Cadernos de Pós-Graduação em Direito Político e Econômico
Published by: Centro de Estudos Acadêmicos Press
ISSN: 1678-2127
Volume 26 *Issue S1*, 2026
Gurugram University Conference : 09-10 April 2026
Website: https://ceapress.org

Algorithmic Warfare and Legal Responsibility: Who is Liable for AI-Driven Military Decisions

Abdul Sakib Majid,
Ph.D. Scholar, Department of Law,
Gurugram University

Submission date 12.04.2026 | Acceptance date: 25.04.2026 | Publication: 29.05.2026

ABSTRACT

The quick adoption of artificial intelligence (AI) into the military system has brought a paradigm change in armed conflict, posing tough questions concerning the legal accountability of AI-assisted decisions. The autonomous and semi-autonomous weapons systems of algorithmic war is also a challenge to the traditional structures of liability in the international law, which assumes human intent, control and foreseeability. The decision of who is responsible in cases of illegal harm is legally unclear as the power of decision-making is transferred more and more to machine-learning systems.

This paper discusses whether the current legal principles are sufficient to resolve the liability of AI-assisted military acts, concentrating on the concepts based on International Humanitarian Law (IHL) and the law of state responsibility. It discusses the issue of states being strictly liable to the acts of AI systems used in armed conflict, despite the minimal or indirect human control over them. The paper further examines the relevance of individual criminal responsibility especially command responsibility in cases where autonomous systems behave unpredictably or are not under the direct control of the military personnel.

One of the main points presented in this paper is the creation of an accountability gap where neither humans nor machines can be held accountable according to the current laws. The paper also addresses the liability that may be held by the private actors such as developers and manufacturers of AI systems and assesses whether the existing legal frameworks can govern the contribution of such technologies to the deployment of AI systems.

Based on the current global discussions under the United Nations, the paper suggests the necessity of a redesigned legal framework that will promote accountability without paralyzing technology. It proposes the introduction of the principles of meaningful human control and improved weapons review procedures to close the gaps in the law. Finally, the research will seek to add to the process of creating a coherent and enforceable framework of responsibility allocation in the age of algorithmic warfare.

Keywords: *Algorithmic Warfare, Legal Responsibility, International Humanitarian Law, Autonomous Weapons Systems, Accountability Gap*

Cadernos de Pós-Graduação em Direito Político e Econômico
Published by: Centro de Estudos Acadêmicos Press
ISSN: 1678-2127
Volume 26 *Issue S1*, 2026
Gurugram University Conference : 09-10 April 2026
Website: https://ceapress.org

I. INTRODUCTION

The pace of the implementation of AI-based systems into warfare has emerged as one of the most controversial topics at the nexus of technology, ethics and international law. Militaries are making investments in capabilities that take advantage of machine learning to conduct surveillance, identify targets, assess threats and, more frequently, autonomously engage. Such developments are offering benefits to operations that are faster, persistent and in certain cases more accurate, but it is increasing the risk that lethal decisions will be made by an opaque, brittle, and hard to predict and manipulate system.[1]

The current systems of law were created based on the premise that it is human beings who make the decisive choice regarding the application of force. The fundamental principles of distinction, proportionality and precaution, postulated by IHL assume human judgement in context and both state responsibility and ICL hold states and natural persons, rather than machines, liable. The connection between human agency and illegal injury can be diluted or lost when AI systems interpose or substitute human decision-making.[2]

The current paper assumes that international law applies to algorithmic warfare but that the manner in which the responsibility is operationalized must be re-examined. Part II describes what algorithmic warfare is and what the key types of military AI used or being developed are. Part III examines the fundamental legal frameworks: IHL, the law of state responsibility and ICL. Part IV builds up the concept of an accountability gap and defines the technical and institutional causes of the gap. Part V focuses on the role of states, commanders and individual developers. Part VI is a survey of current normative developments, especially in the context of the UN. Part VII proposes a reformed, layered regime that is intended to maintain accountability and yet be able to accommodate technologically advanced systems.

II. ALGORITHMIC WARFARE AND MILITARY AI

A. Defining algorithmic warfare

The application of AI, machine learning and associated computational algorithms to execute functions historically performed by humans in military decision-making can be conceptualized as algorithmic warfare. Such functions involve sensor data analysis, pattern identification, target prioritization, or selection, and in a few applications, attack initiation or execution. The concept of war algorithms adopted by Harvard PILAC reflects this change by concentrating on algorithmically calculated decisions influencing armed conflict, executed by constructed systems that are appropriately competent.

