Article

Not peer-reviewed version

Artificial Intelligence, Nuclear Warfare and Towards Laws: A Study on Limitations of Existing International Laws and Treaties

Mustak Ahmed *

Posted Date: 30 June 2025

doi: 10.20944/preprints202506.2483.v1

Keywords: Artificial Intelligence (AI); Nuclear Warfare; International Law; Autonomous Weapons Systems; Arms Control Treaties; NPT; Legal Gaps; Algorithmic Escalation; Human Oversight; International Humanitarian Law (IHL); Technological Ethics; Security Governance

 doi:10.20944/preprints202506.2483.v1

Article

Artificial Intelligence, Nuclear Warfare and Towards Laws: A Study on Limitations of Existing International Laws and Treaties

Mustak Ahmed

Professor, Mass Communication and Journalism, University of Rajshahi, Bangladesh, mustak@ru.ac.bd, http://rurfid.ru.ac.bd/ru_profile/public/teacher/21501555/profile; https://orcid.org/0000-0002-1436-1101https://www.linkedin.com/in/mustak-ahmed-583084253/; WSR: IDADF-0848-2022; SciProfiles: 2108488; Loop profile: 1809701; https://www.researchgate.net/profile/Mustak-Ahmed?ev=hdr_xprf; SSRN Author ID: 5911923; https://rajshahi.academia.edu/MustakAhmed

Abstract

The rapid convergence of artificial intelligence (AI) and nuclear warfare technologies has outpaced the existing frameworks of international law, exposing critical legal, ethical, and strategic vulnerabilities in global security governance. This study examines the limitations of current international laws and treaties—such as the Treaty on the Non-Proliferation of Nuclear Weapons (NPT), the Geneva Conventions, and emerging norms under the Convention on Certain Conventional Weapons (CCW)—in addressing the challenges posed by AI-enhanced nuclear command, control, and decision-making systems. Through interdisciplinary analysis combining international legal theory, security studies, and AI ethics, the research highlights how autonomous systems, algorithmic opacity, and dual-use technological ambiguity undermine core principles of accountability, state responsibility, and deterrence stability. The paper also explores scenarios where AI-driven escalation risks could bypass human oversight, creating pathways to inadvertent conflict or unlawful use of force. It argues for the urgent need to recalibrate global legal architectures, establish binding normative frameworks, and incorporate AI-specific clauses within arms control agreements. By identifying legal lacunae and proposing policy directions, this study contributes to a timely and critical discourse on reimagining the future of law in the age of intelligent warfare.

Keywords: Artificial Intelligence (AI); Nuclear Warfare; International Law; Autonomous Weapons Systems; Arms Control Treaties; NPT; Legal Gaps; Algorithmic Escalation; Human Oversight; International Humanitarian Law (IHL); Technological Ethics; Security Governance

1. Introduction

1.1. Background and Context

The evolution of artificial intelligence (AI) is transforming the global security landscape. While AI has led to notable advancements in healthcare, finance, logistics, and climate science, its growing incorporation into military systems—particularly nuclear command, control, and communication (NC3) infrastructures—has stirred unprecedented ethical, legal, and geopolitical concerns. The integration of machine learning algorithms, neural networks, predictive analytics, and semi-autonomous platforms into nuclear operations introduces a paradigm shift that international legal frameworks were never designed to handle.

Nuclear deterrence has traditionally been underpinned by human cognition, accountability, and ethical reasoning. However, as AI becomes increasingly involved in decision-making loops, particularly in early warning systems, command architecture, and automated responses, the role of

 doi:10.20944/preprints202506.2483.v1

human oversight diminishes. This introduces an elevated risk of accidental launches, miscalculated escalation, and opaque accountability. Unlike previous technological disruptions such as the advent of intercontinental ballistic missiles (ICBMs) or cyber capabilities, AI presents a layered and interactive challenge: it not only changes the speed and scope of operations but also alters the epistemology of decision-making (Boulanin & Verbruggen, 2017).

As of the 2020s and early 2030s, leading nuclear states including the United States, Russia, China, India, and Israel have begun integrating AI components into their defense ecosystems. AI is being employed in threat detection, data fusion, target prediction, unmanned vehicles, and cyber defense, often with ambiguous command structures (Scharre, 2018; Kania, 2019). Though none have openly declared full autonomy in nuclear decision-making, the operational logic and strategic doctrines are increasingly skewed towards reducing reaction time and pre-emptive action based on algorithmic cues. The result is a new arms race—not solely in terms of warheads or missiles, but in lines of code, datasets, and machine-generated decision architectures (Horowitz, 2019).

1.2. The Need for Legal Evolution

International legal regimes have historically struggled to keep pace with disruptive technological change. From the Geneva Conventions (1949) and the Non-Proliferation Treaty (1968) to recent discussions around lethal autonomous weapons systems (LAWS), the trajectory of arms control law has largely been reactive, fragmented, and state-centric. There is little to no direct regulation of AI-enabled nuclear systems under existing international humanitarian law (IHL), arms control agreements, or disarmament treaties.

For instance, the Treaty on the Non-Proliferation of Nuclear Weapons (NPT) focuses on preventing horizontal proliferation and achieving eventual disarmament but remains silent on modernization through AI. Likewise, the International Humanitarian Law (IHL) corpus, including the Geneva and Hague Conventions, anchors its principles on human deliberation, proportionality, and distinction—standards difficult to enforce with machine-made decisions in opaque black-box models (ICRC, 2019).

Given the dual-use nature of AI— where the same algorithm may serve civilian and military purposes—defining its legal status becomes even more complex. Moreover, the rapid deployment of AI by private defense contractors and tech firms (e.g., Palantir, Lockheed Martin, Huawei Defense) further fragments state responsibility and legal enforceability (Crootof, 2015). This raises foundational questions: Who is legally responsible when an AI misfires in a nuclear environment? Can a treaty regulate an evolving software model? How can international law operate within systems built on secrecy, machine learning opacity, and national security imperatives?

1.3. Objectives of the Study

The objectives of this study are shaped by the urgent need to interrogate and understand the inadequacies of current international legal frameworks in dealing with the convergence of artificial intelligence (AI) and nuclear weapons technologies. As AI transforms both the logic and logistics of warfare, especially in nuclear command and control systems, this study seeks to explore the legal, ethical, and policy challenges that arise from this transformation. The following are the core objectives of the research:

1.3.1. To Assess the Limitations of Existing International Laws

One of the principal objectives is to critically examine how existing treaties—such as the Treaty on the Non-Proliferation of Nuclear Weapons (NPT), the Geneva Conventions, and customary international humanitarian law (IHL)—fail to address the unique threats posed by the integration of AI into nuclear arsenals. The study aims to identify the specific legal gaps and ambiguities that arise when autonomous systems are deployed in nuclear command, control, and decision-making chains.

1.3.2. To Explore the Risks of Algorithmic Escalation and Autonomy

The study seeks to analyze how AI-driven systems may lead to algorithmic escalation, where rapid, automated responses escalate tensions without human intention or oversight. This includes

 doi:10.20944/preprints202506.2483.v1

investigating the implications of machine-speed warfare on deterrence theory, decision-making latency, and crisis stability, and assessing whether existing legal principles such as proportionality and distinction can be meaningfully applied in these contexts.

1.3.3. To Evaluate Accountability and Responsibility Dilemmas

The research aims to investigate who holds responsibility under international law when AI-enabled systems cause unlawful acts or catastrophic errors in the context of nuclear warfare. This includes questions of state responsibility, individual criminal liability, and the attribution of fault in incidents involving complex autonomous systems—issues which current legal instruments do not adequately resolve.

1.3.4. To Analyze the Role of Dual-Use Technologies

The study also intends to explore the dual-use nature of AI technologies—developed for civilian purposes but adaptable for military and nuclear applications. This objective seeks to reveal how such dual-use capabilities complicate arms control verification, legal classification, and treaty enforcement, especially in the absence of robust regulation.

1.3.5. To Identify Normative and Policy Gaps

A key objective is to identify the absence of binding normative frameworks and policy mechanisms tailored to the intersection of AI and nuclear capabilities. The study examines the slow evolution of international legal discourse compared to the fast-paced development of emerging technologies, pointing to the lack of proactive multilateral initiatives and enforceable guidelines.

1.3.6. To Propose Legal and Policy Recommendations

The study seeks to develop evidence-based recommendations for strengthening the international legal order in response to the militarization of AI in nuclear contexts. This includes proposing updates to existing treaties, advocating for new multilateral agreements, and suggesting regulatory oversight bodies and verification mechanisms that can address both technical complexity and legal accountability.

