www.ijprems.com
editor@ijprems.com

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
(Int Peer Reviewed Journal)
Vol. 05, Issue 09, September 2025, pp : 804-813

e-ISSN : 2583-1062
Impact Factor : 7.001

# INTEGRATING ARTIFICIAL INTELIGANCE INTO NATIONAL SECURITY: RISKS< BENEFITS, AND POLICY DILEMMAS

**Vephkhvia Grigalashvili**[1]

[1]Associate Professor, International Black Sea University, Tbilisi, Georgia.

E-Mail: vgrigalashvili@ibsu.edu.ge

DOI: https://www.doi.org/10.58257/IJPREMS43881

## ABSTRACT

The integration of Artificial Intelligence (AI) into national security architectures represents a transformative development with far-reaching implications for defense, intelligence, and policy-making. This article critically examines the dual-edged nature of AI in national security, exploring the technological benefits alongside the strategic, ethical, and political dilemmas it introduces. On one hand, AI enhances capabilities in surveillance, cybersecurity, threat detection, decision-making, and autonomous systems, thereby enabling faster responses to complex and dynamic security threats. It facilitates predictive analytics in counterterrorism, improves logistics in military operations, and strengthens border control mechanisms. On the other hand, the adoption of AI brings significant risks: algorithmic bias, loss of human oversight, escalation of autonomous warfare, cyber vulnerabilities, and challenges to democratic accountability. Moreover, the rapid pace of AI innovation outstrips current regulatory and governance frameworks, creating a policy vacuum that may be exploited by both state and non-state actors. This article analyzes key policy dilemmas arising from AI militarization, such as arms race dynamics, civil-military boundaries, and international humanitarian law compliance. It also highlights the uneven global AI capacity, raising concerns about strategic imbalance and normative fragmentation. Drawing on a multidisciplinary perspective, the article calls for a comprehensive and ethically grounded policy framework that balances innovation with accountability, ensuring AI's alignment with democratic values and international security norms. The study concludes by offering recommendations for transparent governance, multilateral cooperation, and adaptive regulation to responsibly integrate AI into national security strategies while mitigating associated risks.

**Keywords:** Artificial Intelligence, National Security, Autonomous Systems, Ethical Risks, AI Governance, Defense Technology.

## 1. INTRODUCTION

The accelerating integration of artificial intelligence (AI) into national security architectures has generated both optimism and alarm. On the one hand, AI promises unprecedented operational advantages—enhanced threat detection, autonomous decision-making, and predictive analytics capable of outpacing human cognition (Khare & Sinha, 2024; Masakowski, 2020). On the other hand, this transformation is unfolding amid significant ethical, legal, and strategic uncertainties (Montasari, 2022; Voeneky et al., 2022). Unlike earlier technological shifts, AI introduces dynamic systems that adapt in real time, operate across cyber-physical domains, and, crucially, challenge long-standing norms of human control, accountability, and warfare governance (Lahmann & Geiss, 2024; Taddeo et al., 2022).

The urgency to assess these developments is compounded by global power competition, where states increasingly view AI capabilities as essential to both deterrence and dominance (Roumate, 2024; Greene, 2023). Yet, national strategies remain uneven, and multilateral governance mechanisms are still nascent or fragmented (Cristiano et al., 2023; Meleouni & Efthymiou, 2024). This article critically examines the benefits, risks, and policy dilemmas arising from AI's integration into national security. It aims to clarify conceptual boundaries, evaluate the empirical landscape, and explore the normative implications of deploying AI in security contexts.

By drawing on interdisciplinary literature published between 2017 and 2025, this study offers a comprehensive framework for understanding AI's transformative impact on defense and intelligence domains. In doing so, it seeks not merely to map technological change, but to assess how such change might be governed responsibly in an increasingly contested international order.

## 2. CONCEPTUAL FRAMEWORK AND LITERATURE REVIEW

Artificial intelligence (AI) broadly denotes computational systems capable of replicating human cognitive functions such as learning, reasoning, perception, and autonomous decision-making (Montasari, 2022). Within national security, AI transcends theoretical algorithms, materialising in operational domains such as intelligence analysis, autonomous weapons, cyber defence, and surveillance (Khare & Sinha, 2024; Yampolskiy, 2018). These applications engage with complex, often adversarial environments, where rapid adaptation and high-stakes decision-making are essential. Thus,

AI's integration necessitates a reconceptualisation of security paradigms that traditionally hinge on human control and accountability.

Lahmann and Geiss (2024) underscore the transformative nature of AI in security, noting its compression of decision-making cycles and the resultant attenuation of human oversight. Roy (2024) highlights that autonomous systems increasingly operate with degrees of independence that challenge conventional doctrines of state sovereignty and international humanitarian law, raising profound ethical and legal concerns. Voeneky et al. (2022) further stress the dual-use dilemma, where civilian AI innovation directly fuels military capabilities, blurring boundaries and complicating governance frameworks. This hybridity exacerbates proliferation risks and regulatory gaps (Cristiano et al., 2023).

