

Between Review, Transfer and Principle: The Limits of Article 36, the Arms Trade Treaty and the Martens Clause in AI Warfare

Dr. Anviksha Pachori*

ABSTRACT

The integration of artificial intelligence (AI) into weapon systems exposes serious accountability gaps in existing international legal frameworks. Article 36 of Additional Protocol I requires states to review new weapons for compliance with international humanitarian law, yet vague definitions and rapid AI integration blur their scope and weaken consistency. Doctrinal analysis reveals uneven state practice and limits standardized review processes amid growing autonomy and adaptability in weapons. The Arms Trade Treaty offers parallel safeguards but struggles with dual-use technologies and classification disputes. The Martens Clause provides interpretive guidance, yet gaps persist, underscoring the urgent need for clearer, adaptive frameworks governing AI-driven warfare

Keywords: Artificial Intelligence, Autonomous Weapons, Military, Technology, International Law

Introduction

*“The genie is out of the bottle. The question is not whether AI will be used in warfare, but whether it will be governed.” – Paul Scharre, *Army of None*.¹*

Artificial intelligence (AI) has emerged as a transformative force in contemporary warfare, posing profound challenges for the governance of weapons systems and the attribution of responsibility under international law. As states increasingly integrate AI into military decision-making, long-standing legal frameworks designed for the human-directed conduct of hostilities are being tested by systems capable of autonomous target selection, analysis, and action. Political leaders have already

*Associate Professor, G D Goenka, School of Law, New Delhi, India. Email-anvikshapachori@gmail.com. Orchid ID: <https://orcid.org/0000-0003-2314-5598>. I would like to express my sincere gratitude to Dr Jonathan Kwik, T.M.C. Asser Institute; Dr Samuel White, Centre for International Law, NUS; and Prof Sean Watts, Lieber Institute, for their invaluable feedback. Views are personal.

¹ Paul Scharre, *Army of None: Autonomous Weapons and the Future of War*, WW Norton & Company, New York and London, 2018.

proclaimed AI superiority as a determinant of global power,² prompting major military actors, including the United States and China, to accelerate investment in AI-integrated weapons systems.³ This rapid militarization of AI has outpaced the development of corresponding legal and institutional mechanisms to ensure accountability. Notably, few nuclear-armed states have also pivoted towards integrating and developing advanced AI to bolster early warning, i.e., Intelligence, Surveillance and Reconnaissance (ISR), which could have consequences for strategic stability and the dominance of nuclear-powered countries.⁴

AI-integrated weapons are already altering the conduct of hostilities in ways that complicate compliance with international humanitarian law (IHL). In the Russia-Ukraine conflict, AI-enabled drones capable of autonomous navigation and target identification have reduced the immediacy of human control over targeting decisions, while Russia's deployment of systems such as the *Abzats*⁵ anti-drone platform exemplifies the mounting reliance on algorithmic threat detection and response. This active deployment of ground robotics platforms on the battlefield of the Russia-Ukraine war has set a trend. More so in the Israel-Hamas conflict, systems including the Iron Dome⁶ operate with a high degree of autonomy in detecting and intercepting threats, while AI-driven decision-support tools, such as *Habsora* and *Lavender* have contributed to target selection through algorithmic analysis of large datasets.⁷ In the same line, South Asia has also witnessed its first AI-integrated drone-led war between India and Pakistan.⁸ India launched drones like Shahpur-II with AI sensors which could identify strategic targets and adjust their trajectory before sending the coordinates to home systems.⁹ In fact, the recent escalation in the US-Iran war illustrates that AI systems such as Project Maven do

² Patrick Caughill, "Vladimir Putin: Country that Leads in AI Development will be the Ruler of the World", *Futurism*, 30 November 2017, available at: <https://futurism.com/vladimir-putin-country-that-leads-in-ai-development-will-be-the-ruler-of-the-world>.

³ Michael C. Horowitz, "Military AI Adoption is Outpacing Global Cooperation", *Council on Foreign Relations*, 11 February 2026, available at: <https://www.cfr.org/articles/military-ai-adoption-is-outpacing-global-cooperation>.

⁴ Vladislav Chernavskikh, *Nuclear Weapons and Artificial Intelligence: Technological Promises and Practical Realities*, Background Paper, Stockholm International Peace Research Institute, September 2024, available at: https://www.sipri.org/sites/default/files/2024-09/bp_2409_ai-nuclear.pdf.

⁵ Christopher McFadden, "Abzats: Russia Deploys AI-powered Tiny Trucks to jam Ukrainian Drones", *Interesting Engineering*, 3 May 2024, available at: <https://interestingengineering.com/innovation/russia-deploys-new-abzat-autonomous-drone-jamming-drone>.

⁶ "Israel-Hamas conflict: What is Iron Dome?", *BBC*, 10 October 2023, available at: <https://www.bbc.com/news/articles/ck5xyz29e47o>.

⁷ Julius Kurek and Björn Laurin Kühn, *Habsora (הבשורה) and Lavender (לַוֵּנְדֵּר) Artificial Intelligence Systems*, Atlantic Forum, 19 September 2024, available at: <https://www.atlantic-forum.com/our-views/habsora-and-lavender-artificial-intelligence-systems>.

⁸ Amoha Basrur, *The Use of Drones Marks a New Phase in India-Pakistan Hostilities*, Observer Research Foundation-Raisina Debates, 13 May 2025, available at: <https://www.orfonline.org/expert-speak/the-use-of-drones-marks-a-new-phase-in-india-pakistan-hostilities>.

⁹ Harshul Sharma, "India vs Pakistan just Fought the First War Coded by Machine and Not Men", *PC Quest*, 13 May 2025, available at: <https://www.pcquest.com/artificial-intelligence/india-vs-pakistan-just-fought-the-first-war-coded-by-machines-not-men-9062672>.

not merely assist military operations but also restructure the temporal and epistemic conditions under which targeting decisions are made, thereby significantly increasing strike operations against Iran.¹⁰ These developments raise pressing questions regarding the allocation of legal responsibility when AI systems influence or constrain human judgement, particularly in relation to the principles of distinction, proportionality, and precaution.

This article sets out to explore whether the existing international legal framework, *i.e.*, Article 36 of Additional Protocol I,¹¹ Arms Trade Treaty¹² (ATT) and Martens clause could meaningfully govern the integration of AI weapons. Article 36 presumes weapons whose behaviour can be assessed *ex ante*; AI-enabled systems evolve after deployment. Autonomous weapons systems, especially integrated with AI, should, like any other weapons systems, be subject to thorough legal review. This paper argues that reviews must extend beyond initial acquisition to encompass software, retraining processes, and operational modifications. One-time static certification is incompatible with evolving technologies. While this approach is already embedded in the International Committee for the Red Cross (ICRC) guidelines,¹³ it should be treated as a legal requirement rather than just as “best practices”. This is discussed in detail under section 2.

The ATT regulates discrete transfers of tangible arms. It assumes that arms transfers occur as discrete, identifiable events. But when a shipment leaves one jurisdiction to enter another, AI-enabled systems disrupt this model through what might be called continuous technological diffusion.¹⁴ Thus, it demands epistemic certainty in a domain defined by technical uncertainty.¹⁵ The more autonomous and adaptive the system, the less meaningful the Treaty’s predictive calculus becomes.¹⁶ In the AI context, compliance with the ATT may satisfy formal legality while leaving substantive humanitarian protection unchanged. That gap underscores the need for

¹⁰ “Inside Palantir’s Maven: The AI Tool Powering US Strikes in Iran”, *The Times of India*, 27 April 2026, available at: <https://timesofindia.indiatimes.com/defence/international/middle-east-war-inside-palantirs-maven-the-ai-tool-powering-us-strikes-in-iran/article-show/130183378.cms>

¹¹ Protocol Additional (I) to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of International Armed Conflicts, 1125 UNTS 3, 8 June 1977 (entered into force 7 December 1978), Art. 36.

¹² Arms Trade Treaty, UNR 67/234B, 2 April 2013.

¹³ International Committee of the Red Cross, *Autonomy, Artificial Intelligence and Robotics: Technical Aspects of Human Control*, Geneva, August 2019, available at: https://www.icrc.org/sites/default/files/document/file_list/autonomy_artificial_intelligence_and_robotics.pdf.

¹⁴ Paul Holtom and Mark Bromley, *Transit and Trans-shipment Controls in an Arms Trade Treaty*, Stockholm International Peace Research Institute, July 2011, available at: <https://www.sipri.org/publications/2011/sipri-background-papers/transit-and-trans-shipment-controls-arms-trade-treaty>.