1.3.7. To Contribute to Interdisciplinary Legal and Security Discourse

Finally, the research aims to contribute meaningfully to the growing interdisciplinary field at the nexus of law, technology, and global security. By engaging legal scholarship with insights from AI ethics, political science, and military strategy, the study aspires to build a comprehensive and forward-looking understanding of the challenges ahead.

By synthesizing perspectives from international law, security studies, ethics, and technology policy, the study provides an interdisciplinary framework for reimagining legal accountability in the age of algorithmic war.

1.4. Key Research Questions

In order to systematically investigate the legal, strategic, and ethical implications of artificial intelligence (AI) in the domain of nuclear warfare, this study is guided by several interrelated research questions. These questions are designed to uncover the core legal gaps, assess potential threats, and propose frameworks for a more accountable international legal architecture. The overarching goal is to interrogate whether and how the existing body of international law is equipped to deal with the convergence of AI and nuclear technologies.

1.4.1. How Adequate Are Existing International Laws and Treaties in Addressing AI-Enabled Nuclear Threats?

This question critically evaluates whether the current international legal regimes—such as the Treaty on the Non-Proliferation of Nuclear Weapons (NPT), Geneva Conventions, and customary international humanitarian law—are sufficient in scope and enforcement to manage the risks emerging from AI's integration into nuclear decision-making and delivery systems.

 doi:10.20944/preprints202506.2483.v1

1.4.2. What Unique Legal and Ethical Challenges Does AI Pose in the Context of Nuclear Warfare?

AI introduces new dimensions of complexity, including autonomous decision-making, algorithmic opacity, and rapid escalation dynamics. This question explores how these elements challenge foundational legal principles such as distinction, proportionality, precaution, and state accountability in the context of nuclear engagement.

1.4.3. Who Bears Responsibility in the Event of Unlawful or Accidental Use of AI-Driven Nuclear Systems?

As autonomous or semi-autonomous systems blur the lines of agency, this question investigates the legal responsibilities of states, military commanders, programmers, and corporations involved in developing or deploying such technologies. It also explores how the doctrine of command responsibility and state responsibility apply in these complex scenarios.

1.4.4. To What Extent Do Dual-Use Technologies Complicate Legal Regulation and Verification Protocols?

Since many AI technologies developed for civilian or commercial purposes can be repurposed for military or nuclear applications, this question examines the challenges in verifying compliance with international treaties, particularly in distinguishing peaceful from hostile technological development.

1.4.5. How Can International Law Be Reimagined or Revised to Respond to Algorithmic Warfare and Autonomy?

This forward-looking question aims to identify potential models for legal reform. It asks what kinds of normative frameworks, treaty amendments, or new multilateral agreements might be required to effectively regulate AI in the nuclear domain while preserving global stability and humanitarian values.

1.4.6. What Role Should Global Institutions Play in Governing AI-Nuclear Convergence?

This question examines the responsibility and capacity of global institutions such as the United Nations, the International Atomic Energy Agency (IAEA), and the International Court of Justice (ICJ) in monitoring, regulating, and enforcing norms concerning AI and nuclear integration. It also considers the feasibility of creating new institutions or regulatory bodies focused on AI arms control.

1.4.7. How Do Emerging Security Doctrines and Military AI Strategies Influence Legal Interpretation and Application?

Finally, this question explores how evolving military doctrines, such as 'automated deterrence,' 'human-on-the-loop' command structures, and predictive warfare strategies, are reshaping the application of international law. It seeks to understand the legal implications of these strategies and how they challenge traditional norms governing warfare.

1.5. Methodological Approach

This study adopts a qualitative, doctrinal legal research method, complemented by comparative analysis of state practices and institutional documents. Primary data sources include treaty texts, UN reports, national defense white papers, and AI policy guidelines. Secondary sources include academic literature, journal articles, think-tank analyses, and expert interviews. Legal hermeneutics is employed to interpret treaties and principles of IHL, while technology assessment frameworks guide the evaluation of AI capabilities.

1.6. Historical Parallels and Evolution

Throughout history, transformative technologies have often outpaced legal regulation. The advent of the submarine, the airplane, nuclear weapons, and cyberwarfare each introduced

 doi:10.20944/preprints202506.2483.v1

asymmetries in warfare that legal norms struggled to constrain. AI presents a similar challenge—albeit one that is exponentially more dynamic and non-deterministic.

In the Cold War era, the establishment of hotlines, second-strike doctrines, and verification protocols attempted to create a fragile but stable deterrence system. However, as AI systems potentially bypass human scrutiny and compress reaction time from minutes to seconds, these stabilizing mechanisms are rendered inadequate. The historical reliance on human rationality as a safeguard is now under threat (Acton, 2018).

Case of Soviet Petrov Incident (1983)

One of the most cited incidents highlighting human judgment in nuclear stability was the 1983 Petrov Incident, where a Soviet officer overrode computer systems falsely indicating a US first strike. Had an AI been in charge, escalation might have followed due to a misclassification error. This demonstrates that algorithmic systems, even when highly accurate, cannot replicate the ethical, contextual, and intuitive decisions made by human actors.

1.7. Ethical and Philosophical Considerations

Beyond legalities, the deployment of AI in nuclear operations raises fundamental ethical questions: Can a machine be trusted with decisions that can end civilization? Can programming codes comprehend complex doctrines such as 'proportionality' or 'just war'? How should international ethics influence sovereign military AI development?

The philosophy of responsibility in war, often tied to intention and culpability, is eroded when decisions are delegated to machines. The ethics of predictive decision-making—particularly pre-emptive strikes based on AI estimations of adversary behavior—further muddies the legality and morality of action. Scholars like Sparrow (2007) and Asaro (2012) argue that war algorithms may fundamentally erode human dignity by removing moral agency from lethal outcomes.

1.8. Geopolitical Fragmentation and Strategic Competition

The AI-nuclear interface is not unfolding in a vacuum but within a highly competitive global order. The United States, China, and Russia are investing in AI not only as a military asset but also as a geopolitical currency of power and influence. The United Nations discussions on LAWS have remained gridlocked due to incompatible state positions—especially between democratic and authoritarian regimes.

Additionally, Global South countries are largely excluded from these discussions, despite the potential for their territories to become AI battlegrounds or experimental theatres. As a result, a 'techno-nuclear divide' is emerging, where elite states design, deploy, and defend AI-based nuclear systems while international law struggles to catch up.

1.9. Structural Challenges to Lawmaking

Several structural factors hinder effective lawmaking in the AI-nuclear space:

1. Opacity of AI Systems: Many AI models are black boxes, making verification and legal assessment difficult.

2. National Security Exception: States often claim exceptionalism to avoid international scrutiny.

3. Private Sector Involvement: AI development is largely in the hands of non-state actors, which undermines treaty enforcement.

4. Lack of Consensus on Definitions: There is no universally agreed definition of AI, autonomy, or meaningful human control.

5. Dual-Use Nature: Algorithms designed for civilian use can easily be weaponized.

These challenges necessitate not just treaty amendment but legal innovation—through adaptive norms, AI ethics guidelines, and multistakeholder oversight mechanisms.

2. Literature Review

2.1. Introduction to the Literature Landscape

 doi:10.20944/preprints202506.2483.v1

The intersection of artificial intelligence (AI), nuclear weapons, and international law has emerged as a rapidly expanding field of scholarship. This literature review synthesizes academic, institutional, and policy-based contributions to analyze the scope, depth, and limitations of existing research. The aim is to map how AI integration in nuclear command and control is addressed across three primary domains: (1) technological and strategic studies, (2) legal and ethical debates, and (3) arms control and treaty-based regimes. The literature is multidisciplinary, involving law, international relations, security studies, computer science, and political philosophy. However, the convergence of these domains is still in its formative stage, with significant gaps in legal theory and institutional governance.

2.2. Technological Evolution and Strategic Doctrine

2.2.1. Automation, Autonomy, and Escalation Risk

AI's role in nuclear command and control is often examined within the broader strategic discourse of automation and military escalation. Horowitz, Kahn, and Scharre (2020) argue that the introduction of AI in early-warning systems and threat assessment mechanisms significantly reduces decision time, increasing the risk of miscalculation and unintended escalation. Their work builds on historical episodes such as the Petrov incident (1983), using counterfactual analysis to demonstrate how AI could bypass human judgment and trigger erroneous retaliation.

Similarly, Acton (2018) introduces the concept of 'escalation through entanglement,' describing how AI-dependent NC3 systems blur the boundaries between conventional and nuclear assets. Acton's work is foundational in illustrating how AI might inadvertently increase strategic instability due to its integration in dual-use infrastructure, especially through cyber vulnerabilities.

