STM JOURNALS

International Journal on
Drones

Volume 2, Issue 2, 2026
July—December
DOI (Journal): 10.37591/IJD

https://journals.stmjournals.com/ijd

Review IJD

# Artificial Intelligence in Autonomous Military Weaponry Systems

Enosh Masih[1,*], Binit Behl[2], Manohar Pandey[3], Anupam Kumar Singh[4], Krishna Yadav[5]

***Abstract***

*The rapid rise of artificial intelligence (AI) has transformed contemporary defence organizations, reshaping military structures, operational strategies, and the design and deployment of modern weapons. AI has significantly influenced the architecture, decision-making processes, and functionality of military systems, particularly in complex and rapidly evolving combat scenarios. This paper provides a comprehensive analysis of the growing use of autonomous military weapon systems, focusing on technological developments, regulatory challenges, governance gaps, and their broader global implications. Through a review of existing academic literature, it examines the integration of AI across diverse military domains, including autonomous aerial systems, robotic soldiers, unmanned naval platforms, and advanced AI-enabled defence technologies. Employing a mixed-method research approach, the study combines regional case studies with comparative national analysis to evaluate AI defence advancements in four major military powers: the United States, China, Russia, and Israel. The findings suggest that while AI offers substantial operational advantages, such as faster response times, enhanced force protection, improved surveillance, and greater targeting accuracy, its deployment also raises significant ethical, legal, and strategic concerns. These include challenges related to accountability for autonomous decisions, algorithmic bias, reduced human oversight, the escalation of conflicts through accelerated decision-making, and potential violations of International Humanitarian Law (IHL). The study argues that existing governance mechanisms remain inadequate to address these emerging risks. It concludes by recommending the establishment of enforceable international regulatory frameworks, the preservation of meaningful human control over critical military decisions, and the development of defence-specific ethical AI governance standards to ensure the responsible and lawful use of autonomous weapon systems.*

**Keywords:** AWS – (Autonomous Weapons Systems), LAWS – (Lethal Autonomous Weapons, Military Technology), IHL – (International Humanitarian Law), defence ethics, artificial intelligence, defence ethics, UAV – (Unmanned Aerial Vehicles)

***Author for Correspondence**
Enosh Masih
E-mail: enoshmasih78@gmail.com

[1,2,3,4,5]B. Tech Student, Department of Information Technology, Department of Computer Science Greater Noida Institute of Technology Greater Noida, Uttar Pradesh, India

Received Date: July 06, 2026
Accepted Date: July 17, 2026
Published Date: July 30, 2026

**Citation:** Enosh Masih, Binit Behl, Manohar Pandey, Anupam Kumar Singh, Krishna Yadav. Artificial Intelligence in Autonomous Military Weaponry Systems. International Journal on Drones. 2026; 2(2): 9–20p.

## INTRODUCTION

Few technological combinations present substantial universal implication as the conjunction of artificial intelligence and modern military dominance. Over the previous two decades defence ministries and armed forces throughout the globe have channelled immense financial and intellectual assets toward engineering AI-empowered autonomous platforms skilled in implementing complexified mission tasks around challenged operational sectors with finite or no real-time human super. Experts and analysts have progressively produced this switch as a groundwork orientation of warfare itself, a modification of comparable historical magnitude to the arrival of explosive force or weapons of mass destruction [1]. Yet the similarity that matters most is also the most disturbing: distinct from earlier revolutionary weapons AI-enabled autonomous systems are equipped with machine-taught targeting reasoning that enables them to locate, select and counteract plans in ways that avoids traditional human thinking process to act therefore an area that raises immediate and unsettled questions about ethical factor, responsibility and the extended trajectory of international conflict.

In such a changing environment, Lethal Autonomous Weapons Systems (LAWS) have a very controversial place. Defined by the International Committee of the Red Cross (ICRC) as platforms that can autonomously identify and interact with targets in the absence of any significant human control over individual strike decisions [2], LAWS represent an unimaginably wide array of technological modalities. These are UAVs that are programmed to deliver precision attacks within a time-sensitive manner, ground robots, which are used to defend the perimeter and neutralize adversaries, crewless ships capable of fighting on the surface and below, and AI-controlled cyber weapons aimed at degrading or destroying computer networks. A large group of nation-states, including the United States, China, Russia, Israel, the Republic of Korea, and the United Kingdom are now developing and testing such technologies on the battlefield, which is raising growing concern in the international policy community, as well as with scholars of ethics, law, and security.

## LITERATURE REVIEW

The academic interest in artificial intelligence as a war tool has become immensely high in the last ten years in line with the pace of technological change and the emerging prominence of ethical and strategic concerns. The literature reviewed in this paper has been grouped into five thematic areas: the evolutionary history of autonomous weapons systems; the AI technologies that support the concept of military autonomy; ethical and legal considerations; strategic and geopolitical analyses; and models of human-machine cooperation in military settings.