[1] Nils Melzer, "Lethal Autonomous Weapons Systems & International Law" 29 ASIL Insights (2025), available at: https://asil.org/insights/volume-29-issue-1/ (last visited on May 2, 2026).
[2] Dustin A. Lewis, Naz K. Modirzadeh and Gabriella Blum, *War-Algorithm Accountability* (Harvard Law School Program on International Law and Armed Conflict, Aug. 31, 2016), available at: https://pilac.law.harvard.edu/war-algorithm-accountability-report (last visited on May 2, 2026).

Cadernos de Pós-Graduação em Direito Político e Econômico
Published by: Centro de Estudos Acadêmicos Press
ISSN: 1678-2127
Volume 26 *Issue S1*, 2026
Gurugram University Conference : 09-10 April 2026
Website: https://ceapress.org

This lens is explicitly broader than weapon platforms to a larger ecosystem of decision-support tools and command-and-control systems. It illuminates the growing algorithmic expression of power in war, where complicated models sieve information, offer courses of action and, in certain instances, give rise to kinetic or cyber effects.

B. Types of AI-enabled military systems

AI-enabled systems in the military domain can be grouped into three broad categories, each with distinct implications for responsibility:

- Autonomous weapon systems (AWS)/lethal autonomous weapon systems (LAWS): The type of system that, once triggered, has the capability to identify and attack targets without any additional human involvement. These bring the most acute accountability issues of all, as the most important select and engage role is transferred to the machine.[3]
- Semi-autonomous or supervised systems: Systems which autonomously execute some functions navigation, target tracking, threat assessment, but must have a human authorize or veto engagement. The question of lawfulness regarding this is to evaluate to what extent the use of algorithmic output is consistent with any meaningful human judgment.
- Non-kinetic AI systems: Decision-support systems, ISR systems and cyber capabilities which do not involve the use of force but have an impact on the manner and timing of the deployment of force, such as through target prioritisation or the production of recommended fire plans.

Within these categories, there are numerous systems that utilize machine-learning models that are trained on massive datasets. These models can be emergent, they may not generalise well when outside training conditions, or they may be susceptible to adversarial manipulation and their outputs may be hard to predict ex ante and to reconstruct ex post. These technical characteristics have a direct impact on the possibility of assigning legal responsibility of wrongful outcomes.

C. Ethical concerns and the centrality of human agency

The most common ethical arguments against autonomous weapons are that it is somehow wrong to leave decisions about life and death to the machines. Critics have stated that machines do not have moral agency, empathy, and the ability to make context-sensitive judgements, and their use threatens to dehumanize the armed conflict. ICRC has emphasized that obligations of IHL are held by humans and not machines and it has urged states to make

[3] Elke Schwarz, "The (im)possibility of meaningful human control for lethal autonomous weapon systems", available at: https://blogs.icrc.org/law-and-policy/2018/08/29/im-possibility-meaningful-human-control-lethal-autonomous-weapon-systems/ (last visited on May 2, 2026).

Cadernos de Pós-Graduação em Direito Político e Econômico
Published by: Centro de Estudos Acadêmicos Press
ISSN: 1678-2127
Volume 26 *Issue S1*, 2026
Gurugram University Conference : 09-10 April 2026
Website: https://ceapress.org

sure that autonomy in weapons systems does not exceed the boundaries that are compatible with humanitarian law and the demands of the common sense.[4]

This ethical discussion contributes directly to the legal discussion of what is meant by meaningful human control, which has become a key organising principle in the discussion of LAWS in the context of the CCW. The following passages analyze the ways in which legal norms that exist address these technological and ethical advancements.

III. EXISTING INTERNATIONAL LEGAL FRAMEWORKS

A. International humanitarian law and targeting rules

IHL governs the manner in which hostilities are to be conducted regardless of the weapons used. The fundamental principles difference, proportionality and precaution are applicable regardless of whether a strike is performed by the human-piloted plane or the autonomous system.[5]

- Distinction requires parties to distinguish at all times between civilians and combatants and to direct attacks only against the latter and other military objectives.
- Proportionality prohibits attacks expected to cause incidental civilian harm excessive in relation to the concrete and direct military advantage anticipated.
- Precautions in attack oblige attackers to take feasible steps to verify that targets are military objectives and to choose means and methods that minimise civilian harm.

