

De Jure Necessity vs. De Facto Improbability? Establishing a New Treaty for Artificial Intelligence in the Military Domain

Zhixiong Huang*
Haokun Yu**

ABSTRACT

The military application of artificial intelligence challenges existing international law, which remains applicable but structurally insufficient to resolve emergent moral dilemmas, accountability gaps, and compliance deficits. Consequently, establishing a new treaty is a de jure necessity. However, this necessity confronts formidable de facto obstacles: the strategic imperatives of major military powers, prevailing scepticism towards treaty-making, and an unfavourable international political environment render the adoption of such an instrument almost de facto improbable. Nevertheless, this article argues that the treaty-making process can facilitate the formation of customary international law, serve as supplementary evidence for legal interpretation, and impose reputational costs on non-party states.

Keywords: AI in the military domain; Military AI; International humanitarian law; New treaty; International law-making process.

1. Introduction

Currently, the deployment of artificial intelligence (AI) in the military domain¹ is becoming a reality. In December 2025, the United Nations General Assembly once again noted the increasing integration of AI into military applications and highlighted its operational use.² Such international attention is well-founded. Reportedly, the United States deployed the Maven Smart System during the conflict

* Zhixiong Huang is the Hongyi Chair Professor at Wuhan University School of Law, and Director of the Wuhan University Institute for Cyber Governance. Email: fxyhzx@whu.edu.cn.

** Haokun Yu is a PhD Candidate in International Law at Wuhan University School of Law. Email: yuhaokun1012@whu.edu.cn.

¹ It should be noted that the terms “Military AI” or “AI in the military domain” often encompass a broad scope of applications, including weapon-related functions. Applications range from decision support systems for military operations to systems that support maritime security through counter-piracy efforts, counter-terrorism operations, and border security. In contrast, autonomous weapon systems (AWS) represent a partially specific application of such technology. While early multilateral discussions primarily focused on AWS, the international community has recently broadened its focus to address the broader implications of AI in warfare. Therefore, while the analysis in this paper remains largely relevant to AWS, the term “Military AI” is primarily used to reflect shifting international concerns and the potential expansion of future discussions. See UNIDIR, “Artificial intelligence in the military domain”, New York, last visited 23 May 2026, available at: <https://disarmament.unoda.org/en/our-work/emerging-challenges/artificial-intelligence-military-domain>.

² UNGA Res. 80/58, 5 December 2025.

with Iran, while the Israel Defense Forces (IDF) utilized the Gospel and Lavender systems during the armed conflict in Gaza, all of which are military AI systems designed to enhance targeting capabilities.³ These developments show that the regulation of military AI is no longer a potential consideration but has become an increasingly pertinent and contemporary issue. While AI's capabilities could offer significant military advantages in many areas, such as supporting military planning and human decision-making,⁴ they also pose profound risks, including increasing the unpredictability of conflicts, accelerating escalation, and causing large-scale destruction.⁵ These inherent risks necessitate robust regulation under international law.

In the realm of international law, academia and practitioners are deeply divided over whether a new treaty is necessary for military AI. Some countries and organizations believe that moral, accountability, and compliance issues compel states to adopt a new treaty to regulate it.⁶ Meanwhile, other countries argue that the existing international law is sufficient to address these challenges.⁷ They further contend that, without sufficient discussion regarding military AI, it is premature for states to establish a treaty to regulate autonomous weapon systems,⁸ and they emphasize that military AI may have potential benefits, such as increasing the efficiency in task assignment and reducing the probability of errors.⁹ This divide has become a focal point within the Convention on Certain Conventional Weapons (CCW) process and reflects the divergent positions of states.

³ See Tara Copp *et al.*, "Anthropic's AI Tool Claude Central to U.S. Campaign in Iran, Amid a Bitter Feud", *The Washington Post*, 4 March 2026, available at: <https://www.washingtonpost.com/technology/2026/03/04/anthropic-ai-iran-campaign/>; Michael N. Schmitt, "Israel – Hamas 2024 Symposium – The Gospel, Lavender, and the Law of Armed Conflict", *Articles of War*, 28 June 2024, available at: <https://lieber.westpoint.edu/gospel-lavender-law-armed-conflict/>.

⁴ Bill Boothby, "AI Warfare and the Law", *International Law Studies*, Vol. 104, 2025, p. 13.

⁵ AutonomousWeapons.org, "The Risks of Autonomous Weapons", last visited on 23 January 2026, available at: <https://autonomousweapons.org/the-risks/>.

⁶ According to Automated Decision Research, approximately 130 states currently support a legally binding instrument, including China, Germany, New Zealand, Switzerland, as well as member countries of the Non-Aligned Movement and the African Group. Conversely, only 12 states explicitly oppose such an instrument, including the United States, Russia, the United Kingdom, Japan, and India. Among international organizations, the ICRC has repeatedly called for the establishment of a legally binding instrument. See AutomatedDecisionResearch.org, "State Positions", last visited on 23 May 2026, https://automatedresearch.org/state-positions/?_state_position_negotiation=no (State Positions Database); ICRC, *ICRC Position on Autonomous Weapon Systems: ICRC Position and Background Paper*, Geneva, 12 May 2021. (ICRC Position)

⁷ United States of America, *Submission of the United States of America*, 22 December 2023, p. 14, available at: [https://docs-library.unoda.org/General_Assembly_First_Committee_-_Seventy-Ninth_session_\(2024\)/78-241-US-EN.pdf](https://docs-library.unoda.org/General_Assembly_First_Committee_-_Seventy-Ninth_session_(2024)/78-241-US-EN.pdf).

⁸ AutomatedDecisionResearch.org, "United Kingdom State Position", last visited on 23 May 2026, available at: https://automatedresearch.org/news/state_position/united-kingdom/.

⁹ Russian Federation, *Working Paper of the Russian Federation "Application of the International Law to Lethal Autonomous Weapons Systems (LAWS)"*, 18 July 2022, available at: https://documents.unoda.org/wp-content/uploads/2022/07/WP-Russian-Federation_EN.pdf.

While existing literature largely focuses on the analysis of legal grounds for establishing a new treaty for military AI (whether pro- or con-treaty),¹⁰ it frequently overlooks the practical impediments to treaty formation. Such as limitations on arms control, inherent skepticism toward new legal instruments, and a fragmented international environment, all of which will significantly impact the treaty's success.¹¹ Meanwhile, the normative and factual value of the treaty-making process is also usually underestimated. Against this backdrop, this article aims to investigate the necessity and probability of establishing a new treaty for artificial intelligence in the military domain through a process-oriented perspective.¹² Based on the analysis of the challenges posed by military AI, Part 2 argues the necessity of establishing a treaty on a *de jure* basis. Part 3 then presents a realistic picture and discusses the *de facto* aspect, stressing the improbability of doing so due to several international political reasons. However, in Part 4, this article concludes on a cautiously optimistic view: the *de facto* improbability does not nullify the *de jure* necessity. In fact, the process of adopting a new treaty can serve as an incubator for customary international law (CIL), facilitate the interpretation of existing international law, and impose reputational costs on major military powers.

2. The De Jure Necessity: The Structural Insufficiency of Existing International Law

This Part examines the consensus and divergence between proponents and opponents of a new treaty from a legal perspective and emphasizes the necessity of the former. In fact, advocating for a new treaty rests not on the inapplicability of existing international law, but rather on its insufficiency. Specifically, the current international legal framework struggles to adequately address the moral dilemmas, accountability gaps, and compliance deficits created by military AI. Furthermore, the indeterminacy of applying existing rules to novel technologies undermines the legal certainty of international law, thereby diminishing the precondition for legal compliance. Collectively, these factors make the establishment of a new treaty *de jure* necessary.

¹⁰ Benjamin Perrin, "Lethal Autonomous Weapon Systems & International Law: Growing Momentum Towards a New International Treaty", *ASIL Insights*, Vol. 29, No. 1, 2025; Laurie R. Blank, "New Treaty Law on Autonomous Weapons? An Opportunity to Reframe the Discourse", *Case Western Reserve Journal of International Law*, Vol. 57, No. 1, 2025, p. 233.

¹¹ For the purpose of this discussion, we define a successful treaty in its traditional sense as a legally binding instrument that secures the broad participation of major military powers and effectively binds their conduct in practice. However, the value of international regulation does not rely on achieving such an ideal outcome. Given its *de facto* improbability (as analyzed in Part 3), Part 4 argues that both the treaty-making process itself and a subsequent treaty with limited adherence would still generate independent value, regardless of the final instrument's formal success.

¹² Many legal theories, such as the New Haven school and International Legal Process schools, view international law as an ongoing, active process. In this study, we share this view. We believe that a treaty's impact goes beyond its written text. Instead, as countries continuously interact, the treaty can deeply influence how customary international law (CIL) is formed, how current laws are understood, and how political and reputational costs work outside the law. See Harold Hongju Koh, "Why Do Nations Obey International Law?", *The Yale Law Journal*, Vol. 106, No. 8, 1997, pp. 2618-2624.

2.1. The Applicability and Limits of Existing International Law

The assertion that a new treaty is required does not negate the applicability of existing international law. Although military AI can be considered novel technology, it does not preclude the applicability of the existing legal framework. This position finds support in the International Court of Justice's *Nuclear Weapon* advisory opinion. The Court held that although nuclear weapons were invented after most of the principles and rules of humanitarian law, "it cannot be concluded from this that the established principles and rules of humanitarian law applicable in armed conflict did not apply to nuclear weapons".¹³ Otherwise, "such a conclusion would be incompatible with the intrinsically humanitarian character of the legal principles in question".¹⁴ From this perspective, what international law primarily governs is the conduct of states and the consequences of their behaviour. Thus, while novel means and methods of warfare may generate unprecedented legal challenges, mere novelty of the tool does not preclude the general application of international law. This jurisprudential logic applies equally to military AI.

Furthermore, persuasive evidence of this applicability can be found in recent state practice. For example, within the United Nations framework, General Assembly Resolutions A/78/241, A/79/62, A/79/239, and A/80/58 all affirmed the applicability of international law, including the Charter of the United Nations, international humanitarian law (IHL), international human rights law, and international criminal law, to autonomous weapons systems and AI in the military domain.¹⁵ Together, this jurisprudence and state practice unequivocally affirm the baseline applicability of existing international law.

What really divides proponents and opponents of a new treaty is not the applicability of existing international law, but its sufficiency. A typical argument against such a treaty is that existing international law could address the legal problems posed by new technologies. However, this article submits that such a view is overly optimistic. Although existing international law can be interpreted or even evolve, challenges posed by military AI are beyond its scope as military AI develops.