### Evolution of Autonomous Weapons Systems

Remote-controlled weapons conceived during the Second World War constitute the earliest antecedents of what has become the modern LAWS paradigm, yet the transformative inflection point arrived with the post-September 11, 2001 scaling of uninhabited aerial vehicle operations in counterterrorism campaigns [3]. Among the first systematic ethical treatments of machine-executed lethality, Asaro [4] advanced the argument that transferring kill authority to automated systems fundamentally violates the moral obligations that warfare imposes on human agents. An alternative position was staked out by Arkin [5], who proposed that robots operating under embedded ethical protocols might actually demonstrate more reliable compliance with humanitarian norms than human soldiers compromised by fear, fatigue or aggression a claim that prompted substantial scholarly pushback. Crucially, Sparrow [6] articulated a structural problem that has anchored the field ever since namely that autonomous weapons generate a diffusion of moral accountability across designers, commanders and machines with the result that no single agent can bear appropriate responsibility for unlawful engagements.

It was subsequently produced what remains the most exhaustive scholarly account of the LAWS trajectory, surveying the progression from operator-in-the-loop remotely piloted aircraft to candidate fully autonomous platforms and cataloguing active programs across all major military establishments [1]. Building on this foundation, it was examined that the operational translation of 'meaningful human control' the normative touchstone of international policy deliberations revealing a troubling gap between states' declared commitments and the actual technical architectures being embedded in weapons development programs [7].

### AI Technologies Enabling Military Autonomy

The machine intelligence powering contemporary LAWS draws from several converging technical disciplines. Variants of reinforcement learning particularly deep Q-network formulations and proximal policy optimization have exhibited notable proficiency in simulated and structured tactical decision environments [8,9]. Advances in computer vision especially deep convolutional architectures such as YOLO and Faster R-CNN have enabled robust realtime object detection and tracking even in visually degraded or densely populated scenes [10]. The embedding of natural language processing capabilities within AI-augmented command and control architectures has in turn opened pathways to more context-sensitive dialogue between human operators and autonomous platforms [11].

The military implications of swarm intelligence documenting how collective multi-agent learning frameworks can orchestrate emergent coordinated behaviours across large heterogeneous fleets of autonomous vehicles was explored [12]. It was scrutinized the dependability of AI judgment in time-sensitive combat tasks identifying persistent performance breakdowns under adversarial input conditions and when systems encounter operating environments that were not represented during training thus a finding with direct implications for battlefield reliability [13]. The vulnerability of AI systems to deliberate adversarial manipulation was examined and extended this analysis to illuminate the specific threat landscape that adversarial machine learning creates for defence applications [14, 15].

**Ethical and Legal Frameworks**

International humanitarian law furnishes the primary normative scaffolding against which autonomous weapons must be assessed and organized around four foundational principles: distinction between lawful targets and protected persons; proportionality between anticipated military gain and foreseeable civilian harm; precaution in the planning and execution of attacks and military necessity. Scholarly debate over whether autonomous systems can give meaningful operational expression to these principles most acutely in the labyrinthine conditions of urban warfare remains deeply contested. Heyns (2013) reporting to the United Nations in his capacity as Special Rapporteur on extrajudicial killings offered a landmark assessment concluding that LAWS as presently constituted cannot satisfy IHL obligations and calling for an anticipatory ban [16].

Roff (2014) contributed the important observation that proportionality judgments require a form of contextualized moral reasoning weighing expected civilian harm against concrete military advantage across a dynamic and unpredictable situational landscape that cannot be reduced to algorithmic computation [17]. Human Rights Watch and the International Human Rights Clinic (2012) presented documented evidence of the humanitarian dangers of fully autonomous weapons in their landmark advocacy report [18]. Liu (2012) assessed the taxonomy and legality of autonomous systems under the laws of armed conflict concluding that existing treaty law is insufficient and that new legal obligations are required [19]. Backstrom and Henderson (2012) similarly interrogated whether inherited legal frameworks are capable of governing emerging weapons technologies and found them materially inadequate [20].

**Strategic and Geopolitical Dimensions**

From the perspective of strategic studies, the operational introduction of LAWS reshapes incentive structures across all three classical dimensions of military competition: deterrence credibility, escalation management and arms acquisition dynamics. Horowitz (2019) contended that AI-enhanced defence systems will reconfigure the competitive logic driving major power rivalry in particular by rewarding early movers in ways that could undermine equilibrium stability [21]. Johnson (2019) focused on the specific dangers arising from AI integration into nuclear command and advisory systems highlighting how automated decision support could compress deliberation windows and amplify the escalatory consequences of targeting errors [22].

Kania (2019) situated China's military AI modernization within the broader conceptual framework of 'intelligentized warfare,' tracing the influence of the 2017 New Generation AI Development Plan and identifying both convergences with and notable departures from prevailing US doctrinal approaches. Geist and Lohn (2018) modelled the consequences of AI-driven acceleration in military decision cycles arguing that compressed operational timelines could undermine the crisis stability architecture that deterrence theory depends upon potentially more severely than Cold War launch-on-warning dynamics. The foundational US strategic rationale was codified by Work and Brimley (2014) who articulated an offset strategy premised on AI-enabled autonomous systems as the primary instrument for restoring eroding American conventional military advantages [23,24,25].

