INTRODUCTION: BACKGROUND AND OVERARCHING THEMES

Samuel White*
Jonathan Kwik**
Danielle Yeow***
Jon Truby****

Background to the Special Issue

For much of modern military history, advances in technology have altered how wars are fought without fundamentally displacing the human decision-maker from the centre of military judgment. The machine gun increased lethality, the aircraft expanded reach, whilst the computer accelerated the processing of information; yet the assessment of military necessity, the weighing of civilian harm and the exercise of legal and moral responsibility remained recognisably human functions. Artificial intelligence (AI) presents a different challenge. Rather than merely extending human capabilities, AI systems increasingly participate in the production of military knowledge itself. They identify patterns, generate recommendations, prioritise targets, allocate resources, and shape the information environment through processes that are often opaque, adaptive, and difficult even for their designers to fully explain. As these systems become integrated into military planning and operations, longstanding assumptions about control, accountability, and legal responsibility are placed under increasing pressure.

This transformation is occurring at a moment when the character of conflict is itself changing (although historians would say, correctly, that the character of conflict is always changing). Contemporary warfare is increasingly data-driven, digitally enabled and fought across interconnected physical, cyber, cognitive, and informational domains. The strategic value of data, algorithms, cloud infrastructure, and computational power now rivals that of many traditional military assets. Military advantage is no longer derived solely from the destruction of an adversary's forces, but also from the ability to collect, process, manipulate, and act upon

* Scientia Senior Lecturer in Military Law and War Studies, UNSW Canberra; Global Fellow at the Centre for International Law, National University of Singapore.
** Researcher in International Law, TMC Asserinstituut (The Hague); Drafting Team, AI Manual Project; Board of Experts member, Asia-Pacific Journal of International Humanitarian Law.
*** Adjunct Senior Research Fellow and Lead, Cyber Law and Governance at the Centre for International Law, National University of Singapore.
**** UNESCO Chair on Artificial Intelligence Law & Sustainability at the Centre for International Law, National University of Singapore.

information at speed and scale. Artificial intelligence sits at the centre of this shift. It promises unprecedented operational efficiencies and decision superiority, while simultaneously generating novel risks that challenge existing legal frameworks and ethical understandings of warfare. Questions of accountability, reliability, transparency, autonomy, and human control are no longer speculative concerns. They have become practical issues confronting military planners, policymakers, and legal advisers in real-world operations.

Recent conflicts have brought these questions into sharper focus, in particular the recent conflict between the United States of America (and its allies) against Iran. The integration of AI-enabled decision-support systems into targeting processes, intelligence analysis, and military planning has demonstrated both the operational potential of these technologies and the difficulties they pose for existing legal regimes. At the same time, the growing strategic importance of data infrastructure, cloud computing services, and commercial technology providers has further blurred traditional distinctions between civilian and military spheres. What was once viewed as a future challenge is now an immediate reality. The law is increasingly required to address not merely autonomous weapons, but an entire ecosystem of AI-enabled military capabilities whose effects extend across the full spectrum of conflict.

AI systems aggregated and processed drone feeds, satellite imagery, signals intelligence, and intercepted communications at speeds beyond human capacity, producing strike recommendations that effectively collapsed what had once been layered, time-consuming approval chains into near-real-time outputs.[1]

Iran's response has, in turn, focused on disrupting the data architecture that enables this form of warfare. In March and April 2026, Iranian forces targeted Amazon Web Services (AWS) data centres in the UAE and Bahrain using drones and missiles, marking the first known instance of a state deliberately striking commercial data infrastructure during an armed conflict.[2] The resulting outages,

---

[1] Hope Hodge Seck, "AI Tool Has 'Saved a Lot of Aircraft' in Epic Fury, AFSOC Chief Says", *Military Times*, 13 May 2026, available at: https://www.militarytimes.com/news/ your-military/2026/05/13/ai-tool-has-saved-a-lot-of-aircraft-in-epic-fury-afsoc-chief-says/ accessed 14 July 2026

[2] Shubham Kaila, Aditya Soni, and Mrinmay Dey, "Iranian Strikes Target AWS Facilities in Bahrain and UAE", *Reuters*, 3 March 2026, available at: https://www.reuters.com/ world/middle-east/amazon-cloud-unit-flags-issues-bahrain-uae-data-centers-amid-iran-strikes-2026-03-02/. accessed 14 July 2026

connectivity failures, and fire damage illustrate how civilian digital systems have become integral to military operations.

