinction, proportionality, and precautions, are applied.[1] These systems do not merely assist human decision-making but shape it, often through processes that are opaque, data-dependent, and difficult to verify. Recent scholarship has therefore focused on questions of meaningful human control, algorithmic accountability, and the limits of applying traditional legal standards within data-driven environments.[2]

Traditional targeting decisions rested on human reasoning, bounded rationality, and contextual judgment. Commanders relied on intelligence synthesis, personal judgment, and legal advice to determine whether a target met the criteria for lawful attack. AI systems, by contrast, can process vast quantities of data, identify patterns, and select targets rapidly and at scale. For example, it has been reported that the Israeli system known as "Lavender" analysed large volumes of data to generate thousands of potential targets, while also exhibiting a significant rate of error.[3] Developments such as these raise important questions: can algorithmic decisions meet the "reasonable commander" standard? How is foreseeability affected when probabilistic outputs are used to make lethal decisions? Does reliance on algorithmic results diminish an individual's moral and legal responsibility?

---

[1] Int'l Comm. of the Red Cross, *Artificial Intelligence and Machine Learning in Armed Conflict: A Human-Centred Approach*, 102 INT'L REV. RED CROSS 463, 470-73 (2020), https://international-review.icrc.org/articles/ai-and-machine-learning-in-armed-conflict-a-human-centred-approach-913.

[2] Wen Zhou & Anna Rosalie Greipl, *Artificial Intelligence in Military Decision-Making: Supporting Humans, Not Replacing Them*, ICRC HUMANITARIAN L. & POL'Y BLOG (Aug. 29, 2024), https://blogs.icrc.org/law-and-policy/2024/08/29/artificial-intelligence-in-military-decision-making-supporting-humans-not-replacing-them/; Jessica Dorsey, *The Erosion of Human(e) Judgement in Targeting? Quantification Logics, AI-Enabled Decision Support Systems and Proportionality Assessments in IHL*, 107 INT'L REV. RED CROSS 1041, 1054-58 (2025).

[3] Lamia Faris, *Algorithmic Targeting in the Iranian-Israeli Confrontation: Technical Realities, Legal Thresholds, and the Boundaries of Human Control*, 14 F1000RESEARCH 1200 (2025), https://f1000research.com/articles/14-1200.

 [ISSN 2581-5369]

This paper argues that AI does not change the normative content of IHL but transforms the epistemic conditions, that is, the way in which its principles are understood under new factual situations. In other words, the law remains the same, but the factual assumptions underlying its application become unstable. As a result, the present legal framework, although sufficient, needs to be reinterpreted, its doctrines clarified, and its enforcement strengthened in order to address the problems of opacity, automation bias, and data dependency arising from the use of AI. This transformation weakens verification under distinction, distorts proportionality assessments, compresses deliberation under precautions, and fragments accountability.[4]

The central challenge, therefore, is not the inadequacy of IHL, but the transformation of the epistemic conditions underlying its application. While the existing legal framework remains formally adequate, it requires reinterpretation and strengthened safeguards to address opacity, automation bias, and data dependency.[5]

II. Background: AI and Military Targeting

AI-enabled defence systems operate at different levels of autonomy, ranging from human-in-the-loop systems that require direct human authorisation, to human-on-the-loop systems in which the human plays a supervisory role, to fully autonomous systems that can select and engage targets independently.[6] At present, most military uses of AI, even in active conflict, remain at the level of AI-enabled decision-support systems (AI-DSS) that enhance human decision-making rather than replacing it.