2.2. Moral Dilemmas

The first primary concern raised by military AI relates to the moral implications of its deployment. At the most fundamental level, opposition to Autonomous Weapons Systems (AWS) rests on the belief that it is morally unacceptable for machines to determine who lives or dies.¹⁶ Compared with the AWS, humans have the moral

¹³ International Court of Justice, *Legality of the Threat or Use of Nuclear Weapons*, Advisory Opinion, ICJ Reports, 8 July 1996, para. 86.

¹⁴ *Ibid.*

¹⁵ UNGA Res. 78/241, 28 December 2023; UNGA Res. 79/62, 10 December 2024; UNGA Res. 79/239, 31 December 2024; UNGA Res. 80/58, 5 December 2025.

¹⁶ ICRC Position, p. 8.

agency to guide their actions.¹⁷ Therefore, the moral elements may make people decide not to use force even if it would be lawful, whereas the AWS does not.¹⁸ However, in the view of other scholars, robots can neither behave morally nor immorally.¹⁹ Only human beings can be inhuman and choose not to comply with international law.²⁰ In their view, when military AI is well-regulated, it may help to facilitate compliance with international law and protect human life. More importantly, they argue that moral concerns do not automatically equate to legal obligations. While morally debatable, there are currently no positive rules explicitly prohibiting or restricting the use of military AI solely on moral grounds.²¹

Pragmatically speaking, this article admits that the application of military AI *per se* may not be immoral and should not be fully forbidden. After all, in the future, AI could perform better than humans in some situations regarding compliance with IHL.²² In that case, applying military AI may promote compliance with international law rather than diminish its influence. However, under other circumstances, the moral risks posed by military AI can also not be overlooked. For instance, when an enemy combatant who is severely wounded and filled with terror fails to exhibit a standard surrender gesture due to fear or physical inability, a human soldier with empathy, morals, and instinct could sensitively perceive the genuine intention to surrender and withhold fire. In contrast, an algorithmically driven system lacking compassion might strictly adhere to its programming and engage the incapacitated target. Even though such an action may formally be consistent with the strict requirements of specific rules under IHL (such as the criteria for *hors de combat*),²³ it reduces the “live or dead” decision to a mere data-processing process. It is an action that puts human dignity aside and arbitrarily takes a human’s life that originally could have been saved.

In this case, the use of military AI challenges the dictates of public conscience established by the Martens Clause of international law. The Martens Clause serves

¹⁷ *Ibid.*

¹⁸ *Ibid.*

¹⁹ Marco Sassóli, “Autonomous Weapons and International Humanitarian Law: Advantages, Open Technical Questions and Legal Issues to be Clarified”, *International Law Studies*, Vol. 90, 2014, p. 313.

²⁰ *Ibid.* at pp. 310, 313-314.

²¹ See Jiawei Chu [褚家玮], *International Law and Autonomous Weapons* [国际法与自主武器], Current Affairs Press [时事出版社], Beijing, 2024, p. 167.

²² See M. Sassóli, above note 19, p.320. Despite this, a counterargument is that the human/AI comparison reflects the assumption of static, hierarchical, and replaceable relations, as well as an oversimplification of the complex socio-technical interactions that characterise AI-enabled targeting practices. Therefore, AI technologies cannot be assessed *in abstracto*. On this point, see Klaudia Klonowska and Taylor Kate Woodcock, “Rhetoric and Regulation: The (Limits of) Human/AI Comparison in Legal Debates on Military AI”, in Bérénice Boutin et.al. (eds.), *Legal, Ethical, and Technical Dilemmas in Military Artificial Intelligence*, T.M.C. Asser Press, Hague, 2026, p.52-70.

²³ Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts (Protocol I), 1125 UNTS 3, 8 June 1977 (entered into force 7 December 1978), Art. 41. (Additional Protocol I)

as a safety net for international law by requiring the application of principles of humanity and public conscience when specific regulations are absent.²⁴ However, the Martens Clause is inherently abstract, and applying it directly to regulate military AI remains difficult.²⁵ When specific positive rules are absent, and some specific means of applying military AI are deemed consistent with international law and could even bring military advantages, states are unlikely to refrain from using it solely on moral grounds. Because the Martens Clause primarily serves as a safety net rather than a detailed standard, the absence of a clear standard in existing international law leaves a dangerous legal vacuum. As aforementioned, the necessity and rationale for a new legally binding instrument is not that existing law is inherently flawed or inapplicable. It is because existing law is insufficient to echo the emerging new technological challenges. Thus, as the Martens Clause requires, a new treaty is necessary not just to debate morality but to translate these moral requirements into specific binding legal restrictions.

2.3. Accountability Gap

Another question raised by military AI is the accountability gap. Accountability demands that someone should be held liable for all wrong committed.²⁶ Thus, it is common to ask: when military AI has violated international law, like war crimes, crimes against humanity, or genocide, who should be responsible?²⁷ If the military AI commits all of the acts, should we allow AI to be the subject to take the responsibility? And is there any practical significance to doing this? In this regard, both Human Rights Watch and the Group of Governmental Experts (GGE) under the CCW framework have been concerned about the accountability issue raised by AWS.²⁸ For instance, the GGE has confirmed that an autonomous weapon system should not be anthropomorphized,²⁹ and IHL only imposes obligations on states, parties to armed conflict, and individuals, but does not impose obligations on machines.³⁰ Instruments as such seem to maintain the states' and individuals' status

²⁴ Under Additional Protocol I, the Martens Clause was expressed as follows: "In cases not covered by this Protocol or by other international agreements, civilians and combatants remain under the protection and authority of the principles of international law derived from established custom, from the principles of humanity and from the dictates of public conscience." *Ibid.*, Art. 1 (2).

²⁵ For instance, Antonio Cassese contends that the Martens Clause is particularly ambiguous. It is very loosely worded and has consequently given rise to a multiplicity of often conflicting interpretations. Antonio Cassese, "The Martens Clause: Half a Loaf or Simply Pie in the Sky?", *European Journal of International Law*, Vol. 11, No. 1, 2000, pp. 187-188.

²⁶ Swati Malik, "Autonomous Weapon Systems: The Possibility and Probability of Accountability", *Wisconsin International Law Journal*, Vol. 35, No. 3, 2018, p. 620.

²⁷ See L. Blank, above note 10, at p. 234.

²⁸ Human Rights Watch, *Mind the Gap: The Lack of Accountability for Killer Robots*, New York, 9 April 2015; Report of the 2018 session of the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems, UN Doc. CCW/GGE.1/2018/3, 23 October 2018, para. 21(c).

²⁹ *Ibid.*, para. 21 (h).

³⁰ Report of the 2019 session of the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems, UN Doc. CCW/GGE.1/2019/3, 25 September 2019, para. 17(b). (GGE 2019 Report)

as the sole responsible entities. However, in practice, this issue remains controversial and complicates the practical allocation of responsibility for two primary reasons.³¹

First, the conduct and the consequences of the AWS are unpredictable.³² Unlike conventional weapons, AWS is based on its trained algorithms rather than human direction, which means the human operator of the machine, in principle, cannot predict the future machine behaviour.³³ Because of this unpredictability, during military operations, even the human operator cannot fully predict the consequences of acts that violate international law. Thereby cutting the actual causal link between the human operator's actions and the AWS's outcomes, creating a significant accountability gap.³⁴ For instance, a troop deploys with AWS, whose accuracy is 99%, to conduct a mission aimed at destroying the enemy's armoured vehicles. However, during the mission, due to an algorithmic hallucination or an undetected problem, the AWS misrecognized a school bus as a target and conducted the attack.³⁵ Who should be responsible for that? If a human soldier committed such an action with intent or at his commander's direction, the soldier and his commander would certainly be held responsible under international criminal law. However, when such an action is conducted by an AWS, a lack of predictability fails to satisfy the strict mental element under international criminal law.³⁶ Specifically, the human commander does not satisfy the "should have known" standard under Article 28 of the Rome Statute,³⁷ nor can the operator of the AWS have "knowledge" under the "ordinary course of events" based on Article 30 in such a situation,³⁸ leaving both the soldiers and their commander unable to take the responsibility.

Second, during the development and operational use of AWS, "the problem of many hands" may also affect the distribution of responsibility.³⁹ This problem is relevant to the development and manufacturing of AWS, which requires the participation of programmers, engineers, different researchers, organizations, and companies.⁴⁰ When a mistake or malfunction occurs with AWS and violates

³¹ It should be noted that the term "responsibility" addressed in the following paragraphs of this section does not refer to "state responsibility" reflected in the ARSIWA and CIL. Instead, it refers to individual responsibility, such as that of commanders, soldiers, programmers, and engineers.

³² J. Chu, above note 21, at p. 171.

³³ See Andreas Matthias, "The Responsibility Gap: Ascribing Responsibility for the Actions of Learning Automata", *Ethics and Information Technology*, Vol. 6, 2004, p. 175.

³⁴ See J. Chu, above note 21, pp. 171-172.

³⁵ See Charles Trumbull, "Future of Warfare and Law Series – Addressing Uncertainty in the Use of Autonomous Weapons Systems", *Articles of War*, 5 November 2025, available at: <https://lieber.westpoint.edu/addressing-uncertainty-use-autonomous-weapons-systems/>.

³⁶ Jonathan Kwik, "Obstacles to Criminal Liability: A Systemic Analysis. In: Lawfully Using Autonomous Weapon Technologies", in *Lawfully Using Autonomous Weapon Technologies*, TMC Asser Press, Hague, 2024, pp. 328-334.

³⁷ Rome Statute of the International Criminal Court, 2187 UNTS 3, 17 July 1998 (entered into force 1 July 2002) (Rome Statute), Art. 28.

³⁸ Rome Statute, Art. 30.

³⁹ See J. Chu, above note 21, pp. 172-174.

⁴⁰ *Ibid.*

international law, it is usually difficult to isolate the responsibility of different participants.⁴¹ Consider a scenario: a company with hundreds of software engineers and programmers developed an AWS with a set of flawed algorithms, but without any intent to commit a war crime, and such a system was extensively tested. Several years later, in an international armed conflict, the frontline commander may use and deploy such AWS in good faith and trust its verified accuracy in previous tests. However, in that real conflict, such a system has malfunctioned multiple times, causing unnecessary civilian casualties. Who should be responsible for that? Under international criminal law, such responsibility is inevitably diluted. Since the result was created by the engineer, programmer, and commander together, none of them, individually, could have violated international criminal law for lack of *mens rea*. Not to mention, after quite a long time, it may be difficult to trace which specific programmers or engineers made such a mistake and attribute the responsibility to them.