Article 36 of Additional Protocol I further provides that new weapons should be legally reviewed by states under Article 36 in which states should establish whether the use of a new weapon, means or method would be outlawed by international law in all or all situations. The ICRC has highlighted that this is strictly true of AWS and other AI-based systems.

Formally, IHL is not technologically discriminating. The use of an AI-enabled system can be legitimate or not, depending on whether it is practically feasible to operate the system in accordance with these principles in the context of each attack. Nevertheless, the lack of transparency in machine-learning models, and unpredictable behaviours, makes ex ante assessment of reliability and discrimination more difficult and may lead to failure of states to meet their precautionary duties.

B. The law of state responsibility

[4] International Committee of the Red Cross, *Autonomous Weapon Systems under International Humanitarian Law*, available at: https://www.icrc.org/sites/default/files/document/file_list/autonomous_weapon_systems_under_international_humanitarian_law.pdf (last visited on May 2, 2026).
[5] Gerald Mako, "Legal Accountability for AI-Driven Autonomous Weapons", *Lieber Institute West Point*, Mar. 9, 2026, available at: https://lieber.westpoint.edu/legal-accountability-ai-driven-autonomous-weapons/ (last visited on May 2, 2026).

Cadernos de Pós-Graduação em Direito Político e Econômico
Published by: Centro de Estudos Acadêmicos Press
ISSN: 1678-2127
Volume 26 *Issue S1*, 2026
Gurugram University Conference : 09-10 April 2026
Website: https://ceapress.org

The International Law Commission's Articles on Responsibility of States for Internationally Wrongful Acts (ARSIWA) codify customary rules on when and how states incur responsibility for breaches of international obligations. The fundamental framework is straightforward: an act or omission is wrongful before international law, when it can be attributed to a state, and it is a violation of an international obligation.[6]

With ARSIWA, the acts of state organs, including armed forces, can be attributed to the state, as well as the acts of organizations that are entitled to exercise the means of state power. The technology employed to conduct the act does not matter on legal grounds: an IHL violation by AWS is prima facie an IHL violation by the state employing it.[7]

The existing doctrine does not acknowledge machines as subjects of international law. No room can therefore exist under existing principles to consider AWS as a distinct bearer of obligations or to rely on its autonomy as a reason to deny attribution. However, the attitudinal complexity of AI systems may influence the argumentation of attribution. States may, as an example, define harm as caused by unexpected malfunction, accident or force majeure, especially when learning systems act in a manner not reasonably predictable when deployed.

C. International criminal law and command responsibility

In relation to war crimes, crimes against humanity and genocide, ICL attributes responsibility to natural persons in the commission of such international crimes. Direct perpetration, ordering, aiding and abetting, indirect co-perpetration necessitates some form of conduct (actus reus) and culpable mental state (mens rea).

The command responsibility offers some type of liability to those in command who were aware or ought to have been aware that subordinates were and/or were going to commit crimes and did not do anything to stop or discipline them. They usually comprise superior-subordinate relationships, efficient control, knowledge (actual or constructive) and failure to perform required and reasonable acts.

These notions are strained in AI-mediated operations. With an autonomous system choosing and attacking a target without a human trigger pull, a human subordinate whose single action can be attributed to the commander may not exist. The question arises as to whether or not programmers, procurement officials, doctrine drafters or commanders may be said to have contributed to the crime with the necessary mens rea, due to the lack of transparency and emergent behaviour of complex models.

Some scholars propose reconceiving "meaningful human control" partly in criminal-law terms: as a set of conditions under which it is fair to attribute principal or derivative criminal

[6] Aisha M. Suleiman, "Legal Accountability for Autonomous Weapon Systems in Counterterrorism under International Law" 13(02) *International Journal of Science and Research Archive* 1604 (2024).
[7] Marta Bo, "Meaningful Human Control over Autonomous Weapon Systems: An (International) Criminal Law Account", *Opinio Juris*, Dec. 18, 2020, available at: https://opiniojuris.org/2020/12/18/meaningful-human-control-over-autonomous-weapon-systems-an-international-criminal-law-account/ (last visited on May 2, 2026).

Cadernos de Pós-Graduação em Direito Político e Econômico
Published by: Centro de Estudos Acadêmicos Press
ISSN: 1678-2127
Volume 26 *Issue S1*, 2026
Gurugram University Conference : 09-10 April 2026
Website: https://ceapress.org

responsibility to individuals who had a duty and a real ability to control the autonomous process.