These examples illustrate how diverse and extensive these technologies have become. The United States' Project Maven employs machine learning to detect and classify objects during intelligence, surveillance, and reconnaissance (ISR) missions. According to reports, Israeli technologies such as Lavender and Gospel use data analytics to generate targeting recommendations. The development of drone swarms and loitering munitions, meanwhile, reflects a trend towards semi-autonomous or fully autonomous engagement capabilities. Such weapons operate through a defined sequence: collecting data, analysing the data algorithmically, recommending a target, confirming it by a human, and then carrying out the strike.[7] Although this appears to maintain human control, the position is not so clear-cut. The sheer volume and

[4] Dorsey, *supra* note 2, at 1054-58; Shin-Shin Hua, *Machine Learning Weapons and International Humanitarian Law: Rethinking Meaningful Human Control*, 51 GEO. J. INT'L L. 117, 126-27 (2019).
[5] Zhou & Greipl, *supra* note 2; Int'l Comm. of the Red Cross, *International Humanitarian Law and the Challenges of Contemporary Armed Conflicts* 33-38 (2024).
[6] *Autonomous Weapons Systems: Law, Ethics, Policy* 25-32 (Nehal Bhuta, Susanne Beck, Robin Geiß, Hin-Yan Liu & Claus Kreß eds., 2016).
[7] Action on Armed Violence, *Kill Codes and Command Lines: Understanding the Rise of Algorithmic Warfare* (2025), https://aoav.org.uk/2025/kill-codes-and-command-lines-understanding-the-rise-of-algorithmic-warfare/.

 [ISSN 2581-5369]

speed of the algorithmic outputs that humans must process can make them so reliant on the machine that actual decision-making authority is effectively transferred to it.

The probabilistic character of AI outputs gives rise to various legal and operational risks. Machine-learning systems depend on the data on which they are trained, and this data may be incomplete, biased, or no longer relevant to target identification. As a result, targeting may proceed on the basis of correlation rather than causation, which can produce systemic misclassification. Such problems are particularly acute in asymmetric conflicts, where distinguishing between combatants and civilians is far from straightforward.[8]

In addition, the psychological effect known as automation bias, whereby humans are strongly inclined to rely on machine-generated information even when it may be erroneous, significantly undermines the human capacity for rational and independent judgment.[9] Dorsey also observes that AI may alter the way proportionality considerations are made through quantification logics, giving greater weight to numerical thresholds at the expense of qualitative assessment, and thereby eroding the human-centred spirit of IHL.[10] Therefore, although AI may improve efficiency, it simultaneously places at risk the very cognitive activities on which legitimate targeting depends.

III. Legal framework: IHL rules on targeting and new weapons

These technological developments must be assessed against the existing normative architecture of international humanitarian law. The core principles governing targeting under IHL are set out in Additional Protocol I (1977), particularly Articles 48, 51, 52, and 57.[11] These provisions establish a normative framework aimed at balancing military necessity against humanitarian considerations.

The principle of distinction requires that parties consistently distinguish between civilians and combatants, and between civilian objects and military objectives, at all times. The principle of proportionality prohibits attacks expected to cause incidental civilian harm that is excessive in relation to the anticipated military advantage. The principle of precautions imposes a duty of constant care, requiring verification of targets and minimisation of civilian harm.

[8] Vasja Badalič, *The Metadata-Driven Killing Apparatus: Big Data Analytics, the Target Selection Process, and the Threat to International Humanitarian Law*, 9 CRITICAL MIL. STUD. 619, 625-28 (2023).
[9] Zhou & Greipl, *supra* note 2.
[10] Dorsey, *supra* note 2.
[11] Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of International Armed Conflicts (Protocol I) arts. 48, 51-52, 57, June 8, 1977, 1125 U.N.T.S. 3.

 [ISSN 2581-5369]

Article 36 further requires states to carry out legal assessments of new weapons, means, or methods of warfare to verify their conformity with IHL.[12] This obligation bears directly on AI-enabled systems, which are potentially new "methods" of warfare even where the physical platforms are unchanged. IHL is generally regarded as technology-neutral, in that its principles apply regardless of technological change.[13] However, this technological neutrality assumes that new technologies operate within the same epistemic framework as traditional ones. AI disrupts this assumption by introducing features such as opacity (decision-making that is a "black box" to humans), probabilistic reasoning, and autonomous adaptation.