To address this accountability gap, many scholars have proposed various solutions, such as expanding the scope of responsibility under international criminal law,⁴² establishing human judgment,⁴³ or establishing “war torts”⁴⁴. However, all those views remain at the proposal stage; and none have been adopted as a rule of international law and have binding force on states and their behaviour. In fact, absent a dedicated treaty and broad international consensus, the realistic chance of these proposals being formally adopted into the current legal framework remains very slim. Admittedly, some elements among them, such as meaningful human control, could serve as an interpretative factor that guides the application of IHL. But at least in attributing criminal responsibility to individuals, such interpretative factors are grossly insufficient. Based on the principle of *nullum crimen sine lege*, rooted in international criminal law, the definition of a crime shall not be extended by analogy, and ambiguities shall be interpreted in favour of the person being prosecuted.⁴⁵ Therefore, merely interpreting existing laws to solve the accountability gaps would be fundamentally in conflict with the established legal framework. For instance, holding a commander or operator liable for unforeseeable violations caused by an AWS would require either extending the “knowledge” standard under Article 30 of the Rome Statute or expanding the scope of commanders’ responsibility under Article 28. Both interpretations would be inconsistent with the rule of strict construction. Therefore, to hold humans accountable for AI’s actions, the international community must establish new rules for the deployment of military AI.

⁴¹ *Ibid.*

⁴² Hin Yan Liu, “Refining Responsibility: Differentiating Two Types of Responsibility Issues Raised by Autonomous Weapons Systems”, in Nehal Bhuta *et al.* (eds), *Autonomous Weapons Systems: Law, Ethics, Policy*, Cambridge University Press, Cambridge, 2016, pp. 325-344.

⁴³ Jack M. Beard, “Autonomous Weapons and Human Responsibilities”, *Georgetown Journal of International Law*, Vol. 45, No. 3, 2014, p. 621.

⁴⁴ See Rebecca Crotoft, “War Torts: Accountability for Autonomous Weapons”, *University of Pennsylvania Law Review*, Vol. 164, No. 6, 2016, pp. 1347-1402.

⁴⁵ Rome Statute, Art. 22.

From this perspective, establishing a new treaty to discuss, enact, and confirm these principles and solutions is necessary.

2.4. Compliance Deficits

The absence of a dedicated treaty also creates significant practical challenges in complying with specific rules of IHL. The ambiguity of terms under existing international law and the uncertainty surrounding their application make compliance difficult and create numerous compliance deficits.

Concerning the legal review of new weapons, it is often posited that Article 36 of Additional Protocol I could serve as a mechanism to solve some of the humanitarian issues of military AI, as it requires the states to determine whether the weapon's employment would be prohibited by the Protocol or by any other applicable rules of international law.⁴⁶ It should be recognized that if the review process stipulated in Article 36 is strictly implemented, it is indeed possible to avoid the deployment of weapons that are incompatible with IHL. However, existing international law has not yet established a complete mechanism for the legal review of autonomous weapons systems.⁴⁷ It does not specify the scope of legal review, nor does it provide standardized substantive or procedural rules, leading to highly divergent state practices.⁴⁸ For instance, whether conventional weapons that are attached with artificial intelligence should be subjected to the weapon reviews is undetermined. Whether a uniform procedure should be applied during the review is uncertain. Whether the means and methods of warfare should be under review is undefined.⁴⁹ This uncertainty allows states to conduct reviews based on subjective preferences and varying standards, rendering Article 36 practically ineffective as a unified safeguard. Furthermore, in practice, most States Parties to Additional Protocol I may choose not to conduct formal weapon reviews.⁵⁰ If this persists, the role of Article 36 as a gatekeeper may be fundamentally undermined.

Admittedly, while Article 36 lacks specific procedural mechanisms for all weapons, this silence creates profound compliance problems when applied to military AI. For example, Article 36 triggers the review obligation upon the study, development, acquisition, or adoption of a “new weapon, means or method of warfare”.⁵¹ Since conventional weapons undergo virtually no substantive changes after deployment, states can meet this obligation by conducting a single weapons

⁴⁶ Additional Protocol I, Art. 36.

⁴⁷ Shouping Li [李寿平], “Improvement and Development of International Legal Regulation of Autonomous Weapon Systems” [自主武器系统国际法律规制的完善和发展], *Law Review* [法学评论], No. 1, 2021, p. 170.

⁴⁸ *Ibid.*

⁴⁹ *Ibid.*

⁵⁰ See James Farrant and Christopher M. Ford, “Autonomous Weapons and Weapon Reviews: The UK Second International Weapon Review Forum”, *International Law Studies*, Vol. 93, 2017, p. 401.

⁵¹ Additional Protocol I, Art. 36.

review prior to their operational use. However, for military AI, the adaptive capabilities gained through machine learning endow it with self-evolving characteristics.⁵² This means it could potentially constitute a “new weapon” or “new method of warfare” at any given moment, necessitating weapon review throughout its entire lifecycle. Against this backdrop, conducting weapon reviews solely prior to deployment fails to meet the objective requirement for continuous scrutiny of military AI.⁵³ To ensure the weapon review mechanism really works and is effectively observed by states, a new treaty needs to refine the specific criteria for the review to provide effective guidance.

At the same time, the aforementioned Martens Clause, as one of the most fundamental principles of IHL, is also challenged by differing interpretations in its application. As early as 2012, Human Rights Watch and the International Human Rights Clinic at Harvard University argued that fully autonomous weapons would likely contravene the Martens Clause because they would run counter to the dictates of public conscience.⁵⁴ Despite such a broad interpretation being NGO-friendly, it is likely to be rejected by some States, as they prefer a narrow view of the Martens Clause and do not deem the principles of humanity or the dictates of public conscience as independent sources of international law.⁵⁵ Against the background of such differences, if a state merely regards the Martens Clause as a protective umbrella, not an independent source of obligation, then it cannot effectively guide the country’s behaviour in military AI. Therefore, the advocacy of a new treaty does not necessarily require states to adopt a broader interpretation of the Martens Clause. Rather, it provides a forum for states to exercise their sovereign rights and translate abstract humanitarian concerns into concrete, negotiated treaty rules.

Both examples demonstrate that different interpretations of the same rules under international law can lead to varying levels of compliance. Some middle and small powers, as well as human rights organizations, are likely to adopt a strict standard to restrict major military powers and ensure that such military AI does not harm them. While some countries with strong military capabilities, because they are better able to deploy military AI in actual military conflict, are more likely to remain silent or take a loose interpretation of the rules meant to regulate them. Such a

⁵² See Damian Copeland *et al.*, “The Utility of Weapons Reviews in Addressing Concerns Raised by Autonomous Weapon Systems”, *Journal of Conflict and Security Law*, Vol. 28, No. 2, 2022, p. 294.

⁵³ As the rolling text notes, “any further development of, or modifications to, existing LAWS that have a significant impact on their target identification, selection, and engagement functions or on their anticipated effects, are reviewed to ensure compliance with IHL.” However, such expressions currently do not have any legally binding force. See GGE on LAWS, “Rolling text”, 18 December 2025, available at: [https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-_Group_of_Governmental_Experts_on_Lethal_Autonomous_Weapons_Systems_\(2026\)/CCW_GGE_LAWS_Rolling_Text_-_status_18_December_2025.pdf](https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-_Group_of_Governmental_Experts_on_Lethal_Autonomous_Weapons_Systems_(2026)/CCW_GGE_LAWS_Rolling_Text_-_status_18_December_2025.pdf).

⁵⁴ Human Rights Watch and International Human Rights Clinic, *Losing Humanity: The Case against Killer Robots*, United States of America, 19 November 2012, p. 4.

⁵⁵ See Tyler D. Evans, “At War with the Robots: Autonomous Weapon Systems and the Martens Clause”, *Hofstra Law Review*, Vol. 41, No. 3, 2013, pp. 723-726.

deduction is evidenced by practice: while the Association of Southeast Asian Nations,⁵⁶ African Group,⁵⁷ and the Non-Aligned Movement⁵⁸ are actively proposing to establish a legally binding instrument, other nations, including the US, Russia, and Israel, emphasize the sufficiency of existing international law while simultaneously developing or deploying AWS in active conflicts. Thus, merely relying on existing international law will lead to multiple and fragmented interpretations of international law and further deepen divergences.⁵⁹ Without any specific action to facilitate the formation of an international consensus, the so-called “applying existing international law” approach will serve as a shield to prevent countries with strong military capabilities from being subject to international law.

2.5. Where Does the Necessity Come From: Legal Certainty as a Precondition for Compliance

From a jurisprudence perspective, legal certainty always plays a significant role in international legal systems.⁶⁰ Although the previous sections have identified specific doctrinal gaps, a more profound crisis lies in the loss of legal certainty. For hundreds of years, despite IHL possessing a degree of inherent constructive ambiguity, it has generally provided a stable and predictable baseline for states and offered a framework to guide them on what to do and what not to do during armed conflicts. However, in the context of rapidly developing military AI, the most basic clarity and predictability appear unattainable, and legal certainty has been eroded by legal indeterminacy.

⁵⁶ “Statement by His Excellency Chhea Keo Ambassador and Permanent Representative of the Kingdom of Cambodia to the United Nations, on behalf of the Association of Southeast Asian Nations (ASEAN)”, ASEAN, for the 80th Session of the General Assembly, 23 October 2025, available at: https://reachingcriticalwill.org/images/documents/Disarmament-fora/1com/1com25/statements/23Oct_ASEAN.pdf.

⁵⁷ “Statement delivered by Nigeria on behalf of the African Group Member States that are High Contracting Parties to Convention on Certain Conventional Weapons”, November 2023, available at: [https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_Meeting_of_High_Contracting_Parties_\(2023\)/Nigeria_-_African_Group_-_2023_MHCP.pdf](https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_Meeting_of_High_Contracting_Parties_(2023)/Nigeria_-_African_Group_-_2023_MHCP.pdf).

⁵⁸ Working paper by the Bolivarian Republic of Venezuela on behalf of the Non-Aligned Movement (NAM) and Other States Parties to the Convention on Certain Conventional Weapons (CCW), 2022, available at: <https://documents.unoda.org/wp-content/uploads/2022/08/WP-NAM.pdf>.

⁵⁹ One might argue that if such divergences exist when states interpret existing international law, they may also exist when the international community is negotiating a new treaty. This is entirely possible due to states’ differing considerations, which will be discussed in Part 3. However, there are still essential differences between establishing a new treaty and interpreting the existing law. Unilateral interpretation permits states to exploit the ambiguity of existing law to preserve their strategic flexibility, which, by its nature, favors states unwilling to regulate military AI. In contrast, the treaty-making process serves as a mechanism that compels states to articulate their legal positions and red lines. This mechanism will help build consensus within the international community and reduce divergences, as will be further discussed in Part 4.