IV. THE "ACCOUNTABILITY GAP" IN ALGORITHMIC WARFARE

A. Concept and manifestations

The accountability gap is the term used to describe cases where legal frameworks are formally applicable but in reality, do not hold any state or individual responsible for damages inflicted by AI-powered systems. It is the gap created between unlawful consequences that cannot readily be attributed to a blame-worthy human action or omission either due to the fact that system behaviour is not reasonably foreseeable, or due to the fact that the blame is diffusely distributed among many agents.[8]

According to the Lieber Institute, LAWS add layers of complexity due to their inscrutable algorithms and unpredictable adaptations, forcing it to be hard to prove intent or negligence on the part of commanders, operators or programmers. Khalil and Raj also point to issue of inexplainability and traceability at their core that creates gaps in accountability as traditional frameworks are challenged by autonomy and machine learning.

B. Technical opacity and evidentiary difficulties

Many machine-learning models are usually black-box: they can give highly accurate outputs, but without giving human-understandable explanations. Following an incident where civilian lives have been lost, the logs of the inputs and outputs can be available to the investigators without being able to understand why the system was mistaken on the classification of a civilian object as a military target.[9]

This opacity has several consequences:

- It complicates proof of foreseeability and knowledge, central to many criminal and civil liability standards.
- It can be exacerbated by proprietary interests or national-security classification, limiting access to training data and model architectures for courts or inquiry mechanisms.
- It may enable states to argue that they exercised due diligence in design and testing, yet could not reasonably have anticipated the particular failure mode that caused harm.

[8] Dustin A. Lewis, Naz K. Modirzadeh and Gabriella Blum, *War-Algorithm Accountability* (Harvard Law School Program on International Law and Armed Conflict, Aug. 31, 2016), available at: https://pilac.law.harvard.edu/war-algorithm-accountability-report (last visited on May 2, 2026).
[9] Amanda Musco Eklund, *Meaningful Human Control of Autonomous Weapon Systems: Definitions and Key Elements in the Light of International Humanitarian Law and International Human Rights Law* (Swedish Defence Research Agency, Feb. 2020).

Cadernos de Pós-Graduação em Direito Político e Econômico
Published by: Centro de Estudos Acadêmicos Press
ISSN: 1678-2127
Volume 26 *Issue S1*, 2026
Gurugram University Conference : 09-10 April 2026
Website: https://ceapress.org

In the absence of plausible access to technical evidence, victims and prosecutors might fail to satisfy burdens of proof, despite a high suspicion of system-level bugs or careless management.

C. Diffusion of agency across the AI life-cycle

The accountability of the behaviour of an AI system in theatre is shared down a long chain of decisions: data curation, model design, training, validation, integration into platforms, doctrinal development, rules of engagement, deployment decisions and real-time supervision. The various actors involved in each stage include engineers, defence contractors, military research units, legal advisers, commanders and operators, which may be in several jurisdictions.[10]

Such spreading agency renders it hard to distinguish one action or omission that is necessary and sufficient to the wrongful result. Rather, damage can be caused by the interplay of a variety of design decisions and operational suppositions. Older principles of complicity and joint criminal enterprise provide a means of dealing with concerted wrongdoing, but it does not sit well with technically intricate, temporally long development pipelines.

V. STATE RESPONSIBILITY, STRICT LIABILITY AND DUE DILIGENCE

A. Technology-neutral baseline

Strictly speaking, ARSIWA and IHL imply that the current rules are technology neutral. A state which has an AWS as part of its military forces has the duty of making sure that its use is in line with IHL and is subject to infractions notwithstanding whether these infractions are caused by human or algorithmic malpractice. The ICRC has stressed that States should be answerable under IHL due to the deployment of an autonomous weapon system and that this is a direct extension of general principles of state responsibility.[11]

In this perspective, it is no normative gap: the issue is implementation. States will be required to undertake strong Article 36 reviews, create doctrine and training specific to AI systems, and to make sure that these systems are utilized just in settings and methods that align with their verified dependability and discrimination.

B. Arguments for strict or heightened responsibility

However, certain academics and NGOs argue that the lack of predictability of learning systems and the inability to assess harm ex post, rather than ex ante, warrants a harsher responsibility regime in relation to harm caused by AWS. The initial intuition is that since machines per se cannot be subject of legal liability, the risk of malfunctions or emergent

[10] *Supra* note 6.
[11] Heather M. Roff, "Meaningful Human Control or Appropriate Human Judgment? The Necessary Limits on Autonomous Weapons" Briefing Paper for delegates at the Review Conference of the Convention on Certain Conventional Weapons (CCW), Geneva, 12-16 December 2016.

behaviour must be internalized by the deploying state as a matter of principle and not let uncertainty water down the sense of accountability.