**Human-Machine Teaming**

Human-machine teaming (HMT) a paradigm in which autonomous systems amplify and assist rather than displace human judgment has crystallized as the dominant policy response to the governance challenges of military AI. Cummings et al. (2019) examined the human-factors dimensions of operating AI-augmented weapons interfaces identifying a persistent tendency toward automation bias in which operators surrender independent judgment in deference to machine-generated recommendations. Endsley (2017) applied situational awareness theory to the problem finding that prevailing HMT design philosophies frequently fail to preserve adequate operator cognizance of the internal state and reasoning processes of autonomous platforms. Roff and Moyes (2016) operationalized the concept of meaningful human control as a set of concrete design requirements, specifying the interface characteristics and procedural safeguards necessary to ensure genuine human authority. Taddeo and Floridi (2018) contributed an ethics-of-information perspective to military AI governance while Russell (2019) argued that embedding human value alignment as an architectural objective in AI system design represents the definitive technical and philosophical challenge of the field with direct relevance to weapons applications [26,27,28,29]

## REVIEW METHODOLOGY

This investigation adopts a mixed-methods research architecture, integrating a systematic literature review, thematic content analysis and a structured comparative case study as its three methodological pillars. The rationale for triangulating across methods derives from the multidisciplinary character of the subject, which encompasses technical, normative, legal, and geopolitical dimensions that no single methodological tradition can fully address subject in a holistic manner as shown in Figure 1.

**Systematic Literature Review**

To guarantee procedural rigor and replicability, the literature review part was conducted in the way of PRISMA (Preferred Reporting Items of systematic reviews and meta-analyses) guidelines. The initial search was through half a dozen major academic repositories, namely IEEE Xplore, ACM Digital Library, JSTOR, Google Scholar, Scopus and Consensus tool. This was complemented by a directed search of institutional gray literature sources including defence ministry policy publications, RAND Corporation publications, Carnegie Endowment publications, Future of Life Institute publications and official ICRC publications. Search string development used Boolean operator logic that combines core conceptual words such as: artificial intelligence AND military, autonomous weapons systems, lethal autonomous weapons, AI AND defence, military robotics, unmanned systems AND combat, machine learning AND warfare. The timeframe of the review was limited to publications that were published in 2009-2026, as this timeframe was a span of the modern epoch of the development of LAWS, but the canonical works of the earlier periods were also included where their significance to the field justified their inclusion as shown in Figure. 2.

Sources were filtered to include based on three positive criteria: substantial involvement with AI or autonomous systems in a military or security context; publication by a peer-reviewed outlet, as a book chapter, as official governmental documentation, or as a product of an established defence or policy research centre; and in English. Not to be considered were strictly technical engineering papers that lacked a strategic or policy aspect, unproven opinion journalism and materials that failed to provide methodological transparency. The initial search sweep found 847 potential sources which were narrowed down to 312 after relevance screening based on abstracts and 78 after text screening using 22 core references that make up the main citation corpus of the paper.

 