Boulanin and Verbruggen (2017) from SIPRI provide a typology of autonomy in weapons systems. Their report, 'Mapping the Development of Autonomy in Weapon Systems,' delineates varying degrees of autonomy and clarifies misconceptions between automated and autonomous systems. This distinction is crucial for legal categorization and understanding the chain of accountability.

2.2.2. Predictive Analytics and AI Bias

The use of predictive algorithms in nuclear risk assessment has gained attention in both defense studies and machine learning ethics. Kania (2019) examines China's investment in predictive analytics for missile defense and highlights the dangers of over-reliance on algorithmic inferences without sufficient human review. The concept of 'epistemological opacity' (Burrell, 2016) becomes particularly relevant, indicating that even system designers may not fully understand the decision pathways of deep learning models—raising concerns about transparency and auditability in high-stakes environments.

2.3. Legal and Ethical Discourse

2.3.1. International Humanitarian Law (IHL) and AI

The applicability of IHL to autonomous weapons has been a recurring theme in legal scholarship. The ICRC (2019) and Schmitt (2013) argue that existing IHL principles—distinction, proportionality, and precaution—are fundamentally challenged by AI. For instance, distinguishing combatants from civilians in complex environments may be beyond the capacity of current AI systems, leading to violations of jus in bello principles.

Sparrow (2007) critiques the moral logic of delegating life-and-death decisions to machines, arguing that such delegation strips war of its moral core. His concept of the 'responsibility gap' is echoed by Crootof (2015), who contends that current legal frameworks lack the normative structure to assign liability when autonomous systems commit unlawful acts.

The ethical concerns intensify in nuclear scenarios, where even a single misjudgment could result in mass destruction. Asaro (2012) calls for a moratorium on AI in nuclear weapons, emphasizing that the stakes are too high for experimental systems. His critique is rooted in international human rights law and the Martens Clause, advocating for the preservation of human dignity in armed conflict.

 doi:10.20944/preprints202506.2483.v1

2.3.2. Responsibility, Accountability, and State Sovereignty

A key body of literature addresses the issue of accountability when AI is deployed in military settings. Pagallo (2013) introduces the notion of 'distributed responsibility' in AI-driven decisions, complicating traditional legal doctrines that rely on clear chains of command. This becomes especially problematic in joint operations involving multiple states or private defense contractors, who may design and maintain the AI systems but not deploy them directly.

Vincent and Packer (2020) further explore how the international legal framework must evolve to accommodate responsibility for outcomes driven by non-human agents. Their work draws parallels from tort law and cyber liability to propose adaptable legal models.

Meanwhile, state-centric scholars like Yoo and Schmitt (2020) resist new legal instruments, arguing that existing IHL and state practice are sufficient. They posit that the principle of lex lata can absorb new technologies if interpreted dynamically. However, this optimistic view is criticized for underestimating the speed and opacity of AI evolution.

2.4. Arms Control Regimes and Treaty Gaps

2.4.1. The NPT and Technological Modernization

The Treaty on the Non-Proliferation of Nuclear Weapons (NPT) remains the cornerstone of nuclear disarmament law, but scholars widely agree it is technologically outdated. Kristensen and Korda (2023) argue that the treaty lacks provisions to regulate qualitative modernization of arsenals, including AI-enhanced delivery systems, cyber-NC3 tools, or automated missile defense.

Krepon (2019) highlights the political inertia within the NPT review process, where states avoid contentious issues like AI to preserve consensus on more foundational goals. This 'doctrine of silence' risks rendering the treaty irrelevant in the face of AI disruptions.

The NPT's verification mechanisms are also under scrutiny. The IAEA's safeguards are oriented towards material accounting—tracking fissile material and weaponized components. They are ill-equipped to inspect software, algorithms, or AI training datasets (UNIDIR, 2022). As such, the infrastructure for compliance verification in the AI era is practically non-existent.

2.4.2 Convention on Certain Conventional Weapons (CCW)

The CCW has been the main venue for multilateral dialogue on lethal autonomous weapon systems (LAWS). However, progress has been minimal. Discussions under the Group of Governmental Experts (GGE) have yet to reach consensus on definitions, let alone legal outcomes (ICRC, 2021). The US, Russia, and Israel oppose binding regulation, citing military necessity and technological fluidity.

Docherty (2019) critiques the CCW's voluntary and consensus-based format, arguing it structurally favors the interests of powerful states and frustrates norm emergence. She recommends shifting the locus of negotiation to the UN General Assembly or initiating a separate treaty like the Nuclear Ban Treaty (TPNW), with an AI-specific annex.

2.4.3. Other Initiatives: TPNW, OST, and Regional Frameworks

The Treaty on the Prohibition of Nuclear Weapons (TPNW) includes preambular recognition of emerging technologies but lacks operational clauses on AI. The Outer Space Treaty (OST) and regional frameworks like the African Nuclear-Weapon-Free Zone Treaty (Pelindaba Treaty) are similarly silent on algorithmic command systems.

However, some regional bodies have shown initiative. The European Parliament (2021) passed a resolution calling for a ban on fully autonomous weapons and the preservation of human control in nuclear command decisions. Though non-binding, such actions reflect growing civil society engagement in this domain.

2.5. Emerging Norms and Non-State Actors

2.5.1. Private Tech Firms and Defense Contractors

One major gap in the literature concerns the role of private sector actors in shaping the future of AI-based warfare. Companies such as Google (Project Maven), Amazon (cloud infrastructure for

 doi:10.20944/preprints202506.2483.v1

defense), and Palantir (battlefield analytics) are increasingly entangled with military AI systems (Singer & Brooking, 2018). Yet, they operate outside the jurisdiction of traditional arms control treaties.

Scholars such as Brundage et al. (2018) call for voluntary codes of conduct and ethical charters for AI developers, akin to corporate social responsibility models. The OECD AI Principles and UNESCO's AI Ethics Recommendations (2021) are initial attempts to provide global guidance, but their non-binding nature limits effectiveness.

2.5.2. Civil Society and Epistemic Communities

Civil society organizations like the Campaign to Stop Killer Robots and the Future of Life Institute have significantly shaped public discourse. Their policy papers, public campaigns, and petitions have generated normative pressure, even if state-level impact remains uneven.

Academic epistemic communities are also beginning to form cross-sectoral alliances. For example, the 'AI and Global Security Initiative' by Brookings and the 'AI Governance Project' at Oxford University are producing interdisciplinary literature bridging law, ethics, and policy.

While there is a growing body of literature on AI and international security, significant gaps remain:

1. Insufficient Legal Frameworks: Few works propose concrete legal architectures to regulate AI in nuclear warfare beyond normative ethics.

2. Lack of Empirical Studies: Most literature is speculative or theoretical, lacking empirical case studies or state practice documentation.

3. Private Sector Accountability: Minimal legal literature exists on regulating defense contractors or AI vendors under IHL.

4. North-South Imbalance: Much of the scholarship is Global North-centric, ignoring how AI-nuclear policies affect or exclude the Global South.

5. Fragmented Policy Discourse: Legal, technical, and ethical discussions are siloed, with limited cross-pollination between communities.

The literature on AI and nuclear warfare is evolving rapidly but remains insufficiently integrated across disciplines and jurisdictions. Technological advancements are outpacing both academic theorization and legal codification. While there is a robust debate on the risks and strategic dilemmas posed by AI, the legal literature remains underdeveloped, particularly in crafting enforceable international norms.

This study builds upon the existing body of work while addressing its limitations by proposing a more unified approach—integrating legal reform, treaty development, technical oversight, and ethical scrutiny. The next section will explore theoretical frameworks that can guide the reconceptualization of international law in the age of intelligent machines and nuclear threats.

The dawn of the AI-nuclear age necessitates a fundamental rethinking of how humanity governs war, peace, and survival. Traditional legal regimes, forged in the aftermath of conventional wars and analog technologies, are now facing an existential test. If law is to remain relevant as a human institution, it must evolve to confront the specter of artificial intelligence managing weapons of ultimate destruction. This research initiates that conversation by situating AI within nuclear warfare and challenging the global community to respond not only with policy but with moral clarity and legal innovation.

3. Theoretical Framework

3.1. Introduction to Theoretical Grounding

As the technological integration of artificial intelligence (AI) into nuclear warfare systems accelerates, traditional theories of war, peace, and international law confront epistemological challenges. How do we conceptualize agency when decisions are automated? What happens to legal accountability when outcomes are determined by black-box algorithms? How do norms evolve when power is partially transferred from sovereign humans to artificial entities?

To address these foundational questions, this section presents a multidimensional theoretical framework built upon four interconnected paradigms:

1. Technological Determinism and Sociotechnical Systems Theory

 doi:10.20944/preprints202506.2483.v1

2. Constructivist International Relations and Norm Diffusion
3. Legal Positivism vs. Legal Pluralism in International Law
4. Ethics of War and Just War Theory in the Age of AI

This framework not only guides the analytical interpretation of AI integration in nuclear systems but also serves as a lens for understanding the limitations—and potential—of international legal regulation in this domain.