It is against this backdrop that this Special Issue is situated. This conflict did not emerge as an isolated technological departure. Rather, it builds on a trajectory visible in earlier uses of AI in the Gaza Strip, where Israeli systems such as "Lavender" and "Gospel" reportedly generated tens of thousands of targeting recommendations from surveillance data which has generated in its wake, concerns about the risks posed by the deployment of such technology[3] and the erosion of civilian protection.[4] It also extends developments associated with Project Maven, the United States military's primary machine learning initiative, which has supported targeting operations across Iraq, Syria, Yemen, and the Red Sea since 2024. At the same time, these developments have sharpened tensions between the ethical positions of commercial AI developers and the operational requirements of military institutions. On the day Operation Epic Fury commenced, Anthropic was reportedly placed on a Pentagon security blacklist after declining to provide unrestricted access to its Claude system, despite its prior integration into strike planning through Project Maven and classified AWS environments.[5]

Taken together, these developments mark a decisive moment. Debates surrounding the ethics, legality, and governance of military AI have moved beyond abstraction and into direct operational practice. What had previously been framed as a prospective challenge is now an immediate and tangible reality, unfolding in real time across interconnected domains of warfare, infrastructure, and commerce.

This Special Issue proceeds from that premise. It brings together a set of contributions to a regional roundtable Workshop in Singapore on 16-17 April 2026, hosted by the National University of Singapore's Centre for International Law and in support of the wider initiative of the *West Point Manual on International Law*

---

[3] Human Rights Watch, Questions and Answers: Israeli Military's Use of Digital Tools in Gaza, 10 September 2024, available at https://www.hrw.org/news/2024/09/10/questions-and-answers-israeli-militarys-use-of-digital-tools-in-gaza#_On_which_grounds. accessed 14 July 2026

[4] Emelie Andersin, "Military Use of Biometrics series – Isreal's Use of AI-DSS and Facial Recognition Technology: The Erosion of Civil Protection in Gaza", *Articles of War – Blog of the Lieber Institute* , 24 October 2025, available at: https://apjihl.org/wp-content/uploads/2026/01/APJIHL-Guidelines-for-Authors.pdf. accessed 14 July 2026

[5] Hayden Field, "Anthropic and U.S. Government to Face Off in D.C. Court Over Blacklisting of AI Company", *CNBC*, 19 May 2026, available at: https://www.cnbc.com/2026/05/19/anthropic-dod-blacklist-court-opening-arguments.htm. accessed 14 July 2026

Applicable to Artificial Intelligence in Warfare ('AI Manual').[6] The papers herein were selected for their novelty and deep understanding of how to engage with the *sui generis* tension that AI brings. To that end, we seek to address issues that find no precedent, such as data poisoning, AI self-discovering perfidious tactics, agency spillovers, and regional data inadequacies. The articles in this issue examine, from a range of disciplinary perspectives, how established frameworks of international humanitarian law (IHL), state responsibility, and military practice are being strained, adapted, or reconstituted in response to these developments. In doing so, the issue seeks not only to diagnose the present moment, but to clarify the contours of the challenges that lie ahead.

Prominent Themes in the Contributions

In discussing the application of the law to any novel technology in warfare, a crucial question invariably arises: What is truly new, and what is merely an amplification or evolution of an existing problem? It is a core question regularly posed by members of the AI Manual's drafting team, and often not one easily answered. Autonomy, for example, is frequently associated with being a unique characteristic of AI systems, but has been a feature of cyber-capabilities for a long time.[7] In contrast, other properties of such systems—such as their stochasticity,[8] opacity,[9] unintuitive failure modes[10] or the ability to creatively find solutions[11]—are unique to AI, and accordingly give rise to novel legal problems and potential legal gaps. Identifying these risks in the first place may be as difficult as mitigating them, often requiring in-depth knowledge of the technology and the setting in which they are to be deployed (military operations).

In this Special Issue, several contributors have identified, analysed and attempted to address *AI-specific challenges and how they impact the application of international law*. Dora Velenczei theorises how goal-based optimisers in AI-enabled

[6] Centre for International Law, National University of Singapore, *West Point AI Manual Workshop Event Report*, Singapore, 16–17 April 2026, available at: https://cil.nus.edu.sg/wp-content/uploads/2026/05/AI-Manual-Workshop-Event-Report.pdf.
[7] Laurent Gisel and Lukasz Olejnik, *The Potential Human Cost of Cyber Operations*, ICRC Expert Meeting 14–16 November 2018, ICRC, 2018, p. 38.
[8] Nathan G. Wood *et al.*, "Stop Saying 'AI'", arXiv, 25 February 2026, p. 23.
[9] Arthur Holland Michel, *The Black Box, Unlocked: Predictability and Understandability in Military AI*, UNIDIR, Geneva, Switzerland, 22 September 2020.
[10] See Jonathan Kwik, *Lawfully Using Autonomous Weapon Technologies*, TMC Asser Press, The Hague, 2024, pp. 114–120.
[11] Stuart J. Russell and Peter Norvig, *Artificial Intelligence: A Modern Approach*, Pearson, London, 2021, pp. 53–54.