⁶⁰ See James R. Maxeiner, “Some Realism About Legal Certainty in the Globalization of the Rule of Law”, *Houston Journal of International Law*, Vol. 31, No. 1, 2008, p. 30.

Legal certainty in military AI is eroded by two forces, both of which make the mere interpretation of existing law incapable of solving the current legal predicament. Firstly, legal certainty is undermined by the fragmented interpretation. While states can give new meaning to existing rules to address technological changes, when multiple interpretations arise on the same issue, none of which has been widely recognized by the international community, there is no clarity in the law to follow.⁶¹ Secondly, legal certainty is affected by the ambiguity of emerging soft law. When numerous proposals have been raised,⁶² none of which has been confirmed as hard law, they cannot create any binding force for states. Meaningful human control is a significant example. It has been addressed in many diplomatic discourses but has not yet been incorporated into the *lex lata*.⁶³ In the absence of a clear treaty definition, states retain excessively broad discretion to interpret this concept, thus making “meaningful human control” lose its due limitation. Meanwhile, such instances also blur the line between legal obligation and political stance, increasing legal ambiguity and decreasing legal certainty. In this regard, merely relying on existing international law is insufficient. In an international society without a central legislative institution, existing law is neither able to actively choose which interpretation is authentic nor to adopt influential new proposals into it.

In this context, one might raise necessary concerns about the consequences of establishing a new treaty, and we must acknowledge their reasonableness. This perspective first contends that the pace of technological development far outstrips the speed of human imagination. Meaning, a new treaty may fail to address emerging issues in the future and create difficulties in defining regulatory content.⁶⁴ Second, this perspective also notes that a new treaty could undermine the existing protective framework of IHL.⁶⁵ This could lead to deviations in the protection afforded to specific actions from the provisions of the Geneva Conventions, or enable non-signatory states to argue that existing IHL does not apply to their military AI technologies.⁶⁶ In other words, such views hold that a new treaty not only fails to

⁶¹ This phenomenon was partially reflected in the “Compliance Deficits” part in the former section.

⁶² See, e.g., REAIM, “REAIM Pathways to Action”, Summit on Responsible Artificial Intelligence in the Military Domain, Spain, 4-5 February 2026, available at: <https://www.exteriores.gob.es/en/REAIM2026/Documents/REAIM%202026%20Pathways%20to%20Action.pdf>; Lieber Institute for Law and Land Warfare, “Manual on International Law Applicable to Artificial Intelligence in Warfare”, last visited 19 May 2026, available at: <https://lieber.westpoint.edu/research/the-west-point-manual/>.

⁶³ “However, the analysis of the concept of meaningful human control in this report shows that even if the concept is developing, it is not sufficiently defined to form the basis of a regulation on autonomous weapon systems at this point.” See 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”, February 2020, p. 49, available at: <https://www.fcas-forum.eu/publications/Meaningful-Human-Control-of-Autonomous-Weapon-Systems-Eklund.pdf>.

⁶⁴ L. Blank, above note 10, pp. 251-252.

⁶⁵ *Ibid.*, p. 252.

⁶⁶ *Ibid.*

resolve future legal uncertainties but also introduces uncertainty in the application of existing law.

Regarding this perspective, this article acknowledges that these concerns pose challenges that must be addressed in the future treaty-making process. However, the article also contends that while these concerns seek to preserve legal stability, they inadvertently perpetuate legal uncertainty. First, the notion that “technological development outpaces human imagination” applies not only to new treaties but also to existing IHL. The aforementioned moral dilemmas, accountability gaps, and compliance deficits stem precisely from the states’ failure to anticipate future technological developments in the past. Therefore, addressing pressing current issues based on the international community’s present understanding of AI technology is far more meaningful than inaction. Simultaneously, this argument also misinterprets the purpose of establishing the treaties. Establishing new treaties does not equate to regulating specific, micro-level AI technologies. Rather, it involves creating international norms aligned with the overarching characteristics of AI technology, such as continuous weapons review and meaningful human control, which can also adapt to future technological transformations to a significant extent.

At the same time, concerns about undermining the existing framework for applying IHL overlook the relationship between general and special rules in international law. The establishment of a new treaty aims to apply IHL specifically in the context of military AI, not to replace general IHL. Even if there are concerns that this issue could serve as an excuse for some states to refuse to apply existing IHL, a new treaty could explicitly incorporate a “savings clause” stipulating that its application does not affect the application of existing IHL.

Based on this background, while existing international law still plays an important role in regulating military AI, establishing a new treaty in this field is *de jure* necessary. It is true that the establishment of a new treaty cannot eliminate all interpretation disputes, because international law is essentially the product of the coordination of will among States.⁶⁷ However, the purpose of establishing a new treaty is not to create a perfect regulatory mechanism overnight, but to clarify the necessary regulatory baseline, address emerging legal issues, and provide the most basic legal certainty. Firstly, a new treaty could offer an authoritative text that, while perhaps still containing compromises, could narrow the space for unilateral interpretation. Secondly, the treaty could actively consider emerging concepts and address new concerns, such as meaningful human control and moral dilemmas, thereby elevating soft law from *lex ferenda* to *lex lata*, filling legal gaps, and resolving legal problems. Only by doing so could states gain clear guidance and direction and comply with international law without ambiguity. Ultimately, the purpose of

⁶⁷ See Gengsheng Zhou (S. R. Chow) [周鲢生], *International Law* [国际法], The Commercial Press [商务印书馆], 2018, pp. 1-10.

establishing a new treaty is to use the precise text agreed upon by states to replace the legal interpretation that can exploit the ambiguity of the law, thereby achieving the rule of law in the field of military AI.

3. The De Facto Improbability: Political Impasse and Majority Follows Minority Dynamic

Although establishing a new treaty for military AI is *de jure* necessary, its establishment has been impeded by various factors. In this part, the article will examine the possibility of a treaty-making process under the CCW and any alternative process beyond it. This article submits that establishing a new treaty for military AI under the CCW mechanism is unlikely. Even if such a treaty were adopted in an alternative mechanism without the participation of the major military powers, it would be unlikely to effectively regulate the application of military AI. Having concluded that such a new treaty is *de facto* improbable, the article will outline the factors that drive this political and legal impasse and present a realist picture of the treaty-making process. Unlike most countries' expectations, the regulation of military AI is a process where the "majority follows the minority".

3.1. Possibility of Establishing a Military AI Treaty Under CCW and Alternative Mechanisms

3.1.1. CCW Mechanism

When considering the establishment of a new treaty for military AI, the CCW process is undoubtedly one of the most influential forums. Historically, this process has adopted five influential protocols regulating different types of conventional weapons ranging from non-detectable fragments and mines to incendiary weapons, laser weapons, and explosive remnants of war.⁶⁸ In 2013, the meeting of States Parties to the CCW decided that the meeting of 2014 would discuss the questions relating to lethal autonomous weapons systems (LAWS),⁶⁹ and would begin its regulatory process for AI in the military domain. Subsequently, the CCW has held annual meetings to discuss autonomous weapons systems,⁷⁰ and established a formal Open-Ended Group of Governmental Experts in 2016.⁷¹ In 2019, the CCW

⁶⁸ Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which may be Deemed to be Excessively Injurious or to Have Indiscriminate Effects, 1342 UNTS 137, 10 October 1980 (entered into force 2 December 1983).

⁶⁹ *Final Report of the Meeting of the High Contracting Parties to the Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to Be Excessively Injurious or to Have Indiscriminate Effects*, UN Doc. CCW/MSP/2013/10, 16 December 2013, para. 32.

⁷⁰ The details of the discussed topic under the CCW mechanism can be found in B. Boothby, above note 4, pp. 405-472.

⁷¹ *Final Document of the Fifth Review Conference of the High Contracting Parties to the Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to Be Excessively Injurious or to Have Indiscriminate Effects*, UN Doc. CCW/CONF.V/10, 23 December 2016, p. 9.

mechanism published eleven guiding principles,⁷² marking a significant milestone and a major consensus within the CCW at the time. However, such guiding principles are non-binding in nature, and since the start of the discussion regarding LAWS in 2013, there has been no mandate to negotiate a treaty.⁷³

Contrary to initial optimism, many factors continue to block treaty negotiations under the CCW process. From the perspective of substantive differences in negotiations, countries have continued to diverge on definitions, technical, security, legal, ethical, and future aspects, none of which have been fully resolved.⁷⁴ Consequently, future negotiations are unlikely to proceed smoothly. Although the CCW process will remain the main forum for multilateral discussions on LAWS, international competition over the development and rule-setting power of LAWS will continue to intensify, and the prospects for LAWS arms control under the CCW process remain bleak.⁷⁵

In fact, after the last CCW protocol was passed in 2003, more than two decades have passed without the adoption of a new protocol, and the future usefulness of the CCW process has been doubted in recent years⁷⁶. The CCW process's weaknesses can be attributed to procedural requirements during new-protocol negotiations. The practice that has evolved during CCW negotiations is that consensus is required before adopting amendments to the Convention or Protocols, or adopting new Protocols.⁷⁷ However, with the growing number of CCW State Parties, achieving such a consensus is challenging.⁷⁸ During the current discussion regarding autonomous weapon systems, while about 130 states actively support the negotiation of a legally binding instrument, a small but decisive twelve states, including the U.S., Russia, India, Japan, South Korea, and Australia, explicitly refuse to adopt a legally binding instrument.⁷⁹ Despite their small number, under the

⁷² See GGE 2019 Report, above note 30.

⁷³ See B. Boothby, above note 4, p. 406.

⁷⁴ The divergences under the CCW process include: (1) Definition Divergence: Should a feasible definition of LAWS be formulated now, and how should it be formulated? (2) Technical Divergence: What are the paths for exploring the characteristics of LAWS, and can strong artificial intelligence arrive, and when? (3) Security Divergence: Will the impact of LAWS be more beneficial or detrimental? (4) Legal Divergence: Does LAWS require new, additional regulation beyond international humanitarian law? (5) Ethical Divergence: Can machines become moral agents? Are machines tools for humans or intelligent agents? (6) Future Divergence: Should we implement preventative prohibition, temporarily allow it, or even encourage the development of LAWS? See Nengwu Xu [徐能武] and Kun Long [龙坤], "The Focus and Trend of Debate on Arms Control of Lethal Autonomous Weapons Systems under the UN CCW Framework" [联合国 CCW 框架下致命性自主武器系统军控辩论的焦点与趋势], *Journal of International Security Studies* [国际安全研究], No. 5, 2019, pp. 124-128.

⁷⁵ *Ibid.*, pp. 128-131.