The expanding literature reveals a nuanced recognition of AI as both an enabler of strategic advantage and a source of novel vulnerabilities. Montasari (2023) and Sfetcu (2024) provide in-depth analyses of AI's role in intelligence agencies, particularly in digital forensics and threat detection, while cautioning against algorithmic biases and adversarial exploitation that may distort operational judgements (Erendor, 2024). Wilner and Atkinson (2023) observe that AI's anticipatory intelligence capabilities offer significant promise but introduce epistemic uncertainties due to opaque "black box" algorithms, complicating command decisions.

Comparatively, national strategies diverge markedly. NATO prioritises "responsible AI" principles centred on transparency, human-in-the-loop control, and ethical compliance (NATO Communications and Information Agency, 2022). Contrastingly, non-Western powers often emphasise rapid AI deployment to enhance deterrence, sometimes at the expense of accountability and normative safeguards (Haney, 2020; Roumate, 2024). This strategic divergence, as De Spiegeleire et al. (2017) argue, risks destabilising global security through uneven technological diffusion and erosion of mutual trust.

Ethical and legal dimensions constitute a substantial focus. Taddeo et al. (2022) and Voeneky et al. (2022) call for robust normative frameworks to maintain human oversight and prevent autonomous systems from making lethal decisions independently. Johnson (2023) critiques the lag between AI development and legal regimes, warning of governance vacuums that undermine state responsibility. Reinhold and Schoernig (2022) frame AI as a "Janus-faced" technology, simultaneously enhancing security and exacerbating risks of arms races and inadvertent conflict.

Cybersecurity literature underscores AI's dual nature: while AI enhances cyber defence through adaptive threat detection (Ventre, 2020), it also introduces vulnerabilities exploitable by adversaries (Roumate, 2024). The weaponisation of AI in information warfare—encompassing disinformation and psychological operations—adds destabilising layers to security environments (Cristiano et al., 2023; Lemieux, 2024).

Surveillance and intelligence augmentation via AI raise civil liberties concerns. Lemieux (2024) and Xu et al. (2023) critically assess how AI-driven surveillance intensifies privacy intrusions and political tensions, revealing the inseparability of technology from governance values.

Strategically, AI emerges as a resource reshaping global power dynamics. Roumate (2021) and Santos Nanni (2024) compare AI's strategic importance to nuclear weapons, enabling states to consolidate technological and geopolitical advantage. This realignment risks entrenching techno-authoritarian models, particularly without democratic safeguards (Girasa, 2020; Fatima Roumate, 2024). The Cambridge Handbook of Responsible Artificial Intelligence (Voeneky et al., 2022) advocates interdisciplinary approaches to navigate these complexities.

National policy responses to AI integration exhibit significant variation. NATO exemplifies a principled approach emphasising ethical standards and alliance interoperability (NATO Communications and Information Agency, 2022), while other actors prioritise military efficacy and strategic competition (Haney, 2020). This fragmentation challenges international norm-setting efforts and underscores the need for flexible governance frameworks accommodating diverse strategic cultures.

Methodologically, research spans technical performance evaluations (Montasari, 2023; Sfetcu, 2024), normative ethics (Taddeo et al., 2022; Johnson, 2023), and strategic foresight (Wilner & Atkinson, 2023; De Spiegeleire et al., 2017). The convergence of these approaches is vital for a holistic understanding of AI's implications in national security, facilitating balanced assessments that integrate empirical data, normative considerations, and geopolitical context.

Scholars also emphasise the growing importance of multilateral cooperation. Meleouni and Efthymiou (2024) warn that unilateral AI arms races exacerbate insecurity, advocating for international frameworks promoting transparency and risk reduction. Lemieux (2024) highlights evolving intelligence-sharing challenges posed by AI-enabled threats, underscoring the tension between collaboration and sovereignty.

Civil society and academia's role in AI governance is increasingly recognised. Jaber (2023) and Fatima Roumate (2024) advocate inclusive governance models integrating diverse stakeholders to enhance accountability and

www.ijprems.com
editor@ijprems.com

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
(Int Peer Reviewed Journal)
Vol. 05, Issue 09, September 2025, pp : 804-813

e-ISSN : 2583-1062
Impact Factor : 7.001

democratic oversight, especially in authoritarian contexts. Such approaches align with the "responsible innovation" paradigm prioritising ethical reflection throughout AI's lifecycle (Voeneky et al., 2022).

Technological challenges such as AI explainability and robustness also dominate discourse. Khare and Sinha (2024) and Wilner and Atkinson (2023) stress that opaque "black box" models pose accountability risks in security applications, reinforcing calls for explainable AI systems tailored to operational needs

## 3. POTENCIAL BENEFITS OF AI INTEGRATION

The integration of artificial intelligence (AI) into national security frameworks presents transformative opportunities that can significantly enhance state defence and intelligence capabilities. At its core, AI's ability to process vast and complex datasets surpasses human analytical capacities, enabling faster and more accurate threat detection and situational awareness. This capability is especially critical given the volume and velocity of information generated in modern conflict and intelligence environments. Machine learning algorithms, for instance, can identify patterns and anomalies within heterogeneous data streams—from satellite imagery to electronic communications—that would otherwise remain obscured (Montasari, 2023; Sfetcu, 2024). This improved analytic capacity supports pre-emptive actions and more effective threat mitigation, vital in an era marked by rapid technological advancement and multifaceted hybrid threats (Khare & Sinha, 2024; Ventre, 2020).