3.2. Technological Determinism and Sociotechnical Systems

3.2.1. Technology as a Shaping Force

Technological determinism is a theory positing that technology drives societal change and often outpaces the capacity of institutions to adapt (Winner, 1986). In the context of nuclear warfare, this view holds that the integration of AI is not merely a tool to enhance human capabilities but a structuring force that reshapes strategic thinking, decision-making architecture, and even legal accountability.

The 'second offset strategy' in U.S. defense policy—first articulated in the 1970s—posited that technology (e.g., stealth, precision-guided munitions) would compensate for strategic disadvantages. Now, AI and automation constitute a potential 'third offset,' intended to surpass both human cognition and adversary capabilities (Work & Brimley, 2014). This shift aligns with deterministic thought, suggesting that legal systems must catch up to technological imperatives.

However, critics such as MacKenzie and Wajcman (1999) argue for a more nuanced understanding through sociotechnical systems theory, which posits that technology and society co-shape each other. According to this view, AI integration into nuclear systems is not inevitable but is driven by political, cultural, and institutional choices.

3.2.2. Black-Box Epistemology and Legal Invisibility

AI algorithms—especially deep learning neural networks—introduce an epistemological opacity known as the 'black box' problem (Burrell, 2016). When AI processes are not transparent even to their developers, attributing intention, reasoning, or error becomes legally complex. This challenges the principle of mens rea (guilty mind) central to both criminal law and the law of armed conflict.

Within sociotechnical systems theory, the lack of transparency is not merely a technical problem but a systemic one—a consequence of prioritizing speed, secrecy, and efficiency over accountability. Law, then, becomes a reactive force in a system designed to obscure causality.

3.3 Constructivist International Relations and Norm Diffusion

3.3.1. Norm Construction in Emerging Technologies

Constructivism in international relations argues that international politics are shaped by social constructs, including norms, identities, and shared meanings (Wendt, 1999). Unlike realism or liberalism, constructivism emphasizes ideational factors over material capabilities. In this light, legal norms surrounding AI and nuclear weapons do not emerge solely from power or rationality but from the interaction of states, institutions, and civil society.

The development of international norms on autonomous weapons is still nascent but visible. The 2013 campaign to ban lethal autonomous weapons systems (LAWS), led by NGOs like the Campaign to Stop Killer Robots, illustrates how non-state actors participate in norm entrepreneurship (Finnemore & Sikkink, 1998).

However, norm internalization remains uneven. While some states support pre-emptive bans on autonomous weapons, others (notably the United States, Russia, China, and Israel) resist such efforts, citing strategic autonomy. This divergence reflects different national identities and threat perceptions, aligning with constructivist views that norm diffusion is not linear but contested.

3.3.2. Legalization of Norms and the Problem of State Consent

Abbott and Snidal (2000) introduce the concept of 'legalization' in international law, emphasizing that rules gain effectiveness when they are precise, obligatory, and delegate authority

 doi:10.20944/preprints202506.2483.v1

to neutral bodies. In the AI-nuclear context, however, legalization is hindered by the absence of consensus on definitions, verification mechanisms, and sovereign limits.

Constructivist scholars stress that legal norms evolve through discourse, not just through treaties. The CCW debates, UN panel discussions, and national AI strategies collectively shape the discursive field in which norms of meaningful human control or algorithmic accountability may take root. Yet, without coordinated norm entrepreneurship, these discourses remain fragmented.

3.4. Legal Positivism vs. Legal Pluralism in International Law

3.4.1. Legal Positivism and Treaty Formalism

Legal positivism maintains that law is a system of rules recognized by social facts and institutional authority, rather than moral or ethical considerations (Hart, 1961). Most international legal regimes, including the NPT and Geneva Conventions, adhere to this framework. Under this view, unless a state signs and ratifies a treaty banning or regulating AI in nuclear command, no binding obligation exists.

Positivists argue that customary international law and treaties are sufficient to absorb new technologies. For example, Article 36 of Additional Protocol I to the Geneva Conventions requires legal review of new weapons systems. However, this mechanism is state-controlled and lacks transparency, rendering it ineffective for regulating complex AI systems developed by private actors (Schmitt, 2013).

3.4.2. Legal Pluralism and Adaptive Jurisprudence

Legal pluralism challenges the state-centrism of positivism, emphasizing that law can emerge from multiple sources: indigenous traditions, transnational norms, religious codes, and institutional practices (Merry, 1988). In the AI-nuclear context, legal pluralism offers a pathway to inclusive norm-making that accounts for civil society, epistemic communities, and even machine ethics as normative agents.

Legal pluralists advocate for 'soft law' mechanisms—guidelines, declarations, and ethical codes—that, while non-binding, can shape behavior and eventually harden into binding rules (Shelton, 2000). The OECD AI Principles, UNESCO's AI Ethics Declaration, and industry standards developed by IEEE or ISO are examples of this emergent normativity.

By embracing pluralism, international law may remain adaptive in the face of unpredictable technological evolution. However, this adaptability may come at the cost of uniformity and enforcement, raising questions about how to reconcile plural legal orders in global security governance.

3.5. Just War Theory and Ethics in the Age of AI

3.5.1. Jus ad Bellum and Strategic AI Dilemmas

Just War Theory is a normative framework for assessing the morality of war, divided into two main branches: jus ad bellum (justice in entering war) and jus in bello (justice in conduct during war). With the introduction of AI, both branches encounter new complexities.

Under jus ad bellum, the principle of last resort and legitimate authority are central. AI systems—especially those deployed in early warning and pre-emptive strike contexts—risk undermining these criteria. Autonomous systems might act on statistical inferences or adversarial data inputs, launching strikes without clear attribution or human deliberation (Roff & Moyes, 2016). Furthermore, predictive algorithms may incentivize preemptive logic, undermining the principle of necessity. If a system forecasts a 70% likelihood of attack based on real-time data, the legal and moral justification for action becomes probabilistic rather than evidentiary—contradicting the traditional doctrine of actual threat.

3.5.2. Jus in Bello and Meaningful Human Control

 doi:10.20944/preprints202506.2483.v1

Under jus in bello, the principles of proportionality and discrimination are severely tested by AI-enabled weapons. In theory, AI could enhance precision, but in practice, its ability to make context-specific moral judgments remains limited. The notion of 'meaningful human control' has emerged as a compromise principle, suggesting that machines may act autonomously but under sufficient human oversight (Ekelhof, 2019).

However, scholars like Sparrow (2016) argue that the ambiguity of this concept allows states to claim compliance while deploying increasingly autonomous systems. As AI capabilities evolve, the threshold of what constitutes 'meaningful' control may erode, rendering the principle symbolic rather than substantive.

3.5.3. Machine Ethics and Computational Morality

Recent literature explores whether machines can be programmed with ethical frameworks. Machine ethics aims to develop AI systems capable of making morally justifiable decisions, drawing from deontological or utilitarian models (Wallach & Allen, 2009). While promising in theory, practical implementation remains elusive due to:

-Context dependency of moral choices, Lack of common values across cultures, Computational limitations in interpreting intent and nuance, even if computational morality becomes viable, who programs the ethics— a question deeply tied to geopolitical interests and ideological variance.

3.6. Integrative Model: A Hybrid Theoretical Approach

Given the complexity of the AI-nuclear-law nexus, no single theory suffices. This study adopts a hybrid theoretical approach that draws on:

-Determinism and sociotechnical systems to explain the material evolution

-Constructivism to track norm emergence and contestation

-Legal pluralism to recognize multi-source lawmaking

-Just war theory and ethics to preserve normative legitimacy

This integrated framework enables a more holistic understanding of how AI is transforming not only the machinery of war but also the moral, legal, and institutional frameworks intended to regulate it.

3.7. Theoretical Limitations and Future Avenues

While the hybrid model offers breadth, it faces certain limitations:

Normativity vs. Realpolitik: Constructivist and ethical theories often struggle to influence state behavior in high-security environments.

Opacity of AI: Theoretical models rely on empirical transparency, which is often missing in classified or proprietary AI systems.

Legal enforcement: Pluralistic and soft law approaches may lack enforcement, especially against powerful actors.

Future theoretical exploration may benefit from cybernetic theories, post-humanist ethics, and systems jurisprudence, all of which examine non-human agency and complex adaptive systems beyond the state-centric paradigm.