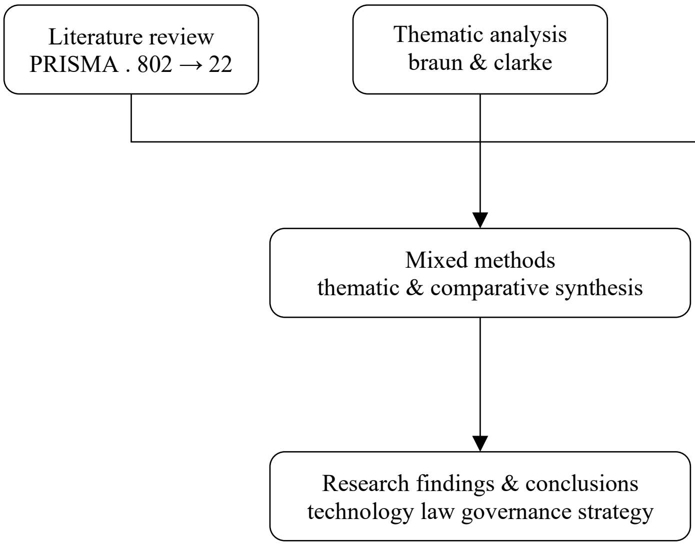

**Figure 1.** Mixed methods design.

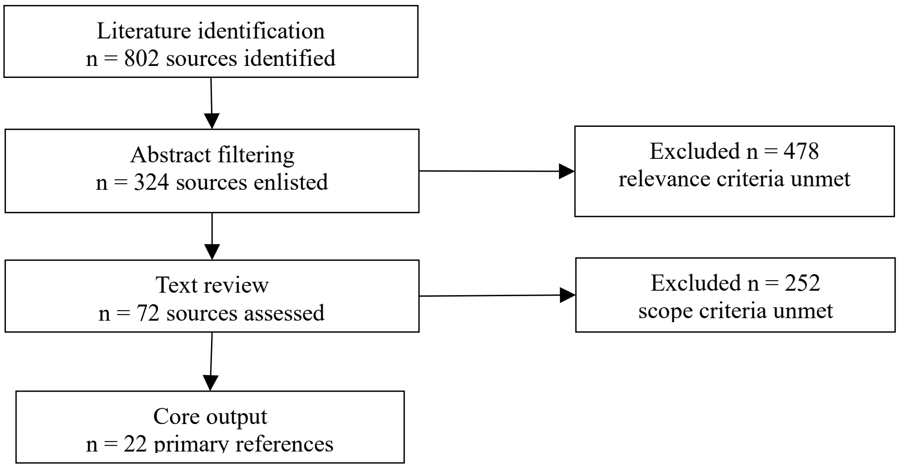

**Figure 2.** PRISMA literature screening.

**Thematic Content Analysis**

Thematic content analysis was performed on the assembled literature corpus following the phased analytical procedure developed by Braun and Clarke (2006). The coding process proceeded through two sequential stages: an inductive open-coding phase during which emergent patterns and categories were identified without imposing a predetermined coding scheme. Additionally followed by an axial coding phase in which cross-cutting relationships among categories were examined and consolidated into an overarching analytical taxonomy. This process yielded four primary thematic domains: (i) technological capability and constraint; (ii) legal compliance and accountability allocation; (iii) strategic stability and competitive arms dynamics and (iv) governance mechanisms and oversight design.

**Comparative Case Study Analysis**

The comparative aspect of the analysis was the four national military AI programs selected on the basis of primacy in the development of global LAWS and their informatively different approaches to governance and ethical constraint; the United States Department of Defence AI Strategy program, the AI integration program of the People's Liberation Army, the National AI Strategy of Russia when applied to defence systems and the autonomous systems portfolio at the Israeli defence establishment. The selection of cases was fashioned to capture variation in both levels of capability and governance orientations. Each case was based on empirical material acquired through the official policy tools, defence budget releases, reports of the specialized defence research institutions, and peer-reviewed program analyses. Each case was analyzed through the analytical framework that was designed in line with four areas of evaluation typically: (1) formally expressed AI military doctrine; (2) autonomous systems programs that are described as active in development or operational deployment; (3) policies that are declared to regulate human control over lethal decisions; and (4) the attitude toward international regulatory processes governing LAWS.

## FINDINGS AND RESULTS

**Table 1. Current State of AI Integration in Military Systems**

The case study analysis of cross-national cases shows a strongly accelerating worldwide trend of AI military assimilation, with a high level of programmatic ambition, technical maturity, and governance philosophy divergence a shown in fig. 3. The US has amassed the most detailed doctrinal basis, defined in its 2018 Defence AI Strategy and the 2020 AI Ethics Principles to the Department of Defence and has invested more than USD 2.5 billion in defence AI annually. Some of the key initiatives that have been identified include Project Maven, the Algorithmic Warfare Cross-Functional Team which is an initiative deploying computer vision intelligence analysis capabilities; the Collaborative Combat Aircraft program which is focused on AI-driven unmanned wingman systems; and the DARPA Offensive Swarm-Enabled Tactics (OFFSET) program which is aimed at exploring coordinated multi-robot tactical autonomy.

The People's Liberation Army of China has been trying to integrate AI in the armed forces at scale across all war fighting domains with the concept of intelligentized warfare and the use of autonomous underwater vehicle systems in anti-submarine warfare, AI augmented command, control systems and unmanned combat aerial vehicles such as the Wing Loong II & CH-7. Human-machine fusion is one of the near-term capabilities listed as a priority in the 2021 military AI roadmap of the PLA. The defence apparatus of Russia has deployed AI-enhanced systems, such as the Uran-9 unmanned ground combat vehicle, which has been battle-tested in Syria and the Poseidon nuclear-capable autonomous torpedo, the most provocatively strategic use of LAWS recorded in this work. The defence industrial sector in Israel has also produced autonomous platforms of international interest such as the Harpy anti-radiation loitering munition which is generally considered one of the first operationally deployed weapons to qualify as a LAWS & the SMASH AI-controlled fire control system to engage direct infantry Table 1.