Beyond enhancing data processing, AI functions as a force multiplier in decision-making processes at strategic and operational levels. Modern defence environments are characterised by compressed timeframes and complex variables, challenging human operators' cognitive limits. AI-driven decision-support systems assist by simulating multiple scenarios, forecasting outcomes, and optimising resource allocation (Roy, 2024). These tools enable policymakers and commanders to consider a wider array of contingencies, leading to more informed and timely decisions (Johnson, 2023). Far from supplanting human judgement, AI acts as an amplifier, reducing cognitive overload and enhancing the precision of strategic planning (Khare & Sinha, 2024; Montasari, 2022).

Operationally, AI contributes to enhanced efficiency and effectiveness through automation. Autonomous platforms such as unmanned aerial vehicles and robotic systems extend surveillance reach while reducing personnel risk in hazardous environments (Sfetcu, 2024). AI also optimises logistical functions by managing supply chains and maintenance schedules, thus bolstering military readiness and resilience (Khare & Sinha, 2024). In cyber defence, AI enables continuous monitoring and rapid response to sophisticated cyber threats, which are increasingly central to national security challenges (Ventre, 2020; Erendor, 2024). These applications significantly shorten reaction times and mitigate damage from cyberattacks.

Intelligence capabilities are further augmented by AI's capacity to automate routine analytical tasks, freeing human analysts to focus on strategic interpretation and contextual understanding (Montasari, 2023; Sfetcu, 2024). The integration of diverse intelligence sources into coherent assessments is facilitated by AI's ability to synthesise large datasets, enhancing the detection and disruption of covert networks and multifaceted threats such as terrorism and cyber espionage (Lemieux, 2024; Greene, 2023). This integration accelerates the intelligence cycle, ensuring more timely and actionable insights.

Another vital benefit of AI integration is its contribution to cybersecurity and systemic resilience. AI's capacity for adaptive threat detection and autonomous response allows defence systems to anticipate and counter cyber intrusions proactively (Ventre, 2020). The automation of recovery protocols following cyber incidents reduces operational disruptions and reinforces the security of critical infrastructure, which increasingly underpins national stability (Khare & Sinha, 2024; Erendor, 2024). Given the growing complexity and scale of cyber threats, AI's role in fortifying cyber defence represents a fundamental enhancement to national security.

Moreover, AI enhances strategic foresight by enabling sophisticated predictive analytics and simulation of geopolitical scenarios. This foresight assists policymakers in identifying emerging risks and adapting defence strategies accordingly (Wilner & Atkinson, 2023; Montasari, 2023). AI applications also support arms control and treaty verification through enhanced monitoring capabilities, fostering greater transparency and trust among states (De Spiegeleire et al., 2017). These functions contribute to the stability of international security architectures and the prevention of inadvertent escalation.

Importantly, the relationship between humans and AI systems in national security is increasingly viewed through the lens of augmentation rather than replacement. Emphasising human-in-the-loop frameworks, scholars argue for maintaining human oversight over critical decisions, particularly those involving lethal force (Voeneky et al., 2022; Taddeo et al., 2022). This approach balances AI's operational advantages with ethical and legal imperatives, preserving accountability and preventing automation from eroding normative standards (Roy, 2024). Human-machine

teaming leverages AI's strengths in processing and analysis while ensuring that ultimate control and judgement rest with human operators.

In sum, AI integration offers multifaceted benefits across threat detection, decision-making, operational efficiency, intelligence augmentation, cybersecurity, strategic foresight, and human-machine collaboration. These advantages have the potential to redefine national security capabilities, enabling states to respond more swiftly and effectively to an increasingly complex threat environment. However, the realisation of these benefits depends on robust governance, technological reliability, and ethical safeguards, which must be addressed to ensure that AI enhances security without compromising fundamental principles.

## 4. RISKS AND CHALANGES

While the integration of artificial intelligence (AI) into national security systems promises significant advantages, it simultaneously introduces a complex array of risks and challenges that require careful examination. These risks span technical vulnerabilities, ethical dilemmas, legal ambiguities, strategic instability, and governance shortcomings. A comprehensive understanding of these challenges is essential for developing responsible policies that harness AI's benefits while mitigating potential harms.

A primary technical concern involves the reliability and robustness of AI systems deployed in security contexts. As Montasari (2023) and Sfetcu (2024) highlight, AI algorithms—especially those based on machine learning—are vulnerable to adversarial attacks that manipulate input data to produce erroneous outputs, potentially misleading decision-makers. Such attacks threaten the integrity of intelligence assessments, autonomous weapon targeting, and cyber defence mechanisms (Erendor, 2024). Additionally, the "black box" nature of many AI models, noted by Khare and Sinha (2024) and Wilner and Atkinson (2023), limits transparency and explainability, making it difficult for operators to verify or challenge AI-generated decisions. This opacity undermines trust in AI systems and complicates attribution of accountability when failures or unintended consequences occur (Johnson, 2023).

