International Journal of Disaster Risk Management • Vol. 8. No. 1 • https://doi.org/10.66050/xhmjy003

Publisher: Scientific-Professional Society for Disaster Risk Management

International Journal of Disaster Risk Management

Review article

Strategic Risks of AI-Enabled Automation: Implications for Crisis Management and Systemic Resilience

Dejan Vuletić[1]

[1] Strategic Research Institute, University of Defence, Veljka Lukića Kurjaka 33, 11042 Belgrade, Serbia.
dejan.vuletic@mod.gov.rs

* Correspondence: dejan.vuletic@mod.gov.rs; tel.: +381 63 610 744

Received: 3 January 2026 Revised: 27 February 2026; Accepted: 1 March 2026; Published: 3 March 2026.

ABSTRACT

Artificial intelligence is increasingly embedded in military decision-making systems, transforming not only operational performance but the structural conditions under which risk is produced and governed. This article develops a conceptual socio-technical analysis of AI-enabled automation through the lenses of systemic risk, risk governance, and resilience. Rather than treating AI as a discrete tool, the paper conceptualizes automation as a structural intervention in high-risk decision infrastructures. Four analytical findings are advanced. First, AI-mediated workflows redistribute practical control by shaping what information is salient, how threats are classified, and which actions appear feasible. Second, automation compresses decision time, reducing space for deliberation and increasing reliance on algorithmic outputs under uncertainty. Third, tighter coupling and accelerated interaction across domains enable localized errors or manipulations to propagate into broader escalation dynamics. Fourth, distributed human–machine architectures contribute to responsibility diffusion, weakening the alignment between accountability and effective decision influence. The article argues that prevailing "human-in-the-loop" safeguards are insufficient when control is reduced to technical override rather than institutional capacity for judgment and accountability. Effective integration of AI in high-risk environments requires governance arrangements that preserve human judgment as a strategic resource for crisis management and systemic resilience.

KEYWORDS

Artificial intelligence; military decision-making; risk governance; socio-technical systems; escalation risk.

1. Introduction

Artificial intelligence is rapidly integrating into military operations. This development is not merely technological; it represents a systemic transformation with implications for risk, crisis management, and institutional governance. Contemporary military systems are becoming deeply embedded in complex, tightly coupled socio-technical infrastructures in which algorithmic processes mediate critical functions such as intelligence analysis, targeting, operational planning, and com-

Academic Editor:
Prof. Dr. Vladimir M. Cvetković

e-ISSN2620-2786

Vuletić, D. (2026). Strategic Risks of AI-Enabled Automation: Implications for Crisis Management and Systemic Resilience, 8(1), 131-146.

mand support. In such environments, the central question is no longer whether AI improves performance. The key issue is how integration reshapes the emergence, distribution, and governance of risk in high-stakes decision systems. In this article, socio-technical systems are understood as integrated configurations of technical infrastructures, human actors, organizational practices, and institutional arrangements that jointly shape how decisions are produced, interpreted, and governed.

In contemporary military practice, AI systems already support a range of operational functions. These include multi-source intelligence fusion that integrates satellite imagery, signals intelligence, and open-source data; automated target recognition and object classification; predictive logistics optimization for maintenance and supply chains; cyber threat detection and anomaly identification; and decision-support interfaces that synthesize operational data under severe time pressure. While these systems expand analytical capacity and responsiveness, they also increase interdependence across operational layers. By linking previously distinct domains through shared data architectures and accelerating information flows, AI-enabled processes intensify system coupling, compress decision cycles, and create new pathways through which localized anomalies, misinterpretations, or adversarial manipulations may escalate into broader strategic consequences. Recent studies in disaster risk management show that artificial intelligence and related digital technologies increasingly support disaster prediction, situational awareness and response coordination, thereby transforming operational decision environments (Hanspal & Behera, 2024, pp. 135–136). The integration of advanced technologies such as artificial intelligence, sensor networks and communication platforms has become a central element of modern protection and rescue systems and contributes to more effective disaster response and risk management (Jovičić, Gostimirović, & Milašinović, 2024, pp. 111–114).