For example, the principle of distinction relies on the assumption that targets can be identified with a high degree of reliability. AI systems, however, may identify individuals by their behavioural patterns or indirectly through metadata rather than by their direct participation in hostilities. Likewise, the principle of proportionality depends on the ability of commanders to predict and weigh potential collateral damage, yet AI-generated predictions may be uncertain or confined to a particular context. Precautions, a requirement that depends on careful verification, can be compromised within rapid decision-making cycles.

The ICRC has accordingly emphasised the importance of meaningful human control, stressing that even in the case of machine-operated weapons, targeting decisions must remain human-centred.[14] This is consistent with the Martens Clause,[15] which underscores the relevance of the principles of humanity and the dictates of public conscience to the interpretation of IHL. Hence, although the existing legal principles remain applicable, their implementation will differ, and they must be interpreted in light of the expanding range of AI-enabled contexts.

IV. STATE PRACTICE AND MILITARY DOCTRINE

There are both similarities and differences in states' approaches to regulating AI-enabled targeting. It is widely accepted that IHL continues to apply and provides a sufficient basis for addressing such systems. However, states differ on how AI should be incorporated into military operations. The United States, for example, in its Department of Defense Law of War Manual, indicates that autonomous systems may be lawful provided their use accords with IHL principles.[16] It notes that accountability rests on human operators and commanders. In a similar

[12] *Id.* art. 36.
[13] 1 *Customary International Humanitarian Law* 3-8 (Jean-Marie Henckaerts & Louise Doswald-Beck eds., 2005).
[14] Int'l Comm. of the Red Cross, *ICRC Position on Autonomous Weapon Systems* 2-4 (May 12, 2021), https://www.icrc.org/en/document/icrc-position-autonomous-weapon-systems.
[15] Protocol I, *supra* note 11, art. 1(2).
[16] U.S. Dep't of Def., *Department of Defense Law of War Manual* § 6.5.9 (rev. ed. Dec. 2016).

 [ISSN 2581-5369]

vein, the United Kingdom emphasises the necessity of human intervention and accountability in its armed forces' doctrine.[17]

Other states, such as France, and members of NATO call for the creation of responsible-use frameworks that emphasise operational guidelines rather than strict prohibitions. By contrast, various states and civil-society groups have advocated clear bans or moratoriums on fully autonomous weapon systems.

Discussions at the global level within the Convention on Certain Conventional Weapons (CCW) have shown the regulatory environment to be far from unified.[18] While almost all participants accept that some degree of human control and human responsibility should frame the chain of decisions from beginning to end, they have been unable to reach agreement on the introduction of legally binding measures. These divergences reflect a shared underlying tension between the advancement of military capability and the protection of human life. One prominent feature of state practice, however, is the acknowledgment that legal responsibility should not be assigned to machines, however technologically advanced they may be.

V. Practical use-cases and case studies

A. Case study 1: the Israel-Hamas conflict (2023-24)

AI-enabled targeting systems used during the Israel-Hamas conflict offer a compelling example. According to some sources, systems such as "Lavender" generated extensive lists of potential targets based on the analysis of metadata.[19] While the system's accuracy was estimated at about 90 per cent, this still meant that thousands of misidentifications were possible.[20]

This raises several legal concerns. First, the use of metadata alone for targeting may result in over-inclusive targeting, blurring the distinction between combatants and civilians. Second, the rapid generation of targets may leave insufficient time for verification and thereby compromise precautionary obligations. Third, the large scale of operations can amplify the effects of any systemic errors.