Ethical dilemmas are especially pronounced in military applications. The prospect of autonomous weapons systems capable of selecting and engaging targets without direct human intervention raises profound questions about the delegation of life-and-death decisions to machines (Taddeo et al., 2022; Roy, 2024). Scholars such as Voeneky et al. (2022) and Reinhold and Schoernig (2022) warn that removing humans from critical loops may erode moral responsibility and violate international humanitarian law principles. Moreover, Johnson (2023) stresses the challenges of aligning AI's operational logic with ethical norms, given that AI systems may lack the capacity for contextual judgement or empathy.

Legal ambiguity further complicates AI's integration into national security. Existing international law and domestic legal frameworks were not designed with autonomous technologies in mind, creating "governance vacuums" that hinder clear rules of engagement and accountability mechanisms (Meleouni & Efthymiou, 2024; Lemieux, 2024). For example, the attribution of liability in the event of AI-induced errors or unlawful actions remains contested, with implications for both state responsibility and individual criminal liability (Fatima Roumate, 2024; Johnson, 2023). These gaps risk undermining the rule of law and international stability, especially if states exploit legal uncertainties to develop or deploy controversial AI capabilities without adequate oversight.

Strategically, AI introduces risks of escalation and destabilisation. The rapid decision-making enabled by AI compresses reaction times, increasing the likelihood of miscalculation or accidental conflict (De Spiegeleire et al., 2017; Wilner & Atkinson, 2023). As Haney (2020) and Fatima Roumate (2024) point out, AI-driven arms races may incentivise states to adopt riskier postures, heightening tensions in already volatile regions. Moreover, the diffusion of AI technologies to non-state actors or rogue regimes raises proliferation concerns, with potential for asymmetric threats that are harder to predict or counter (Cristiano et al., 2023; Santos Nanni, 2024).

Cybersecurity challenges pose a paradoxical dilemma. Although AI can enhance cyber defence, its increasing complexity creates new vulnerabilities exploitable by sophisticated adversaries (Ventre, 2020; Erendor, 2024). AI systems themselves may become targets of adversarial manipulation, data poisoning, or software exploitation, compromising national security infrastructure. Roumate (2024) further warns that AI can be weaponised in information warfare, enabling large-scale disinformation campaigns and psychological operations that undermine social cohesion and democratic governance (Lemieux, 2024).

The intersection of AI and surveillance also raises significant privacy and human rights concerns. Lemieux (2024) and Xu et al. (2023) emphasise that AI-enabled mass surveillance expands states' capacity for intrusive monitoring, often without adequate legal safeguards or transparency. This risks eroding civil liberties and fueling authoritarian practices, particularly where democratic accountability is weak or absent (Fatima Roumate, 2024; Jaber, 2023). These tensions highlight the broader political and normative challenges in balancing security imperatives with fundamental rights.

Another challenge lies in governance and regulation. As Montasari (2022) and Meleouni and Efthymiou (2024) observe, AI's rapid technological evolution outpaces the development of effective national and international regulatory frameworks. Fragmented policy responses—exemplified by differing approaches between Western alliances prioritising "responsible AI" principles and states focused on rapid military advantage complicate efforts to establish shared norms (NATO Communications and Information Agency, 2022; Haney, 2020). The dual-use nature of AI, where civilian innovations are repurposed for military use, further obscures regulatory boundaries and enforcement (Voeneky et al., 2022; Girasa, 2020).

The opaque supply chains and complex ecosystems supporting AI development also introduce challenges in verification and compliance. De Spiegeleire et al. (2017) highlight difficulties in monitoring AI arms development, while Santos Nanni (2024) stresses the need for transparency mechanisms to prevent clandestine escalation. Without such mechanisms, trust deficits among states may deepen, increasing strategic uncertainty and reducing prospects for arms control.

Finally, the human dimension presents persistent challenges. Despite technological advances, human operators remain integral to security decision-making, yet they face issues of over-reliance on AI, skill degradation, and ethical dissonance (Roy, 2024; Johnson, 2023). Ensuring that personnel maintain critical judgement and accountability while effectively collaborating with AI systems requires comprehensive training, doctrinal adaptation, and cultural change within security institutions (Taddeo et al., 2022).

In conclusion, while AI's integration into national security holds promise, the risks and challenges it presents are multifaceted and profound. Addressing technical vulnerabilities, ethical quandaries, legal gaps, strategic instabilities, and governance deficiencies demands coordinated, interdisciplinary efforts. Only through robust oversight, international cooperation, and continuous ethical reflection can the security community mitigate AI's risks while leveraging its transformative potential.

## 5. POLICY DILEMMAS

The integration of artificial intelligence (AI) into national security raises profound policy and governance dilemmas that challenge traditional frameworks of regulation, accountability, and international cooperation. While AI technologies offer significant strategic advantages, their rapid development and complex nature expose gaps and tensions in existing governance mechanisms, calling for urgent, multidimensional responses. This section critically analyses the principal dilemmas policymakers face in managing AI's national security implications.

A central dilemma stems from the tension between the imperative for innovation and the necessity for responsible governance. States and defence establishments are under intense pressure to harness AI's potential swiftly to maintain strategic advantage, as underscored by Haney (2020) and Wilner and Atkinson (2023). This competitive imperative often drives accelerated deployment of AI capabilities, sometimes at the expense of thorough ethical vetting, transparency, and oversight (NATO Communications and Information Agency, 2022). Conversely, as Voeneky et al. (2022) and Taddeo et al. (2022) emphasise, establishing principled constraints and accountability is essential to prevent misuse, unintended escalation, and violations of international law. Balancing speed and prudence remains a persistent governance dilemma, where failures on either side risk strategic or normative harm.