From a systemic risk perspective, the incorporation of AI into military decision-making must be understood in light of the dynamics of complex and tightly coupled systems. In this article, systemic risk refers to risk that emerges from structural interdependence, tight coupling, and nonlinear interaction within socio-technical systems, rather than from isolated component failure. Risk, in this view, is not an external disturbance to be eliminated, but a structural feature of system architecture arising from interactions among technical, organizational, and institutional components. Perrow's (1984) theory of normal accidents provides a foundational account of these dynamics, demonstrating that in highly interconnected socio-technical systems, failures are not exceptional events caused by isolated human or technical errors, but structurally produced outcomes of system interaction. Applied to AI-enabled military infrastructures, this perspective suggests that automation reconfigures the conditions under which cascading failures and unintended interactions emerge as normal features of system behavior.

Leveson (2011) argues that safety in complex socio-technical systems cannot be reduced to component reliability or individual human performance, but must be conceptualized as a problem of system-level control. Increasing automation does not inherently reduce risk. Instead, it redistributes and transforms it by altering how authority, information, and decision rights are structured across human and technical components. In military contexts, where decisions unfold under conditions of uncertainty, adversarial interaction, and political consequence, such transformations have direct implications for crisis stability and escalation management.

The growing reliance on AI-enabled decision support systems also reflects broader shifts in how contemporary risks are governed. Risk governance scholarship emphasizes that modern risks are characterized by complexity and uncertainty, rendering traditional technocratic models of prediction and control increasingly inadequate (Renn, 2008; Renn, Klinke, & van Asselt, 2011). Effective governance therefore depends less on accurate prediction than on institutional capacities for coordination, reflection, and accountability when knowledge is contested and system behavior is evolving.

Within globally networked systems, these challenges are further amplified by interdependence and nonlinearity. Helbing (2013) demonstrates that contemporary risks increasingly arise from the structural coupling of interconnected systems, in which small disturbances can trigger cascading effects across multiple domains. Military AI systems, embedded in global information networks and tightly coupled operational architectures, exhibit precisely these characteristics. Their performance and failure modes cannot be fully understood in isolation, as they interact dynamically within complex socio-technical systems. Contemporary disaster risk environments increasingly require integrated risk analysis and planning approaches capable of identifying high-risk conditions across complex socio-technical settings (Ulal, Saha, Gupta, & Karmakar, 2023, pp. 82–83).

In response to such systemic vulnerabilities, resilience has emerged as a complementary paradigm to traditional risk management. In this article, resilience is understood primarily in the sense of resilience engineering and governance resilience, emphasizing institutional capacities for monitoring, responding, learning, and adapting under conditions of stress and uncertainty. Rather than focusing exclusively on prediction and prevention, resilience highlights the capacity of socio-technical systems to absorb disturbances, adapt to changing conditions, and recover from unexpected shocks (Hollnagel et al., 2006; Linkov & Trump, 2019; Milenković, Cvetković, & Renner, 2025, p. 80). Applied to AI-enabled military systems, this perspective implies that the central challenge is not to eliminate risk through perfect automation, but to design institutional structures capable of maintaining control, learning from anomalies, and preventing cascading failures. Taken together, these strands of scholarship suggest that AI-enabled military decision-making should be analyzed not primarily as a problem of technological capability, but as a problem of systemic risk and governance. This article advances a strategic, governance-oriented analysis of AI-enabled automation in military operations. Rather than offering a technical assessment of specific AI applications, it conceptualizes AI as a systemic intervention in high-risk decision infrastructures.

The core argument of this article is that AI-enabled automation represents not merely a technological enhancement of military capability, but a structural transformation in governance that redistributes control, responsibility, and interpretive authority within socio-technical systems. The paper advances four theoretical findings: (1) automation redistributes practical control across human–machine networks; (2) it compresses decision time and reshapes judgment; (3) it amplifies systemic risk through tighter coupling and acceleration; and (4) it contributes to responsibility diffusion within distributed decision architectures. Together, these findings position AI as a governance transformation rather than a neutral operational tool. Recent debates on AI oversight in defense institutions similarly emphasize that the integration of artificial intelligence requires accountable governance structures, institutionalized review mechanisms, and clearly defined responsibility frameworks (Defense Innovation Board, 2019; Scharre, 2018). Recent multilateral assessments further identify malfunction risk, bias, opacity, and accountability gaps as central governance challenges in military AI systems, emphasizing the need for life-cycle oversight and precautionary safeguards (Unal & Richard, 2024).