¹⁵ Tsvetelina Van Bentham, “Battlefield Uncertainties and the Use of AI”, *International Law Studies*, Vol. 107, 2026, available at: <https://digital-commons.usnwc.edu/cgi/viewcontent.cgi?article=3129&context=ils>.

¹⁶ Lethal Autonomous Weapons Systems: Report of the Secretary-General, A/79/88, 1 July 2024, available at: <https://docs.un.org/en/A/79/88>.

design-oriented, continuous accountability mechanisms, which are advanced later in this paper. Crucially, this shift toward design-stage responsibility does not require the invention of new legal norms. It flows directly from existing IHL principles.

The cumulative analysis of this paper leads to a clear conclusion that Article 36 reviews alone cannot anticipate adaptive behaviour. Export controls cannot track intangible diffusion. The Martens Clause cannot substitute for enforceable rules. What remains is not a regulatory vacuum but a regulatory reorientation; compliance must migrate upstream, from procedure to architecture, from review to design. IHL-by-Design provides this reorientation. It translates enduring humanitarian principles into technical and institutional constraints capable of operating at the speed and complexity of AI-enabled warfare. Without such integration, existing frameworks risk devolving into formalities that legitimate deployment while leaving civilians exposed to algorithmically mediated harm. In short, if states remain responsible for the conduct of autonomous systems under international law, they must also be responsible for how those systems are built. The age of AI warfare, therefore, does not demand more norms and restrictions, but law embedded in its code and more transparency.

1. Autonomy of Weapon as a legal Disrupter

When IHL conceptualized weapons, it perceived weapons as a discrete tangible object whose effect and modes of operation were sufficient to have *ex-ante* legal evaluation. This assumption is embedded in Article 36 of the Additional Protocol I¹⁷ which requires states to assess whether “a new method and means of warfare” would be prohibited under international laws. Based on article 36, it underpins core regulatory mechanism which presumes that a weapon’s design, function and foreseeable effects can be assessed prior to deployment. AI enabled weapon systems fundamentally disrupt this premise. Unlike conventional weapons, their legal salient features do not reside primarily in their physical or kinetic performance, but in their complex computational decision-making processes that mediate perception, select, target, and engage with other objects and humans. Their behaviour is not solely determined at the point of manufacture but is continuously shaped by the data, updates, behaviour, observance and interactions.¹⁸ As ICRC reports observed, autonomy in weapon systems shifts the locus of control from direct human decision making to algorithmic functions that operates in dynamic environments.¹⁹ Consequently, the legal “object” of regulations shifts from a weapon thing to a decision-making process embedded within a broader system. This reconceptualization raises immediate challenges for IHL, which lacks doctrinal tools

¹⁷ Additional Protocol I, above note 11, p. 4.

¹⁸ Aparna Krishna Bhat, “The Evolution of AI: From Foundations to Future Prospects”, *IEEE Computer Society*, 11 March 2025, available at: <https://www.computer.org/publications/tech-news/research/evolution-of-ai>.

¹⁹ ICRC Guidelines, above note 13, p. 4.

to regulate weapons whose most consequential features lie in the decision-making process rather than in their kinetic effects alone.

Under IHL, foreseeability is central to the lawful application of the principles of distinction and proportionality. Commanders are required to anticipate the effect of attacks, assess civilian harm and adjust conduct accordingly. On one hand, traditional weapons regulation assumes that such anticipation is grounded in weapons' known behaviour, on the other hand, AI-enabled systems, particularly those incorporating machine learning (ML), undermine this premise by introducing probabilistic and adaptive behaviour²⁰ that cannot be fully specified in advance. While proponents of Military AI often argue that such systems remain "predictable within parameters", this technical reassurance does not resolve the legal problem from an IHL perspective, a partial predictability is insufficient where legal responsibility hinges on reasonably foreseeable consequences. It is pertinent to note that increasing autonomy risks creating a "responsibility gap"²¹ insofar as human operators lack sufficient understanding of how a system will behave in concrete operational settings. In the case of the ongoing contemporary wars, Russia and Ukraine are deploying loitering munitions such as drones that can patrol an area for extended periods and engage targets with varying degrees of autonomous guidance. The Iranian-designed *Shahed-136*, widely deployed by Russian forces, operates as a pre-programmed loitering munition capable of autonomously conducting surveillance and strike missions, akin to a long-range cruise missile, within an approximate 150 km radio-control range.²² Similarly, Israel's defence forces have operationalized sophisticated AI systems such as Lavender, which automatically process vast amounts of surveillance and intelligence data to identify individuals and structures for targeting.²³ It maintains a large database of individuals linked to prospective militant groups, reportedly recommends effective targets using AI decision-support systems (AI-DSS), and aids in complex decision-making by processing large volumes of relevant information.

Critics argue that this AI-driven process blurs the distinction between "autonomy" and "intentions". Especially where human review is also limited or perfunctory, approaching this from a legal viewpoint, terms like "understanding" and "intentions" can be narrowly defined and have a huge space for implicit

²⁰ Major Vincenzo Gallitelli, "Artificial Intelligence-Enabled Military Decision-Making Process: The Forgotten Lessons on the Nature of War", *Journal of Advanced Military Studies*, Vol. 16, No. 2, 2025, p. 99.

²¹ Report of the Special Rapporteur on Extrajudicial, Summary or Arbitrary Executions, Chistof Heyns, UN Doc. A/HRC/23/47, 9 April 2013.

²² Bintari Rahmanita, "Specifications of Iran-Made Shahed-136 'Kamikaze' Drone", *TEMPO English*, 16 February 2025, available at: <https://en.tempo.co/photo/126274/specifications-of-iran-made-shahed-136-kamikaze-drone>.

²³ Emelie Andersin, "The Use of the 'Lavender' in Gaza the law of Targeting: AI-Decision Support Systems and Facial Recognition Technology", *Journal of International Humanitarian Legal Studies*, Vol. 16, 2025, p. 336.

assumptions.²⁴ IHL adopts a definition of intent that incorporates “awareness” that occurs in the ordinary course of events.²⁵ An act can be reckless or intentional, but will only exist when the system engages in or deals with conduct voluntarily triggered by an algorithm, leading to the attack of a civilian, not capable of distinguishing it from a military objective, leading to the death or injury of an innocent. In other words, Samuel and Brendon have defined it as an alignment problem in their paper, suggesting “a concept in the field of AI that describes the challenge of aligning the goals and behaviour of AI systems with the designer or operator’s intention”.²⁶ In this use case, AI doesn’t merely assist human judgement by drastically narrowing the universe of targetable entities based on algorithmic inferences, it reconfigures the targeting process itself. When AI output significantly influences strike decisions, the legal burden shifts.

1.1. The Fragmentation of Agency and Dilution of Responsibility

The fundamental disruption introduced by AI-enabled weapons lies in the fragmentation of agency. At present, IHL and International Criminal Law (ICL) are premised on a relatively linear chain of responsibility where engineers build weapons, States deploy them, commanders decide, and operators execute. AWS disrupts this sequence by distributing decision-making across verticals from engineers, data trainers, procurement authorities, to algorithmic processes at machine speed. It not only obscures the question of responsibility but also blurs the line between the combatants and the role of civilians in a conflict.²⁷ Critics have argued, the delegation of critical functions to machines complicates the attribution of conduct and intent,²⁸ particularly where unlawful outcomes arise from emergent system behaviour rather than explicit human instruction. This does not entirely dismiss the State’s responsibility under the law of State responsibility; however, it significantly weakens the practical mechanism of accountability, especially in the absence of transparency regarding system design and operational parameters.

AI-enabled weapons systems also disrupt the temporal assumptions underpinning weapons regulation. Article 36 reviews are typically conducted at a discrete point prior to adoption or deployment on the assumption that the weapon’s characteristics will remain materially constant. AI systems, however, evolve through software updates, retaining and contextual learning often after deployment. As the

²⁴ International Criminal Tribunal for former Yugoslavia, *Prosecutor v. Zejnir Delalic*, Case No. IT-96-21-T, Judgement, 1 November 1996.

²⁵ Magdalena Pacholska, “Military Artificial Intelligence and the Principle of Distinction: A State Responsibility Perspective”, *Israel Law Review*, Vol. 56, No. 1, p. 3. 2022, available at: <https://doi.org/10.1017/S0021223722000188>.

²⁶ Samuel Hartridge and Brendan Walker-Munro, “Autonomous Weapons Systems and the AI Alignment Problem”, *Journal of International Humanitarian Legal Studies*, Vol. 16, 2025, p. 38.

²⁷ Andreas Wenger and Simon J.A. Mason, “The Civilianization of Armed Conflict: Trends and Implications”, *International Review of the Red Cross*, Vol. 90, 2012, p. 872.