The dual-use nature of AI technologies complicates governance further. As Montasari (2022) and Girasa (2020) note, civilian AI innovations frequently cross into military applications, blurring the lines between commercial and defence sectors. This overlap hampers regulation and export controls, as technologies diffuse rapidly beyond traditional arms control frameworks (Cristiano et al., 2023; Sfetcu, 2024). Policymakers struggle to monitor and manage this diffusion, particularly given the globalised and digitised nature of AI development. Lemieux (2024) highlights the challenges in differentiating peaceful uses from potentially destabilising military applications, complicating verification and compliance efforts.

Another critical governance dilemma arises from divergent national and international approaches to AI regulation. Western alliances such as NATO foreground ethical standards, human oversight, and interoperability, encapsulated in principles of "responsible AI" (NATO Communications and Information Agency, 2022; Wilner & Atkinson, 2023). In contrast, other states prioritise rapid capability development and strategic deterrence, sometimes with less emphasis on transparency or ethical constraints (Haney, 2020; Fatima Roumate, 2024). This fragmentation inhibits the formation of common norms or binding agreements, raising the risk of arms races and normative erosion (De Spiegeleire et al., 2017; Santos Nanni, 2024). As Meleouni and Efthymiou (2024) argue, reconciling these divergent approaches is essential for global stability but remains highly challenging.

Legal ambiguities further exacerbate governance difficulties. Existing international law frameworks—crafted before AI's emergence—struggle to accommodate autonomous systems, particularly regarding accountability for unlawful

acts or violations of humanitarian law (Johnson, 2023; Lemieux, 2024). Fatima Roumate (2024) stresses the absence of clear liability regimes when AI systems malfunction or operate autonomously, creating "accountability gaps" that undermine enforcement. This gap complicates both state responsibility and individual criminal liability, raising questions about how to hold actors accountable for AI-enabled actions in conflict.

The governance of AI also faces challenges related to transparency and explainability. Khare and Sinha (2024) and Wilner and Atkinson (2023) highlight that the opacity of many AI systems hinders meaningful oversight and informed decision-making. This "black box" problem is particularly problematic in defence contexts, where understanding the basis for AI-generated recommendations or actions is vital for trust, legal compliance, and risk management. Without greater explainability, commanders may either over-rely on AI or dismiss it outright, both outcomes posing risks (Montasari, 2023).

Cybersecurity concerns introduce another layer of governance complexity. While AI enhances cyber defence, it also increases attack surfaces and vulnerability to adversarial manipulation (Ventre, 2020; Erendor, 2024). Roumate (2024) notes that malicious actors can exploit AI's weaknesses for disinformation, influence operations, and cyberattacks, complicating defensive governance. Managing this dynamic requires integrated policies that address both technological and human factors, balancing security with civil liberties (Lemieux, 2024).

The role of non-state actors and private industry compounds governance dilemmas. Jaber (2023) and Voeneky et al. (2022) emphasise that much AI innovation occurs outside state control, with commercial companies and academia driving research and development. This decentralisation challenges governments' ability to regulate or steer AI technologies effectively, especially given competing economic and strategic interests. Collaborative governance models that include diverse stakeholders are increasingly advocated to ensure accountability and ethical standards (Fatima Roumate, 2024; Taddeo et al., 2022).

International cooperation remains both necessary and difficult. Meleouni and Efthymiou (2024) argue that fragmented governance risks an AI arms race, destabilising the international system. Yet, trust deficits, geopolitical rivalries, and different values hinder consensus on standards or arms control measures (De Spiegeleire et al., 2017; Santos Nanni, 2024). Lemieux (2024) and NATO Communications and Information Agency (2022) suggest confidence-building measures, transparency initiatives, and norm development as pragmatic steps, but achieving effective multilateral governance will require overcoming entrenched strategic competition.

Finally, there is a pressing need to address the human dimension in AI governance. Roy (2024) and Johnson (2023) stress that despite automation, human judgement remains indispensable. Training, doctrinal adaptation, and cultural shifts within security organisations are essential to ensure responsible human-machine teaming. This includes recognising cognitive biases, preventing overreliance on AI, and maintaining ethical vigilance. Failure to integrate human factors risks undermining both operational effectiveness and normative commitments (Taddeo et al., 2022).

In summary, policy and governance dilemmas around AI in national security revolve around balancing innovation with ethical responsibility, managing dual-use challenges, bridging divergent national approaches, clarifying legal frameworks, ensuring transparency, addressing cybersecurity risks, involving diverse stakeholders, fostering international cooperation, and integrating human factors. Addressing these interlinked dilemmas requires coordinated, interdisciplinary, and multilevel strategies to realise AI's potential while safeguarding security, legality, and ethics.