This article is organized as follows. Section 2 (Methods) outlines the conceptual and socio-technical methodological approach that informs the analysis. Section 3 (Results) presents four core analytical findings on the systemic effects of AI-enabled automation on military decision-making and crisis governance. Section 4 (Discussion) examines the implications of these findings across six thematic subsections: military judgment as a strategic concept, algorithmic decision-making in military contexts, speed and time compression, strategic and operational risks of automation, the limitations of human-in-the-loop models and control beyond the loop, and implications for risk governance and crisis management. Section 5 (Conclusion) synthesizes the main insights and highlights the importance of institutional arrangements that preserve human judgment as a resource for stability, accountability, and systemic resilience.

2. Methods

This study adopts a qualitative, conceptual research design grounded in interdisciplinary theory and socio-technical systems analysis. The article does not employ empirical data collection or statistical modeling. Instead, it develops an analytical framework through theoretical synthesis of military studies, systems theory, risk governance, and crisis management. The methodological approach is based on three complementary components. The analytical findings presented in this study should be understood as theory-driven conceptual propositions derived from interdisciplinary synthesis, rather than as empirically tested results.

First, a conceptual analysis is used to clarify key constructs such as military judgment, algorithmic decision-making, systemic risk, and human control. Conceptual analysis allows the study to examine how meanings of judgment, control, and risk are constructed across different scholarly traditions and how these meanings shape interpretations of AI-enabled military systems.

Second, the article employs a theoretical synthesis of socio-technical systems literature. This synthesis integrates insights from Perrow's theory of tightly coupled systems, Leveson's systems-theoretic safety model, and organizational perspectives on how actors make sense of complex situations and gradually drift into failure. These frameworks provide the analytical lens through which AI-enabled automation is examined as a systemic intervention in high-risk decision infrastructures rather than as a discrete technological innovation.

Third, the study applies a risk governance perspective to interpret the strategic implications of algorithmic decision-making. Rather than assessing risk in probabilistic or technical terms, risk is conceptualized as an emergent property of institutional arrangements, control architectures, and organizational practices. This allows the analysis to focus on how AI reshapes the temporal, organizational, and political conditions under which military decisions are made and contested.

The analytical procedure follows a structured interpretive process in which key theoretical arguments are first identified in the relevant literature and then compared and connected in order to generate higher-level analytical categories, such as time compression, control redistribution, automation bias, and diffusion of responsibility. These categories are subsequently applied to contemporary military decision-making contexts to derive systemic insights into strategic risk and crisis governance. By combining conceptual analysis with interdisciplinary theoretical synthesis, the study aims to provide transferable analytical insights relevant both to military operations and to broader fields of disaster risk management and crisis governance.

This study does not rely on primary empirical data, case-based analysis, or statistical modeling. It is a conceptual and theory-driven inquiry aimed at clarifying the systemic implications of AI-enabled automation. The literature selection was guided by a governance-oriented research logic rather than exhaustive technical coverage. Priority was given to foundational works in systems theory (tight coupling, control structures), resilience engineering, risk governance scholarship, and organizational failure research. The analytical synthesis was structured around thematic keywords including "systemic risk," "AI governance," "socio-technical systems," "automation bias," "time compression," and "escalation dynamics." The objective was not to evaluate model-level performance or technical robustness of AI systems, but to develop an interdisciplinary analytical framework applicable to high-risk decision environments.

3. Results

The analysis yields four theory-driven analytical findings that clarify the structural effects of AI-enabled automation on military decision-making and crisis governance. First, the integration of AI systems restructures military decision-making by redistributing control across human–machine networks. Rather than operating as neutral decision-support tools, algorithmic systems actively shape what information is considered relevant, how threats are classified, and which options are presented as viable. This produces decision environments in which algorithmic outputs become primary reference points for human actors. As a result, the practical locus of decision authority shifts. For example, an AI-enabled intelligence fusion system may classify ambiguous sensor data as indicating hostile movement. Once this classification is integrated into shared operational dashboards, it can quickly become the epistemic reference point for subsequent decisions across command levels. Even if the original signal was uncertain or incomplete, the algorithmic framing shapes what is treated as credible threat information, thereby influencing action before alternative interpretations are fully explored.