weapon systems may discover methods to fool enemies by (unlawfully) misrepresenting themselves as protected entities—potentially in violation of the prohibition on perfidy.[12] Samuel White and Daniel Harker further raise the dangers of data poisoning (a tactic where adversaries exploit a vulnerability unique to AI, by compromising the AI system as it is being trained)[13] and how these may cause erratic or dangerous behaviours in the affected AI systems, possibly in violation of international law. Anviksha Pachori raises in her contribution the problem of AI's continued evolution through updates, modifications and patches,[14] which generally conflicts with obligations under IHL that often assume static, relatively unchanging capabilities, and suggests that a legal principle of 'IHL-by-Design' can potentially address those tensions.

Fears of AI systems overstepping their boundaries, amongst others due to the issues described above, have led to several propositions to *'code in' legal or ethical limitations*.[15] Sze-Hong Lam, however, strongly questions in his contribution whether such proposals are feasible. Through an in-depth analysis of various sources of indeterminacies that lie at the very core of IHL, Lam argues that integrating the laws of war into AI only succeeds in a somewhat utopian vision of IHL that does not comport to realities. Indeed, if coding in legal limitations directly is an unreliable way to ensure legal compliance, it falls to the users and developers of the AI systems to ensure that this is the case. Through his *dignitary design* concept, Sergey Sayapin advances a comprehensive, lifecycle-spanning framework for embedding respect for human dignity throughout the design, governance, policy, and deployment practices of military AI systems.

Multiple contributions demonstrate that obligations that are the most evident to IHL-lawyers as ways to constrain conduct during warfare—such as the rules related to proportionality and precautions in Art. 57(2) of Additional Protocol I—may not be available to address AI-specific problems. Many use-cases of AI such as

---

[12] For a similar argument on how this may be possible with AI-based cyber defence agents, see also Jonathan Kwik and Adriaan Wiese, "Can AI Teach Itself to Abuse IHL-Protected Indicators?", *Lieber Institute West Point*, 5 December 2025.

[13] See Micah Goldblum *et al.*, "Dataset Security for Machine Learning: Data Poisoning, Backdoor Attacks, and Defenses", 18 December 2020.

[14] Klaudia Klonowska, "Techno-Legal Tinkering in War: AI Decision-Support Systems and International Humanitarian Law", Doctoral thesis, University of Amsterdam, 18 March 2026, pp. 26–27.

[15] E.g. Ronald C. Arkin, Patrick Ulam, and Brittany Duncan, *An Ethical Governor for Constraining Lethal Action in an Autonomous System*, Georgia Institute of Technology Atlanta Mobile Robot Lab, 2019.

decision-support systems (AI-DSS) fall below the threshold of an attack, and poisoning attacks such as those discussed by White and Harker are likely to be used as a grey zone tactic, and executed before an armed conflict even commences.[16] As a result, many authors also take to analysing *complementary legal rules and instruments* as a means to regulate military AI systems. Pachori's contribution which focuses on Additional Protocol I Article 36 reviews and the Arms Trade Treaty, and the Martens Clause has received particular attention in this Issue, with White and Harker, Pachori, and Sayapin all dedicating time to analyse whether and how the "principles of humanity and the dictates of public conscience" can be operationalised to drive the protection of civilians when AI systems are involved.

In addition to contributions exploring how existing law can be applied to military AI systems, the possibility and feasibility of *new governance and norm or treaty-building* has also been explored extensively in this Issue. Zhixiong Huang and Haokun Yu in this regard advance a dual argument: that existing law is insufficient to properly regulate military AI systems ('de jure necessity'), but that several prevailing factors render the possibility of a comprehensive treaty that also binds the Great Powers very unlikely ('de facto improbability'). They nevertheless end with a constructive note, noting the normative benefit of nevertheless pursuing the treaty-making process in the longer term. Bagus Jatmiko echoes this view, also encouraging international engagement and discourse on military AI as an important initiative that States, and in particular middle-power States such as those in Southeast Asia, should pursue.

Mei-Ching Liu, however, issues a word of warning on how challenging such initiatives may prove in practice. In her contribution, she extensively documents and analyses the discourse at the Group of Governmental Experts (GGE) in Geneva surrounding the singular word 'lethal' in LAWS (Lethal Autonomous Weapon Systems) between 2013 and 2026. From her exploration on the way States have used, developed and occasionally instrumentalised the notion of lethality, she highlights how even small semantic or conceptual disagreements in the diplomatic process can entail long-term legal and practical consequences.