⁷⁶ William H. Boothby, *Weapons and the Law of Armed Conflict*, 2nd ed., Oxford University Press, Oxford, 2016, p. 18.

⁷⁷ *Ibid.*, p. 99.

⁷⁸ *Ibid.*, p. 18.

⁷⁹ About the State Position Collection, see State Positions Database.

consensus requirements of the CCW, their refusal plays a veto role. Without a dramatic shift in position, the establishment of a new treaty through the CCW is improbable.

3.1.2. *Alternative Mechanism Beyond the CCW*

For the reasons above, the international community is also exploring alternative processes outside the CCW, aiming to negotiate a legally binding instrument. In 2022, the Human Rights Watch and the International Human Rights Clinic at Harvard Law School published a research report titled “An Agenda for Action: Alternative Processes for Negotiating a Killer Robots Treaty” to promote such an alternative mechanism.⁸⁰ The report suggested that the alternative mechanism should have the following characteristics: a common purpose of creating new law, a model of decision-making by majority vote, a clear and ambitious deadline, and an inclusive approach.⁸¹

When the CCW process is unworkable for negotiating a legally binding instrument, establishing a new treaty through an alternative process would be a smarter option and more likely to succeed. As Human Rights Watch observed, the Mine Ban Treaty, the Convention on Cluster Munitions, and the Treaty on the Prohibition of Nuclear Weapons all offer valuable precedents for establishing a treaty on military AI.⁸² Taking the Convention on Cluster Munitions (CCM) as an instance: after consistently failing to reach a consensus under the CCW mechanism,⁸³ countries and organizations concluded that the existing negotiation framework was disproportionately influenced by major military powers, resulting in delays in developing a dedicated protocol on cluster munitions.⁸⁴ Therefore, after 2006, those countries and organizations that strongly advocated for a ban began seeking to establish an international treaty outside the framework of the CCW.⁸⁵ After the launch of the Oslo Process in February 2007, they finally reached a Convention in 2008.⁸⁶ Compared to the long, fruitless CCW negotiation process, the Oslo Process took only about 2 years and resulted in a treaty. The advantages of the alternative

⁸⁰ See Human Rights Watch and International Human Rights Clinic, *Agenda for Action: Alternative Processes for Negotiating a Killer Robots Treaty*, United States of America, November 2022.

⁸¹ *Ibid.*, p. 4.

⁸² *Ibid.*, p. 15-25.

⁸³ Zhixiong Huang [黄志雄] and Lin Fan [范琳], “The 2008 Convention on Cluster Munitions under the Trend of Humanization of International Law” [国际法人本化趋势下的 2008 年《集束弹药公约》], *Law Review* [法学评论], No. 1, 2010, p. 69.

⁸⁴ *Ibid.*

⁸⁵ *Ibid.*

⁸⁶ Zhixiong Huang [黄志雄], “Non-Governmental Organizations and the Convention on Cluster Munitions in 2008: Participation, Influence and Revelation” [非政府组织与 2008 年《集束弹药公约》: 参与、影响及启示], *Wuhan University International Law Review* [武大国际法评论], Vol. 13, 2008, pp. 351-352.

mechanism stand out. Compared to previous successful experiences, the regulation of military AI faces a similar situation. Both scenarios enjoy sufficient support for such a treaty. Regarding AWS regulation, approximately 130 countries support the establishment of a legally binding instrument.⁸⁷ Therefore, despite the blocking elements, achieving a binding treaty under an alternative mechanism in the long term is technically feasible.

However, feasibility must not be confused with the treaty's success. Notwithstanding the possibility of a new treaty being adopted, using such a treaty *per se* to effectively regulate the international community is another question. The fundamental problem of an alternative mechanism is that the key stakeholders of military AI may be absent. Historical precedent on nuclear weapon control provides stark evidence on this issue. For instance, despite ninety-five States having signed the Treaty on the Prohibition of Nuclear Weapons,⁸⁸ none of the recognized nuclear-weapon states have participated in it. Instead, they firmly and directly state their opposition to the Treaty on the Prohibition of Nuclear Weapons and reaffirm the application of the Treaty on the Non-Proliferation of Nuclear Weapons.⁸⁹ This precedent reminds us that, when the same logic applies to the regulation of military AI, the countries that are most likely to own and use AWS are also highly unlikely to participate in a potential convention on military AI. If this is the case, only countries that do not possess AWS would be bound by the treaty, while those that control the technology would operate without restriction; this outcome would render the convention *de facto* futile.

3.2. The Underlying Logic of Weapons Regulation and Treaty-Making

Several factors explain the *de facto* improbability of establishing the treaty. This section examines these obstacles from three perspectives: the strategic imperatives of major military powers, prevailing scepticism towards treaty-making among Western countries, and the current unfavourable international order and political environment.

3.2.1. The Strategic Imperatives of Major Military Powers

From a historical perspective, Hans J. Morgenthau has elaborated on the limitations and the fate of arms control. As an effort to strengthen international peace by increasing military stability, it has failed repeatedly with conventional weapons and

⁸⁷ State Positions Database, above note 6.

⁸⁸ UNODA, "Treaty on the Prohibition of Nuclear Weapons, Participants", last visited 24 May 2026, available at: <https://treaties.unoda.org/t/tpnw/participants>.

⁸⁹ The Ministry of Foreign Affairs of the Russian Federation, "Joint Statement by China, France, Russian Federation, United Kingdom and United States", 29 October 2018, available at: https://www.mid.ru/en/foreign_policy/news/1577389/.

has succeeded only in the nuclear field.⁹⁰ Such control of nuclear weapons is made possible in theory because the ability of major nuclear powers has reached its optimum of assured destruction, beyond which it is irrational to go.⁹¹ By contrast, in conventional realms, it has failed because the quantity and disposition of such weapons will directly affect the distribution of military power and military advantage, thereby undermining effective regulation.⁹²

Artificial intelligence, as a civil-military dual-use technology, exhibits characteristics closer to those of conventional weapons than to those of nuclear weapons. Compared to nuclear weapons that guaranteed deterrence and mutual assured destruction, military AI can provide sophisticated decision support and conduct precision strikes directly. These characteristics mean that the deployment of military AI directly dictates battlefield supremacy. Given its disruptive potential to alter the balance of power, states are locked in a security dilemma.

Unwillingness to establish a treaty in related fields is by no means unique to one country. On the contrary, it reflects the strategic considerations among all technologically advanced countries. For example, although China has declared its support for the establishment of a new treaty,⁹³ the structure of the international system determines that, without the commitment of its military adversaries, no great power, including China, Russia, and the US, will give up such a significant military advantage and join such a treaty directly. Instead, all of these countries are integrating AI into their militaries and trying to significantly increase their military advantages.⁹⁴ Under this situation, if a group of states, such as the US and its allies,

⁹⁰ Hans J. Morgenthau, *Politics among Nations: The Struggle for Power and Peace*, 7th ed., revised by Kenneth W. Thompson and W. David Clinton, McGraw-Hill/Irwin, New York, 2005, p. 423.

⁹¹ *Ibid.*

⁹² *Ibid.*

⁹³ Indeed, such support for establishing a binding instrument was not unconditional. As Ambassador Shen Jian stated regarding the facilitation of a legally binding instrument: "It deserves further discussions on these specific matters. It is recommended that the GGE further elaborate the interpretation, application, and implementation of specific rules of IHL, taking into consideration the nature of emerging military technologies, their modes of operation, and the operational environment." See UNODA, "Ambassador SHEN Jian's Remarks at the First Session of the 2024 CCW Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapon Systems (GGE on LAWS)", 2024, available at: [https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-_Group_of_Governmental_Experts_on_Lethal_Autonomous_Weapons_Systems_\(2024\)/GGE_LAWS_-_China_-_General_statement_-_English_Copy.pdf](https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-_Group_of_Governmental_Experts_on_Lethal_Autonomous_Weapons_Systems_(2024)/GGE_LAWS_-_China_-_General_statement_-_English_Copy.pdf); Permanent Mission of the People's Republic of China to the UN Office at Geneva and Other International Organizations in Switzerland, "Statement by Ambassador SHEN Jian at the Second Session of the 2024 CCW Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapon Systems", 27 August 2024, available at: https://geneva.china-mission.gov.cn/eng/dbtxwx/202408/t20240827_11480159.htm.

⁹⁴ Priyesh Mishra, *et al.*, *Code, Command, and Conflict: Charting the Future of Military AI*, Harvard Kennedy School Belfer Center for Science and International Affairs, Cambridge, MA, 12 December 2025, pp. 5-6, available at: <https://www.belfercenter.org/research-analysis/code-command-and-conflict-charting-future-military-ai>.

explicitly refuses to participate in this treaty, it would be extremely dangerous for their adversaries, such as China and Russia, to accept such unilateral constraints.

Consequently, applying Morgenthau's explanation to military AI could help explain why several countries are unwilling to participate in its regulation. On the one hand, since none of the states can ascertain their adversaries' strategic intentions, developing military AI that could enhance their own capabilities, rather than restricting them, becomes the only viable solution for some states. In this context, unilateral constraints will only increase the vulnerability of states. Meanwhile, for nations leading in military capabilities, the motivation behind developing military AI extends beyond mere strategic defence. Instead, they seek to cultivate an absolute advantage through asymmetric means and widen the gap in military superiority over other nations. Therefore, the "security dilemma" and the pursuit of absolute hegemony still exist in the era of AI. Without substantial changes to the global security situation, the development, deployment, and application of military AI as a strategic imperative for major military powers are unlikely to be directly constrained by treaty law.

3.2.2. *Scepticism Towards Treaty-Making among Western Countries*

Another reason for the treaty remaining improbable may stem from several Western countries' scepticism toward treaty-making in the technology realm. Such a sceptical stance is primarily held by some of the technologically advanced states led by the US and the UK.⁹⁵ Beyond the realm of military AI, such scepticism was also evident in cyberspace governance.⁹⁶ On the surface, such scepticism stems from the concern that slowly evolving international rules are unable to keep pace with rapid

⁹⁵ The countries mentioned here primarily include Australia, Estonia, Israel, Japan, Poland, the Republic of Korea, the United Kingdom, and the United States.

