

ARTIFICIAL INTELLIGENCE AND IRREGULAR WARFARE: STRATEGIC GOVERNANCE AND INSTITUTIONAL INTEGRATION

Maurice DAWSON, Samson QUAYE, Hrijuta CHAKRABORTY

Illinois Institute of Technology, Chicago, USA
mdawson2@illinoistech.edu

Abstract: Artificial intelligence is reshaping irregular warfare by accelerating decision-making, expanding cyber and information operations, and transforming competition below the threshold of conventional conflict. AI-enabled systems strengthen intelligence processing, anomaly detection, predictive analytics, surveillance, cyber defense, and influence operations, but they also introduce governance, ethical, legal, and escalation risks in ambiguous operational environments. This paper examines how military and security institutions can integrate AI across strategic, doctrinal, operational, technical, and regulatory domains while preserving accountability and legitimacy. Building on scholarship in AI-enabled cybersecurity, cognitive security, digital twin resilience, and national security governance, the study develops a five-layer institutional integration model: strategic alignment, doctrinal adaptation, operational embedding, data fusion, and human-AI governance. The model is applied to four irregular warfare contexts: the Russia-Ukraine conflict, cyber and information operations, extremist networks, and digitally dependent critical infrastructure. Findings indicate that AI can serve as a force multiplier when embedded within coherent institutional structures, but adoption without coordinated oversight increases risks of accountability erosion, algorithmic opacity, miscalculation, and normative destabilization.

Keywords: Artificial Intelligence, Irregular Warfare, Strategic Governance, Cyber Operations, Human-AI Governance

1. Introduction

Artificial Intelligence has become a defining feature of contemporary strategic competition [1]. States increasingly rely on these systems to enhance intelligence analysis, cyber defense, logistics, surveillance, decision support, and information operations. In irregular warfare environments, where ambiguity, influence, deniability, and political legitimacy often matter more than conventional military superiority, AI can produce disproportionate strategic effects.

Irregular warfare includes political warfare, influence operations, proxy activity, insurgency, counterinsurgency, cyber operations, sabotage, economic coercion, and competition within civilian information

ecosystems. These activities often occur below the threshold of open armed conflict. AI can support such operations by accelerating targeting, amplifying narratives, detecting vulnerabilities, and enabling persistent engagement. However, these same capabilities increase the potential for miscalculation, civilian harm, accountability gaps, and escalation.

This paper addresses three research questions: How should defense and security institutions govern AI integration in irregular warfare? What institutional structures are required to balance operational effectiveness with accountability and restraint? How do the governance requirements differ across irregular warfare domains? The objective is to develop a practical institutional integration

model that connects strategy, doctrine, operations, data fusion, and oversight, and to evaluate that model across four representative case contexts.

The paper proceeds as follows. Section 2 presents the methodological approach. Section 3 situates AI within strategic competition. Section 4 characterizes the AI-supported irregular warfare environment. Section 5 addresses doctrinal evolution. Section 6 presents the five-layer institutional integration model. Sections 7 through 10 apply the model to four case contexts. Section 11 draws lessons from AI-powered IoT cybersecurity research. Section 12 addresses governance, risk, and ethics. Section 13 presents policy implications. Section 14 proposes an implementation framework. Section 15 discusses findings, and Section 16 concludes.

2. Methodological Approach

This article uses a conceptual, policy-oriented approach based on a narrative synthesis of literature on military AI oversight, strategic competition, irregular warfare, cybersecurity, digital twins, and cognitive security. It develops an institutional model rather than testing a quantitative hypothesis. The literature was selected because it addresses the strategic, doctrinal, operational, technical, and accountability dimensions of AI integration in security environments.

The four case contexts were selected because they represent distinct AI-relevant irregular warfare environments: battlefield adaptation in the Russia-Ukraine conflict, cyber and information operations, extremist network activity, and critical infrastructure vulnerability. Together, these cases allow the five-layer model to be assessed across state conflict, digital operations, non-state actor activity, and civilian infrastructure resilience.