Suggestions include more extreme liability of particular types of systems (e.g. full autonomy of offensive weapons in dynamic settings), or rebuttable inferences of responsibility which can only be overcome by demonstrating that the damage was solely caused by truly exogenous events. These models would guarantee victims a readily available means of reparation and would establish powerful motivations to hard testing, cautious deployment and investment in clarification.

Those opposing strict liability warn it may stifle innovation in systems that may lessen civilian fatalities compared to current weapons or promote over secrecy and underreporting. The compromise is to reinforce due-diligence requirements and consider grave shortcomings in design, testing or implementation as exacerbating elements in responsibility evaluations.

C. Due diligence and risk-management duties

Even without adopting strict liability, states clearly owe positive duties of due diligence in relation to high-risk technologies. Applied to military AI, these duties can be articulated as including obligations to:

- Conduct rigorous testing and validation under realistic, complex and adversarial conditions before deployment.
- Ensure that systems are only used within operational envelopes for which their performance has been reliably demonstrated.
- Incorporate mechanisms for human supervision, intervention and abort, consistent with meaningful human control.
- Monitor system performance in theatre and suspend or modify use if unexpected behaviours or reliability degradations are detected.
- Maintain audit logs and technical documentation sufficient to enable post-incident reconstruction and accountability.

Lack of such responsibilities can constitute negligence or recklessness that involves state responsibility despite the inability to recreate the precise algorithmic path to harm.[12]

VI. INDIVIDUAL CRIMINAL RESPONSIBILITY, COMMAND RESPONSIBILITY AND DEVELOPERS

A. Adapting criminal-law concepts to AI-mediated operations

[12] Human Rights Watch, "Submission to the United Nations Secretary-General on Autonomous Weapons Systems", available at: https://www.hrw.org/news/2024/05/06/submission-united-nations-secretary-general-autonomous-weapons-systems (last visited on May 2, 2026).

Cadernos de Pós-Graduação em Direito Político e Econômico
Published by: Centro de Estudos Acadêmicos Press
ISSN: 1678-2127
Volume 26 *Issue S1*, 2026
Gurugram University Conference : 09-10 April 2026
Website: https://ceapress.org

In the situations where AI is utilized by human actors to perpetrate crimes, conventional ICL doctrine continues to be relevant. A commander who intentionally uses an AWS to make indiscriminate attacks on civilian targets, such as, may be an indirect perpetrator or an orderer of war crimes.

The more difficult ones occur when damage is caused by a misclassification or unforeseen behaviour in systems which the commanders and engineers thought could be used lawfully. It is not easy to prove intent or knowledge in such situations. Some have suggested technology-specific adaptations, such as recognizing that acceptance of certain systemic risks e.g., deploying in environments known to exceed the system's tested capacities may satisfy the mental elements of recklessness or dolus eventualis

Khalil and Raj suggest such innovations as the introduction of such a requirement as intent-specific to technologies that specifically covers algorithmic decision-making and the introduction of new strict-liability regulations in instances of deployment failures of autonomous weapons. These concepts are somewhat controversial although they exemplify efforts to reconcile criminal-law concepts, with AI-mediated risk.

B. Re-thinking command responsibility through meaningful human control

A powerful school of thought connects meaningful human control to circumstances of fair criminal responsibility attribution. In this perspective, a duty to command ought to be ascribed whereby a superior was clearly obliged and had a realistic capacity to control or suspend the autonomous process but was not in fact so doing.

This implies that:

- The "subordinate" in AI contexts may be conceptualised as the unit or team responsible for operating or overseeing the system, rather than the system itself. Commanders must ensure those units have the expertise and authority to intervene.
- The "should have known" standard encompasses awareness of AI-specific risks documented in testing reports, Article 36 reviews and performance data. Access to such information may put commanders on notice of substantial risks.
- Duties to prevent include ensuring meaningful human control at relevant stages of the targeting cycle authorisation, supervision, intervention and review and establishing clear procedures for suspending use when anomalies are detected.