This case highlights a broader and more structural problem. Where decision-making is conducted through algorithmic targeting at scale, the margin of error is not merely an isolated incident but a systemic feature. Even a low error rate, when applied to thousands of targets, can

[17] U.K. Ministry of Def., *The Joint Service Manual of the Law of Armed Conflict*, JSP 383, paras. 5.32-5.33 (2004).
[18] *Report of the 2019 Session of the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems*, U.N. Doc. CCW/GGE.1/2019/3, annex IV (Sept. 25, 2019).
[19] Action on Armed Violence, *supra* note 7.
[20] Faris, *supra* note 3; Yuval Abraham, *'Lavender': The AI Machine Directing Israel's Bombing Spree in Gaza*, +972 Mag. (Apr. 3, 2024), https://www.972mag.com/lavender-ai-israeli-army-gaza/.

produce significant unlawful harm. This raises the question whether meaningful verification is possible and whether existing safeguards under IHL are sufficient. More specifically, it suggests that compliance cannot be evaluated solely on the basis of individual decisions, but must also account for the aggregate consequences of algorithmic targeting.

B. Case study 2: the Ukraine conflict (2022-24)

The Ukraine conflict illustrates the extent to which AI is being used in high-intensity warfare. AI-enabled intelligence, surveillance, and reconnaissance (ISR), drone swarms, and data analytics have substantially enhanced situational awareness and the accuracy of target identification.[21] However, these technologies also reflect the extent to which the basis for targeting decisions is shaped by algorithms.

Indeed, in both cases AI operates as a force multiplier, enhancing the performance of armed forces while also raising legal and ethical challenges. AI opens up new possibilities, but it can simultaneously increase risk, which is ultimately a question of whether effective safeguards are in place. These examples show that AI goes beyond the role of a mere efficiency tool and instead modifies the very nature of decision-making authority in modern warfare.

VI. Legal uncertainties and gaps

Crucially, it should not be assumed that greater data and computational power will necessarily lead to better compliance with IHL; in some cases, expanded targeting capabilities may instead cause legal errors to occur on a larger scale and at greater speed. The use of AI-enabled targeting systems introduces significant legal uncertainty in the following respects.

A. Distinction and target definition

One difficulty with AI systems is that they may identify and target individuals on the basis of patterns rather than actual evidence of combatant status. This may increase the number of protected persons placed at risk, thereby softening the distinction between civilians and combatants.[22]

B. Proportionality, foreseeability, and the illusion of precision

The principle of proportionality is particularly destabilised by AI-enabled targeting, as it depends on the ability of commanders to foresee and evaluate harm. Traditionally, such evaluation is qualitative, context-dependent, and grounded in human insight. AI tools, by

[21] Action on Armed Violence, *supra* note 7.
[22] Badalič, *supra* note 8, at 625-28.

 [ISSN 2581-5369]

contrast, increasingly convert such evaluations into probabilistic outputs and numerical thresholds, thereby inducing what may be termed the "illusion of precision."[23]

Although these tools appear to bring greater scientific rigour to the analysis, they may at the same time mislead as to the very normative character of proportionality, which cannot be reduced to a simple objective computation of two quantities, namely civilian harm and military advantage, but is a matter of value judgment shaped by context, experience, and legal reasoning. The attempt to convert these aspects into data-driven metrics may, while ostensibly increasing precision, conceal uncertainty and give rise to undue confidence in outputs that remain contingent and incomplete.[24]

Moreover, the problem is aggravated by the unforeseeability of the consequences of using AI. Such systems are based on historical data and modelling assumptions that may be unable to predict indirect, cumulative, or long-term harm, especially in complex urban environments. Substantial degradation of essential facilities such as electricity, water, or medical services may set in motion a series of humanitarian consequences extending well beyond the immediate effects of the strike. In consequence, reliance on algorithmic outputs can lead to systematic underestimation of civilian harm and a disproportionate emphasis on military advantage when determining proportionality.[25]