4. Research Methodology

4.1. Introduction

The fusion of artificial intelligence (AI) with nuclear weapons systems represents a technological and legal inflection point that demands robust, interdisciplinary methodologies. This section outlines the philosophical orientation, research design, data collection techniques, and analytical strategies employed in examining the limitations of international law in governing AI-nuclear integration. The objective is not merely to catalog legal gaps but to interrogate the epistemic, institutional, and strategic dynamics that impede legal evolution in this high-risk technological domain.

This methodology section adopts a qualitative, doctrinal, and exploratory research design, drawing from legal hermeneutics, international relations analysis, and technology assessment frameworks.

 doi:10.20944/preprints202506.2483.v1

The complexity of AI in nuclear systems, often shielded by secrecy and dual-use ambiguity, precludes quantitative measurement in many cases. Thus, the research emphasizes interpretation, triangulation, and case-based synthesis to extract patterns, contradictions, and legal-ethical tensions.

4.2. Research Philosophy and Ontological Assumptions

The research is anchored in a constructivist and critical realist ontology, which recognizes that legal and technological realities are co-constructed through discourse, norms, and institutional power (Searle, 1995; Bhaskar, 2008). This ontological approach resists deterministic models and instead focuses on how meaning, agency, and legitimacy are produced in the AI-nuclear-law nexus. The epistemological stance is interpretivist, privileging subjective understanding, historical context, and normative intent over statistical generalization.

4.3. Hypotheses

Guided by the objectives outlined in Section 1, this study explores the following research questions:

1. What legal risks and uncertainties are introduced by integrating AI into nuclear command-and-control systems?

2. How do existing international treaties (e.g., NPT, CCW, Geneva Conventions) address or neglect AI-related nuclear risks?

3. What mechanisms (legal, political, or ethical) can be proposed to fill regulatory and accountability gaps?

While not hypothesis-driven in the quantitative sense, the research presumes that existing legal instruments are insufficiently equipped to regulate AI in nuclear systems and that a multifaceted governance model is needed.

4.4. Legal Doctrinal Methodology

The core of this research relies on doctrinal legal analysis, which involves the systematic examination of legal texts, precedents, and interpretive practices. Primary sources include international treaties (NPT, CCW, TPNW), UN General Assembly resolutions, International Court of Justice (ICJ) opinions, and ICRC commentaries. Secondary sources comprise journal articles, legal monographs, and expert commentary.

The doctrinal method follows four analytical steps:

-Exegesis: Close reading of treaty texts and interpretive documents.

-Comparative Analysis: Cross-national comparisons of treaty implementation.

-Teleological Interpretation: Understanding legal norms in light of their purposes.

-Legal Gap Identification: Systematic mapping of unregulated domains.

This method allows the identification of silent zones, ambiguities, and contradictions within and across treaties.

4.5. Case Study Approach

To contextualize the doctrinal analysis, a multiple-case study method is employed. Case studies include the United States, China, Russia, and Israel—four nuclear-armed states with known AI military investments. Each case examines:

a) National AI and defense policy documents

b) NC3 infrastructure modernization programs

c) Legal review processes for new weapons (e.g., Article 36 protocols)

d) Public and parliamentary debates (where available)

e) Data for each case is triangulated from think tank reports (e.g., SIPRI, CNAS, Brookings), official documents, media investigations, and expert interviews.

4.6. Expert Virtual Interviews and Qualitative Insights

To enhance empirical grounding, semi-structured virtual interviews were conducted with 12 experts, including:

 doi:10.20944/preprints202506.2483.v1

a) International legal scholars b) Arms control negotiators c) Military AI researchers d) Ethicists and technologists

Virtual interview protocols were designed to extract views on legal adequacy, accountability structures, technological risks, and regulatory innovation. Thematic analysis (Braun & Clarke, 2006) was used to code and cluster responses into key categories: transparency, responsibility, human control, treaty reform, and geopolitical dynamics.

4.7. Technology Assessment Framework

Given the technical opacity of AI systems, the study incorporates technology assessment (TA) frameworks to evaluate their military-nuclear applications. Drawing on frameworks developed by OECD, UNESCO, and RAND Corporation, the following dimensions are assessed:

Functionality: What tasks are AI systems performing (e.g., early warning, targeting)?

-Autonomy: Degree of human oversight in operational decisions.

-Transparency: Auditability and interpretability of AI systems.

-Dual-Use Nature: Civilian-military overlap.

Vulnerability: Susceptibility to hacking, spoofing, or bias.

These dimensions help evaluate the fitness of legal categories like 'weapons,' 'command systems,' or 'human operator' in the context of machine learning architectures.

4.8. Normative and Ethical Analysis

Beyond legal text, the study includes normative ethical analysis, especially through just war theory, international human rights law, and machine ethics. This is necessary to address:

-The morality of delegating life-and-death decisions to AI

-Proportionality in algorithmically-determined force

-Due process and dignity under automated warfare

-Philosophical texts, policy papers, and advisory documents (e.g., from the Vatican, UNESCO, UNIDIR) inform this analysis.

4.9. Limitations of the Methodology

While comprehensive, this methodology has limitations:

-Access to classified data: Most nuclear AI systems are state secrets.

-Reliance on secondary sources: Primary fieldwork in nuclear command centers is unfeasible.

-Technological fluidity: AI systems evolve faster than legal research cycles.

-Western-centric data bias: Most sources come from Western think tanks and academia.

-Mitigation strategies include triangulation, regional consultation, and transparency in assumptions.

4.10. Ethical Considerations in Research

-The research adheres to ethical standards for qualitative inquiry:

-Informed Consent: All interviewees were briefed and consented.

-Anonymity: Identities are pseudonymized unless interviewees opted for attribution.

-Sensitivity to National Security: No classified or harmful information was pursued or published.

-Interdisciplinary Ethics: Ethical norms from law, sociology, and computer science were integrated.

4.11 Synthesis of Methodology

This methodology blends doctrinal legal analysis with case studies, expert interviews, and technology assessment to construct a holistic view of AI's disruption of nuclear law. The approach allows for deep textual interpretation, empirical illustration, and normative evaluation—essential for understanding a phenomenon that defies disciplinary boundaries.

As AI and nuclear weapons converge, legal scholars, ethicists, and policymakers must innovate their research methods to keep pace. This section has detailed a transdisciplinary methodology tailored to the evolving threats and challenges of AI-integrated warfare. In the next section, this framework is applied to real-world examples to uncover how international law is responding—or failing to respond—to this new strategic reality.

 doi:10.20944/preprints202506.2483.v1

5:. Case Study Analysis

5.1. Introduction

This section applies the interdisciplinary research framework developed in Section 4 to a set of strategically selected case studies: The United States, China, Russia, and Israel. Each of these states has either acknowledged or is believed to be integrating artificial intelligence (AI) into their nuclear command, control, and communications (NC3) systems. They represent different legal traditions, strategic cultures, and degrees of transparency, offering a comparative vantage point to assess how international law is coping with emerging AI-nuclear threats.

The case study method serves multiple functions. First, it grounds abstract legal and ethical debates in empirical realities. Second, it enables cross-national pattern recognition regarding legal accountability, strategic doctrine, and treaty interpretation. Third, it exposes the diversity of approaches to AI governance within nuclear infrastructures, reflecting the fragmented and pluralistic nature of global arms control regimes.

5.2. United States

5.2.1. Strategic Context

The United States remains the most technologically advanced nuclear power and the global leader in AI research and military applications. The 2018 Department of Defense (DoD) Artificial Intelligence Strategy and subsequent AI-enabled modernization of its NC3 architecture underscore the centrality of AI to future deterrence capabilities (DoD, 2018)

5.2.2. NC3 and AI Integration

According to a RAND Corporation report (Boulanin & Verbruggen, 2017), U.S. NC3 systems are increasingly incorporating machine learning tools for early warning, threat detection, and cyber defense. Project Maven, initially aimed at drone footage analysis, has reportedly evolved into broader applications, including predictive threat assessments.

5.2.3. Legal Review Mechanisms

The United States maintains a structured process for weapons review under Article 36 of Additional Protocol I. However, AI-specific reviews remain classified. The Department of Defense has issued the 3000.09 directive outlining autonomy in weapon systems, emphasizing human control, but its implementation in nuclear contexts is opaque (DoD, 2012).

5.2.4 Normative Discourse and Accountability

U.S. strategic discourse privileges technological dominance and strategic ambiguity. In congressional hearings, officials resist binding AI regulations, preferring ethical guidelines developed in collaboration with industry (CNAS, 2020). This reflects a commitment to soft law rather than formal international legal obligations, echoing realist and exceptionalist traditions in U.S. foreign policy.

5.3. Russia

5.3.1. Strategic Context

Russia views AI as a force multiplier in both conventional and nuclear domains. The 2014 Military Doctrine and 2020 Foundations of State Policy on Nuclear Deterrence mention the role of automated systems in nuclear operations, though specifics are classified.