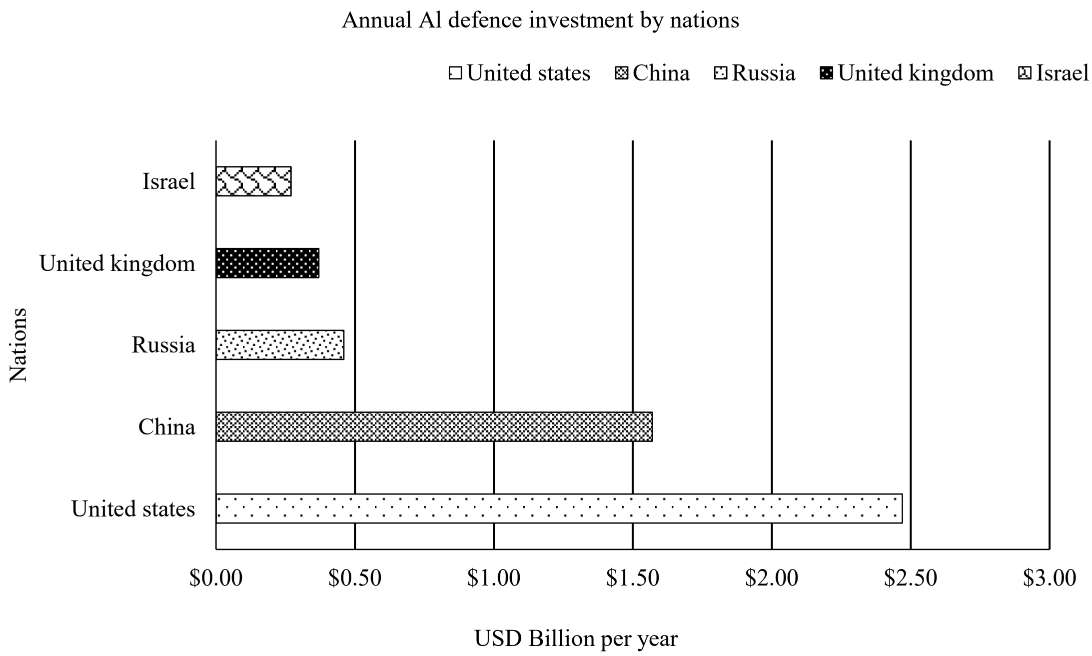


**Figure 3.** AI Defence investment.

| Nation | Al military doctrine | Key autonomous platforms | Human control policy | International regulatory stance |
|---|---|---|---|---|
| United States | Third Offset Strategy, DoD Al Strategy 2018, Al Ethics Principles 2020 | Program Project Maven (CV/AI), Collaborative Combat Aircraft, OFFSET Swarm | Responsible Al authorization, Human-in-loop selection formally declared | CCW GGE member, Hinders binding ban treaty |
| China | Intelligentized Warfare Doctrine, 2017 National Al Development Plan | Wing Loong II UCAV, CH-7 Stealth Drone, Auto no mo us Underwater Vehicles | Human-machine fusion generalization, Limited public clarity | Participates in GGE, Repels legally binding restrictions |
| Russia | National Al Strategy 2019, Conventional Force Modernization | Uran-9 Unmanned Armed AUV Torpedo ground Vehicle; Poseidon Nuclear- | Minimal declared restraints, Limited disclosure | Counters binding LAWS treaty |
| Israel | Capability-led, no formally published doctrine | Harpy Loitering Munition; SMASH AI Fire Control System | Operational stationing Confidential control, framework | Actively withstands ban treaty |

**Table 1.** Case study matrix – defence analytical dimensions across 4 leading nations.

**Technological Capabilities and Limitations**

Thematic analysis of the literature reviewed shows that there are three main categories of AI capabilities that actively contribute to the progress of the LAWS program. Computer vision systems based on deep convolutional neural networks have demonstrated excellent target detection in structured testing conditions but exhibit quantitative accuracy loss in dynamically competitive, visually perturbed or adversarial perturbed environments. Reinforcement learning tactical decision architectures have demonstrated higher performance than human levels in simulated enclosed problem domains - including aerial combat simulations but always face generalization failures in the open-ended distribution-shifted conditions inherent in real battlefield environments. Multi-agent coordination technologies have facilitated more and more complex group behaviours and the DARPA OFFSET program has offered a test case example of 250 coordinated air and ground autonomous agents performing integrated tactical actions.

The literature brings out a collection of basic performance constraints that cannot be swept off as incremental engineering problems. These are: high levels of brittleness during adversarial manipulation of inputs at a low level; proven failure to consistently distinguish lawful combatants and shielded civilians in high-density urban operations; vulnerability to GPS spoofing and electromagnetic interference attacks; decision opaqueness in deep learning models that frustrates post-hoc accountability and checking law compliance and the challenge of instantiating contextually sensitive ethical constraints in statistical learning models. All of these constraints taken together mean that the disparity between the current state of AI and the reliability criteria required by international humanitarian law is not only technical in nature but indicative of architectural constraints of the current AI paradigm that the current engineering breakthroughs are unlikely to bridge.

**Strategic Stability Implications**

Strategic analysis sheds light on several ways in which the development of LAWS will harm the stability of the international security environment. The ability of autonomous strike systems to make machine-speed engagements may gradually remove the time interval in which human beings might deliberate in crisis, significantly increasing the risk that a miscalculation or the failure of algorithms to identify errors triggers escalatory processes. The opaque nature of AI-based decision logic makes it more challenging to assess adversary intent which adds new aspects of strategic uncertainty potentially disrupting the credibility of deterrence. Political economy of autonomous weapons use - whereby the lack of own-force casualties reduces domestic political cost of launching armed action can also systematically decrease the decision rule at which states can invoke military as opposed to diplomatic reactions with consequences to aggregate conflict prevalence. Worst of all, recorded incorporation of AI elements in nuclear command advisory systems in Russian programs and the potential likelihood of similar Chinese ones suggest a scenario where AI-induced false positive target evaluation leads to a nuclear level of escalation.

The analysis of the dynamics of arms competition shows that the incentive terms of the competitive development of AI military by major powers are so strong and unresponsive to unilateral restraint. The first-mover advantage logics, the dual-use nature of general-purpose AI technologies that complicates arms control verification and the ongoing lack of binding treaty restraints all contribute to a situation where individual states who choose restraint to take strategic risk will not trigger restraint in others the traditional logic of a security dilemma. This observation underpins the unequivocal view that governance responses sufficient to the problem should be in the form of binding multilateral legal instruments and not voluntarily binding ethical commitments or statements of norm.