3. Artificial Intelligence and Strategic Competition

AI has become a core instrument of national power [2]. Major powers view AI superiority

as essential to military modernization, economic competitiveness, intelligence advantage, and deterrence [3]. National security strategies increasingly emphasize data infrastructure, advanced computing, semiconductor supply chains, cyber resilience, and AI workforce development.

In irregular warfare, AI supports persistent engagement by improving the ability to monitor political instability, social unrest, infrastructure vulnerabilities, and adversarial influence campaigns. Predictive analytics can identify emerging risks. Natural language processing can monitor narratives across social media and open-source platforms. Computer vision can support surveillance and battlefield awareness. Machine learning can assist cyber defense by identifying anomalous patterns in network traffic. Generative AI can produce or amplify persuasive messages at scale.

These capabilities blur the boundary between peace and conflict, creating deterrence and escalation challenges because AI-based operations may be continuous, deniable, and difficult to attribute. The relationship between AI and strategic competition is also evident in critical infrastructure and Industrial IoT environments, where hyperconnected systems create efficiency gains but also introduce new attack surfaces. Cognitive security approaches are necessary because adversaries may not only attack systems directly; they may manipulate data flows, decision models, or operator trust [4].

4. Irregular Warfare and the AI-Enabled Operating Environment

Irregular warfare depends on ambiguity, adaptation, and influence. AI changes this environment by expanding the speed and reach of indirect methods, allowing actors to identify target audiences, automate deception, model social behavior, exploit cyber vulnerabilities, and coordinate distributed operations [5].

The AI-supported irregular warfare environment has four defining features.

First, it is data-intensive: actors compete to collect, process, and interpret data from sensors, social platforms, networks, commercial systems, and open sources. Second, it is cognitively oriented: the objective is often to shape beliefs, behaviors, trust, and legitimacy. Third, it is infrastructure-dependent: civilian systems such as communications networks, power grids, financial systems, and industrial control systems become targets or conduits for influence and disruption.

Fourth, it is institutionally complex: military organizations, intelligence agencies, private firms, technology vendors, civil society, and international partners all influence the operating environment. AI magnifies both opportunity and risk in this environment. It can help defenders identify adversarial activity earlier, automate triage, and support decision-making. Yet adversaries can use the same methods to probe defenses, generate synthetic media, and manipulate perceptions. This dual-use nature requires oversight structures that are flexible enough to support innovation but strong enough to prevent uncontrolled escalation or illegitimate use.

5. Doctrinal Evolution in Irregular Warfare

Irregular warfare doctrine has evolved from population-centric counterinsurgency toward broader competition frameworks that include cyber operations, information warfare, proxy activity, gray-zone operations, and resilience. AI accelerates this doctrinal evolution because it changes the methods through which influence, surveillance, and disruption can be conducted.

Doctrinal coherence requires alignment between strategic objectives, operational concepts, professional norms, and legal obligations. AI challenges this alignment because it redistributes authority between humans and machines. Professional military education and training must include AI literacy, covering model limitations, data

quality concerns, bias, adversarial manipulation, automation bias, and human-machine teaming.

Doctrinal adaptation should incorporate lessons from AI-powered cybersecurity research. Models that perform well in isolated settings may fail in heterogeneous environments [6]. Defense AI systems trained in one environment may not generalize to another operational, cultural, linguistic, or technical context. Doctrine should therefore require cross-domain validation before deployment.

6. Five-Layer Institutional Integration Model

This paper proposes a five-layer institutional integration model for AI in irregular warfare. The model is grounded in a review of national security governance literature [1, 7], AI and strategic competition scholarship [2, 3], cybersecurity and IoT defense research [6], digital twin and critical infrastructure studies [8], and cognitive security frameworks [4]. Each layer addresses a distinct oversight requirement, and together they form a coordinated architecture for responsible AI integration.

Layer 1: Strategic Alignment ensures that AI adoption is connected to national security priorities, risk tolerance, legal obligations, and alliance commitments. AI initiatives should be linked to defined strategic objectives such as improving cyber resilience, strengthening deterrence, protecting critical infrastructure, or countering disinformation. Senior leaders must define where AI can be used, where human approval is required, and where automation should be prohibited.