The concern, therefore, is not merely that AI can make mistakes, but that it alters the entire paradigm for perceiving and determining the level of harm. If lethal operations are conducted on the basis of probabilistic information, the threshold of what amounts to "excessive" harm may be altered inadvertently, raising the prospect of non-compliance with the proportionality norm or, at the very least, its effective "recalibration" through technology.[26]

C. Human control, automation bias, and the risk of superficial oversight

The requirement of meaningful human control has emerged as a central safeguard in debates on AI-enabled warfare. Yet it remains debatable whether such control is effective in practice. Many systems formally retain a human "in the loop", but the volume, speed, and complexity of the data produced by algorithms may reduce human control to a formality rather than a meaningful safeguard.[27]

---

[23] Dorsey, *supra* note 2.
[24] *Id.*; Int'l Comm. of the Red Cross, *supra* note 1, at 470-73.
[25] Dorsey, *supra* note 2; Yéelen Geairon, *Deciding Under Algorithms: Artificial Intelligence and the Protection of Civilian Infrastructure in Armed Conflict*, ICRC HUMANITARIAN L. & POL'Y BLOG (Mar. 12, 2026), https://blogs.icrc.org/law-and-policy/2026/03/12/deciding-under-algorithms-artificial-intelligence-and-the-protection-of-civilian-infrastructure-in-armed-conflict/.
[26] Hua, *supra* note 4, at 139-40; Dorsey, *supra* note 2.
[27] *Id.*.

 [ISSN 2581-5369]

At a rapid operational tempo, human operators receive machine-generated recommendations that are both numerous and produced within very short time-frames, making it difficult to scrutinise them properly. The tendency of decision-makers to rely on algorithmic outputs that they regard as objective, even where those outputs are in fact uncertain, is a manifestation of automation bias and a warning that this problem may escalate.[28]

Ultimately, the issue is not that human involvement is wholly removed, but that the human decision-making function is reduced to mere approval. Such cursory verification is at odds with the qualitative assessment that is the hallmark of IHL and exposes the weaknesses of current conceptions of human control in ensuring compliance.[29]

D. The accountability gap and the fragmentation of responsibility

AI-enabled targeting systems expose significant limitations in existing frameworks of legal responsibility. IHL and international criminal law are premised on human agency and therefore require identifiable actors who possess both the intent and the control over the conduct in question. In algorithmic settings, however, decisions are made by an entire network of developers, data analysts, commanders, and operators.[30]

The involvement of multiple parties in this way makes it difficult to attribute responsibility in cases of unlawful harm. Where outcomes stem from system behaviour that is opaque or even unpredictable, it is difficult to determine whether liability lies in design, deployment, or use. The fact that many machine-learning systems function as a "black box" makes it even harder to reconstruct decision-making processes after the event.[31]

Furthermore, the doctrine of command responsibility is premised on the assumption that a commander exercises effective control over subordinates, an assumption that is severely tested when decisions are made through autonomous or semi-autonomous systems. Likewise, establishing intent or negligence is far from straightforward where the harmful result arises from a system's probabilistic or emergent behaviour.[32]

This produces a structural accountability gap, in which responsibilities are so dispersed that it is unclear to whom they should be attached. State responsibility remains relevant in theory, but

[28] Zhou & Greipl, *supra* note 2.
[29] Dorsey, *supra* note 2.
[30] Gerald Mako, *Legal Accountability for AI-Driven Autonomous Weapons*, LIEBER INST. W. POINT (Mar. 9, 2026), https://lieber.westpoint.edu/legal-accountability-ai-driven-autonomous-weapons/.
[31] *Id.*; Hua, *supra* note 4, at 126-27.
[32] Mako, *supra* note 30.

 [ISSN 2581-5369]

this framework quickly reaches its limits when addressing harms that arise from complex human-machine interactions.[33]

E. Article 36 reviews

Moreover, conventional review frameworks can scarcely examine the distinctive features of algorithmic systems in sufficient depth. Factors such as data bias, opacity, and temporal variation are precisely the elements that make such review especially difficult.[34] Disregarding them is likely to result in non-compliance with the law, which may ultimately lead to the misuse of even the most advanced technologies, technologies that could otherwise make warfare more selective and less harmful.