²⁸ Michael N. Schmitt and Jeffery S. Thurnher, “Out of the loop: Autonomous Weapon Systems and Law of Armed Conflict”, *Harvard National Security Journal*, Vol. 4, No. 2, 2013, p. 231.

ICRC emphasized that systems capable of modifying their behaviour raise questions about whether legal review should be understood as a continuous obligation rather than a one-off procedural step.²⁹ The temporal fluidity challenges the binary distinction between lawful and unlawful weapons. A system that appears compliant at the design stage may behave differently in a specific operational environment, raising questions about when legal responsibility crystallizes and whether compliance can be meaningfully assessed at a single time.³⁰ This misalignment does not imply a legal vacuum, but it does necessitate a shift in regulatory focus from static weapon assessments to process-based, continuous and design-oriented regulation. Without such a shift, mechanisms like Article 36 risk functioning as symbolic safeguards rather than effective tools of accountability.

2. Review of Weapons under Article 36 and State Practices

It is an established norm by the International Court of Justice (ICJ) in its *Advisory Opinion on the Legality and Threat or Use of Nuclear Weapons* that, “IHL permeates the entire law of armed conflict and applies to all forms of warfare and to all kinds of weapons, those of the past, those of the present and those of the future.”³¹ The ICJ also noted that this prohibition ensures that “States do not have unlimited freedom of choice of means in the weapons they use,”³² and declared the rule as one of the cardinal principles contained in the texts constituting the fabric of IHL. Legal reviews are frequently portrayed as a foundation of the domestic implementation of IHL. Although Article 36 imposes a binding obligation on States Parties to determine the legality of new weapons, means, and methods of warfare, the practical effectiveness of this mechanism remains uneven and heavily dependent on national discretion.³³ In the absence of uniform international standards, independent oversight, or reporting obligations, Article 36 reviews often risk becoming opaque procedural exercises rather than meaningful accountability mechanisms.³⁴ Only a limited number of States, including the Netherlands, Sweden, Switzerland, and Australia, have publicly disclosed aspects of their weapons review procedures, including their potential application to autonomous or AI-enabled weapon

²⁹ Kathleen Lawand, *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*, ICRC, Geneva, 2006.

³⁰ Jonathan Kwik, “Iterative Assessment for Military Artificial Intelligence (AI) Systems” in Bérénice Boutin, Taylor Kate Woodcock and Sadjad Soltanzadeh (eds), *Legal, Ethical, and Technical Dilemmas in Military Artificial Intelligence*, T.M.C. Asser Press, The Hague, 2026, available at: <https://doi.org/10.1007/978-94-6265-759-5>.

³¹ International Court of Justice, *Legality of the Threat or Use of Nuclear Weapons*, Advisory Opinion, *ICJ Reports* 1996, 8 July 1996, para. 86.

³² *Ibid.*, para 78.

³³ Additional Protocol I, above note 11, p. 4; K. Lawand, above note 29, p. 6.

³⁴ Dan Saxon (ed.), *International Humanitarian Law and the Changing Technology of War*, Martinus Nijhoff, Leiden, 2013.

systems.³⁵ Such disclosures, however, are largely confined to general descriptions of institutional processes and legal criteria, rather than substantive assessments of specific systems or technologies. Consequently, while transparency in the conduct of legal reviews is frequently invoked as a confidence-building measure, the restricted visibility into review methodologies, assumptions, and outcomes significantly constrains their capacity to ensure effective compliance with treaty and customary IHL governing the means and methods of warfare,³⁶ particularly in the context of increasingly autonomous and adaptive weapons systems.

Article 36 of Additional Protocol I is often portrayed as the foundation of lawful weapons regulation in emerging technological contexts. It obliges states to determine, during the “research, development, acquisition or adoption” of a new weapon, whether its use should be prohibited under international law. This obligation is preventive in nature, seeking to internalize compliance with IHL before harm occurs. However, this preventive logic rests on a critical but under-examined assumption that the weapon under review is stable and knowable characteristics to permit a meaningful legal assessment *ex ante*.³⁷ AI-enabled weapons directly undermine this assumption. Unlike conventional weapons, whose operational parameters remain largely fixed after deployment, AI-driven systems may alter behaviour through learning, environmental interaction or software modification. As a result, invoking the concept of “meaningful human control” cannot be determined in AWS as a range of potential future behaviours many of which cannot be fully anticipated at the point of review. The problem, therefore, is not that Article 36 is ignored by many States but that it is asked to perform a function for which it was never designed. ICRC has also acknowledged that meaningful human control functions more as a policy aspiration than a legal standard, offering limited guidance for concrete review decisions.³⁸

Perhaps the most significant limitation of Article 36 in the context of AI lies in temporal rigidity. Weapons reviews are typically conducted once, prior to the

³⁵ Netherlands Ministry of Defence, Advisory Report on Autonomous Weapon Systems (2016); Also See, “Statement by Sweden at the CCW GGE on Lethal Autonomous Weapons (LAWS)”, *UNODA*, 21-25 September 2020, available at: https://documents.unoda.org/wp-content/uploads/2020/09/200921-Anforande-LAWS-human-control_.pdf; A “compliance -based” approach to Autonomous Weapon System, CCW/GGE.1/2017/WP.9, 10 November 2017, available at: [https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-_Group_of_Governmental_Experts_\(2017\)/2017_GGEonLAWS_WP9_Switzerland.pdf](https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-_Group_of_Governmental_Experts_(2017)/2017_GGEonLAWS_WP9_Switzerland.pdf); The Australian Article 36 Review Process, CCW/GGE.2/2018/WP.6, 30 August 2018, available at: [https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-_Group_of_Governmental_Experts_\(2018\)/2018_GGE%2BLAWS_August_Working%2Bpaper_Australia.pdf](https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-_Group_of_Governmental_Experts_(2018)/2018_GGE%2BLAWS_August_Working%2Bpaper_Australia.pdf).

³⁶ International Committee of the Red Cross, “Autonomous Weapon Systems: Implications of Increasing Autonomy in the Critical Functions of Weapons”, 1 September 2017, available at: <https://www.icrc.org/en/publication/4283-autonomous-weapons-systems>; Rebecca Crootof, “The Killer Robots Are Here: Legal and Policy Implications”, *Cardozo Law Review*, Vol. 36, 2015, p. 1837.

³⁷ Additional Protocol I, above note 11, p. 4.

³⁸ ICRC, above note 13, p. 8.

deployment and rarely revisited unless a weapon undergoes substantial modification. On the other hand, AI-embedded systems are evolving continuously. This evolution may materially affect how the system identifies targets, prioritizes threats, detects or executes force. The legal consequences of this are stark i.e. when a weapon deemed lawful at time T1 may behave illegally at time T2; the weapon being properly reviewed but the illegality not being foreseeable. The critics have argued that treating legality as a static attribute of a weapon is increasingly untenable³⁹ where the decisive features of the system are dynamic.

Without a doctrine of continuous review or post-deployment accountability, Article 36 risks becoming disconnected from operational realities of AI-governed weapons. The shortcomings of Article 36 in relation to AWS should not be understood as failures of implementation alone; rather, they expose a structural inadequacy rooted in the provisional reliance on stability, foreseeability and one-time review temporally fixed assessments. These assumptions are incompatible with technologies that re-distribute decision-making authority across machines, humans and time. But this does not render Article 36 irrelevant; rather, it necessitates an evolution in how it is understood and applied. It just requires an evolution of its own understanding and application. In the further sections, the paper captures the reconceptualization of Article 36's function from one-time legality to a continuous, design-oriented responsibility approach.

2.1. Procedural legitimacy and constraint under Article 36

State practice relating to Article 36 of Additional Protocol I demonstrates a striking paradox where, on one hand, the obligation to conduct national weapon reviews has achieved near-universal normative acceptance, but on the other hand, with AWS, its practical operation remains fragmented, opaque and uneven. The cumulative evidence suggests that Article 36 functions less as a robust constraint on unlawful or destabilizing weapons and more as a procedural assurance mechanism whose effectiveness depends heavily on domestic capabilities to perform the evaluation and the willingness of the nation. As an aggravating factor the complexities of AWS are growing each day, the more complex the architecture of the system the harder and more expensive it becomes to perform separate tests and evaluations.⁴⁰ The complexity of the environment in which the system is intended to operate is another important variable that further muddles and increases the cost of testing and evaluation which in turn is directly linked with the budget of the country and its capacity to sustain that cost of further analysis. The current assessment under Article 36 is neither practical nor financially prudent for AI-enabled AWS using the existing

³⁹ Marco Sassoli, *International Humanitarian law: Rules, Controversies and Solutions*, 2nd ed., Edgar Elgar Publishing, Northampton, 2019.