When there are red flags over system reliability and commanders deploy AI in different settings that have not been checked, their omissions can meet the requirements of command responsibility in instances when crimes occur.[13]

C. Liability of designers, engineers and corporate actors

[13] *Ibid.*

Cadernos de Pós-Graduação em Direito Político e Econômico
Published by: Centro de Estudos Acadêmicos Press
ISSN: 1678-2127
Volume 26 *Issue S1*, 2026
Gurugram University Conference : 09-10 April 2026
Website: https://ceapress.org

Military AI developers and manufacturers are more likely to be subject to domestic law contracts and tort regulations as opposed to international criminal law. But in other cases, they might be guilty as aiders and abettors or co-perpetrators when they knowingly or carelessly design systems in such a manner that violations of IHL is a foreseeable result.

An example would be that training a targeting algorithm on biased data that systematically falsely identifies particular civilian patterns as hostile or deliberately shutting off safeguards to be as lethal as possible, might form the basis of liability should the system be subsequently applied to perpetrate crimes. Khalil and Raj imagine the mechanisms of collective responsibility of multi-actor autonomous operations, as the project teams, corporate leadership, and military clients co-create the system behaviour.[14]

In practice, however, such prosecutions are improbable except in extreme cases due to the high evidentiary hurdles, jurisdictional constraints and the requirement to prove that a particular error was caused negligently and not willfully. More encouraging is the enhancement of civil and regulatory regimes with safety, transparency and reporting requirements to corporate actors.[15]

VII. RESPONSIBILITY OF PRIVATE ACTORS AND REGULATORY APPROACHES

A. Product liability, tort and regulatory standards

Legal systems in countries will already have legal frameworks of product-liability and negligence that, in principle, may be used to apply to the harm that autonomous weapons or decision-support systems may cause. Courts could take into account the failure of developers to adhere to reasonable standards in design, testing or caution about constraints.

However, there are a number of reasons why such mechanisms cannot be as effective in the context of armed conflicts: state and contractor sovereign immunity, the secretive nature of military technologies, the barriers to foreign victims in a court of law, and the challenge of disaggregating the role of particular design decisions in producing battlefield outcomes.

As a result, regulatory strategies that include export controls, certification schemes, and safety standards play a more significant role. States are able to mandate companies to facilitate the Articles 36 reviews, technical documentation, enable auditability functions and include meaningful human control in system designs as procurement or export licences.[16]

[14] Heather M. Roff, "Meaningful Human Control or Appropriate Human Judgment? The Necessary Limits on Autonomous Weapons" Briefing Paper for delegates at the Review Conference of the Convention on Certain Conventional Weapons (CCW), Geneva, 12-16 December 2016.
[15] Marta Bo, "Meaningful Human Control over Autonomous Weapon Systems: An (International) Criminal Law Account", *Opinio Juris*, Dec. 18, 2020, available at: https://opiniojuris.org/2020/12/18/meaningful-human-control-over-autonomous-weapon-systems-an-international-criminal-law-account/ (last visited on May 2, 2026).
[16] *Supra* note 3 at 15.

Cadernos de Pós-Graduação em Direito Político e Econômico
Published by: Centro de Estudos Acadêmicos Press
ISSN: 1678-2127
Volume 26 *Issue S1*, 2026
Gurugram University Conference : 09-10 April 2026
Website: https://ceapress.org

B. Soft law, industry self-regulation and scrutiny governance

The Harvard PILAC framework of war-algorithm accountability highlights a third axis of scrutiny governance: monitoring, oversight and norm-making mechanisms that are not necessarily based on formal liability. These are professional codes of conduct, transparency campaigns and civil-society campaigns, as well as independent monitoring bodies.

Consortia within the industry, and military ministries have started to express principles of responsible military AI reliability, predictability, traceability, human oversight which, although not binding, are capable of affecting practice and expectations of due diligence. Such a practice, particularly over time, can solidify into established norms or guide the application of the legal duty presently in existence.[17]

VIII. MEANINGFUL HUMAN CONTROL AND ENHANCED WEAPONS REVIEW

A. Conceptualising meaningful human control

"Meaningful human control" has become a focal concept in CCW deliberations, ICRC commentary and academic literature. Although no universally accepted definition exists, several common elements emerge:

- Humans must understand the system's functioning, capabilities and limitations.
- They must have sufficient information about the operational context to make informed decisions.
- They must retain the ability to supervise, intervene and override or abort operations.
- The effects of the system's actions must be reasonably foreseeable for human decision-makers.