VII. Counterarguments and rebuttals

The prevailing view is that IHL is technology-neutral and therefore readily adaptable to AI-powered systems without any change to its principles or doctrines. This view, however, overlooks both the fact that AI changes how decisions are made and the way it gives rise to new kinds of uncertainty.

Equally, the contention that AI yields greater precision fails to account for the biases inherent in the system or for the manner in which errors may multiply.[35] AI may improve certain operational indicators, but this does not guarantee adherence to humanitarian principles. Furthermore, the assumption that legal-responsibility provisions remain unaffected is far from accurate, given the difficulty of identifying who should be held accountable when failures occur in highly complex AI systems.

VIII. Ethical and policy implications

Beyond legal considerations, AI-enabled targeting raises serious moral concerns. Delegating the decision to kill to machines may lead to the erosion of human control and accountability.[36] It may also undermine public trust in the military and call into question the very legitimacy of armed conflict. The risk of states entering into an AI arms race, of increasing dependence on technology companies, and of a lack of transparency are among the core problems associated with the growing use of AI. There is therefore a clear need both to apply international law and to develop ethical guidelines.

---

[33] *Id.*.
[34] Int'l Comm. of the Red Cross, *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* 9-15 (2006).
[35] Dorsey, *supra* note 2, at 1058-62.
[36] Int'l Comm. of the Red Cross, *International Humanitarian Law and the Challenges of Contemporary Armed Conflicts* 33-38 (2024).

 [ISSN 2581-5369]

IX. Recommendations and normative reforms

The following reforms are not aspirational but operationally necessary to preserve the integrity of IHL in algorithmic environments.

A. Re-centring human judgment: mandatory meaningful human control

At the core of lawful targeting lies human judgment. States must therefore ensure that AI systems do not replace, but support, human decision-making at the critical stages of target selection and engagement, through the preservation of meaningful human control.[37]

Meaningful human control should not be understood as the mere formal presence of a human operator, but as substantive involvement, in which the human decision-maker understands the basis of the AI's recommendation, has sufficient time and authority to question or override it, and exercises independent judgment rather than passive approval.

This reform is essential to preserve the "reasonable commander" standard under IHL.[38] Without it, accountability risks being diluted and decision-making reduced to automated execution. Embedding meaningful human control is therefore not a technical preference but a legal necessity for compliance with distinction, proportionality, and precaution.

B. Reimagining Article 36 reviews: from weapons to algorithms

Existing Article 36 weapons-review mechanisms must evolve to address the distinctive characteristics of AI systems.[39] These systems differ fundamentally from traditional weapons because of their reliance on data, their adaptability, and their probabilistic outputs.

States should expand these reviews to include algorithmic transparency, that is, an understanding of how the system reaches its conclusions, at least at a functional level even if not fully explicable; data integrity and bias assessment, evaluating whether training data may produce systematic misidentification of targets;[40] and contextual testing, ensuring that systems are tested across diverse operational environments rather than under controlled conditions alone.

This shift transforms Article 36 reviews from a hardware-focused exercise into a dynamic, lifecycle-based evaluation of both software and data ecosystems. Such an approach is essential to prevent the deployment of systems whose risks emerge only in real-world use.

---

[37] Int'l Comm. of the Red Cross, *supra* note 14, at 2-4.
[38] Henckaerts & Doswald-Beck, *supra* note 13, at 3-8.
[39] Protocol I, *supra* note 11, art. 36.
[40] Badalič, *supra* note 8, at 625-28.

 [ISSN 2581-5369]

C. Ensuring explainability and auditability of AI systems

For IHL to remain enforceable, decisions leading to the use of force must be traceable and reviewable. Accordingly, AI systems used in targeting must incorporate mechanisms that allow explainability, so that commanders can understand why a target was flagged, and auditability, so that systems maintain records of inputs, outputs, and decision pathways.