Layer 2: Doctrinal Adaptation translates strategy into operational concepts, training standards, rules of engagement, and professional norms. AI must be incorporated into doctrine in ways that clarify authority, responsibility, and limitations. Doctrine should identify which AI capabilities support intelligence, cyber defense, influence assessment, logistics, planning,

and decision support, and must address human-machine teaming requirements.

Layer 3: Operational Embedding involves integrating AI tools into command systems, intelligence workflows, cyber defense platforms, logistics networks, and mission planning processes. Key challenges include ensuring that AI tools are interoperable, secure, validated, and aligned with mission needs. Operational embedding should be accompanied by testing, red teaming, and continuous performance monitoring.

Layer 4: Data Fusion and Intelligence Integration address the combination of information from sensors, cyber logs, social media, geospatial sources, intelligence reporting, and partner networks. This layer requires data provenance standards, access controls, audit trails, and validation procedures. Digital twin environments are especially relevant here because they depend on real-time data streams and introduce risks from adversarial data manipulation [8].

Layer 5: Human-AI Governance is the foundation of responsible AI integration. It includes legal review, ethical oversight, audit mechanisms, accountability structures, validation protocols, and escalation controls. This layer ensures that AI-supported capabilities remain subject to institutional judgment and democratic control [7]. It should address automation bias, explainability, proportionality, and the potential for civilian harm.

The four cases are analyzed using the same five-layer model. Each case is considered in terms of strategic alignment, doctrinal adaptation, operational embedding, data fusion, and human-AI governance. This allows the cases to be compared systematically rather than treated as separate examples.

7. Case Study 1: Russia-Ukraine and AI-Enabled Irregular Warfare

The Russia-Ukraine conflict illustrates how AI-supported capabilities can shape irregular and conventional operations simultaneously [9]. The conflict has involved drone warfare, open-source

intelligence, cyber operations, information campaigns, electronic warfare, and rapid battlefield adaptation. These tools have assisted in imagery analysis, pattern detection, targeting workflow support, narrative monitoring, and network defense. The conflict also demonstrates that AI systems are vulnerable to deception, electronic warfare, data manipulation, and supply chain constraints. A system that performs well in laboratory conditions may degrade in contested environments, reinforcing the importance of cross-domain validation (Layer 2) and continuous performance monitoring (Layer 3). The involvement of commercial platforms in military operations also creates a governance challenge because accountability frameworks designed for military systems do not automatically extend to private-sector data ecosystems, creating gaps in Layer 5.

8. Case Study 2: Cyber Operations and Information Warfare

Cyber operations and information warfare are among the most direct areas where AI affects irregular warfare. AI can improve vulnerability discovery, malware detection, intrusion detection, anomaly analysis, and automated response. Offensive actors can use AI to automate reconnaissance, craft phishing messages, generate malicious code, and identify weak points in digital infrastructure.

Information warfare is similarly transformed. Generative AI can create persuasive text, images, audio, and video at scale. These capabilities reduce the cost of influence operations. However, AI-supported cyber and information operations create accountability and escalation concerns [10]: a deepfake may trigger political instability, false attribution may provoke retaliation, and an automated cyber defense tool may respond incorrectly. This case particularly stresses Layers 1 and 5 of the model: strategic alignment must define clear escalation thresholds, while human-AI oversight must require review before

automated offensive or influence action is taken.

9. Case Study 3: Extremist Networks and Non-State Actors

AI lowers barriers to entry for non-state actors [11]. Extremist networks can use automation for propaganda, recruitment, translation, target research, and operational security. Generative AI may enable extremist actors to produce large volumes of content, create synthetic personas, or evade moderation systems.

This creates a regulatory challenge for states and platforms. Automated content moderation may be useful, but it can produce false positives, suppress legitimate speech, or fail across languages and cultural contexts. This case illustrates the interaction between Layers 2 and 4: doctrine must define lawful, proportionate counter-extremism methods, while data fusion across law enforcement, intelligence, and platform sources requires legal frameworks that constrain misuse. Human oversight of AI flags remains essential to prevent over-removal and civil liberties violations.