⁴⁰ "Autonomous Weapon Systems: Technical, Military, Legal and Humanitarian Aspects", *International Review of the Red Cross*, 28 March 2014, available at: https://www.icrc.org/sites/default/files/document/file_list/4221-002-autonomous-weapons-systems-full-report.pdf.

testing and evaluation method.⁴¹ It is paramount to develop new methodologies for testing and evaluating new weapons. Some states, such as the United States, are currently conducting important research and development efforts to address the issue. For AI-enabled weapons, a system should report on its learning and modifications throughout its lifecycle, allowing the user to obtain a new review mechanism.⁴² At the moment, the creation of such “self-aware” or “self-explainable” AWS remains at the stage of research itself.

2.2. State Practices Developed under Article 36

Although at the level of principle, the contestation on Article 36 is minimal. A large majority of participating States affirm that they are legally bound to review new weapons, means and methods of warfare. Few countries treat the obligation as reflective of customary international law.⁴³ Institutionally, State practice reveals three distinct models of practices. First groups comprise of countries such as Australia, the Netherlands, Germany, Norway among others who have established relatively permanent structured review mechanisms that integrate legal, technical, operational and sometimes medical expertise.⁴⁴ These systems tend to embed legal advice early in the procurement and development cycle and in some cases life-cycle assessment. Such institutionalization enhances professionalism and predictability of the weapons in general. The second group relies on *ad hoc* or procurement-driven assessments, conducting reviews only when acquisition decisions arise.⁴⁵ Although these countries are formally compliant, such models risk becoming reactive rather than preventive. The purpose of Article 36 being an *ex-ante* safeguard is to scrutinize legality during research and development yet this episodic review may simply rationalize decisions already made for strategic reasons.⁴⁶ Finally, a third model

⁴¹ Alan Backstrom and Ian Henderson, “New Capabilities in Warfare: an Overview of Contemporary Technological Developments and the Associated Legal and Engineering Issues in Article 36 Weapon Reviews”, *International Review of the Red Cross*, Vol. 94, No. 886, 2012, p. 483.

⁴² Vincent Boulanin, Laura Bruun and Netta Goussac, *Autonomous Weapons Systems and International Humanitarian Law, Identifying Limits and the Required Type and Degree of Human-Machine Interaction*, Stockholm International Peace Research Institute, June 2021, available at: https://www.sipri.org/sites/default/files/2021-06/2106_aws_and_ihl_0.pdf.

⁴³ ICRC, “A Guide to the Legal Review of New Weapons, Means and Methods of Warfare: Measures to Implement Article 36 of Additional Protocol I of 1977”, *International Review of the Red Cross*, Vol. 88, No. 864, 2006.

⁴⁴ *Ibid.*, p. 931.

⁴⁵ Rosie Cavdarski, Lauren Sanders & Rain Liivoja, “Legal Reviews of Autonomous Weapon Systems: Report on Submissions to the GGE on LAWS (version 2.0)”, *Law and the Future of War Research Group*, The University of Queensland, Australia, 2023, available at: <https://doi.org/10.14264/1c0275b>.

⁴⁶ Vincent Boulanin, *Implementing Article 36 Weapon Reviews in the Light of Increasing Autonomy in Weapon Systems*, SIPRI Insights on Peace and Security, November 2015, available at: <https://www.sipri.org/sites/default/files/files/insight/SIPRIInsight1501.pdf>.

provides little or no detail about their review mechanism, despite affirming adherence to the obligation.⁴⁷

The effectiveness of Article 36 is strongly based on institutional design, procedural standards and formalized mechanisms. The prime purpose of review is to assess compatibility with core principles of IHL i.e. distinction, proportionality, precaution, necessity of attack and other treaty-based restrictions. The first group of states have embedded review with these set standards. Australia provides one of the most detailed and structured examples. All weapons intended for operational use must undergo testing. The review produces one of three determinations such as “cleared”, “cleared subject to conditions” or “not cleared for operational use”.⁴⁸ Australia has also introduced an annual national course to certify legal advisers responsible for conducting reviews. This model treats review as a preventive governance rather than a *post hoc* justification. The Netherlands adopts an equally institutionalized approach; its review process is explicitly multidisciplinary, involving legal, technical, medical and environmental expertise, culminating in final ministerial approval.⁴⁹ Such integration reflects an understanding that compliance cannot be reduced to purely doctrinal analysis but must incorporate operational realities and ethical risk assessment.

Germany emphasizes structural independence and procedural safeguards; their reviews are formalized through Central Service regulations and conducted by qualified legal, technical and operational experts who maintain hierarchical independence from both developers and end users.⁵⁰ It also recognizes that programming modifications that alter the system's or ML's functioning require renewed review. By foregrounding independence and re-review, Germany addresses the conflict of static review and the problem of one-time approval. Other states have adopted a more declaratory posture. The UK characterizes legal reviews as “gatekeepers”, ensuring that weapons capable of lawful use reach the battlefield, yet provides comparatively little detail on procedures or outcomes.⁵¹ Many states differentiate between sharing general processes and withholding substantive findings;

⁴⁷ Damian Copeland, Rain Liivoja and Lauren Sanders, “The Utility of Weapons Reviews in Addressing Concerns Raised by Autonomous Weapon Systems”, *Journal of Conflict and Security Law*, Vol. 28, No. 2, p. 285.

⁴⁸ *Ibid.*

⁴⁹ Vincent Boulanin and Maaïke Verbruggen, *SIPRI Compendium on Article 36 Reviews*, Stockholm International Peace Research Institute, December 2017, available at: https://www.sipri.org/sites/default/files/201712/sipri_bp_1712_article_36_compendium_2017.pdf.

⁵⁰ Federal Foreign Office, “German Commentary on “Operationalizing all Eleven Guiding Principles at a National Level as Requested by the Chair of the 2020 Group of Governmental Experts (GGE) on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems (LAWS) within the Convention on Certain Conventional Weapons (CCW)”, *UNODA*, 24 June 2020, <https://documents.unoda.org/wp-content/uploads/2020/07/20200626-Germany.pdf>.

⁵¹ “United Kingdom: Item 5 – Topic 5: Legal Reviews”, *UNODA*, 9 March 2023, available at: https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_Group_of_Governmental_Experts_on_Lethal_Autonomous_Weapons_Systems_%282023%29/UK_Intervention_Item_5_Topic_5_Legal_Reviews.pdf.

they neither share the criteria nor reject the decision post-review, which could be publicly verifiable. Nineteen European Union member states and the ICRC support real information sharing and the exchange of best practices.⁵² Argentina, among others, have argued that transparency is necessary to reduce disparities between technologically advanced States and procurers with limited evaluative capacity. A few countries, like China and Russia, have also resisted stronger transparency obligations.⁵³

2.2.1. State Practice and AWS

These structural limitations are magnified in the case of AWS. Unlike conventional weapons, autonomous systems challenge the traditional logic of review because their behaviour may evolve through software updates, data inputs, machine learning, or contextual deployment rendering one-time static review and approval as inadequate. Thirteen States recognize this, a growing number of States emphasize the need for life-cycle review, reliability testing, and reassessment following modifications. Some explicitly argue that changes to programming or learning functions should trigger a renewed review.⁵⁴ Norway highlights that programming changes the system's functioning and necessitates reassessment.⁵⁵ Similarly, Germany treats algorithm or behavioural changes as triggers for fresh legal scrutiny. Yet Pakistan contends that the nature of AI makes continuous review effectively impossible.⁵⁶ This disagreement underscores a deeper tension that Article 36 was designed for discrete weapon objects, whereas AWS demands continuous, system-level oversight. Without procedural adaptation, review risks becoming obsolete precisely where scrutiny is most needed. In short, State practice demonstrates that Article 36 has succeeded in entrenching a global expectation that weapons be reviewed but has not yet succeeded in ensuring that such reviews function as genuine barriers to unlawful or destabilizing systems. The norm is procedurally entrenched but substantively fragile. AWS development exposes this fragility and makes clear that, unless review mechanisms evolve toward greater consistency and accountability, Article 36 may legitimize rather than regulate the next generation of weapons.

⁵² “Convention on Certain Conventional Weapons – EU Statement on Lethal Autonomous Weapons Systems (LAWS)”, *European Union External Action*, 13 November 2017, available at: https://www.eeas.europa.eu/node/35508_fr.

⁵³ Approaches of the Russian Federation to the Issue of Emerging Technologies in the Area of Lethal Autonomous Weapons Systems: Submitted by the Russian Federation, CCW/GGE.1/2024/WP.2, 1 March 2024.

⁵⁴ Thirteen states and the ICRC consider that modifications to AWS require renewed review.