## 6. CASE STUDIES AND COMPARATIVE ANALYSIS

The deployment and governance of artificial intelligence (AI) in national security vary significantly across global powers, shaped by their unique strategic objectives, political systems, technological capabilities, and ethical frameworks. This section provides a comparative analysis of AI integration in national security among key actors—the United States, the European Union, China, Russia, and Israel—highlighting how divergent approaches illuminate broader governance and policy dilemmas. Examining these cases offers valuable lessons on the complexities of balancing innovation, risk, and ethical responsibility in the AI security domain.

The United States remains a leading innovator in AI applications for national security, with extensive investments spanning autonomous systems, intelligence analysis, cyber defence, and command and control. Montasari (2023) and Khare and Sinha (2024) detail how the U.S. Department of Defense established the Joint Artificial Intelligence Center (JAIC) to centralise AI development and operational integration, reflecting a strategic imperative to maintain technological superiority. This initiative prioritises embedding ethical principles such as human oversight, accountability, and transparency, aligning with international humanitarian law (Wilner & Atkinson, 2023). Nonetheless, Johnson (2023) and Greene (2023) underscore persistent challenges, including difficulties in integrating AI into existing complex military hierarchies and the risk of overdependence on "black box" algorithms that resist

www.ijprems.com
editor@ijprems.com
INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
(Int Peer Reviewed Journal)
Vol. 05, Issue 09, September 2025, pp : 804-813
e-ISSN : 2583-1062
Impact Factor : 7.001

explainability. These challenges highlight the delicate balance the U.S. attempts to strike between rapid innovation and responsible governance.

In contrast, the European Union (EU) adopts a more governance-centric and cautious approach, prioritising ethical standards, human rights, and transparency alongside security applications. Meleouni and Efthymiou (2024) describe the EU's AI Act as a comprehensive regulatory framework designed to govern "high-risk" AI uses, including those in defence, with rigorous accountability and risk management requirements. The EU also emphasises international cooperation and multilateralism in AI governance, seeking to set global norms grounded in democratic values and human-centric AI (Lemieux, 2024; NATO Communications and Information Agency, 2022). However, Haney (2020) and Wilner and Atkinson (2023) note that the EU's regulatory caution may slow deployment, potentially impacting strategic competitiveness against faster-moving states. Nevertheless, the EU model's emphasis on embedding normative considerations into policy offers a blueprint for balancing innovation with ethical governance.

China presents a markedly different paradigm, characterised by rapid AI deployment driven by a state-led model of technological sovereignty and authoritarian governance. Fatima Roumate (2024) and Girasa (2020) document China's robust investments in AI for military modernisation, cyber capabilities, and comprehensive surveillance systems. The concept of "civil-military fusion" is central to China's strategy, facilitating rapid translation of civilian AI advances into military applications (Sfetcu, 2024). As Xu et al. (2023) highlight, China leverages AI-enabled surveillance extensively for domestic security, ensuring regime stability and control, which diverges significantly from Western ethical frameworks emphasising privacy and civil liberties. Roumate (2021) argues that China's focus on speed and operational efficacy over normative constraints reflects a security model that privileges technological dominance. This approach not only challenges international governance norms but also intensifies strategic competition in AI-enabled defence technologies.

Russia's approach to AI in national security is characterised by pragmatism, strategic opportunism, and a focus on hybrid warfare tactics. Haney (2020) and Santos Nanni (2024) note that Russia strategically targets AI investments to enhance cyber operations, information warfare, and autonomous systems that exploit adversaries' vulnerabilities. Russia's AI programmes, although more limited in scale compared to the U.S. and China, are tailored to asymmetric conflict environments where rapid, covert, and disruptive capabilities offer leverage (Cristiano et al., 2023). Reinhold and Schoernig (2022) observe that Russia's governance of AI technologies lacks transparency and ethical oversight, reflecting broader geopolitical considerations prioritising operational advantage over normative governance. This opacity complicates efforts to establish global standards and increases the risks of unchecked escalation.

Israel offers a distinct model of AI integration in national security, characterised by an innovative ecosystem that fosters close collaboration between defence institutions and a vibrant private tech sector. Montasari (2023) highlights Israel's pioneering use of AI in intelligence collection, autonomous defence systems, and cyber operations, often driven by battlefield-tested technologies rapidly adapted to evolving threats. Khare and Sinha (2024) describe Israel's operational pragmatism and agility in deploying AI-enabled systems, supported by a culture of innovation and security needs shaped by a volatile regional environment. However, Lemieux (2024) raises concerns regarding oversight and ethical issues, particularly around pervasive surveillance and targeted operations, underscoring tensions between strategic necessity and normative considerations. Israel exemplifies the challenges of balancing rapid technological adaptation with accountability in AI governance.

Comparing these diverse national approaches reveals key patterns and tensions. The United States and European Union illustrate the ongoing tension between innovation speed and governance rigor: the U.S. prioritises rapid deployment with emerging ethical frameworks, while the EU emphasises comprehensive regulation at the cost of slower technological adoption (NATO Communications and Information Agency, 2022; Wilner & Atkinson, 2023). China's state-driven, authoritarian model contrasts sharply with democratic norms, prioritising rapid AI militarisation and pervasive surveillance over transparency and human rights (Fatima Roumate, 2024; Girasa, 2020). Russia's focus on hybrid warfare and opaque governance highlights risks associated with limited oversight and strategic ambiguity (Haney, 2020). Israel's hybrid civil-military innovation ecosystem demonstrates how strategic pressures drive AI adoption while raising questions about governance sufficiency (Montasari, 2023; Lemieux, 2024).