10. Case Study 4: Digital Dependency and Critical Infrastructure

Irregular warfare increasingly targets critical infrastructure. Power systems, water

systems, ports, transportation networks, healthcare systems, and communications networks depend on digital platforms and connected devices. Digital twins intensify this challenge by allowing organizations to simulate, monitor, and optimize physical systems; but if an adversary manipulates the data feeding a digital twin, decision-makers may act on false representations of reality. Dawson, Vassilakos, Moore, and Szakonyi [8] argue that digital twin ecosystems require zero-trust architectures, behavioral analytics, continuous surveillance, and cooperative governance.

Cognitive security for Industrial IoT also matters here. Industrial environments are not only technical systems; they are decision environments in which operators must trust sensor data, dashboards, alerts, and AI recommendations. This case most heavily loads Layers 3, 4, and 5: operational embedding must include ICS/OT-specific AI tools, data fusion must detect and resist manipulation of digital twin state data, and human-AI oversight must preserve operator authority over physical actuation decisions.

Table 1 synthesizes the four case contexts by mapping each case to the five-layer institutional integration model and identifying the main governance implications

Table 1. Case-Study Analytical Matrix: Five-Layer Model Application

Case Study	L1: Strategic Alignment	L2: Doctrinal Adaptation	L3: Operational Embedding	L4: Data Fusion	L5: Human-AI Governance
Russia-Ukraine	AI integral to national security strategy; contested sovereignty objectives	Doctrine rapidly adapted to drone warfare, ISR, and open-source intelligence	Commercial imagery, AI targeting workflows, real-time battlefield analytics	Multi-source fusion: OSINT, satellite, cyber logs; data integrity risks	Ad hoc oversight; accountability gaps when commercial platforms involved
Cyber & Information Ops	AI-based operations must align with rules of engagement	Doctrine must govern automated response,	Intrusion detection, NLP monitoring,	Real-time threat intelligence fusion; risk of	Human review required before automated

	and escalation thresholds	attribution standards, and deepfake policy	generative content tooling embedded in SOC workflows	data poisoning and attribution	offensive or influence action
Extremist Networks	Counter-extremism aligned with civil liberties; cross-agency mandate required	Moderation doctrine balances detection effectiveness with false-positive risk	AI content moderation, synthetic persona detection, multilingual NLP	Platform, law enforcement, and intelligence data fusion; legal constraints on sharing	Human review of AI flags; oversight of automated takedowns to prevent over-removal
Critical Infrastructure / Digital Twins	Zero-trust mandated at national level; resilience tied to strategic deterrence	ICS/OT doctrine integrates behavioral analytics and continuous monitoring protocols	Digital twin monitoring, anomaly detection on ICS/OT, AI-driven maintenance alerts	Sensor streams, digital twin state data, cyber logs; AI-manipulation is core adversary method	Operator cognitive security training; human authority retained over physical actuation

Source: Authors' own elaboration

11. Cybersecurity Lessons from AI-Powered IoT Defense

The MDPI study on AI-powered cybersecurity models provides four lessons directly applicable to irregular warfare [6]. First, model performance must be evaluated across diverse environments because a model that performs well on one dataset may not generalize to contested, culturally diverse, or technically heterogeneous operational conditions. Second, AI-supported security should be layered: tiered deployment involving edge, gateway, and cloud-level analysis maps onto the intelligence, cyber defense, and operational decision-support tiers in these environments. Third, AI systems require continuous monitoring because threat environments change, adversaries adapt, and model drift can degrade effectiveness. Fourth, AI should be integrated with human expertise, since technical performance metrics are necessary but not sufficient; human analysts, operators, lawyers, commanders, and policymakers

must interpret system outputs within context.