⁵⁵ “Submission by Norway on Lethal Autonomous Weapons Systems”, *UNODA*, 24 May 2024, available at: https://docs-library.unoda.org/General_Assembly_First_Committee_-_Seventy-Ninth_session_%282024%29/78-241-Norway-EN.pdf.

⁵⁶ “Pakistan”, *Automated Decision Search*, available at: https://automatedresearch.org/news/state_position/pakistan/.

3. The Arms Trade Treaty and the Disaggregation of Responsibility

The ATT is only a decade-old treaty, albeit with 20 years in the making.⁵⁷ It is presented as a complementary safeguard to IHL compliance, operating upstream of battlefield use by regulating the international transfer of conventional arms. The safeguards and/or safety mechanisms address the transfer of weapons to another country. Thus, the most ATT can do, in its most ideal form, is stop a country with weapons from giving them to another country which is perhaps aiming to violate human rights with such weapons. In addition, the international arms trade accounts for only a fraction of total military expenditure globally.⁵⁸ However, if one were to look at conflicts, for instance, the Yemen Conflict and Israel-Hamas war, the importance of optimal application of ATT resurges. Saudi Arabia used weapons on schools, hospitals, markets and weddings, even contributing to famine-like situations; while Israel has been using air-dropped bombs and artillery, also destroying medical facilities and refugee camps.⁵⁹ Both countries are importers of weapons and indulge in the arms trade to facilitate their operations. Both imported about 70% of their arms from the United States of America.⁶⁰ This raises the question of the supplier's responsibility for these apparent human rights violations.

The role of AI and the rise of AWS further complicate the question of responsibility. The first documented use of AWS was in 2021 during the Libyan War.⁶¹ So, even though the research has been in place for such technology and automation since the 1950s,⁶² the use is fairly recent but steady for both AI and automation in weaponry.⁶³ Now, AI and automation may be complementary but have distinct roles to play in terms of responsibility.

⁵⁷ Peter Woolcott, "Arms Trade Treaty: Introductory Note", *Audiovisual Library of International Law*, 2 April 2013, <https://legal.un.org/avl/ha/att/att.html>.

⁵⁸ Xiao Liang *et al.*, *Trends in World Military Expenditure, 2024*, Stockholm International Peace Research Institute, April 2025, available at: https://www.sipri.org/sites/default/files/2025-04/2504_fs_milex_2024.pdf.

⁵⁹ "Press Release: New Study Shows Israel Air-dropped 2000lb Bombs Within Lethal and Damage Ranges of Hospitals in Gaza", *FXB Center for Health and Human Rights at Harvard University*, 10 October 2024, available at: <https://fxb.harvard.edu/blog/2024/10/10/new-study-shows-israel-air-dropped-2000lb-bombs-within-lethal-and-damage-ranges-of-hospitals-in-gaza/>.

⁶⁰ "US Arms Exports", *Campaign Against Arms Trade*, 22 January 2026, available at: <https://caat.org.uk/data/countries/united-states/us-arms-exports/>.

⁶¹ Majumdar Roy Choudhury *et al.*, Letter dated 8 March 2021 from the Panel of Experts on Libya established pursuant to resolution 1973 (2011) addressed to the President of the Security Council, S/2021/229, 8 March 2021, available at: <https://documents-dds-ny.un.org/doc/UNDOC/GEN/N21/037/72/PDF/N2103772.pdf?OpenElement>.

⁶² Goddy Uwa Osimen, John Anegbode and Temidayo Oladipo, "Autonomous Warfare: Exploring the Consequences of AI-Powered Weapons System on Global Peace and Security", *Power Tech Journal*, Vol. 49, No. 2, 2025, p. 1458.

⁶³ *Ibid.*

3.1. Risk Knowledge in a world of software, components, and updates

The first structural difficulty concerns the treaty's object of the regulation. ATT applies to the "conventional weapons" listed under article 2(1), categories historically tied to physical platforms and munitions.⁶⁴ AI enabled capabilities, by contrast are often embedded in software updates, training datasets, algorithms and cloud-based services rather than in tangible hardware. This raises an important question of "what exactly is being transferred?" If the decisive targeting capabilities of an AWS reside in codes rather than in the airframe or munition, export control that only focuses on the physical system would regulate the shell, not the substance. As the Stockholm International Peace Research Institute (SIPRI) has observed, emerging military AI technologies frequently fall outside traditional arms classifications⁶⁵ because they resemble dual-use software rather than weapons per se.

The very fundamental obligation of the ATT is to assess the intended use of the weapons before green-lighting the trade. The ATT provides for this obligation within its text under Article 6 and 7. Article 6 places an obligation on the states to not authorise or stop the transfer if it has "knowledge" of any untoward use such as to commit genocide, crimes against humanity, grave breaches of the Geneva Conventions, attacks directed against civilians and/or other war crimes.⁶⁶ Article 7, on the other hand, provides for the overriding risk standard.⁶⁷ States are obligated, under Article 7, to assess the potential that the arms/items would undermine peace and security, and/or be used to commit or facilitate serious violations of international humanitarian law and human rights law, and acts of terrorism and transnational organized crime to which the exporting state is a party.⁶⁸ When this risk outweighs other considerations, the exporting State must refuse to transfer.

Both articles provide for the minimum knowledge or justified belief required before the law assigns responsibility or triggers a legal consequence. Both determine when a State must refuse to authorise an arms transfer. This is the epistemic threshold of ATT. This requirement to assess risk and deny when there's knowledge of untoward consequences at the time of transfer, on the surface, seems good enough. But the advent and involvement of AI and automation pose a significant issue: how do you evaluate the risks which may evolve across time and actors?

The ATT, in its existing text, presumes identifiable weapons and the time of transfer, assessable future risk and concentrated responsibility, which perhaps works

⁶⁴ Arms Trade Treaty, UNR 67/234B, 2 April 2013, Art. 2 (1).

⁶⁵ Yasmin Afina and Giacomo Persi Paoli, "Governance of Artificial Intelligence in the Military Domain: A Multi-Stakeholder Perspective on Priority Areas", *UNIDIR*, 5 September 2024, available at: <https://unidir.org/publication/governance-of-artificial-intelligence-in-the-military-domain-a-multi-stakeholder-perspective-on-priority-areas/>.

⁶⁶ Arms Trade Treaty, UNR 67/234B, 2 April 2013, Art. 6.

⁶⁷ Arms Trade Treaty, UNR 67/234B, 2 April 2013, Art. 7.

⁶⁸ *Ibid.*

well for conventional arms. But with software systems, iterative updates, modular components and learning systems, the assumptions upholding the effectiveness of “risk knowledge” become inconsistent at best, and ineffective at worst. Risks, now, are not assessable *ex ante*. With AI and automation, the risks may evolve after the export is already done, knowledge may not even be available, harm may emerge through integration, causation may not be linear, and most importantly, the attribution of responsibility may be fragmented at best, and non-existent at worst.

To explain why AI supply chains defeat this linear risk assessment, we must examine what it assumes and how AI supply chains operate. Linear risk assessment is a simple approach in which causation flows in a straight line and the risk assessment occurs at a single identifiable decision point. It assumes that there is an identifiable object, stable functionality, single transfer of technology, traceable causation and, again, centralised responsibility. AI supply chains, however, are update-driven not static, modular not monolithic, and recursive not stagnant. Firstly, AI supply chains being update driven implies they get updated, undergo training and retraining, and adjust and improve over time. Secondly, AI supply chains being modular imply that they consist of separate modules rather than a single technology. Thus, each module can be developed, exported and classified differently.⁶⁹ Recursive AI supply chains imply that they are adaptable and can incorporate feedback from deployment into future behaviour.

The update-driven, modular, and recursive nature of AI supply chains undermines temporal stability, disrupts responsibility, and undermines causal traceability. Risks, then, cannot be reasonably foreseen and the knowledge is abstract, if it is even available. This is what the scholars refer to as “temporal misalignment”⁷⁰ between authorisation and operation.⁷¹

3.2. Dual-use AI and the collapse of the conventional arms boundary

Conventional arms control law concerns itself with tangible, military-design objects with the primary object or operation being killing/injuring.⁷² Non-arms refers to civilian goods, services and/or technologies. This differentiation or boundary collapses when AI has dual-use. Dual-use AI refers to artificial intelligence systems which can slide between benign and harmful uses depending on context, integration,

⁶⁹ Aspen Hopkins *et al.*, “On Deployment: AI Supply Chains”, *Social Science Research Network*, 3 April 2013, available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4789403.