5.3.2 NC3 and AI Integration

Russian NC3 modernization includes the Perimeter system (colloquially known as 'Dead Hand'), which already incorporates autonomous retaliation capabilities under extreme conditions. Analysts suggest that AI is being integrated to enhance redundancy and survivability (Acton, 2018).

5.3.3. Legal and Institutional Framework

 doi:10.20944/preprints202506.2483.v1

Russia is a signatory to key treaties like the NPT and CCW but has opposed efforts to ban lethal autonomous weapons systems. Its domestic law does not include transparent Article 36 review mechanisms. Instead, weapons development is governed by classified military-industrial directives.

5.3.4. Normative Resistance and Disinformation

Russia frequently employs strategic disinformation and information warfare, complicating legal transparency. It has used international forums to deflect regulatory efforts, often framing them as Western attempts to stifle technological parity (UNODA, 2021). This reflects a strategic posture of legal minimalism and norm resistance.

5.4. China

5.4.1. Strategic Context

China's rapid militarization of AI is visible in its dual-use policy framework. The State Council's 2017 AI Development Plan sets goals for military-civil fusion, integrating AI into defense and nuclear systems (Kania, 2019).

5.4.2. AI-Nuclear Integration

Chinese scholars suggest that AI is being deployed in early-warning systems, satellite image processing, and possibly nuclear submarine communications. However, ambiguity remains over whether China has adopted an autonomous or semi-autonomous posture for nuclear launch decisions (Lewis & Xue, 2020).

5.4.3. Legal Doctrine and Treaty Interpretation

China supports some multilateral disarmament efforts but emphasizes national sovereignty. It rejects pre-emptive bans on autonomous weapons and has not committed to AI-specific legal regimes. Legal scholars in China often adopt a defensive posture, framing international law as a vehicle for hegemonic containment.

5.4.4. Epistemic and Technocratic Governance

China's governance approach relies heavily on technocratic management rather than legal oversight. Ethics guidelines issued by the Ministry of Science and Technology highlight AI safety and controllability, but lack binding enforcement (Zeng et al., 2020).

5.5. Israel

5.5.1. Strategic Ambiguity and Technological Innovation

Israel has not officially confirmed its nuclear arsenal but is widely acknowledged to possess it. Its defense sector, particularly in cyber and AI domains, is highly advanced, often in partnership with private firms and academic labs (Sachs, 2019).

5.5.2. NC3 and Cyber-AI Overlap

Given its strategic reliance on cyber capabilities, Israeli NC3 systems are believed to integrate AI for cyber defense and automated decision-support. The Iron Dome's targeting system, while not nuclear, exemplifies high levels of algorithmic control and provides a model for future nuclear applications.

5.5.3. Legal Framework and Secrecy

Israel has not ratified the NPT and is not a party to the TPNW or CCW. Its legal framework for weapons review is opaque, and national security justifications often override transparency. However, internal ethical debates, particularly within academia and civil society, have pressured the government to address ethical use of AI.

5.5.4. Civil-Military Fusion and Dual-Use Dynamics

 doi:10.20944/preprints202506.2483.v1

Like China, Israel exhibits strong civil-military integration. Many AI innovations emerge from private firms that later receive military contracts. This complicates accountability and blurs the lines between civilian and military AI, making regulatory oversight difficult.

5.6 Comparative Insights and Cross-Case Patterns

5.6.1. Legal Ambiguity as Strategic Asset

All four states maintain varying degrees of legal ambiguity regarding AI integration in nuclear systems. This ambiguity allows operational flexibility but weakens international legal normativity. Legal pluralism exists, but without cohesive enforcement mechanisms.

5.6.2. Soft Law and Ethical Guidelines

Each state, with the exception of Russia, has engaged in developing ethical or soft-law principles. However, these are not substitutes for binding treaties. They function more as reputation management tools than accountability instruments.

5.6.3. State Sovereignty and Treaty Resistance

There is a broad resistance to binding international norms on AI from major powers. Sovereignty, technological advantage, and strategic uncertainty serve as common justifications for rejecting multilateral regulation.

5.6.4. Epistemic Communities and Civil Society

Despite governmental opacity, epistemic communities—academics, think tanks, and advocacy networks—play crucial roles in norm advocacy and public awareness. In democracies like the U.S. and Israel, these communities have influenced public debate. In authoritarian settings, their influence is limited but growing.

5.7. Summary of Case Study Findings

The U.S. promotes soft law, maintains doctrinal reviews, and integrates AI cautiously.
Russia favors strategic opacity and deterrence through ambiguity.

China emphasizes technocratic control and dual-use integration.

Israel blends secrecy with innovation, guided by an entrepreneurial-military complex.

These case studies demonstrate the inadequacy of existing international legal frameworks to effectively regulate AI in nuclear command systems. The diversity of national practices, absence of transparency, and legal resistance from dominant states suggest a crisis of norm enforcement.

5.8. Iran

5.8.1. Strategic Context

Iran occupies a contentious position in global nuclear politics. Although it is a signatory to the Treaty on the Non-Proliferation of Nuclear Weapons (NPT), its nuclear ambitions and enrichment programs have triggered international sanctions and diplomatic confrontations. Iran officially denies the pursuit of nuclear weapons, yet Western intelligence and international watchdogs remain skeptical of its long-term intentions (IAEA, 2022).

5.8.2. AI Integration in Military and Strategic Domains

Iran has shown increasing interest in AI development for military use, particularly in unmanned aerial vehicles (UAVs), cyber warfare, and surveillance systems (Cordesman, 2021). While there is no conclusive evidence that Iran has integrated AI into nuclear command-and-control (NC3) structures, its development of autonomous and semi-autonomous platforms indicates a trajectory toward AI-enabled defense systems. This trajectory, combined with Iran's opaque governance model and secretive nuclear infrastructure, raises concerns about future AI-nuclear convergence.

5.8.3. Legal and Normative Frameworks

 doi:10.20944/preprints202506.2483.v1

Iran asserts its compliance with international law through its NPT obligations and cooperation with the International Atomic Energy Agency (IAEA). However, its domestic legal system lacks transparency in weapons development, and there are no publicly known legal review processes for new AI military technologies. Iran has not ratified or endorsed efforts to ban lethal autonomous weapons, nor does it participate actively in AI-focused disarmament forums.

5.8.4. Geopolitical Isolation and Legal Exceptionalism

Due to sanctions and political isolation, Iran has adopted a defensive legal posture, often framing international legal norms as tools of Western coercion. This contributes to a form of legal exceptionalism in which international obligations are interpreted through national security imperatives. As a result, Iran operates in a dual environment of formal legal adherence and practical legal resistance (Esfandiary & Fitzpatrick, 2016).

5.8.5. Risk Amplification through AI and Strategic Ambiguity

Iran's deployment of AI-enabled cyber tools—such as Stuxnet-like retaliatory capabilities—and potential dual-use AI research elevates the risk of strategic miscalculation. In the absence of confidence-building measures or transparency protocols, the integration of AI into Iran's military apparatus could be misinterpreted by adversaries, especially Israel and the United States, as steps toward autonomous nuclear escalation.

5.8.6. Civil-Military Research Nexus

Iranian universities and military-linked research centers collaborate on AI development. This civil-military fusion raises challenges for export controls and verification regimes. Without robust oversight, dual-use AI technologies may blur the line between civilian research and military applications, complicating global non-proliferation efforts.

Iran represents a complex test case for international law in the AI-nuclear domain. While formally compliant with certain treaties, its lack of transparency, isolation from global AI governance forums, and ambiguous military research trajectory make it a potential hotspot for future legal and strategic crises.

6. Discussion and Policy Recommendations

6.1. Introduction

Building on the empirical insights from Section 5, this section presents a critical discussion on the implications of integrating artificial intelligence (AI) into nuclear command and control (NC3) systems. It synthesizes findings from the doctrinal, ethical, and geopolitical dimensions explored in previous sections and proposes actionable policy recommendations to address the legal and normative gaps that currently exist. The discussion is guided by the central thesis of this study: that international law, as it stands, is ill-equipped to govern the complex, opaque, and rapidly evolving domain of AI-nuclear integration.

6.2. The Legal Vacuum and Strategic Risk Multiplication

AI introduces new dynamics into nuclear strategy, including automation bias, opacity in decision-making, and vulnerability to cyber interference. Existing legal regimes—such as the Treaty on the Non-Proliferation of Nuclear Weapons (NPT), the Geneva Conventions, and the Convention on Certain Conventional Weapons (CCW)—do not adequately regulate these dynamics. While Article 36 of Additional Protocol I provides a basis for legal review of new weapons, it is limited by state discretion, lacks transparency, and does not address AI's unique challenges such as algorithmic unpredictability or black-box decision-making (Schmitt, 2013).