12. Governance, Risk, and Ethics

Governance is the central requirement for responsible AI integration in irregular warfare [1]. Primary risks include algorithmic opacity, automation bias, data manipulation, privacy violations, escalation errors, accountability gaps, and overreliance on automated recommendations. These risks affect operational legitimacy, alliance trust, civilian protection, and strategic stability. Recent defense and security scholarship also emphasizes that AI adoption in warfare requires careful attention to operational reliability, legal accountability, and the changing character of modern conflict [12, 13].

AI governance should include several institutional mechanisms: centralized AI oversight bodies that coordinate strategy, doctrine, procurement, testing, and compliance; legal and ethical review before deployment; documentation of model purpose, data sources, limitations, and

known failure modes; adversarial testing and operationally realistic scenario evaluation; and maintained human decision authority for high-risk applications. Ethics must be operationalized rather than treated as a statement of principles, and international coordination is necessary because these operations cross borders.

13. Policy Implications

The five-layer model yields six policy implications. First, AI oversight should be centralized to establish institutional standards, approval processes, technical requirements, and accountability mechanisms, while remaining shaped by operational realities at the mission level. Second, AI literacy should become a core professional requirement for military and security personnel. Third, AI tools should be validated across operational contexts before mission use, addressing the generalization problem identified in AI-powered IoT cybersecurity research [6]. Fourth, critical infrastructure AI systems should use zero-trust principles, behavioral analytics, continuous monitoring, and cooperative governance because digital twins and Industrial IoT systems connect cyber activity to physical consequences [8]. Fifth, human judgment must remain central in high-risk applications, including targeting, influence operations, critical infrastructure control, and automated response. Sixth, public-private and allied coordination should be strengthened because much of the AI-relevant operating environment is owned or operated by private firms, and such threats cross national borders.

14. Proposed Institutional Implementation Framework

Figure 1 presents a phased implementation framework for translating the five-layer model into operational practice. The framework consists of six phases.

First, assessment involves identifying existing AI use cases, institutional capacities, data dependencies, and accountability gaps. Second, risk

classification categorizes AI systems as low, medium, or high risk based on their operational function, level of autonomy, potential for civilian harm, and escalation consequences. Third, governance design establishes review procedures, documentation requirements, audit mechanisms, approval authorities, and accountability structures for each risk category.

Fourth, operational testing evaluates AI systems in realistic environments, including adversarial scenarios, cross-domain transfer, and degraded-data conditions. Fifth, deployment and monitoring provide continuous oversight through incident reporting, performance review, model drift detection, and periodic reauthorization. Sixth, lessons learned and adaptation incorporates findings from exercises, operations, failures, and emerging threats so that institutional controls evolve alongside technological change and adversary behavior.

In this way, the framework creates a lifecycle approach to responsible AI management that connects assessment, risk classification, governance design, operational testing, deployment and monitoring, and lessons learned and adaptation into a continuous process for responsible AI integration in irregular warfare. The five-layer model identifies the governance priorities, while the six-phase framework shows how they can be implemented.

Figure 1. Proposed Institutional AI Governance Implementation Framework.

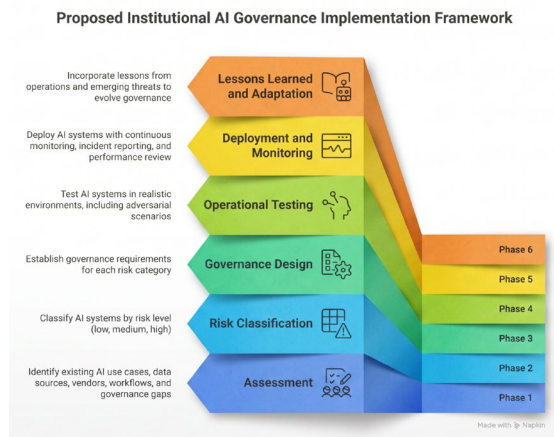


Figure 1. Proposed Institutional AI Governance Implementation Framework

15. Discussion

AI integration in irregular warfare presents a paradox: the same capabilities that improve speed, awareness, and adaptability can also increase opacity, instability, and overconfidence. The five-layer institutional integration model helps resolve this tension by linking innovation to oversight. The analytical matrix in Table 1 demonstrates that the model's governance requirements are not uniform across cases. Strategic alignment and doctrinal adaptation are critical in cyber and information operations, where escalation risks are highest. Data fusion and operational embedding dominate in the critical infrastructure case, where adversarial manipulation of digital twin state data creates direct physical risk. Human-AI governance is stressed across all four cases but is most acute in contexts involving influence operations and automated responses.