⁷⁰ Wenjing Lou *et al.*, “Temporal Misalignment Attacks against Multimodal Perception in Autonomous Driving”, *ArXiv*, 1 October 2025, available at: <https://arxiv.org/html/2507.09095v2>.

⁷¹ Sze Hong Lam, “Will Treaty Following AI Collapse into the dichotomy between ‘Apology’ and ‘Utopia’”? *Asia Pacific Journal of International Humanitarian Law Special Issue on Artificial Intelligence and Warfare*, 2026, pp.48-81.

⁷² Frank Sauer, “Autonomy in Weapons Systems and the Struggle for Regulation”, *Center for International Governance Innovation*, 28 November 2022, available at: <https://www.cigionline.org/articles/autonomy-in-weapons-systems-and-the-struggle-for-regulation/>.

and intent without any change to their constitution.⁷³ Thus, being able to be used for both legitimate civilian purposes and for military, security or coercive applications. Examples of dual-use AI include facial recognition technologies and predictive analytics or logistics. At present, there is a great number of potential applications of biometrics used in the military domain; it can be used throughout the full range of operations, including combat, security, and peace support operations. In general, biometrics allows verification of personal identity and identification of enemy threats with a higher degree of confidence when personal enters military facilities.⁷⁴ Military power is no longer concentrated in tangible arms, but spread across algorithms, data and upgrades that could become non-traditional weapons.

Algorithms are capable of being used as weapons. Algorithms can be used to select targets, prioritise threats, optimise strike decisions and/or automate surveillance.⁷⁵ Algorithms may not be traditionally weapons but can be operated just like one. Secondly, data enables identification, tracking, prediction and classification of populations. Data also is not a traditional weapon but can transform people into precise targets, making violence administratively scalable.⁷⁶ Thirdly, to understand upgrades as weapons requires a bit more understanding. Upgrades are continuous, service-based, framed as support and innovation and delivered across borders digitally. No transfer tangibly occurs, no weapon changes hands, and no review is triggered with each update. However, with each upgrade, the lethality increases. An upgrade in itself is not a weapon. Upgrades of AI systems are sites where lethal capacity is introduced, intensified, or transformed, outside the object-based logic of arms control law. Moving from ontology to governance now. Modularisation refers to breaking down a system into separable components, each developed, sold and transferred independently.⁷⁷ Responsibility and governance then become convoluted and invisible. Modularisation enables legal evasion. AI distributes capability across modules. There is no single transfer to trigger arms laws. States and/or entities

⁷³ Jayshree Pandya, "The Dual-Use Dilemma of Artificial Intelligence", *Forbes*, 7 January 2019, available at: <https://www.forbes.com/sites/cognitiveworld/2019/01/07/the-dual-use-dilemma-of-artificial-intelligence/>.

⁷⁴ Marten Zwanenburg, "Know They Enemy: The Use of Biometrics in Military Operations and International Humanitarian Law", *International Law Studies Stockton Centre for International Law*, Vol. 97, 2021, US Naval War College, available at: <https://digital-commons.usnwc.edu/cgi/viewcontent.cgi?article=2986&context=ils>.

⁷⁵ Anthony Downey, "The Alibi of AI: Algorithmic Models of Automated Killing", *Digital War*, Vol. 6, No. 9, 2025, available at: <https://link.springer.com/article/10.1057/s42984-025-00105-7>.

⁷⁶ Lucy Suchman *et al.*, "Tracking and Targeting: Sociotechnologies of (In)security", *Science Technology and Human Values Journal*, Vol. 42, No. 6, 2021 available at: https://journals.sagepub.com/doi/full/10.1177/0162243917731524?casa_token=ayRc299r33cAAAAA%3A6Jc7GYyyBvyJISrbkCsxjUnMhJmYrMQioCsajxsr-hXG9ZKJ-Y1BHOIm7v-rUGV5qwdsvqaOvKm.

⁷⁷ Paolo Trucco *et al.*, "Modularisation as a System Life Cycle Management Strategy: Drivers, Barriers, Mechanisms and Impacts", *International Journal of Engineering Business Management*, Vol. 11, 2019, available at: <https://journals.sagepub.com/doi/epub/10.1177/1847979018825041>.

engage in licensing agreements, instead of direct transfers. Modularisation, then, does not violate but bypasses the existing laws.⁷⁸

The IBM case perfectly encapsulates the complexity of dual-use AI, modularisation and legal evasion. IBM's punch-card tabulation systems, supplied to and serviced within Nazi Germany, were not weapons in any conventional sense.⁷⁹ They were data-processing machines, commercially marketed, ostensibly administrative, and entirely compatible with ordinary civilian use. Yet they became infrastructural instruments of persecution.⁸⁰ What rendered them consequential was not their material form, but their integration into a broader architecture of state violence. The harm did not arise from a single transfer event. It was sustained through modular components: machines, punch cards, technical servicing, updates, and maintenance. Each of these could be framed as commercially routine. Responsibility was therefore distributed across functions rather than concentrated in a discrete act of arming. No rifle crossed a border; no bomb was supplied.

This structural dynamic mirrors the contemporary problem of dual-use AI. Contemporary examples include Amesys selling their internet inspection system to Gaddafi's Libya, which was used to monitor dissidents, journalists and activists.⁸¹ Executives of Amesys and Nexa were indicted for complicity in torture.⁸² However, the transfer of technology never triggered any ATT compliance. Algorithms, datasets, and software upgrades are rarely lethal in isolation. Their significance lies in their embedding, recombination, and iterative enhancement within operational systems. Modularisation allows each component to appear neutral, while its cumulative integration materially reshapes the capacity for surveillance, targeting, or repression. Legal scrutiny that focuses only on finished weapons or discrete export events risks overlooking this distributed architecture of enablement.

3.3. Export control without accountability

To revisit the case of Amesys and Libya, Amesys is a technology company based in France. Amesys supplied its technology to Libya and Egypt. Even though there was ATT at the time, the technological transfer was not scrutinised. ATT compliance, as it is, is procedural, not substantive. The authorisation that is given under ATT provides for *ex ante* obligations for the exporting state which are, as discussed, insufficient in light of AI systems. Public international law generally also tolerates

⁷⁸ Neil Cooper, "What's the point of arms transfer control?", *Contemporary Security Policy*, Vol. 27, No. 1, 2006, p. 118.

⁷⁹ Harry Murphy, "Dealing with the Devil: The Triumph and Tragedy of IBM's Business with the Third Reich", *The History Teacher*, Vol. 53, No. 1, 2019, p. 171.

⁸⁰ *Ibid.*

⁸¹ MEE Staff, "French Spy-tech Firms Accused of 'Complicity in Torture' in Libya and Egypt", *Middle East Eye*, 22 June 2021, available at: <https://www.middleeasteye.net/news/egypt-libya-french-spy-tech-firms-accused-complicity-torture>.

⁸² *Ibid.*

due diligence standards not strict liability.⁸³ Accountability with AI systems and autonomous weapons is distributed and harder to uphold. It is distributed across technical, legal and temporal functions, a problem of too many chefs. With such diffusion of responsibility among developers, exporters and states, the law is not violated but the legality is restructured. Traditional attribution doctrines such as state responsibility, complicity, aiding and abetting struggle with cumulative causation. This becomes apparent in the IBM case as well. Even though the involvement was documented and has never been denied, the company was never held accountable.⁸⁴ To draw on Ulrich Beck's organised irresponsibility, ignorance is not the absence of knowledge but a systemically generated condition.⁸⁵

The ATT focuses on conventional arms, the assessment of foreseeable risk, and the denial of supply if there is an overriding risk. With AI systems, three problems arise: firstly, compliance under ATT is event-based, whereas harm from AI systems is process-based; secondly, ethical risks may be known but legally insufficient to deny supply under Article 7; and thirdly, modularisation enables formal compliance while functionally evading control. These features allow ATT compliance to operate as a procedural rather than a mechanism of substantive containment. Once an export authorisation is granted, the Treaty does not impose a continuing duty to reassess evolving system behaviour, even where subsequent software updates, retraining, or system integration materially alter the risk profile.⁸⁶ The regulatory gaze remains fixed on the moment of transfer, while the locus of harm shifts to post-export modification and use. Compliance is thus temporally satisfied even as risk migrates.

In the context of AI-enabled systems, risks associated with autonomy, algorithmic bias, or downstream repurposing are often foreseeable in an ethical sense but remain insufficiently determinate to meet the Treaty's legal threshold.⁸⁷ This creates a space in which States may acknowledge serious concerns while still lawfully authorising transfers, thereby fostering ethical awareness without legal consequences. Modularisation further entrenches this gap. By disaggregating capability across software, data, hardware, and ongoing services, AI supply chains

⁸³ Mark Taylor *et al.*, "Human Rights Due Diligence: The Role of States", *Human Rights in Business*, December 2012, available at: <https://humanrightsinbusiness.eu/wp-content/uploads/2015/05/De-Schutter-et-al.-Human-Rights-Due-Diligence-The-Role-of-States.pdf>.