These divergences complicate the prospects for establishing universally accepted norms or treaties governing AI in national security. Meleouni and Efthymiou (2024) argue that bridging these divides requires reconciling competing strategic interests and value systems to reduce risks of destabilising arms races and conflict escalation. Lemieux (2024) and Santos Nanni (2024) stress that confidence-building measures, transparency initiatives, and incremental norm-building offer practical pathways to cooperation. Nevertheless, as De Spiegeleire et al. (2017) caution, geopolitical rivalry and mistrust pose significant barriers to effective multilateral governance.

www.ijprems.com
editor@ijprems.com

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
(Int Peer Reviewed Journal)
Vol. 05, Issue 09, September 2025, pp : 804-813

e-ISSN : 2583-1062
Impact Factor : 7.001

Beyond leading powers, the diffusion of AI capabilities to smaller states and non-state actors further complicates the security environment. Cristiano et al. (2023) highlight that AI-enabled cyber and information warfare tools empower actors with limited conventional military strength, undermining traditional deterrence models and increasing conflict unpredictability. This proliferation intensifies the urgency for adaptable governance frameworks capable of addressing a broad spectrum of actors and threats (Sfetcu, 2024).

In conclusion, the comparative analysis of AI integration in national security across major geopolitical actors reveals a fragmented landscape shaped by differing governance philosophies, strategic imperatives, and technological capacities. While AI offers transformative security advantages, governance structures struggle to keep pace with rapid technological change, complicating efforts to manage risks and uphold normative principles. Understanding these diverse experiences provides critical insights for policymakers seeking to navigate the complex intersection of innovation, ethics, and security in the AI era. Moving forward, interdisciplinary and multilateral approaches will be essential to reconcile competing interests and mitigate the risks inherent in the militarisation of artificial intelligence.

## 7. CONCLUSION

The integration of artificial intelligence (AI) into national security represents one of the most significant technological and strategic developments of the contemporary era. This transformation offers unprecedented opportunities to enhance defence capabilities, improve intelligence analysis, and strengthen cybersecurity. However, it simultaneously introduces a constellation of complex risks and governance challenges that require careful and nuanced management. The dynamic interplay between the promise and peril of AI underscores the imperative for a balanced, responsible approach to its development and deployment within national security frameworks.

Throughout this study, it has become evident that AI's potential benefits are vast and multifaceted. From augmenting human decision-making with superior data processing and predictive analytics to automating routine tasks and enhancing operational efficiency, AI reshapes the capabilities of defence institutions. The technology's ability to operate at speed and scale enables states to respond more rapidly to emerging threats, improving situational awareness and threat mitigation. Moreover, AI enhances cyber defence and resilience by providing adaptive, real-time protection against increasingly sophisticated cyberattacks. These advances collectively have the capacity to redefine the strategic landscape, offering states new tools to safeguard their national interests in an era of rapid technological change.

Yet, these opportunities come with significant caveats. The deployment of AI in national security is accompanied by profound ethical, legal, and strategic dilemmas. The challenges of ensuring transparency and accountability in AI decision-making are acute, particularly where autonomous systems might operate with limited human oversight. The risk of malfunction, adversarial manipulation, or unintended consequences in high-stakes environments poses critical questions about reliability and trust. Furthermore, the acceleration of AI-driven arms races could exacerbate geopolitical tensions, increasing the likelihood of miscalculation or inadvertent conflict. These risks emphasize the need for robust governance structures that can safeguard against both technological failures and strategic instability.

The comparative examination of leading global actors illustrates the diversity of approaches to AI integration and governance in national security. While some states prioritise rapid technological advancement and strategic dominance, others emphasise ethical constraints, regulatory oversight, and international cooperation. These differences reflect broader political cultures, strategic doctrines, and institutional capacities, complicating efforts to establish common norms or binding agreements. The fragmentation of governance approaches not only poses challenges for international stability but also risks undermining public trust and normative legitimacy in the use of AI for defence purposes.

Addressing these multifaceted challenges requires an integrated, interdisciplinary approach that bridges technical innovation, ethical reflection, and policy development. It is imperative that states and international actors develop governance frameworks that uphold human dignity, accountability, and respect for international law while fostering innovation and operational effectiveness. Ensuring meaningful human control over AI systems, particularly in lethal or coercive applications, remains a critical principle to mitigate ethical and legal concerns. In parallel, transparency and explainability must be enhanced to build trust among operators, policymakers, and the public.

Moreover, international cooperation and multilateral dialogue are essential to navigate the complex security implications of AI. Confidence-building measures, joint research initiatives, and the development of shared standards can help reduce mistrust and prevent destabilising arms races. While geopolitical rivalries pose formidable obstacles, establishing forums for dialogue and norm-setting can create pathways for collaboration that balance national interests with collective security. Efforts to integrate smaller states and non-state actors into governance frameworks will also be crucial to manage the proliferation and misuse of AI technologies globally.