The indicative risk scores being discussed were obtained by means of qualitative analysis of five fundamental academic sources. The analytical foundation of the theory of escalation risk and decision timeline compression was offered by Geist and Lohn (2018), whereas the deterrence instability scores were informed by Horowitz (2019) by analyzing the restructuring of competitive logic between major powers using AI. Johnson (2019) provided AI integration in nuclear command advisory systems as the main evidence of nuclear escalation scores. The China-specific scores of all the dimensions were based on the analysis of the PLA intelligentized warfare doctrine by Kania (2019) and the most comprehensive comparative basis which Scharre (2018) provided, polling active LAWS programmes in all four states. Collectively, these sources provided the qualitative evidence based on which the relative national risk profiles were built on a relative 0 to 10 scale to draw the purposes of Figure 4 [16,17,24].

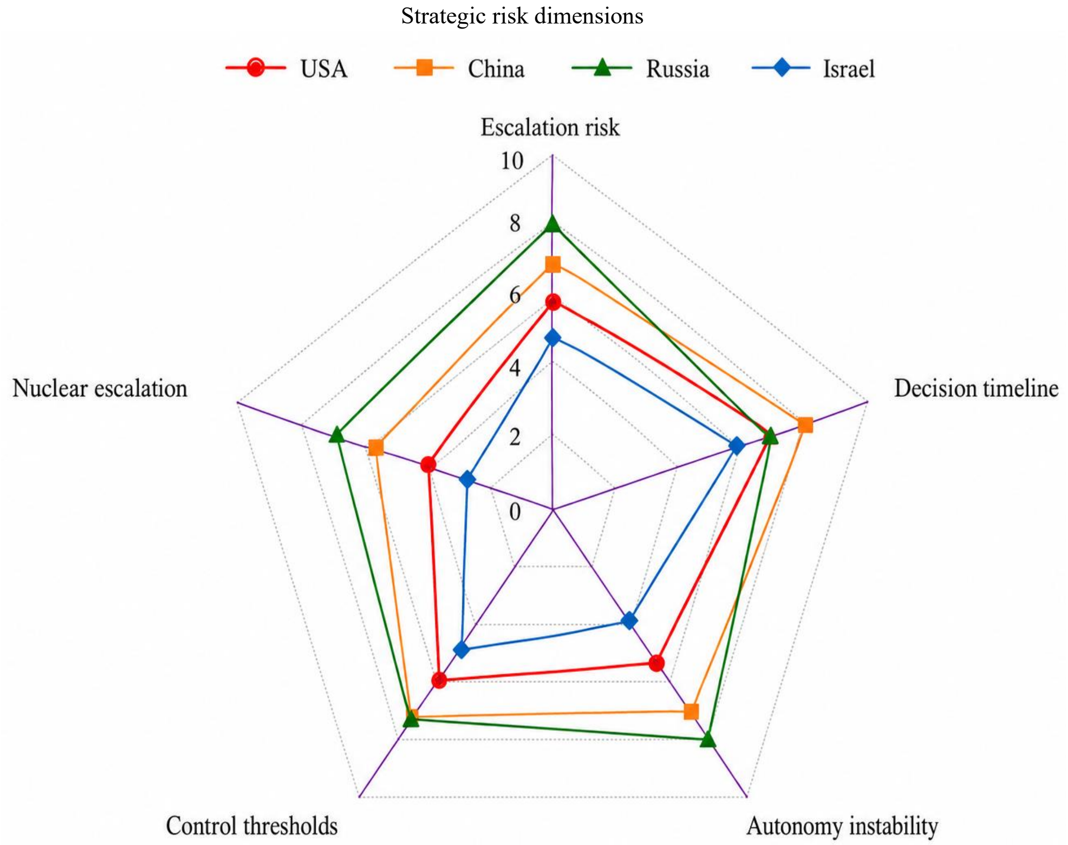


**Figure 4.** Strategic risk profile of 4 nations across five dimensions.

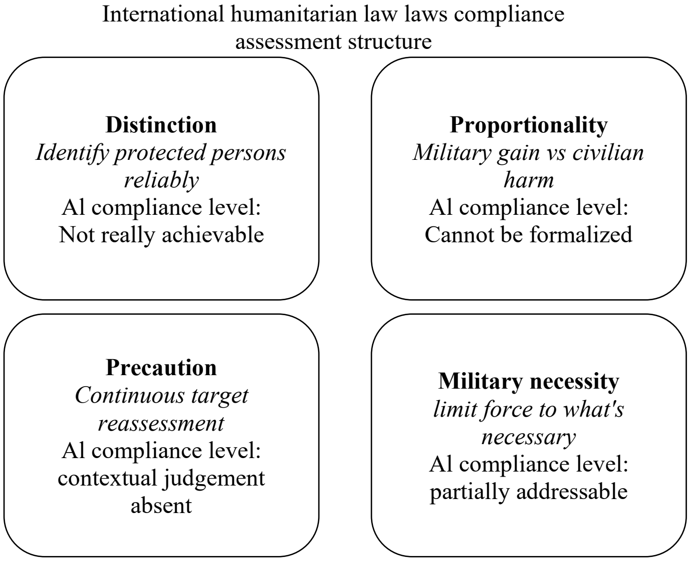


**Figure 5.** Mapping of LAWS Compliance against four primal IHL principles (red→orange→yellow→green depicts order of compliance in decreasing order).