⁹⁶ For example, in 2011, the US vice president expressed his opposition to the establishment of an international legal instrument and stressed the applicability of existing rules; the same view was also reflected in the U.S. "International Strategy of Cyberspace" of that year. See The White House, "VP's Remarks to London Cyberspace Conference", 1 November 2011, available at: <https://obamawhitehouse.archives.gov/the-press-office/2011/11/01/vps-remarks-london-cyberspace-conference>; The White House, *International Strategy for Cyberspace: Prosperity, Security, and Openness in a Networked World*, Washington, DC, 2011, p. 9.

technological change.⁹⁷ To avoid such consequences, some Western countries argue for the interpretation of existing international law.⁹⁸

However, at a deeper level, such scepticism is rooted in the contestation over international rule-making power among countries. To understand this logic, one can draw a direct parallel to the historical struggles in the negotiation process of international treaties on cybercrime. Historically, major Western powers promoted the application of the Budapest Convention, spearheaded by the Council of Europe,⁹⁹ and took a negative attitude towards the UN Convention Against Cybercrime, which involved deep participation from China, Russia, and the global south. The basic attitude of Western countries toward the Convention against Cybercrime can be reflected in their diplomatic records. For instance, the US deputy ambassador commented that the resolution on the Convention against Cybercrime will “undermine international cooperation to combat cybercrime”.¹⁰⁰ The EU also announced that the Convention against Cybercrime will “lower the standards for protecting human rights and fundamental freedoms, increase the digital divide, and endorse state control of the Internet”.¹⁰¹ On the contrary, despite the global south countries also having internal differences on specific terms of the treaty, these supporters advocate for a more inclusive global process. This divergence reflects that countries have a stronger preference for the framework in which they are deeply involved. The preferences for the existing framework on military AI also reflect Western countries’ desire to shape international rules through existing platforms and interpretive methods, where they have historically held dominant influence.

At the same time, on a practical level, this scepticism toward new treaties enables Western states to steer the evolution of existing international norms in a

⁹⁷ In 2013, the United Kingdom expressed its view that “The United Kingdom does not believe that attempts to conclude comprehensive multilateral treaties, codes of conduct or similar instruments would make a positive contribution to enhanced cybersecurity for the foreseeable future. The complex and comprehensive nature of any binding agreement across the entirety of a cyberspace that is evolving at “net speed” means that it could not be effective or command widespread support without many years, possibly decades, of painstaking work on norms of behaviour and confidence-building measures to build up the necessary understanding and trust among signatories and to ensure that they can be reliably held to account for their adherence to their commitments.” See UN General Assembly, *Developments in the field of information and telecommunications in the context of international security, Report of the Secretary-General*, A/68/156, 16 July 2013, pp. 18-19.

⁹⁸ The UK further noted that “the efforts of the international community should be focused on developing common understandings on international law and norms rather than negotiating binding instruments that would only lead to the partial and premature imposition of an approach to a domain that is currently too immature to support it.” See *Ibid.*

⁹⁹ See Delegation of the European Union to the Council of Europe, “EU Statement in support of the Council of Europe Convention on Cybercrime”, 15 January 2020, available at: https://www.eeas.europa.eu/delegations/council-europe/eu-statement-support-council-europe-convention-cybercrime_en?s=51. (EU Statement)

¹⁰⁰ See Edith M. Lederer, “UN Gives Green Light to Draft Treaty to Combat Cybercrime”, *The Associated Press*, 28 December 2019, available at: <https://apnews.com/general-news-79c7986478e5f455f2b281b5c9ed2d15>.

¹⁰¹ See EU Statement, above note 99.

direction favourable to their interests by championing soft law norms. For instance, in the realm of cyberspace governance, the currently established “norms of responsible state behaviour in Cyberspace” largely reflect the strategic priorities of Western states.¹⁰² In contrast, issues of critical concern to many non-Western nations, such as cyber terrorism and cyber surveillance, are virtually absent from these existing norms.¹⁰³

Therefore, beyond the explicit justifications offered in national position statements, Western countries’ scepticism towards new treaties can be seen as a strategy to leverage their dominance in the international rule-making process. The ultimate goal is to shape the future development of existing international law. Regarding the question of whether a new treaty should be established for military AI, it is crucial to recognize the underlying reality: the scepticism maintained by Western and certain other states is, in essence, a manoeuvre to utilize alternative mechanisms to guide the evolution of international law toward their preferred outcomes.

However, it should also be admitted that, unlike in cyberspace governance, divergences exist in the approach to regulating military AI within Western countries. Despite countries like Germany, France, Canada, and Austria adopting a strategy similar to that of the US and UK in cyberspace governance, they are more active in supporting a legally binding instrument or taking a more prudent view on military AI¹⁰⁴ instead of merely opposing a new treaty. This situation likely arises because cyber operations normally occur below the threshold of use of force, and European countries are hence more likely to value their strategic flexibility. In contrast, military AI is directly relevant to lethal attacks and human lives, triggering the red line of IHL directly. Therefore, for countries like France and Germany, the concern for basic humanitarian principles outweighs the concern for flexible regulation. As these countries stressed in 2022, the international community should take a “two-tier approach” within the UN framework, prohibit autonomous lethal weapons outside human control, while strictly regulating others to ensure meaningful human control.¹⁰⁵ Meanwhile, Germany even directly supported the establishment of a new treaty.¹⁰⁶ Therefore, although the overall prospect of establishing a new treaty remains uncertain, such internal divergences objectively increase the chances of its successful establishment.

¹⁰² Zhixiong Huang [黄志雄], “Norms of Responsible State Behavior in Cyberspace: Origins, Impacts and Responses” [网络空间负责任国家行为规范：源起、影响和应对], *Contemporary Law Review* [当代法学], No. 1, 2019, pp. 67-68.

¹⁰³ Of course, it cannot be denied that the norms of responsible state behavior in cyberspace themselves possess considerable rationality and merit. Moreover, the broad scope and non-binding nature of such norms make non-Western countries like China and Russia willing to participate in them. See *Ibid.*

¹⁰⁴ See State Positions Database, above note 6.

¹⁰⁵ See *Working paper Submitted by Finland, France, Germany, the Netherlands, Norway, Spain and Sweden*, UN Doc. CCW/GGE.1/2022/WP.7, 9 August 2022.

¹⁰⁶ State Positions Database, Germany, above note 6.

3.2.3. *Current Unfavourable International Order and Political Environment*

Third, the prevailing international order and political environment are unfavourable for the establishment of a new treaty. From an international political perspective, the intensification of geopolitical rivalry has substantially shaped an environment that incentivizes the deployment of AI weaponry rather than its restriction. Geopolitically, the international arena is currently defined by active regional “hot wars” involving the use of military AI. In the Russia-Ukraine conflict, both parties have leveraged AI-enabled drones and decision-support systems to gain significant battlefield advantages.¹⁰⁷ Similarly, in the Israel-Palestine conflict, the “Gospel” system has significantly accelerated the generation of targets for the Israeli military.¹⁰⁸ Furthermore, during recent US military operations against Iran, military AI systems such as Project Maven have been extensively deployed to identify targets.¹⁰⁹ The demonstrated operational utility of such applications is likely to cause states facing potential future conflicts to adopt a prudent approach toward any restrictions on military AI.

From a macro perspective, the contemporary international environment also no longer provides a foundation for establishing arms control treaties. The erosion of the rule-based international order by the great powers in the past decade has undermined the foundation of international cooperation. For example, as a dominant global actor, the United States’ recent disregard for international norms has precipitated a degradation of the legal order, ushering the international community into an era akin to the “collapse of rites and music” (*Li Beng Yue Huai*, 礼崩乐坏).¹¹⁰ Since the inauguration of the Trump Administration in 2025, a series of flagrant violations of international law, ranging from the detention of the

¹⁰⁷ Abduljalil Abdurasulov, “The New AI Arms Race Changing the War in Ukraine”, *BBC*, 10 October 2025, available at: <https://www.bbc.com/news/articles/cly7jrez2jno>.

¹⁰⁸ Harry Davies, et al., “‘The Gospel’: how Israel uses AI to select bombing targets in Gaza”, *The Guardian*, 1 December 2023, available at: <https://www.theguardian.com/world/2023/dec/01/the-gospel-how-israel-uses-ai-to-select-bombing-targets>.

¹⁰⁹ Michael Brown, “The First AI War: How the Iran Conflict Is Reshaping Warfare”, *Forbes*, 30 March 2026, available at: <https://www.forbes.com/sites/mikebrown/2026/03/30/the-first-ai-war-how-the-iran-conflict-is-reshaping-warfare/>.

¹¹⁰ Derived from The Analects, the concept of “the collapse of rites and music” (*Li Beng Yue Huai*, 礼崩乐坏) describes a state of social disorder and normative decay. During China’s Zhou dynasty (1046–256 BCE), waning quasi-central authority caused established ritual institutions to lose their binding force. Today, although lacking a central authority comparable to the Zhou court, the international community relies on international law as its normative framework. As the efficacy of international law declines, the contemporary global order increasingly mirrors this historical period of “the collapse of rites and music”.

Venezuelan President¹¹¹ to assertions of sovereignty over Greenland,¹¹² have triggered widespread scepticism regarding the “rules-based international order” and the very efficacy of international law. As Canadian Prime Minister Carney warned at the 2026 World Economic Forum: “the rules-based order is fading... the strong do what they can, and the weak must suffer what they must.”¹¹³ As the international community increasingly shifts from idealism to realism, the efficacy of international law—inherently devoid of coercive enforcement—becomes progressively limited.

It should be admitted that some new multilateral treaties have been negotiated in recent years. The Convention against Cybercrime is a prominent one. Facing widespread opposition, the treaty still passed in 2024 and was opened for signature in 2025. At first glance, such an instance seems to provide the possibility to establish a new treaty in an unfriendly environment. But in essence, the Convention against Cybercrime is a treaty that authorizes states to cooperate against non-state actors, while the potential treaty on military AI might constrain state power. When states’ perception of international relations becomes more realist and cooperation on enforcement might exist in the future, requiring states to unilaterally disarm would be more difficult. Consequently, attempts to employ international law to constrain state power have become increasingly difficult, rendering a new treaty on military AI *de facto* improbable.

3.3. The “Majority Follows Minority” Dynamic in Military AI

Under the cumulative impact of the aforementioned factors, despite the sustained efforts of the international community, the prospect of establishing a treaty or achieving effective regulation through one remains bleak. This reality presents a unique landscape in the regulation of military AI: while the vast majority of states support a binding treaty, the opposition of a small minority of countries renders such a treaty improbable to conclude or incapable of having substantive legal force. In this sense, the regulation of military AI has devolved into a process where “the majority follows the minority”.

Within the CCW framework, the requirement for consensus nominally grants every state the right to have its will respected. However, in substance, it accords a *de facto* veto to major military powers capable of influencing the outcomes of progress. Under alternative mechanisms, even if a treaty secures widespread

¹¹¹ Michael N. Schmitt, et al., “International Law and the US Military and Law Enforcement Operations in Venezuela”, *Just Security*, 5 January 2026, available at: <https://www.justsecurity.org/127981/international-law-venezuela-maduro/>.