This article contributes a five-layer institutional integration model that links strategic alignment, doctrinal adaptation, operational embedding, data fusion, and human-AI governance for irregular warfare. Its main contribution is showing how AI oversight must be connected to operational use, especially in below-threshold, digitally

mediated, and civilian-connected security environments.

This model differs from existing frameworks in two respects. It integrates governance explicitly with operational embedding rather than treating oversight as a separate compliance function. It also applies systematically across irregular warfare domains, including non-state actor contexts and civilian infrastructure environments, rather than focusing exclusively on conventional military applications. Future research should develop quantitative metrics for each model layer, test the framework across additional cases, examine allied interoperability of responsible AI standards, and apply the model to hybrid warfare in the Indo-Pacific region and AI-supported economic coercion.

16. Conclusion

Artificial intelligence is transforming irregular warfare by accelerating decision cycles, expanding influence operations, enhancing cyber capabilities, and reshaping institutional processes. This paper developed a five-layer institutional integration model connecting strategic alignment, doctrinal adaptation, operational embedding, data fusion, and human-AI governance. Applied to four case contexts through the analytical matrix in Table 1, the model shows that governance requirements vary in emphasis across irregular warfare domains but are universally necessary.

The paper argues that AI integration must be treated as a governance challenge rather than a purely technical modernization effort. Defense and security institutions must preserve human judgment, legal accountability, ethical oversight, and strategic restraint while leveraging technological advantage. The future of irregular warfare will be shaped not only by which actors possess AI capabilities, but by which institutions can govern them responsibly, adaptively, and legitimately.

References

- [1] Boulanin, V., Verbruggen, M., & Goussac, N. Responsible military use of artificial intelligence. Stockholm International Peace Research Institute. 2020.
- [2] Horowitz, M. C., Scharre, P., & Chalk, P. Artificial intelligence and international security. RAND Corporation. 2023.
- [3] Rickli, J. M., & Mantellassi, F. Artificial intelligence in warfare. Routledge. 2023.
- [4] Jha, S., Solanki, V. K., & Dawson, M. (Eds.). Cognitive security for Industrial IoT: AI-driven approaches to Industry 4.0. Taylor & Francis. 2026. doi:10.1201/9781003609773.
- [5] Jones, L. The future of warfare is irregular. Fletcher Forum of World Affairs. 2022. 46: 107–120.
- [6] Quaye, S., Khan, A. H., Tapre, K., & Dawson, M. AI-powered cybersecurity models for training and testing IoT devices. Applied Sciences. 2025. 15(24): 13073. doi:10.3390/app152413073.
- [7] UNIDIR. Artificial intelligence and military decision making. United Nations Institute for Disarmament Research. 2023.
- [8] Dawson, M., Vassilakos, A., Moore, S., & Szakonyi, A. Beyond the mirror: Digital twin cybersecurity for national resilience. International Conference Knowledge-Based Organization. 2025. 31(1): 87–95. doi:10.2478/kbo-2025-0010.
- [9] Cummings, M. L. Artificial intelligence and the future of warfare. Chatham House. 2021.
- [10] Lin, H. Cyber-enabled information warfare and the ethics of AI manipulation. Ethics and Information Technology. 2022. 24(3): 211–225. doi:10.1007/s10676-021-09592-4.
- [11] Davis, A. L. Artificial intelligence and the fight against international terrorism. American Intelligence Journal. 2021. 38(2): 63–73.
- [12] Iqbal, S., Rizvi, S. W. A., Haider, M. H., & Raza, S. Artificial intelligence in security and defense. International Journal of Human and Society. 2023. 3(4): 341–353.
- [13] Chalagashvili, I. AI era in modern warfare. SSRN. 2024. doi:10.2139/ssrn.4813807.