**Ethical and Legal Challenges**

A legal analysis of the interaction of advanced LAWS and international humanitarian law reveals that there are deep structural incompatibilities that cannot be addressed without substantive changes to treaties. Adherence to the principle of distinction requires that weapons systems have the ability to distinguish and reliably detect legitimate military targets and individuals deserving of protective classification a task whose complexity in dense urban conflicts systematically surpasses the reliable operating space of existing AI target recognition systems. The proportionality rule requires a moral estimation of the expected magnitude of civilian casualties versus the expected military utility of a strike which includes political, circumstantial and temporal factors that cannot be computed quantitatively by any current AI algorithm. Compliance with precautionary duties dictates that target legitimacy must always be reassessed by attacking systems as situational variables change during an engagement, assuming it has the ability to dynamically re-appraise situations as it is currently assumed that no deployed autonomous platform is capable of doing so.

Sparrow (2007) responsibility gap rears its head in numerous instances as the most stubborn governance issue in this field. In cases of an unlawful targeting decision in an AI-enabled system, moral and legal responsibility is distributed along a chain of actors software architects, training data curators, commanding officers and the system itself without being wholly attributable to any particular actor in a way sufficient to meet the criteria of international criminal law. This lack of accountability is especially egregious in deep learning systems, the internal inference mechanisms of which cannot be understood even by the people who wrote them. The net conclusion is that the Geneva Conventions, the Hague Regulations and other customary IHL rules are structurally incapable of controlling LAWS as they exist and that special legal tools are needed to bridge the gap (Figure 5).

**Governance and Policy Landscape**

The evaluation of the international governance environment demonstrates the existence of a gaping and growing chasm between the velocity at which LAWS development should proceed and the capacity of multilateral regulatory processes to build an institution. The Convention on Certain Conventional Weapons Group of Governmental Experts on LAWS has met regularly since 2014 but has so far been unsuccessful in producing legally binding output with the key military powers especially the United States, Russia, China, Israel and the United Kingdom, continually opposing the drafting of binding prohibitions. A campaign named Campaign to Stop Killer Robots is a coalition of the civil society of over 180 non-governmental organizations which has encouraged over 30 states to support a preemptive ban treaty, yet it lacks the involvement of the major LAWS-developing states such a tool would not make a significant impact. The AI regulatory framework of the European Union not intended to cover military use proposes a risk-based principles of governance that offers conceivably transferable conceptual architecture to a future LAWS-specific international tool.

## CONCLUSION

This paper has accumulated an in-depth and multidimensional description of artificial intelligence's increasing fusion into autonomous military weapons systems. Incorporative mode exploration, such as the technical realness, legal commitments, moral requisites, strategic outcomes and governance essentials that this development disseminates. Combining all the information that has undergone peer review depicts a picture of rapid rate development that is progressing at a pace that is not in line with regulatory readiness. In this way, constructing a hazardous asymmetry in the middle of what AI-enabled weapons are technically capable of doing and what is approved of by international humanitarian law and foundational ethical principles is constructed.

The academic and empirical history brought to bear in this case decisively testifies to the conclusion that the AI technologies today present in the autonomous weapons systems is not adequately equipped to ensure a steady adherence to the fundamental requirements of IHL in the field of realistic combat. However spectacular in engineered test settings, target recognition and discrimination systems have long-known failure modes when moved to the unpredictable, adversarial combat conditions of real warfare. The dispersal of moral and legal responsibility in the distributed chains of weapons development AI is a structural issue that cannot be addressed in the form of minor improvements in the current legal frameworks. The strategic risks, such as the compression of decision timelines, destabilized deterrence, lower threshold on conflict initiation, and the nuclear escalation conditions that are posed by the AI-assisted command systems, are all dire and remain unrecognized by a significant portion of policy actors.

Based on such results, there are five policy prescriptions. The negotiation of a legally binding multilateral treaty instrument creating a prohibition on the development and combat deployment of fully autonomous weapons systems that have no verified meaningful human control operationally construed as the need to affirmatively authorize human control before each individual targeting decision should be made should be treated with priority by the international community. States having active AI-military programs are to put in place formal transparency mechanisms, including the capability disclosures, as well as requiring incident reporting procedures, to facilitate the confidence-building measures and verification infrastructure on which successful arms control relies. Defence institutions are advised to establish statutory independent AI ethics review committees, similar in power and remit to existing biosafety and nuclear weapons regulatory bodies, with the remit of exposing autonomous weapons systems to legal and ethical scrutiny before they are given a go-ahead. Large-scale further government funding is necessary in technical research on AI interpretability, robustness of systems, and resistance to adversarial attacks, to systematically reduce the gap between AI capability and the reliability needs of IHL. Lastly, the international standards organizations are to establish binding technical specifications in human-machine interface design in weaponry to ensure that meaningful human control is coded structurally in platform architecture and not just promoted as operational doctrine.