Equally important is the recognition of the human dimension in AI integration. The successful adoption of AI in national security depends not only on technological capabilities but also on the skills, judgement, and ethical commitment of human operators. Comprehensive training, doctrinal adaptation, and cultural change within defence institutions are necessary to ensure that AI functions as an effective partner rather than an unaccountable substitute. Human-machine teaming should be designed to harness the complementary strengths of both, with humans retaining ultimate responsibility for critical decisions.

In conclusion, the integration of artificial intelligence into national security is a double-edged sword, offering transformative benefits alongside profound challenges. Its success hinges on a delicate balance between harnessing AI's innovative potential and managing its risks through responsible governance, ethical foresight, and international collaboration. As AI technologies continue to evolve, the security community must remain vigilant and adaptive, constantly reassessing policies and practices to ensure that technological progress aligns with fundamental values and long-term stability.

The path forward demands a comprehensive strategy that integrates technical expertise, policy innovation, ethical deliberation, and diplomatic engagement. Only through such a multifaceted approach can the international community hope to realise the promise of AI-enhanced security while safeguarding against its perils. The stakes are high: the manner in which states choose to develop and govern AI in national security will shape not only the future of warfare and defence but also the broader contours of international peace, justice, and human rights in the decades to come.

## 8. REFERENCES

[1] De Spiegeleire, S., Maas, M., & Sweijs, T. (2017). Artificial intelligence and the future of defense: Strategic implications for small- and medium-sized force providers.

[2] Erendor, M. E. (Ed.). (2024). Cyber security in the age of artificial intelligence and autonomous weapons.

[3] Girasa, R. (2020). Artificial intelligence as a disruptive technology: Economic transformation and government regulation.

[4] Greene, J. A. (2023). National security and artificial intelligence.

[5] Grigalashvili V., (2025), Artificial Intelligence in Higher Education: A New Era of Smart Learning

[6] Grigalashvili V., (2025), ARTIFICIAL INTELLIGENCE IN PUBLIC ADMINISTRATION:

[7] Grigalashvili V., (2025), Bridging Governance and Technology: AI in E-Governance

[8] Haney, B. S. (n.d.). Applied artificial intelligence in modern warfare & national security policy.

[9] Hunnewell, B. (2025). National security concerns for artificial intelligence and civilian critical infrastructure.

[10] Imam, I. (2021). Role of artificial intelligence in defence strategy: Implications for global and national security.

[11] Jaber, W. (Ed.). (2023). Artificial intelligence in the age of nanotechnology.

[12] Johnson, J. (2023). Artificial intelligence and national security.

[13] Khare, V. S., & Sinha, A. (Colonel). (2024). Artificial intelligence and national security.

[14] Lahmann, H., & Geiss, R. (Eds.). (2024). Research handbook on warfare and artificial intelligence.

[15] Lemieux, F. (2024). Intelligence and state surveillance in modern societies: An international perspective.

[16] Mallick, P. K. (2024). Artificial intelligence, national security and the future of warfare.

[17] Masakowski, Y. R. (2020). Artificial intelligence and global security: Future trends, threats and considerations.

[18] Meleouni, C., & Efthymiou, I.-P. (2024). The use of artificial intelligence (AI) in national security: Defining international standards and guidelines.

[19] Montasari, R. (2022). Artificial intelligence and national security.

[20] Montasari, R. (2023). Applications for artificial intelligence and digital forensics in national security.

[21] Morgan, F. E., Boudreaux, B., Lohn, A. J., Ashby, M., Curriden, C., Klima, K., & Grossman, D. (2020). Military applications of artificial intelligence: Ethical concerns in an uncertain world.

[22] Nanni, J. P. S. (2024). The use of artificial intelligence as a national defense strategy. [

[23] NATO Communications and Information Agency. (2022). Responsible AI in defense: Principles and practice.

[24] IISS – The International Institute for Strategic Studies. (2023). Artificial intelligence in intelligence services.

[25] Reinhold, T., & Schoernig, N. (2022). Armament, arms control and artificial intelligence: The Janus-faced nature of machine learning in the military realm.

[26] Roy, K. (Ed.). (2024). Artificial intelligence, ethics and the future of warfare: Global perspectives.

[27] Roumate, F. (2021). Artificial intelligence and digital diplomacy: Challenges and opportunities.

[28] Roumate, F. (2024). Artificial intelligence and the new world order: New weapons, new wars and a new balance of power.

[29] Sfetcu, N. (2024). Artificial intelligence in intelligence agencies, defense and national security. Taddeo, M., Ziosi, M., Tsamados, A., Gilli, L., & Kurapati, S. (2022). The use of artificial intelligence (AI) in national security: Defining international standards and guidelines.

[30] Ventre, D. (2020). Artificial intelligence, cybersecurity and cyber defence.

[31] Voeneky, S., Kellmeyer, P., Mueller, O., & Burgard, W. (2022). The Cambridge handbook of responsible artificial intelligence: Interdisciplinary perspectives.

[32] Wilner, A., & Atkinson, R. (2023). Artificial intelligence and national defence: A strategic foresight analysis.

[33] Xu, L., Qereshniku, E., Hazari, H., & Edwards, M. (2023). Understanding national security threats enabled by artificial intelligence: Implications for CSIS.

[34] Yampolskiy, R. (2018). Artificial intelligence safety and security.