[16] See Michael N. Schmitt, "Big Data: International Law Issues Below the Armed Conflict Threshold", in L.A. Dickinson and E.W. Berg (eds.), *Big Data and Armed Conflict: Legal Issues Above and Below the Armed Conflict Threshold*, Oxford University Press, 14 March 2024.

Challenges in *ensuring legal accountability* for the use of military AI systems has been a dominant concern for more than a decade of international discourse.[17] Fan Yang theorises that the phenomenon of 'agency spillover', where machine agency deviates from human presets, lies at the root of difficulties to assign accountability under multiple legal regimes, including IHL, State responsibility and individual criminal liability. Yang nevertheless adds a constructive leg to his critique by recommending that the law should instead consider human-machine hybrid agency as the paradigm when assessing legal accountability for AI systems, which overcome many obstacles he identified. Drexx Laggui, subsequently, adds a practical perspective by considering how States and victims may practically collect and audit proof of misconduct by military AI. He argues that AI systems are prone to leaving digital evidence trails and records that, given the proper forensic procedures and mechanisms, can be obtained and safeguarded for use in legal proceedings.

Finally, the NUS Workshop was also conceived as an opportunity to discuss problems and concerns arising from the development and use of military AI *from an Asia-Pacific and Southeast Asian lens*—a region that, in a field of discourse mostly examined through Western and Chinese perspectives, has received relatively little scholarly attention.[18] Erick Nielson Javier in his contribution conducted a thorough analysis of the state of military AI in the Southeast Asian region, which features unique strategic and security challenges. Looking at both the national and regional interests of Southeast Asian States in military AI, but also shortcomings in their national strategies or acquisition processes, he argues that these States need to seriously consider and consolidate what ways forward they wish to pursue. Finally, in advancing a similar argument, Bagus Jatmiko juxtaposed several problems unique to AI to the Southeast Asian region. He argues that a lack of regional-representative datasets and other factors (such as opacity and human-machine dynamics like sycophancy),[19] which may exacerbate the lower technical expertise of prospective users, could raise novel risks if AI are integrated carelessly into the region, and recommends active engagement with regional and international processes to mitigate these risks.

---

[17] Netta Goussac and Magdalena Pacholska, *The Interpretation and Application of International Humanitarian Law in Relation to Lethal Autonomous Weapon Systems*, UNIDIR, Geneva, 2025, pp. 30–33.

[18] See also Ishmael Bhila, "Strained missions : the diplomatic dilemmas of small states from the Global South in *the area of autonomous weapons systems", University of Malta. Islands and Small States Institute, 2024.*

[19] See Jonathan Kwik, "Digital Yes-Men: How to Deal With Sycophantic Military AI?", *Global Policy*, Vol. 16, No. 3, 2025.

Concluding Thoughts

The contributions gathered in this special issue extend beyond their individual insights. While by no means exhaustive, taken together they illuminate both the promise and complexity of AI-enabled military systems, as well as the legal and ethical challenges posed by accountability frameworks that were not designed with such systems in mind. The papers have sought to address the central question of how international legal disciplines can remain relevant or evolve in order to remain effective and operational under conditions of technical opacity and accelerated decision-making. The lessons they offer are both cautionary and constructive. They counsel against the temptation of assuming that legal constraints can simply be incorporated into systems whose modes of learning and adaptation resist such determinism, while equally counselling against legal paralysis in the face of genuine complexity. Meaningful progress will require sustained, cross-disciplinary engagement among lawyers, military commanders, policy decision-makers, diplomats, corporate actors, academia and communities most directly exposed to risks.

The Asia-Pacific and Southeast Asian perspectives advanced in several papers also offer timely and substantive contributions to a discourse that has too often reflected the strategic interests and technical assumptions of a narrow set of actors. The potential for cross-regional fertilisation of ideas is considerable, and characteristics particular to the Southeast Asian context illuminate challenges that have wider global resonance. We hope this collection will spur other similar intra-regional discourse.

In closing, recent developments—among them the emergence of frontier AI systems of sufficient capability to prompt significant international concern such as Anthropic's Mythos model[20]—serve as a measured reminder that the pace of technological development continues to strain the deliberative capacity of existing international law. The concerns are especially urgent as the strategic effects of military AI will not be distributed evenly. Data asymmetries, procurement constraints, and uneven technical capacity may determine how risks materialise in practice. Looking ahead, an effective and durable governance and accountability framework will require universal principles accompanied by granular regional

[20] Paul Mozur and Adam Satariano, "Anthropic's New A.I. Model Sets Off Global Alarms", *New York Times*, 22 April 2026, available at: https://www.nytimes.com/2026/04/22/technology/anthropics-mythos-ai.html.

inquiry. The latter can identify prioritisation of issues to be addressed, where common norms require context specific implementation, and where regional perspectives and experiences can in turn shape global norm-making.