⁸⁴ H. Murphy, above note 79, p. 19.

⁸⁵ Maria Galantino, "Organised Irresponsibility in the Post-Truth Era: Beck's Legacy in Today's World at Risk", *Italian Sociological Review*, Vol. 12, No. 8S, 2022, available at: <https://www.proquest.com/openview/a1f67cfd911f2e2741b7e59304d145a1/1?pq-origsite=gscholar&cbl=1796337>.

⁸⁶ Paul Beijer, *Taking Stock of the Arms Trade Treaty: Application of the Risk-assessment Criteria*, Stockholm International Peace Research Institute, August 2021, available at: <https://www.sipri.org/publications/2021/policy-reports/taking-stock-arms-trade-treaty-application-risk-assessment-criteria>.

⁸⁷ "A Hazard to Human Right: Autonomous Weapons Systems and Digital Decision-Making", *Human Rights Watch*, 28 April 2025, available at: <https://www.hrw.org/report/2025/04/28/a-hazard-to-human-rights/autonomous-weapons-systems-and-digital-decision-making>.

allow each transaction to be assessed in isolation, none of which appears to trigger the prohibitive logic of Article 6 or 7.⁸⁸ The cumulative system, however, may enable forms of surveillance, targeting, or repression that the Treaty was designed to prevent. Responsibility is thereby diffused across actors and transactions, while no single authorisation bears the weight of the eventual harm. In this configuration, export control remains formally intact but normatively hollow. The ATT succeeds in regulating compliance at the level of authorisation, yet fails to ensure ethical or legal containment of harm arising from iterative, distributed, and adaptive technological processes. The result is not regulatory absence, but a regime in which legality and accountability increasingly diverge.

4. The Martens Clause: Ethical Corrective or Normative Illusion?

When technological developments outpace positive legal regulation, international lawyers frequently invoke the Martens Clause as a residual normative safeguard. First articulated in the 1899 Hague Convention II and reiterated in Additional Protocol I, the clause provides that in cases not covered by specific treaty provisions, civilians and combatants remain protected by “the principles of humanity and the dictates of the public conscience.”⁸⁹ It is essentially a residual clause meant to fill in the gaps created by the passage of time and the evolution of humanity and its technology.

In contemporary debates on AWS, the Clause has re-emerged as a doctrinal refuge. Faced with uncertainty regarding the adequacy of Article 36 reviews and the limited applicability of export controls under the ATT, commentators and States alike cite the Martens Clause as evidence that autonomous weapons cannot escape humanitarian restraint merely because positive law is indeterminate.⁹⁰ Considering that the clause upholds principles of humanity and public conscience, it is argued to be capable of dealing with technological novelty and innovation in warfare. Even the ICJ, in its opinion, referred to the Martens Clause as an effective means of addressing the rapid evolution of military technology.⁹¹ However, AI warfare seems far out of its reach.

⁸⁸ Vincent Boulanin, Kolja Brockmann and Luke Richards, *Responsible Artificial Intelligence Research and Innovation for International Peace and Security*, Stockholm International Peace Research Institute, November 2020, available at: https://www.sipri.org/sites/default/files/2020-11/sipri_report_responsible_artificial_intelligence_research_and_innovation_for_international_peace_and_security_2011.pdf.

⁸⁹ Hague Convention II (1899) Preamble; Additional Protocol I (1977), art. 1(2); “Until a more complete code of the laws of war has been issued, the High Contracting Parties deem it expedient to declare that, in cases not included in the Regulations adopted by them, the inhabitants and the belligerents remain under the protection and the rule of the law of nations, as they result from the usages established among civilized peoples, from the laws of humanity and the dictates of public conscience”.

⁹⁰ Rob Sparrow, “Ethics as a Source of Law: The Martens Clause and Autonomous Weapons”, *Humanitarian Law & Policy*, 14 November 2017, available at: <https://blogs.icrc.org/law-and-policy/2017/11/14/ethics-source-law-martens-clause-autonomous-weapons/>.

⁹¹ Advisory opinion, above note 31, p. 7.

Holistically, the clause is reactive rather than anticipatory, reflecting one influential yet highly contested interpretation of the Clause in contemporary IHL. Under this view, the Clause does not itself create new substantive prohibitions in advance of technological development; it operates as a residual interpretive safeguard once gaps in treaty law become apparent. However, this interpretation is not universally accepted. A persuasive argument is that Clause cannot, in itself, ensure state compliance in the context of autonomous weapon systems, which requires engaging with competing understandings of the Clause and explaining why the reactive reading is ultimately more convincing. Scholars have interpreted this provision in three ways: as a mere rhetorical statement, as an interpretive principle guiding the application of existing laws, or as an independent normative source capable of generating legal restraints beyond positive treaty law. Antonio Cassese's seminal article remains foundational because it maps these divergent interpretations and demonstrates that State practice has never conclusively settled the Clause's precise legal status.⁹² Those advocating a proactive or anticipatory interpretation argue that the clause enables IHL to respond to emerging technologies before treaty law crystallizes. This position is especially visible in debates concerning AWS, where commentators and civil society actors invoke the Clause to argue that autonomous weapons remain constrained even in the absence of explicit treaty prohibition.⁹³

Such approaches effectively treat the Clause as anticipatory because it purportedly enables legal restraint prior to formal legal codification. Yet this expansive reading faces serious doctrinal and practical difficulties. Firstly, State practice demonstrates little acceptance of the idea that the Martens Clause independently generates binding prohibitions. Cassese himself cautioned that although the Clause exerts moral and interpretive influence, States have generally resisted recognizing it as a self-sufficient source of concrete obligations. The International Court of Justice has likewise referred to the Clause primarily as an "effective means" of addressing humanitarian considerations rather than as an autonomous rule-creating mechanism, particularly in its Nuclear Weapons Advisory Opinion.⁹⁴ This suggests that the Clause supplements existing law rather than replacing the need for positive legal regulation. Secondly, the concepts of "humanity" and "public conscience" remain deeply indeterminate. In the AWS context, States sharply disagree over whether autonomous weapons inherently offend either standard. While some States advocate pre-emptive prohibition, others maintain that existing IHL is sufficient and continue investing heavily in

⁹² Antonio Cassese, "The Martens Clause: Half a Loaf or Simply Pie in the Sky?", *European Journal of International Law*, Vol. 11(1), 187, 2000.

⁹³ Rebecca Crootof, "The Killer Robots Are Here: Legal and Policy Implications", *Cardozo Law Review*, Vol. 36(2), p. 1837, 2015.

⁹⁴ International Court of Justice, *Legality of the Threat or Use of Nuclear Weapons*, Advisory Opinion, *ICJ Reports* 1996, 8 July 1996, para. 86.

autonomous military technologies.⁹⁵ Such disagreement undermines the claim that the Clause can function anticipatorily as a clear legal constraint. A norm whose content depends on contested moral evaluations is ill-suited to ensure consistent State compliance absent subsequent treaty development or consolidated customary practice. Accordingly, the stronger interpretation is that the Martens Clause operates reactively rather than anticipatorily. It becomes relevant once legal lacunae emerge and assists courts, tribunals, and States in interpreting humanitarian principles in situations insufficiently addressed by positive law.

There is nothing wrong with a clause being effective *ex post*, while the ATT is working *ex ante*. However, in the context of AI warfare, the gap the Martens clause was intended to address in the evolution of military technology is not adequately addressed by its content. Martens Clause cannot govern and or regulate AI warfare, it can only address the aftermath. Martens Clause, thus, cannot be over-relied upon to deal with modern military technology. The clause does not specify the level of accountability, how it must be enforced, or how the people will remain protected.⁹⁶ It is imperative to note that principles of humanity and public conscience remain undefined as well, perhaps intentionally. The more detrimental aspect is the lack of enforcement mechanisms. Martens clause works through interpretation, a guideline, within diplomatic discourse. No obligations are placed through the clause. A doctrinal link could be established under Article 36 read with the Clause.⁹⁷ It could be argued that this review under Article 36 could be informed by the Martens Clause. To explain, it means a review of new weapons should ensure that they are in line with the principles of humanity and public conscience.

5. From Compliance to Design: Reimagining State Responsibility through IHL-by-Design

IHL traditionally regulates weapons after they are designed, at the moment of use, and/or at the time of trade. The law explicitly regulates compliance and the aftermath with a presumption of concrete, tangible harm and concentrated responsibility. AI weapons topple the laws and regulations back on their heads. State responsibility gets diffused and dismantled with respect to AI warfare, under the existing IHL framework. The international order then breaks down: the powerful get more controlling, and the weak get weaker.