Without these features, it becomes nearly impossible to assess compliance with IHL or to assign responsibility in cases of unlawful harm. Explainability and auditability are therefore not merely technical ideals but foundational requirements for accountability and the rule of law in warfare.

D. Institutionalising post-strike review and accountability mechanisms

The use of AI in targeting necessitates a robust post-strike accountability framework. States should mandate comprehensive logging of AI-assisted decisions (including data inputs and system outputs), independent post-operation reviews to assess accuracy, proportionality, and compliance, and feedback loops to correct systemic errors and improve future performance.

Such mechanisms ensure legal traceability in the use of force.[41] They also create an iterative learning process that reduces repeated errors. Without such safeguards, AI systems risk becoming opaque instruments of force, shielded from scrutiny and accountability.

E. Developing international guidelines on AI in warfare

Given the far-reaching cross-border consequences of AI-powered warfare, international harmonisation of standards is highly desirable, and at times urgent, particularly within a framework such as the Convention on Certain Conventional Weapons (CCW).[42]

States should aim at non-binding guidelines or codes of conduct, shared standards on human control, transparency, and accountability, and confidence-building measures, including the sharing of information and best practices.

Such guidelines would help harmonise state practice, reduce regulatory fragmentation, and prevent a "race to the bottom" in military AI deployment. They would reinforce the principle that technological advancement must remain subordinate to humanitarian considerations.

F. Proposed normative formulation

To reinforce these principles, the following formulation is proposed: "States should ensure that any AI-operated targeting system is subject to meaningful human oversight and transparency measures, accompanied by the possibility of immediate human intervention where the system's

[41] Int'l Comm. of the Red Cross, *supra* note 33, at 9-15.
[42] Convention on Certain Conventional Weapons, *supra* note 17, annex IV.

 [ISSN 2581-5369]

decisions are doubtful or inaccurate. The deployment of such systems must remain under periodic legal review to verify their compliance with the principles of distinction, proportionality, and precaution under international humanitarian law."

Taken together, these proposals aim to fortify the current IHL framework rather than to replace it. The intention is not to hinder technological progress but to ensure that the employment of novel technologies does not violate the fundamental principles of humanitarian protection. States can benefit from technology provided that the legal safeguards governing AI are incorporated at the stages of development, deployment, and monitoring.

X. Conclusion

"Algorithmic warfare" alters the way wars are fought, not by discarding IHL but by changing the very conditions under which it applies. AI-driven weapons systems disrupt the epistemic foundations of the main IHL principles, namely distinction, proportionality, and precautions, by reducing transparency, introducing probabilistic reasoning, and accelerating decision-making. This paper has argued that the key concern is not any inadequacy of the law itself, but its potential misuse in a technologically driven environment. Trusting outcomes based on algorithmic calculation lowers the "reasonable commander" standard, hampers accountability, reduces transparency in decision-making, and can give rise to situations that escape meaningful control. Compliance is therefore secured by reinstating human discretion as the principal locus of responsibility, rebalancing accountability mechanisms, and reforming the Article 36 review process to address algorithmic systems.

The real challenge is not the reform of IHL but ensuring that, in a world of intelligent machines, the law remains guided by human conscience and legal principle, while accountability is preserved throughout the process. The central claim of this paper is that what should concern us is not the inadequacy of the law, but the high probability of its misuse in technologically mediated environments. Reliance on algorithmic outputs may lower the "reasonable commander" standard, fragment accountability, and normalise decision-making practices that are non-transparent and difficult to control. Compliance is therefore best secured by reinstating human discretion as the principal locus of responsibility, rebalancing accountability mechanisms, and reforming the Article 36 review process to address algorithmic systems.

*****

 [ISSN 2581-5369]