The case studies underscore this legal vacuum. All examined states—especially the nuclear-armed powers—resist binding regulations on AI integration in military systems, preferring national sovereignty, strategic ambiguity, or soft law approaches. This resistance contributes to what can be termed strategic risk multiplication, wherein technological advancement outpaces legal and

 doi:10.20944/preprints202506.2483.v1

normative safeguards, exacerbating the likelihood of miscalculation, escalation, and legal impunity.

6.3 Ethical Challenges and Normative Ambiguities

The ethical concerns raised by AI in NC3 systems are manifold. Delegating life-and-death decisions to machines raises profound moral questions about human dignity, accountability, and the just conduct of war. The principle of 'meaningful human control'—while increasingly popular—remains vaguely defined and inconsistently applied across jurisdictions (Ekelhof, 2019).

From a just war theory perspective, AI complicates both jus ad bellum and jus in bello. Predictive algorithms may incentivize preemptive strikes, undermining the principle of necessity. In combat scenarios, AI's inability to interpret context or intent jeopardizes the principles of distinction and proportionality. These ethical challenges demand more than technical fixes; they require sustained normative engagement and inclusive global discourse.

6.4. Fragmentation of Governance Regimes

The current governance landscape is marked by fragmentation. Treaties like the NPT focus on state-to-state nuclear relations, while newer frameworks such as the OECD AI Principles or UNESCO's AI ethics guidelines target civilian AI development. There is no comprehensive regime that addresses the intersection of AI and nuclear weapons.

This fragmentation is not accidental but reflective of divergent geopolitical interests. The United States and its allies promote ethical AI frameworks but resist binding restrictions. Russia and China emphasize strategic stability but oppose Western-dominated norm-setting. Emerging powers like Iran adopt a defensive posture, citing sovereignty and anti-hegemonic principles.

The result is a patchwork of overlapping but non-binding instruments, none of which provide enforceable accountability mechanisms for AI-related violations in nuclear contexts.

6.5. Comparative Legal Analysis: Converging Gaps

Despite their strategic differences, the case studies reveal converging legal and institutional gaps:

Absence of Transparent Article 36 Review: None of the nuclear states maintain a transparent, independent, and public-facing legal review process for AI-based military systems.

Weak Normative Commitment to Human Control: Even where guidelines exist (e.g., DoD Directive 3000.09 in the U.S.), the principle of human oversight is narrowly defined and subject to operational exceptions.

Soft Law Preference: Ethical principles are often invoked in place of binding obligations, creating legal ambiguity.

Lack of Multilateral Engagement: Major powers resist comprehensive disarmament or regulatory frameworks that could constrain their technological edge.

6.6. Policy Recommendations

6.6.1. Establish an AI-Nuclear Governance Framework

A dedicated international framework is needed to govern AI applications in nuclear systems. This framework should:

Be anchored in the UN Charter and international humanitarian law.

Include mechanisms for verification, reporting, and compliance.

Provide a platform for multistakeholder engagement, including states, civil society, and the private sector.

6.6.2. Strengthen Article 36 Review Mechanisms

States should institutionalize independent, transparent legal review bodies for all new weapons systems involving AI. International bodies such as the ICRC or the UN Institute for Disarmament Research (UNIDIR) could provide oversight or technical assistance.

6.6.3. Clarify and Operationalize 'Meaningful Human Control'

 doi:10.20944/preprints202506.2483.v1

A normative consensus is needed on what constitutes meaningful human control. This includes: Setting minimum standards for human oversight in decision-making loops.

Prohibiting fully autonomous nuclear launch systems.

Integrating human rights principles into military AI design.

6.6.4. Promote Dual-Use Research Oversight

Given the civil-military overlap in AI development, national governments should strengthen export controls and dual-use oversight mechanisms. Academic institutions and private tech companies should be held accountable through ethical review boards and compliance audits.

6.6.5. Expand Multilateral Dialogue and Confidence-Building Measures

Track-II diplomacy and informal dialogues between nuclear states should be expanded to address AI risks. Confidence-building measures could include:

Joint exercises with human-in-the-loop protocols.

Data-sharing on AI-related incidents or near misses.

Multilateral declarations on the non-deployment of fully autonomous nuclear systems.

6.6.6. Support Norm Entrepreneurship by Non-State Actors

Civil society organizations, epistemic communities, and international NGOs play a critical role in norm diffusion. Funding, recognition, and participatory platforms should be provided to these actors to shape discourse and pressure reluctant states.

6.7. Anticipating and Mitigating Future Risks

AI technologies evolve rapidly, often beyond the predictive capacity of current regulatory models. Foresight tools such as scenario planning, red-teaming, and technology impact assessments should be embedded within national and international governance architectures.

Moreover, AI's convergence with other emerging technologies—quantum computing, synthetic biology, and advanced cyberwarfare—necessitates an integrated approach to arms control. Siloed regulation will fail to capture the systemic risks posed by cross-domain technological fusion.

The integration of AI into nuclear weapons systems represents a critical juncture for international law, ethics, and security policy. The current legal architecture—fragmented, state-centric, and ethically underdeveloped—is inadequate for the challenges ahead. This section has outlined key discussion points and policy pathways that could realign the governance of AI-nuclear systems with principles of international law, ethical responsibility, and global security.

Implementing these recommendations will require political will, epistemic humility, and normative innovation. The stakes—human survival and the integrity of the international order—could not be higher.

7. Conclusions

The preceding sections have critically examined the convergence of artificial intelligence (AI) and nuclear weapons governance through interdisciplinary lenses—legal, strategic, ethical, and technological. Drawing upon doctrinal analyses, empirical case studies, expert interviews, and normative evaluations, this research has demonstrated that the integration of AI into nuclear command, control, and communication (NC3) systems poses urgent and unprecedented challenges to international legal regimes. This concluding section synthesizes the key findings of the study, evaluates their implications for global governance, and outlines strategic pathways for future research, advocacy, and policymaking.

International law, as codified in the Non-Proliferation Treaty (NPT), the Geneva Conventions, and the Convention on Certain Conventional Weapons (CCW), was not designed to address the algorithmic complexity, dual-use ambiguity, and speed of technological evolution introduced by AI. Article 36 of Additional Protocol I, though theoretically relevant, is inconsistently applied and often opaque in nuclear-armed states. This research identifies a systemic legal lag, whereby the evolution

 doi:10.20944/preprints202506.2483.v1

of military AI systems far outpaces regulatory mechanisms, rendering existing laws functionally obsolete in critical domains.

As the case studies of the United States, Russia, China, Israel, and Iran show, the leading states in AI-nuclear integration maintain deliberate strategic ambiguity. This opacity is often justified through appeals to national security, technological sovereignty, and deterrence stability. However, it also reflects institutional resistance to international accountability. States prefer soft-law measures, ethical codes, and non-binding principles over legally enforceable treaties, thereby eroding the normative strength of international humanitarian law and arms control regimes.

AI in nuclear systems introduces profound ethical dilemmas: Should machines be allowed to make life-and-death decisions? Can algorithmic warfare conform to the principles of distinction and proportionality? What happens when states deny accountability by blaming machine failure? This research concludes that the prevailing moral frameworks—just war theory, Kantian ethics, and human rights law—struggle to address the black-box nature and predictive orientation of AI systems. The delegation of moral responsibility to machines or opaque algorithms challenges the very foundation of ethical warfare.

The governance landscape is fragmented. Military treaties, ethical AI guidelines, export control regimes, and human rights frameworks all operate in silos. Moreover, geopolitical tensions—especially between the U.S., China, and Russia—impede multilateral regulatory efforts. Non-alignment and legal pluralism exacerbate normative inconsistency, leaving large swaths of AI-nuclear development unregulated or self-regulated. Iran and Israel exemplify how legal exceptionalism—through either disobedience or non-participation—creates further friction in collective governance.

Despite state-level resistance, epistemic communities, NGOs, and academic institutions have emerged as critical actors in shaping norms, proposing treaty frameworks, and conducting risk assessments. Track II diplomacy, people-to-people exchanges, and ethics coalitions provide fertile grounds for policy innovation, though they lack coercive power. The gap between norm entrepreneurship and enforcement mechanisms remains wide.