The final form that AI will assume in military organizations will not be predetermined by technological determinism; rather, it will be written by conscious political decisions of states, armies and the global community in the nearest future. Given the choices made during this tiny window, there will be the creation of institutional precedents and technical path dependencies whose reverberations will be experienced in the nature of armed conflict generations to come. It is hoped that this study may help to enlighten such decisions with all the analytical force of evidence--and with the clear understanding that permitting machines to have a say in taking human life, without any human say, is a step beyond which there is no easy retrace.

 

**REFERENCES**

1. Arkin RC. Governing lethal behaviour in autonomous robots. Boca Raton (FL): CRC Press; 2009.
2. Asaro P. On banning autonomous weapon systems: human rights, automation, and the dehumanization of lethal decision-making. Int Rev Red Cross. 2012;94(886):687-709.
3. Backstrom A, Henderson I. New capabilities in warfare: an overview of significant legal and ethical challenges for the US military. Int Humanit Law Stud. 2012;3(1):214-41.
4. Bode I, Watts T. Meaning-less human control: Lessons from air defence systems for the debate on autonomous weapons. One Earth Future; 2021. Available from: https://doi.org/10.18289/OEF.2021.050
5. Brundage M, Avin S, Clark J, Toner H, Eckersley P, Garfinkel B, et al. The malicious use of artificial intelligence: Forecasting, prevention, and mitigation. Future of Life Institute; 2018. Available from: https://arxiv.org/abs/1802.07228
6. Cummings ML. Artificial intelligence and the future of warfare. London: Chatham House; 2017.
7. Cummings ML, Roff HM, Cukier K, Parakilas J, Bryce H. Artificial intelligence and international affairs: Disruption anticipated. London: Chatham House; 2019.
8. Endsley MR. From here to autonomy: lessons learned from human-automation research. Hum Factors. 2017;59(1):5-27.
9. Geist E, Lohn AJ. How might artificial intelligence affect the risk of nuclear war? Santa Monica (CA): RAND Corporation; 2018. Available from: https://doi.org/10.7249/PE296
10. Goodfellow I, Bengio Y, Courville A. Explaining and harnessing adversarial examples. In: Proceedings of the International Conference on Learning Representations (ICLR); 2015. Available from: https://arxiv.org/abs/1412.6572
11. Gallois F, Haner J. Drone swarms: the future of aerial combat? J Strateg Stud. 2020;43(1):100-26.
12. Heyns C. Report of the Special Rapporteur on extrajudicial, summary or arbitrary executions. New York: United Nations General Assembly; 2013. Report No.: A/68/382.
13. Horowitz MC. When speed kills: lethal autonomous weapon systems, deterrence and stability. J Strateg Stud. 2019;42(6):764-88.
14. Human Rights Watch, International Human Rights Clinic. Losing humanity: The case against killer robots. New York: Human Rights Watch; 2012.
15. International Committee of the Red Cross. ICRC position on autonomous weapons systems. Geneva: International Committee of the Red Cross; 2021. Available from: https://www.icrc.org/en/document/icrc-position-autonomous-weapon-systems
16. Johnson J. Artificial intelligence and future warfare: implications for international security. Def Secur Anal. 2019;35(2):147-69.
17. Kania EB. Artificial intelligence in China's revolution in military affairs. J Strateg Stud. 2019;44(4):515-42.
18. Liu HY. Categorization and legality of autonomous and remote weapons systems. Int Rev Red Cross. 2012;94(886):627-52.
19. Mnih V, Kavukcuoglu K, Silver D, Rusu AA, Veness J, Bellemare MG, et al. Human-level control through deep reinforcement learning. Nature. 2015;518(7540):529-33.
20. Redmon J, Farhadi A. YOLOv3: An incremental improvement. arXiv [Preprint]. 2018. Available from: https://arxiv.org/abs/1804.02767
21. Roff HM. The strategic robot problem: lethal autonomous weapons in war. J Mil Ethics. 2014;13(3):211-27.
22. Roff HM, Moyes R. Meaningful human control, artificial intelligence and autonomous weapons. Briefing paper prepared for the Informal Meeting of Experts on Lethal Autonomous Weapons Systems (LAWS); 2016; Geneva, Switzerland. Geneva: United Nations; 2016.
23. Russell S. Human compatible: Artificial intelligence and the problem of control. New York: Viking; 2019.
24. Scharre P. Army of none: Autonomous weapons and the future of war. New York: W. W. Norton & Company; 2018.

 

25. Silver D, Schrittwieser J, Simonyan K, Antonoglou I, Huang A, Guez A, et al. Mastering the game of Go without human knowledge. Nature. 2017;550(7676):354-59.
26. Singer PW. Wired for war: The robotics revolution and conflict in the 21st century. New York: Penguin Press; 2009.
27. Sparrow R. Killer robots. J Appl Philos. 2007;24(1):62-77.
28. Taddeo M, Floridi L. How AI can be a force for good. Science. 2018;361(6404):751-52.
29. Work B, Brimley S. 20YY: Preparing for war in the robotic age. Washington (DC): Center for a New American Security; 2014.

 