¹¹² OHCHR, “Greenland: UN Experts Urge United States to Respect International Law and Right to Self-Determination”, 14 January 2026, available at: <https://www.ohchr.org/en/press-releases/2026/01/greenland-un-experts-urge-united-states-respect-international-law-and-right>.

¹¹³ Mark Carney, “Davos 2026: Special address by Mark Carney, Prime Minister of Canada”, *World Economic Forum*, 20 January 2026, available at: <https://www.weforum.org/stories/2026/01/davos-2026-special-address-by-mark-carney-prime-minister-of-canada/>.

participation, it represents merely a “symbolic victory”. As long as the few states possessing actual military AI capabilities refuse to join, the treaty’s capacity to alter battlefield realities will approach zero. Consequently, the influence of major military powers over the regulation of military AI remains disproportionately immense. In the absence of binding hard law, middle and small states eager to regulate military AI are constrained by these major powers and may be compelled to accept soft-law norms imposed by them. In this regard, the United States-led Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy is a prominent example.¹¹⁴ On the one hand, the US explicitly refuses to establish a legally binding new treaty. On the other hand, it is trying to adopt a political declaration to transfer its national will into normative proposals to increase its rule-making power in the military AI, and call on more than fifty middle and small states to endorse it.¹¹⁵ This allows major powers to seize the agenda-setting power and ensure the field remains under their dominance.

In this context, the “majority follows minority” dynamic underscores that state power remains the determining factor in whether a treaty can be established and effectively implemented. If the broader community of nations seeks to reclaim the initiative in the regulatory game and surmount the obstacles posed by the minority, they must adopt strategic approaches. This will be the focus of the discussion in Part 4.

4. From De Jure Pursuit to De Facto Law: The Treaty Process as an Incubator, Evidence, and Reputational Costs

Despite the improbability of establishing a treaty to effectively regulate military AI, this does not render the pursuit of such an instrument futile. In fact, this article adopts a cautiously optimistic view regarding the normative value of the treaty-making process. In a landscape defined by the “majority follows minority” dynamic, the majority of states can leverage the treaty-making process to reclaim the initiative in the international rule-making process through three methods. First, the treaty-making process itself could serve as an incubator for the formation of CIL. Second, the treaty text could serve as substantive evidence influencing the interpretation of existing international law. Finally, the international community could impose constraints through legal and reputational costs on non-party states. Through these approaches, both treaty participants and non-participants will be effectively regulated within the evolving regulatory framework for military AI.

¹¹⁴ US Department of State, “Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy”, 2023, <https://www.state.gov/bureau-of-arms-control-deterrence-and-stability/political-declaration-on-responsible-military-use-of-artificial-intelligence-and-autonomy>. (Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy)

¹¹⁵ *Ibid.*

4.1. The Treaty as an Incubator: Facilitating the formation of CIL

The treaty-making process may facilitate the formation of CIL. Admittedly, grounded in the fundamental rule that *pacta tertiis nec nocent nec prosunt*, a treaty cannot impose direct binding obligations on non-party states.¹¹⁶ However, this does not imply that such treaties are devoid of influence on non-parties. According to Article 38 of the Vienna Convention on the Law of Treaties (VCLT), treaty provisions reflecting CIL remain binding on all states in that capacity.¹¹⁷ Under specific circumstances, treaty provisions may even transform into general international law.¹¹⁸ Through this mechanism, the core norms under the military AI treaty, like meaningful human control, have the potential to generate or crystallize into CIL and exert universal binding force upon non-party states.

The formation of CIL depends on two fundamental elements: state practice and *opinio juris*. In terms of state practice, as the ICJ elaborated in the *Asylum Case*, constant and uniform usage is necessary for the formation of CIL.¹¹⁹ Meanwhile, *opinio juris* requires that such practice was an act in a way that they believed themselves legally compelled to do so.¹²⁰ These requirements dictate that the emergence of CIL in the field of military AI must go beyond the mere political behaviour, and must demonstrate widespread, representative, and consistent practice.¹²¹ In this regard, the treaty-making process and the resulting text could promote the development of CIL in two ways.

First, the treaty-making process would facilitate the states to articulate their legal positions on military AI regulation, thus generating relevant state practice with *opinio juris*. To date, the international consensus in this field is still limited to political

¹¹⁶ Vienna Convention on the Law of Treaties, 1155 UNTS 331, 23 May 1969 (entered into force 27 January 1980), Art. 34. (Vienna Convention on the Law of Treaties)

¹¹⁷ Vienna Convention on the Law of Treaties, above note 116, Art. 38.

¹¹⁸ As the ICJ elaborated in the *North Sea Continental Shelf Case*: “A norm-creating provision which has constituted the foundation of, or has generated a rule which, while only conventional or contractual in its origin, has since passed into the general corpus of international law, and is now accepted as such by the *opinio juris*, so as to have become binding even for countries which have never, and do not, become parties to the Convention. There is no doubt that this process is a perfectly possible one and does from time to time occur: it constitutes indeed one of the recognized methods by which new rules of customary international law may be formed.” See ICJ, *North Sea Continental Shelf Cases (Federal Republic of Germany v. Denmark; see Federal Republic of Germany v. Netherlands)*, Judgment, ICJ Report 1969, para. 71 (ICJ, *North Sea Continental Shelf Cases*); see Giorgio Gaja, “1969 Vienna Convention Article 38 Rules in a treaty becoming binding on third States through international custom”, in Olivier Corten *et al.* (eds), *The Vienna Conventions on the Law of Treaties: A Commentary*, 2nd ed., Oxford University Press, Oxford, 2011, p. 954-956.

¹¹⁹ International Court of Justice, *Colombian-Peruvian Asylum Case*, Judgment, ICJ Report 1950, pp. 276-277.

¹²⁰ Malcolm N. Shaw, *International Law*, 9th ed., Cambridge University Press, Cambridge, 2021, p. 74.

¹²¹ *Report of the International Law Commission of its Seventieth Session*, U.N. Doc. A/73/10, p. 120, Conclusion 8 (2018). (ILC 2018 Report)

declarations¹²² and lacks the “accepted as law” characteristic, thereby obstructing the formation of CIL. Although CIL may gradually develop in the absence of treaties, this process will undoubtedly be slow, given the prevailing caution among states.¹²³ In this context, the treaty-making process would accelerate CIL’s generation and crystallization. In fact, the treaty-making process or draft treaty provisions may themselves give rise to State practice and evidence of acceptance as law.¹²⁴ If the international community can effectively leverage this mechanism and initiate the treaty negotiation process, the textual formulations negotiated by states during treaty negotiation can become state practice reflecting their *opinio juris*, thereby transcending mere political commitments. By providing multilateral forums for states to articulate their practices and *opinio juris*, the treaty-making process offers a practical pathway for the eventual formation of CIL.

Secondly, the treaty-making process could pressure minority opposing states to publicly articulate their legal positions, reducing the strategic ambiguity that they intentionally maintain and facilitating the generation and crystallization¹²⁵ of CIL. Under the current framework, AI is regulated solely through the interpretation of existing laws. If a nation wishes to avoid international oversight, it need only remain silent at no cost or employ strategic ambiguity regarding specific legal standards to achieve its objectives. For instance, for rules such as “meaningful human control”, a state may repeatedly affirm its applicability while avoiding clarification of the specific content. However, the formal treaty-making process would fundamentally change this phenomenon. During the negotiation of formal treaty provisions, each article drafted may be relevant to future state obligations and state responsibility. If states completely refuse to participate in the process of expressing legal positions both within and outside the treaty process, once specific treaty provisions crystallize into CIL, their silence may be regarded as implied consent, thereby compelling them to accept the rules they opposed. Therefore, if a draft treaty explicitly requires AWS to be subject to human control before launching each specific attack, opposing states would have to express their legal objections to protect their military advantages and clearly state the legal standards they support. Thereby helping to establish a “baseline consensus” on military AI within the international community and serving as the starting point for the generation of CIL.

¹²² Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy, above note 114.

¹²³ In fact, the formation of CIL is inherently time-consuming. Under the best of circumstances, CIL may be created in as little as 10 to 15 years. See Matthijs Maas, “International Law Does Not Compute: Artificial Intelligence and the Development, Displacement or Destruction of the Global Legal Order”, *Melbourne Journal of International Law*, Vol. 20, No. 1, 2019, p. 50.

¹²⁴ ILC 2018 Report, above note 121, p. 143, Commentary to Conclusion 11, para. 1.

¹²⁵ Crystallization means that the custom was *in statu nascendi* when the treaty was drafted, but became a customary rule subsequently. See Marina Fortuna, “Interpretation of Customary Rules by Reference to Treaties and General Principles of Law”, in Marina Fortuna *et al.* (eds), *Customary International Law and Its Interpretation by International Courts: Theories, Methods and Interactions*, Cambridge University Press, Cambridge, 2024, pp. 269-270.

Nevertheless, the transformation of treaty provisions into CIL is subject to several important limitations. During this process, while the treaty accepted by a sufficient number of states is required,¹²⁶ it should also include all states that have the opportunity and possibility to apply the rule in question.¹²⁷ In particular, the participation of “specially affected states” remains crucial in this process.¹²⁸ If this holds true, the subsequent practice of non-party states could substantially affect the formation of CIL. If the specially affected opposing states, such as the US, Russia, and the UK, continue to oppose the provisions of a new treaty, the formation of CIL may face significant obstacles. However, building on the treaty’s broad international support, and constrained by considerations of international reputation and concerns over the treaty’s stigmatizing effects, they may opt to align their practices with the treaty’s provisions in practice.¹²⁹ In such circumstances, the treaty’s provisions themselves will be widely emulated by states, thereby facilitating the formation of CIL.¹³⁰ Even if these specially affected states persistently object to the formation of the CIL, the accumulated state practices and *opinio juris* under the treaty would still lay a robust foundation for future normative development. More importantly, even if these countries cross the red line drawn by the treaty in practice, the treaty ensures they can no longer do so under the guise of legal ambiguity, thereby forcing them to bear the full political responsibility of their actions. Therefore, rather than directly generating CIL, the treaty-making process may function as an incubator that provides a suitable environment and accelerates the formation of state practice and *opinio juris*.

Ultimately, while acknowledging the complexity of the issue, a forward-looking perspective must be maintained in regulating military AI. Historically, the establishment of international legal regimes governing domains such as warfare, navigation, trade, and human rights has spanned decades or even centuries, evolving into the frameworks we have today through long-term development. In contrast, the regulation of military AI, since the CCW mechanism’s intervention in 2013, has a history of only a bit more than a decade, placing it in the nascent stage of technical regulation. Compared to the historical experience of international legal regulation in other areas, the path to regulating military AI has just begun. Therefore, the process of establishing customary law through the formulation of new treaties remains a crucial long-term normative pathway that requires persistent international efforts.