This research advocates a shift from state-centric legal positivism to legal constructivism, wherein law is understood not just as binding rules but as evolving discourses shaped by power, norms, and institutions. Such an approach is essential for dealing with the fluidity of AI technologies and the polycentric nature of their development. Traditional legal categories such as 'weapon,' 'combatant,' 'command,' and 'control' must be redefined. AI challenges these boundaries through its ability to simulate cognition, alter decision-making hierarchies, and obscure human intent. International law must embrace adaptive jurisprudence, capable of interpreting emerging phenomena like autonomous agents, human-machine teams, and cyber-kinetic hybrids. Doctrinal legal analysis should be complemented by structured technology assessments. Legal scholars, computer scientists, and ethicists must collaborate to evaluate new AI systems for legality, ethical soundness, and strategic risk. This interdisciplinary synthesis is no longer optional but essential.

The international community must prioritize the development of a binding treaty framework addressing AI in NC3 systems. Such a treaty should prohibit full autonomy in nuclear launch decisions, require transparent human-in-the-loop structures, and institutionalize independent legal reviews. The framework should also mandate periodic reporting and peer-review mechanisms.

Existing institutions like the International Atomic Energy Agency (IAEA), United Nations Institute for Disarmament Research (UNIDIR), and the International Committee of the Red Cross (ICRC) must be reformed and empowered to monitor AI integration in military systems. Legal innovation should focus on anticipatory governance, risk modeling, and scenario-based regulation.

Diplomatic forums such as the UN General Assembly, the Conference on Disarmament, and the G20 should be leveraged to institutionalize AI risk diplomacy. This includes red lines, transparency protocols, and early warning agreements. States should commit to annual AI-risk disclosures and participate in cooperative threat assessments.

AI governance should not be left to states alone. Democratic inclusion of civil society, indigenous voices, and marginalized communities is vital for building legitimacy and foresight. Participatory

 doi:10.20944/preprints202506.2483.v1

regulation and public interest technologies must shape the norms of AI-nuclear integration.
This study opens several avenues for future research:

-Comparative legal analyses between regional organizations (e.g., EU vs. ASEAN) on AI-military policy.

-Deep ethnographies of AI weapon development teams to understand ethical decision-making cultures.

-Simulation-based modeling of AI-induced nuclear escalation scenarios.

-Exploration of post-human legal theory and non-anthropocentric accountability models.

7.2. Final Reflections

We are at a liminal moment in human history. The convergence of AI and nuclear weapons—two of the most powerful and unpredictable forces—demands a collective moral awakening, legal transformation, and political commitment. This research has shown that while the risks are monumental, so too are the opportunities for innovation, cooperation, and the reinvention of global governance.

A legal system that cannot govern the future is one that enables catastrophe. It is imperative that scholars, policymakers, technologists, and citizens act now—not after the fact. Law must rediscover its anticipatory function. Ethics must reclaim its place in statecraft. And humanity must decide whether it is willing to entrust survival to code.

References

1. Acton, J. M. (2018). Escalation through entanglement: How the vulnerability of command-and-control systems raises the risks of an inadvertent nuclear war. *International Security*, 43(1), 56–99.
2. Abbott, K., & Snidal, D. (2000). Hard and soft law in international governance. International Organization, 54(3), 421–456.
3. Asaro, P. (2012). On banning autonomous weapon systems: Human rights, automation, and the dehumanization of lethal decision-making. *International Review of the Red Cross*, 94(886), 687–709.
4. Boulanin, V., & Verbruggen, M. (2017). Mapping the development of autonomy in weapon systems. SIPRI.
5. Bhaskar, R. (2008). *A realist theory of science*. Routledge.
6. Boulanin, V., & Verbruggen, M. (2017). Mapping the development of autonomy in weapon systems. Stockholm International Peace Research Institute.
7. Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101.
8. Brundage, M., Avin, S., Clark, J., Toner, H., Eckersley, P., Garfinkel, B., ... & Amodei, D. (2018). The malicious use of artificial intelligence: Forecasting, prevention, and mitigation. Future of Humanity Institute.
9. Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society*, 3(1), 1–12.
10. Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. Big Data & Society, 3(1), 1–12.
11. Cordesman, A. H. (2021). Iran's Military Forces and Warfighting Capabilities: The Threat in the Northern Gulf. CSIS.
12. Crootof, R. (2015). The Killer Robots Are Here: Legal and Policy Implications. Cardozo Law Review, 36(4), 1837–1915.
13. Department of Defense. (2012). Directive 3000.09: Autonomy in Weapon Systems. U.S. DoD.
14. Department of Defense. (2018). Summary of the 2018 Department of Defense Artificial Intelligence Strategy.
15. Docherty, B. (2019). Stopping killer robots: Country positions on banning fully autonomous weapons and retaining human control. Human Rights Watch.
16. Ekelhof, M. (2019). Moving beyond semantics on autonomous weapons: Meaningful human control in operation. Global Policy, 10(3), 343–348.
17. Esfandiary, D., & Fitzpatrick, M. (2016). Iran's Nuclear Program and International Law: From Confrontation to Accord. IISS.

 doi:10.20944/preprints202506.2483.v1

18. Finnemore, M., & Sikkink, K. (1998). International norm dynamics and political change. International Organization, 52(4), 887–917.
19. Hart, H. L. A. (1961). The concept of law. Oxford University Press.
20. Horowitz, M. C. (2019). Artificial Intelligence and the Future of Warfare. Belfer Center for Science and International Affairs.
21. Horowitz, M. C., Kahn, L., & Scharre, P. (2020). Artificial intelligence and international stability: Risks and confidence-building measures. Center for a New American Security.
22. International Atomic Energy Agency. (2022). Verification and Monitoring in the Islamic Republic of Iran in Light of United Nations Security Council Resolution 2231 (2015). IAEA.
23. International Committee of the Red Cross. (2019). Autonomous weapon systems: Implications of increasing autonomy in the critical functions of weapons. Geneva: ICRC.
24. Kania, E. B. (2019). Battlefield Singularity: Artificial Intelligence, Military Revolution, and China's Future Military Power. Center for a New American Security.
25. Kristensen, H. M., & Korda, M. (2023). Status of World Nuclear Forces. Federation of American Scientists.
26. Pagallo, U. (2013). The laws of robots: Crimes, contracts, and torts. Springer.
27. Lewis, J., & Xue, L. (2020). China's strategic posture in the age of AI. *Science and Global Security*.
28. MacKenzie, D., & Wajcman, J. (1999). The social shaping of technology (2nd ed.). Open University Press.
29. Merry, S. E. (1988). *Legal pluralism. Law & Society Review*, 22(5), 869–896.
30. OECD. (2019). OECD Principles on Artificial Intelligence. Organisation for Economic Co-operation and Development.
31. Roff, H. M., & Moyes, R. (2016). Meaningful human control, artificial intelligence and autonomous weapons. Article 36 & International Committee for Robot Arms Control.
32. Sachs, N. (2019). Cybersecurity and Israel's national security strategy. Israel Journal of Foreign Affairs, 13(3), 231–244.
33. Scharre, P. (2018). Army of None: Autonomous Weapons and the Future of War. W. W. Norton & Company.
34. Schmitt, M. N. (2013). Autonomous weapon systems and international humanitarian law: A reply to the critics. *Harvard National Security Journal*, 4(1), 1–45.
35. Searle, J. R. (1995). *The construction of social reality*. Simon and Schuster.
36. Shelton, D. (2000). Commitment and compliance: The role of non-binding norms in the international legal system. Oxford University Press.
37. Singer, P. W., & Brooking, E. T. (2018). LikeWar: The weaponization of social media. Houghton Mifflin Harcourt.
38. Sparrow, R. (2007). Killer robots. *Journal of Applied Philosophy*, 24(1), 62–77.
39. Sparrow, R. (2016). Robotic weapons and the future of war. In Ethics and Emerging Technologies (pp. 311–323). Palgrave Macmillan.
40. UNESCO. (2021). Recommendation on the Ethics of Artificial Intelligence. United Nations Educational, Scientific and Cultural Organization.
41. UNIDIR. (2020). The weaponization of increasing autonomy in future warfare. United Nations Institute for Disarmament Research.
42. UNIDIR. (2022). Artificial Intelligence and the Future of Deterrence: A Survey of Issues. United Nations Institute for Disarmament Research.
43. Zeng, Y., Lu, E., & Huangfu, C. (2020). Linking AI principles. Nature Machine Intelligence, 2(1),
44. Vincent, R. J., & Packer, J. (2020). State responsibility and autonomous weapon systems. *Journal of Conflict and Security Law*, 25(3), 357–385.
45. Wallach, W., & Allen, C. (2009). *Moral machines: Teaching robots right from wrong*. Oxford University Press.
46. Wendt, A. (1999). *Social theory of international politics*. Cambridge University Press.
47. Winner, L. (1986). The whale and the reactor: A search for limits in an age of high technology. University of Chicago Press.
48. Work, R., & Brimley, S. (2014). *20YY: Preparing for war in the robotic age*. Center for a New American Security.

 doi:10.20944/preprints202506.2483.v1