The influential account of the ICRC has established three important components: human control and capability to intervene and deactivate; predictability and reliability; and relevant operational constraints. Systems which can modify their own objectives or evolve in such a manner as to subvert previous evaluations, are, in this sense, necessarily incompatible with significant human control.[18]

B. Operationalising meaningful human control

Operationalising meaningful human control requires context-specific criteria. These include limits on:

[17] Human Rights Watch, "Submission to the United Nations Secretary-General on Autonomous Weapons Systems", available at: https://www.hrw.org/news/2024/05/06/submission-united-nations-secretary-general-autonomous-weapons-systems (last visited on May 2, 2026).
[18] Elke Schwarz, "The (im)possibility of meaningful human control for lethal autonomous weapon systems", available at: https://blogs.icrc.org/law-and-policy/2018/08/29/im-possibility-meaningful-human-control-lethal-autonomous-weapon-systems/ (last visited on May 2, 2026).

- The types of targets: many argue that autonomous engagement should be restricted, if allowed at all, to clearly defined military objects such as fixed infrastructure, not individuals.
- The environments: dynamic, cluttered urban settings may be incompatible with existing recognition capabilities.
- The temporal and spatial scope of autonomous operation: longer and wider autonomy windows reduce the possibility of effective human oversight.
- Human–machine interface design: interfaces must convey information in ways that support comprehension rather than overwhelm operators.

This can be encoded in rules of engagement, fire-control doctrine and technical architectures such as the human confirmation of a target category, or by providing a human-in-the-loop or human-on-the-loop configuration with real intervention capability.[19]

C. Strengthening Article 36 weapons reviews

Article 36 reviews offer an important institutional site of incorporating significant human agency into the state practice. In the case of AI systems, such reviews should extend beyond the conventional assessment of blast radius or accuracy to review the performance of the algorithm, training data, robustness and human-machine interaction.[20]

Recommended enhancements include:

- Testing under realistic, including adversarial, conditions that reflect operational environments.
- Evaluation of performance across the full range of envisaged conditions, not only idealised scenarios.
- Systematic assessment of training data for biases that might affect distinction and proportionality.
- Review of user interfaces and control mechanisms to ensure that claimed human control is actually meaningful in practice.
- Mandated logging, explainability and post-deployment monitoring plans.

By refusing to approve systems that fail these enhanced reviews, states can prevent particularly problematic forms of autonomy from entering service.

IX. NORMATIVE DEVELOPMENTS UNDER THE UNITED NATIONS AND BEYOND

[19] International Committee of the Red Cross, *Autonomous Weapon Systems under International Humanitarian Law*, available at: https://www.icrc.org/sites/default/files/document/file_list/autonomous_weapon_systems_under_international_humanitarian_law.pdf (last visited on May 2, 2026).
[20] Amanda Musco Eklund, *Meaningful Human Control of Autonomous Weapon Systems: Definitions and Key Elements in the Light of International Humanitarian Law and International Human Rights Law* (Swedish Defence Research Agency, Feb. 2020).

Cadernos de Pós-Graduação em Direito Político e Econômico
Published by: Centro de Estudos Acadêmicos Press
ISSN: 1678-2127
Volume 26 *Issue S1*, 2026
Gurugram University Conference : 09-10 April 2026
Website: https://ceapress.org

A. CCW GGE on LAWS

Since 2013, the Group of Governmental Experts on LAWS of the CCW has been the main multilateral venue of the debate on autonomous weapons. Guideline principles have been embraced by the GGE, which state that IHL applies wholly to all weapons systems, that the decision-making authority on the use of weapons should be left to humans and that accountability should be upheld.[21]

States have argued over regulatory alternatives including legally enforceable treaties with prohibitions and restrictions to political statements and interpretations. Although the debate on the subject is divided, many states, the UN Secretary-General and civil-society alliances are increasingly supporting the banning of fully autonomous weapons that are not under meaningful human control and ensure strict control of other types of autonomous systems.

B. General Assembly and Secretary-General initiatives

Recent UN General Assembly resolutions have urged states to reflect on a normative and operational framework on LAWS and asked the Secretary-General to report on views and recommend a way forward. These processes have been employed by human rights watch and others to point out the gaps in accountability and to urge treaty-based prohibitions and positive, strong obligations, such as meaningful human control and victim-based remedies.[22]

Fully autonomous weapons have been described as being morally repugnant and politically unacceptable by the Secretary-General who has urged states to negotiate legally binding instruments to ban them. Such statements are not binding, but may have normative influence on state practice and expectations.

C. Regional and national approaches

National principles of responsible military AI focusing on human responsibility, reliability, traceability and governability have been implemented in some states and regional organizations. Others have indicated their support of moratoria or prohibition of specific types of autonomous weapons. These various efforts would serve as best practice laboratories yet may otherwise become fragmented and regulatory arbitrage without coordination.