¹²⁶ Gary L. Scott and Craig L. Carr, “Multilateral Treaties and the Formation of Customary International Law”, *Denver Journal of International Law & Policy*, Vol. 25, No. 1, 1996, p. 72.

¹²⁷ ILC 2018 Report, above note 121, p.136, Commentary to Conclusion 8, para. 3.

¹²⁸ ICJ, *North Sea Continental Shelf Cases*, above note 118, para. 73.

¹²⁹ The reputational cost for the states will be analysed in part 4.3.

¹³⁰ See Michael P. Scharf, “Accelerated Formation of Customary International Law”, *ILSA Journal of International & Comparative Law*, Vol. 20, No. 2, 2014, p. 320.

4.2. The Treaty as Evidence: Codification as an Effective Guide for Interpretation

Beyond its direct binding force on party states, the new treaty will serve as a reference for interpreting existing international law, providing an effective guide to its application, and thereby imposing influence on non-party states. To be specific, currently, numerous rules in the field of IHL are codified in treaty texts within the Geneva Conventions and their Additional Protocols. As new weapons emerge, treaty texts require necessary “evolutionary interpretation” based on subsequent agreements and practices to maintain their relevance in new scenarios.¹³¹ In this context, the new convention on military AI will provide a reference for the evolutionary interpretation of existing IHL.

During this process, it should be acknowledged that establishing “subsequent agreement” and “subsequent practice” under Vienna Convention on the Law of Treaties (VCLT) Article 31 imposes a high threshold, requiring that such practice reflect the common understanding of the parties as a whole regarding the interpretation of the treaty.¹³² Thus, because the new treaty on military AI may not reflect the agreement of all State Parties to the Geneva Conventions, it cannot constitute a subsequent agreement or practice under VCLT Article 31. However, the practice regarding the new treaty could, in its legal sense, constitute other subsequent practice under VCLT 32, serving as a supplementary means of interpretation for the Geneva Conventions.¹³³ In its nature, the new treaty could be seen as evidence of states’ widespread understanding of their obligations regarding military AI under IHL. In practice, to clarify ambiguous or obscure provisions in the Geneva Conventions, the interpreter could refer to the treaty on military AI. Without distorting the substance of the rules, such supplementary materials may be invoked to clarify the specific implications of existing IHL when applied to the context of military AI. Through this approach, even though some states are not parties to the new treaty, most are signatories to the Geneva Conventions and will thus be influenced by this practice. Should they significantly deviate from the new treaty’s provisions in their use of military AI, they may face a greater burden of proof to justify the rationality and legality of such actions.

This interpretative function has particular significance for some broad international rules, such as the Martens Clause. According to the International Law Commission’s opinion, the Martens Clause provides that even in cases not covered

¹³¹ Such an interpretation practice was early reflected in the WTO U.S. - Shrimp cases. See WTO Appellate Body, *United States - Import Prohibition of Certain Shrimp and Shrimp Products*, WT/DS58/AB/R, 12 October 1998, paras 129-130.

¹³² See ILC 2018 Report, above note 121, pp. 20, 28, Commentary to Conclusion 2 para. 9, Commentary of Conclusion 4 para 4.

¹³³ See *Ibid.*, pp. 20-21, 33-34, Commentary to Conclusion 2 para. 9, Commentary to Conclusion 4 paras. 23-25; *Report of the International Law Commission on the Work of its Sixteenth Session*, UN Doc. A/5809, 11 July 1964, p. 204, para. 13.

by specific international agreements, civilians and combatants remain under the protection and authority of the principles of international law derived from established custom, from the principles of humanity, and from the dictates of public conscience.¹³⁴ However, as mentioned in Part 2, there is no universally accepted interpretation of the Martens Clause.¹³⁵ Therefore, when international courts are trying to apply such standards to military AI and adjudicating whether a state's application of AWS is consistent with the public conscience, they need to seek references to identify the specific interpretation of that clause. If many states have declared in the new treaty that meaningful human control would be the critical element for complying with the principle of humanity, then such normative expression would affect the courts and state understanding of the Martens Clause, and the new treaty could effectively reduce interpretive discretion.

4.3. Beyond the Law: Imposing Reputational Costs on Non-Party States

A potential treaty on military AI would not only generate legal effects but also impose inescapable reputational costs, thereby promoting treaty compliance even among non-party states. Although the concept of reputational costs has historically remained somewhat ambiguous, international law scholars have consistently identified it as a pivotal mechanism for promoting compliance with international law.¹³⁶ As research suggests, in the realm of national security, reputational cost plays a significant role, as in other fields.¹³⁷ States that did not shape the current framework of international law face high strategic costs in opting to violate international security law norms perceived as not in their interests.¹³⁸ In the field of military AI, based on the simplified cost-benefit theory analysis, if the state pays more reputational cost when it is against the treaty than it gains, then the state is more likely to act consistently with the treaty.¹³⁹ So far, the absence of a treaty governing military AI means we may lack empirical means to analyse the reputational costs it generates. Nevertheless, past international law practices offer valuable insights for predicting the reputational costs that new treaties may incur.

Historically, reputational cost was frequently considered when establishing an arms control treaty and has yielded effective results. Not only in the Convention on Cluster Munitions are reputational costs designed as a part of the stigmatizing

¹³⁴ *Report of the International Law Commission on the Work of its Forty-Sixth Session*, UN Doc. A/49/10, 22 July 1994, p. 317, Commentary to Article 29, para. 3.

¹³⁵ Rupert Ticehurst, "The Martens Clause and the Laws of Armed Conflict", *International Review of the Red Cross*, Vol. 37, No. 317, 1997, p. 126.

¹³⁶ See Rachel Brewster, "Unpacking the State's Reputation", *Harvard Journal of International Law*, Vol. 50, No. 2, 2009, p. 269.

¹³⁷ See Andrew C. Blandford, "Reputational Costs Beyond Treaty Exclusion: International Law Violations as Security Threat Focal Points", *Washington University Global Studies Law Review*, Vol. 10, No. 4, 2011.

¹³⁸ *Ibid.*, p. 726.

¹³⁹ See Andrew T. Guzman, *How International Law Works: A Rational Choice Theory*, Oxford University Press, Oxford, 2008, pp. 74-75.

mechanism,¹⁴⁰ but also in the Mine Ban Treaty, reputational costs substantially facilitated the ban on mines.¹⁴¹ Specifically, the Mine Ban Treaty gives the signatory states a solid moral and political basis for criticizing other states' ownership or use of these weapons, increasing the cost for opposing states of using them.¹⁴² Hence, when the Mine Ban Treaty was adopted, it resulted in an abrupt change of the UK's position on mines. The humanitarian concerns even affected its domestic politics, which pushed the UK to join the treaty, and resulted in the United States being isolated from its NATO allies.¹⁴³ Eventually, the Mine Ban Treaty successfully forced countries like China, India, and the US to make a substantial change in their policy.¹⁴⁴

Regarding the potential treaty on military AI, we should admit that military AI could provide much more military utility than cluster munitions and mines, which could affect the efficacy of reputational costs. However, this article advocates for the regulation of military AI, rather than a complete ban akin to those on cluster munitions and landmines. In this sense, if a state can regulate certain aspects of military AI without compromising its core military interests while simultaneously gaining additional reputational benefits, it can take such action based on its rational choice. Therefore, the mechanism by which reputational costs influence the regulation of military AI does not directly ensure that states fully control such technologies. Instead, it indirectly prompts states to make choices aligned with international expectations by increasing or decreasing these costs. A potential treaty on military AI serves as an amplifier in this process.

5. Conclusion

The military application of artificial intelligence challenges the international legal system, splits states, as well as scholars, between advocating for a new treaty and relying on existing international law. The opponents of establishing a new treaty argue that existing international law is sufficient, while the proponents contend that the new problems posed by military AI necessitate a new treaty. This article examines the potential problems posed by military AI and argues that the moral dilemmas, accountability gaps, and compliance deficits would necessitate the establishment of a legally binding instrument on a *de jure* basis to diminish legal uncertainty.

However, from a *de facto* perspective, establishing a new treaty in the near future faces too many obstacles to be feasible, making the establishment of a treaty

¹⁴⁰ ICRC, Convention on Cluster Munitions, Geneva, 2008, Foreword, p. 6.

¹⁴¹ Arms Control Association, "The Ottawa Convention at a Glance", 2024, available at: <https://www.armscontrol.org/factsheets/ottawa-convention-glance>.

¹⁴² Kenneth R. Rutherford, "The Evolving Arms Control Agenda: Implications of the Role of NGOs in Banning Antipersonnel Landmines", *World Politics*, Vol. 53, No. 1, 2000, pp. 109-110.

¹⁴³ *Ibid.*, p. 100-101.

¹⁴⁴ *Ibid.*, n. 131.

that can directly regulate major powers almost improbable. Under the CCW process, there has been no mandate to negotiate a treaty, and the consensus requirements for adopting one give the major military powers *de facto* veto power, making it almost impossible to negotiate a new treaty. Under the alternative process, even though a treaty can be established like the Convention on Cluster Munitions, it is doubtful to what extent such a treaty *per se* would effectively regulate its opposing states. From a realist perspective, since military AI is a strategic imperative for major military powers, expecting these countries to restrict technologies that would significantly increase their battlefield advantages is irrational. Under the landscape of contestation over international rule-making power, the universal scepticism among Western states makes the establishment of such a treaty even harder. At the macro level, the intensification of geopolitical rivalry and an unfavourable political environment also serve as negative factors affecting treaty-making. All those elements contribute to the international regulation of military AI devolving into a “majority follows minority” process.

Despite obstacles mentioned, the article still holds a cautiously optimistic note regarding the establishment of a new treaty on military AI. It argues that the *de facto* improbability does not nullify the *de jure* necessity. Instead, the very pursuit of a new treaty initiates a dynamic legal process. In the long term, such a legal process would impose legal and non-legal force on the party and non-party states through three approaches, all of which would facilitate the international regulation of military AI. Firstly, the new treaty could serve as an incubator for CIL. The treaty-making process will promote the generation and crystallization of CIL through enhancing state practice and *opinio juris*. Secondly, the treaty text could serve as evidence of existing international law during legal interpretation, thereby diminishing the possibility of arbitrary interpretation of the existing law by treaty opponents. Finally, beyond the legal effects that it could have on non-party states, it could also create reputational costs that make states think twice before openly acting inconsistently with it.