⁹⁵ ICRC, “ICRC Position on Autonomous Weapon Systems”, 12 May 2022; and CCW Convention, Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems (GGE on LAWS), “Draft articles on autonomous weapon systems—prohibitions and other regulatory measures on the basis of international humanitarian law (IHL)”, Working paper submitted by Australia, Canada, Japan, the Republic of Korea, the United Kingdom and the United States, CCW/GGE/1/2023/WP.4, 6 Mar. 2023, p. 1.

⁹⁶ R. Sparrow, above note 90.

⁹⁷ Natalia Jevglevskaja, “Interpreting Article 36: The Object of Review” in *International Law and Weapons Review*, Cambridge University Press, Cambridge, 2021, p. 87.

In technology and AI governance discourse, “ethics-by-design” and “safety-by-design” are approaches that aim to build normative constraints into systems at the point of design, rather than relying solely on user behaviour, post-deployment regulation, or *ex post* liability.⁹⁸ IHL-by-design refers to embedding legal constraints into the architecture of weapons systems, rather than relying solely on *post hoc* compliance.⁹⁹ So, IHL-by-design implies a legal obligation of states, not a discretionary engineering choice.¹⁰⁰

5.1. Re-engineering Article 36 and ATT Obligations

Article 36 imposes an *ex-ante* obligation, that applies during study, development, acquisition or adoption and requires review against all applicable international law, including IHL. It is apposite to note, that Article 36 does not limit itself to physical weapons and does not restrict the review to a one-time assessment. At least, not explicitly. ATT obligations address the trade in arms in a similar manner.

A very important reframing of the classification of weapons, inclusive of AI innovations, is to treat them as evolving systems. This is not just acceptable but defensible.¹⁰¹ Article 36 traditionally assumes weapons are stable, predictable and static after deployment, rendering a single review sufficient.¹⁰² AI systems break this assumption apart. AI-enabled weapons are update-driven, modular and adaptive. The re-engineering of Article 36 is proposed to be interpretive. The existing wording of Article 36 could be utilised for the re-engineering. The use of the words “study, development” could be interpreted to include retraining and upgrading of AI systems; “acquisition or adoption” to include integration into new platforms and “new weapon, means or method” to include evolved functional capabilities, not just objects. So, substantial software updates or capability expansions then retrigger review under Article 36.

There must be AI-specific metrics to address the risks associated with AI systems. Conventional weapons’ risks assume that a weapon’s harmful potential is materially fixed and that its effects can be assessed through relatively stable, observable characteristics. Risks associated with AI-systems, however, are systemic, opaque and probabilistic.¹⁰³ These characteristics directly affect core IHL principles, including distinction, proportionality, and precautions in attack, as codified in

⁹⁸ Lee Bygrave, “Security by Design: Aspirations and Realities in a Regulatory Context”, *Oslo Law Review*, Vol. 8, No. 3, 2022, p. 126.

⁹⁹ R. Sparrow, above note 91, p. 24.

¹⁰⁰ Giovanna Maletta, “Seeking a Responsible Arms Trade to Reduce Human Suffering in Yemen”, *Italian Journal of International Affairs*, Vol. 56, No. 1, 2021, p. 73.

¹⁰¹ Matthis Maas, “Framing a Global AI Regime Complex in Five Steps”, in *Architectures of Global AI Governance: From Technological Change to Human Choice*, Online Edition, Oxford Academic, Oxford, 2025, p. 497.

¹⁰² V. Boulanin, above note 46, p. 12.

¹⁰³ Samuel Carey, “Regulating Uncertainty: Governing General-Purpose AI Models and Systemic Risk”, *European Journal of Risk Regulation*, Vol. 1, 2025, p. 1.

Articles 48, 51, 52, and 57 of Additional Protocol I.¹⁰⁴ The ability to distinguish between civilians and combatants depends on reliable target identification; proportionality assessments require foreseeability of incidental harm; and precautions in attack demand the capacity to anticipate and mitigate risks. Where system behaviour is probabilistic, opaque, or context-dependent, traditional metrics are insufficient to evaluate compliance with these obligations.

While Article 36 addresses the regulation of weapons internally, the ATT addresses externalities. ATT deals with the supply and/or export of weapons. AI-enabled systems have software updates, cloud services, algorithms, data pipelines, etc. None of which qualify as arms, munitions or discrete export events. However, they still are lethal, have autonomy and enable repression.¹⁰⁵ The regulation of exports or trade, thus, must move beyond the tangible. It has to be capability-based. Articles 6 and 7, dealing with risk knowledge and the principle of overriding risk, should be used to review and re-assess whenever “capability” changes. Without extending export controls beyond the tangible, states or supplying entities get to evade responsibility. This shift allows for the accountability gap to be filled.

5.2. Implications for future AWS governance

The Convention on Certain Conventional Weapons (CCW)¹⁰⁶ regulates specific categories of conventional weapons considered to cause unnecessary suffering or indiscriminate harm. It has become the primary multilateral forum for discussions on LAWS.¹⁰⁷ The debates consider whether autonomous weapons should be banned, regulated, or addressed under existing IHL.¹⁰⁸

The Hard Law has limitations regarding AWS governance.¹⁰⁹ AI systems undermine existing hard law because they are rapidly evolving, dual-use, and dependent on updates. Hard law, then, is obsolete, easily circumvented and outdated with respect to AI systems. Adaptive regulations, on the other hand, are not only compatible but ideal for AI systems. It refers to continuous review mechanisms, design-stage legal constraints, iterative compliance obligations, and post-deployment oversight. The suggestion to re-engineer Article 36 or to continue its ongoing review is an example of adaptive regulation. The impasse between the two highlights a

¹⁰⁴ Additional Protocol I, above note 11, p. 4.

¹⁰⁵ V. Boulanin, K. Brockmann and L. Richards, above note 88 p. 22.

¹⁰⁶ UNODA, “Convention on Certain Conventional Weapons”, *UN-iLibrary*, March 2016, available at: <https://www.un-ilibrary.org/content/books/9789210578967>.

¹⁰⁷ Josephine Jackson, “Mapping the Lethal Autonomous Weapons Debate: An Introduction”, *Ethics & International Affairs*, Vol. 37, No. 3, 2023, p. 254.

¹⁰⁸ Charukeshi Bhatt and Tejas Bharadwaj, “Understanding the Global Debate on Lethal Autonomous Weapons Systems: An Indian Perspective”, *Carnegie India*, 30 August 2014, available at: <https://carnegieendowment.org/russia-urasia/research/2024/08/understanding-the-global-debate-on-lethal-autonomous-weapons-systems-an-indian-perspective>.

¹⁰⁹ Damian Copeland, “How does International Law Regulate AWS?”, in *A Functional Approach to the Legal Review of Autonomous Weapon System*, Brill, 2025, p. 41.

broader tension between hard law and adaptive regulation in the governance of autonomous weapons. While adaptive regulations seem ideal, strictly adaptive regulations would also risk fragmentation, uneven implementation, and erosion of legal accountability. Hard law offers clarity, legitimacy and enforceability. Thus, it is reliable and stable. AWS governance requires both flexibility and responsiveness, along with functional relevance, which are offered by adaptive regulations.

6. Concluding Remarks

IHL rests on the premise that decisions about life and death remain attributable to identifiable human or institutional actors who can be held responsible. If lethal decision-making is increasingly delegated to opaque or adaptive systems without corresponding mechanisms of control and review, this premise weakens. The analysis demonstrates that the problem confronting IHL is not the absence of regulation, but the misalignment between inherited legal architectures and contemporary technological realities.

Article 36 reviews should be reconceived as continuous and lifecycle-based. Reviews must extend beyond initial acquisition to encompass software updates, retraining processes, and operational modifications. Static certification is incompatible with adaptive technologies. This approach is already implicit in ICRC guidance and should be treated as a legal requirement rather than best practice. Article 36 presumes weapons whose behaviour can be assessed *ex ante*, AI-enabled systems evolve after deployment. The ATT regulates discrete transfers of tangible arms, and military AI diffuses through software, data, and distributed supply chains. The Martens Clause preserves humanitarian values but lacks operational precision. Each framework retains normative relevance, even in the absence of a binding treaty on AWS; shared interpretive principles and technical benchmarks can harmonise state practice and reduce regulatory arbitrage. Ultimately, the challenge posed by AI-enabled warfare is not whether machines can act lawfully, but whether states will accept responsibility for how those machines are built. The future credibility of international humanitarian law depends on that answer.