X. TOWARDS A RECONFIGURED LEGAL APPROACH

A. Reaffirming state responsibility and victim-centred remedies

[21] Nils Melzer, "Lethal Autonomous Weapons Systems & International Law" 29 ASIL Insights (2025), available at: https://asil.org/insights/volume-29-issue-1/ (last visited on May 2, 2026).
[22] Human Rights Watch, "Submission to the United Nations Secretary-General on Autonomous Weapons Systems", available at: https://www.hrw.org/news/2024/05/06/submission-united-nations-secretary-general-autonomous-weapons-systems (last visited on May 2, 2026).

Cadernos de Pós-Graduação em Direito Político e Econômico
Published by: Centro de Estudos Acadêmicos Press
ISSN: 1678-2127
Volume 26 *Issue S1*, 2026
Gurugram University Conference : 09-10 April 2026
Website: https://ceapress.org

A central pillar of any reconfigured approach must be the reaffirmation that states remain fully responsible for uses of force undertaken by or on their behalf, regardless of the degree of autonomy of the systems used. This suggests:

- A presumption that harm caused by deployed AI weapons is attributable to the deploying state, subject to narrow exceptions.
- A duty to provide effective remedies and reparation for unlawful harm, potentially through specialised claims mechanisms with simplified evidentiary requirements where AI systems are involved.
- A commitment not to invoke the unpredictability of AI as a defence where states voluntarily chose to deploy such systems in conditions of substantial uncertainty.

Such measures would ensure that victims are not left without recourse, while preserving incentives for states to internalise the full costs of algorithmic warfare.[23]

B. Codifying meaningful human control as a binding constraint

Although aspects of meaningful human control can be made out of the current IHL, there is a good argument to codify a clear duty to ensure meaningful human control over the use of force, particularly in a future treaty on LAWS. Such requirement would specify minimum requirements of legal autonomy that would cover restrictions on target types, environments, timeframes, and necessary human monitoring abilities.

Both operationalization of ethical issues and clear legal foundation of evaluating responsibility in situations where systems are exploited to effectively marginalize human judgement would be achieved by embedding meaningful human control as a major rule of conduct.[24]

C. Clarifying due-diligence and command-responsibility standards

To narrow the accountability gap, due-diligence obligations and command-responsibility standards must be interpreted in light of AI-specific risks. This entails:

- Interpretive guidance from bodies such as the ICRC and international courts on how precaution, proportionality and command responsibility apply to AI-mediated operations.
- Military manuals and doctrine that spell out AI-specific precautions, including mandatory red-teaming, adversarial testing and continuous monitoring.
- Training for commanders and legal advisers to ensure they can meaningfully assess AI capabilities and limitations.

[23] *Supra* note 6.
[24] *Supra* note 3.

Leaders who ignore established constraints, use systems that are not within their verified envelopes or do not provide meaningful human control ought to be construed as violating their responsibilities, and they should be held liable in cases where criminal acts are committed.

D. Strengthening corporate responsibility and technical governance

Since the central role in creating military AI belongs to the private actors, states ought to reinforce regulatory frameworks, which demand safety, transparency, auditability and compliance with IHL throughout all the design stages. Export restrictions, certification programs and procurement policies can be very potent tools to ensure that systems delivered to militaries incorporate significant levels of human control and accountability.

Simultaneously, the weaker mechanisms may strengthen formal law, by developing the cultures of accountability in the AI and defense sectors, through the industry codes, professional ethics standards, and independent audits.[25]

XI. CONCLUSION

Pressure points in a legal architecture constructed around human agency, intention and control are revealed in algorithmic warfare. The law of state responsibility and IHL are applicable, yet their practical functioning is difficult due to the opaque, probabilistic and distributed decision processes. Unattended, these challenges have a potential to create an accountability gap, where victims of AI-mediated harm find it hard to locate the responsible duty-bearers and seek redress.

An integration of responsibility in algorithmic warfare should be layered in a coherent framework. It must renew state accountability and victim-based solutions; apply individual criminal and command-responsibility principles to AI-relevant risks; and apply regulatory and soft-law requirements to the entire life cycle of military AI, with or without its own developers and manufacturers. At the centre of this framework is the idea that not algorithms but humans should be responsible of using force. Meaningful human control, review of weapons, transparency and auditability in both the legal and technical architectures provide a way to balance technological innovation with the long-term commitments of international law to human dignity and responsibility.

[25] *Ibid.* at 20.