Second, AI-enabled automation generates a condition of time compression that reduces the temporal space available for interpretation, deliberation, and judgment. As decision cycles accelerate, human involvement shifts from substantive evaluation toward procedural validation of algorithmic recommendations. This temporal transformation increases reliance on automated systems under precisely those conditions of uncertainty, ambiguity, and strategic consequence where interpretive judgment is most critical. In a time-critical operational setting, a commander reviewing an AI-generated targeting recommendation may focus primarily on confirming procedural compliance rather than substantively reassessing the system's output. Especially if the model has demonstrated high historical accuracy, the cognitive burden of re-evaluation under time pressure can encourage deference to the algorithm. Judgment thus shifts from critical interpretation toward rapid validation.

Third, the analysis demonstrates that AI systems function as strategic risk amplifiers within complex socio-technical infrastructures. Automation increases coupling, speed, and interdependence across operational domains. As a result, localized anomalies can spread quickly and evolve into systemic disturbances. Errors, misclassifications, or adversarial manipulations are more likely to cascade across interconnected systems, producing escalation dynamics that exceed the capacity of localized control mechanisms. Consider a scenario in which adversarial actors introduce subtle distortions into data streams feeding an automated detection system. Even minor input manipulation may shift model outputs sufficiently to trigger defensive reactions. In tightly coupled, high-speed systems, localized distortions can spread across operational domains before humans detect the anomaly.

Fourth, AI-enabled decision-making contributes to a diffusion of responsibility and accountability within military organizations. As decision-relevant functions are distributed across human operators, algorithmic models, and organizational routines, it becomes increasingly difficult to locate responsibility for specific outcomes. Formal accountability remains attached to human actors, while effective influence over decisions is exercised through opaque socio-technical processes. When a decision outcome emerges from layered algorithmic filtering, automated prioritization, and multiple human approvals, responsibility formally remains with identifiable commanders. However, no single actor may have exercised full interpretive control over the decision trajectory. This diffusion between influence and accountability illustrates how socio-technical architectures can obscure responsibility without eliminating it.

Taken together, these findings indicate that AI-enabled automation does not primarily reduce uncertainty or risk, but transforms their structural conditions. Automation shifts decision-making from bounded human judgment toward distributed socio-technical control systems characterized by acceleration, opacity, and tight coupling. The result is not enhanced predictability, but increased systemic vulnerability and escalation sensitivity.

4. Discussion

The findings of this analysis have significant implications for how AI-enabled military systems are conceptualized, governed, and evaluated. Rather than confirming dominant narratives that portray AI as a tool for increasing rationality, efficiency, and control, the results suggest that automation fundamentally reshapes the architecture of risk in high-stakes decision environments. The discussion that follows elaborates these implications by examining judgment, algorithmic mediation, time compression, systemic risk amplification, and the limits of loop-based control.

From a risk governance perspective, the redistribution of control across human–machine networks challenges traditional assumptions about command responsibility and decision authority. Military governance frameworks presuppose identifiable human agents capable of exercising judgment and bearing responsibility for outcomes. However, AI-enabled systems blur these boundaries by embedding decision-relevant functions within opaque and distributed socio-technical processes. This creates a structural mismatch between responsibility and control, in which humans remain accountable for outcomes that are increasingly shaped by algorithmic mediation.

The observed effects of time compression further complicate crisis management and escalation control. While speed is often framed as a strategic advantage, the analysis indicates that acceleration undermines institutional capacities for sensemaking, coordination, and interpretation. In crisis contexts, where misinterpretation and errors can have irreversible consequences, the reduction of deliberative space increases the likelihood of reactive and fragile decision-making.

The identification of AI as a strategic risk amplifier aligns with systemic risk theories that emphasize tight coupling and nonlinear interaction as core sources of vulnerability. Automation intensifies these dynamics by linking multiple operational domains through fast and opaque control structures. This transforms localized anomalies into potential systemic crises and increases the probability of unintended escalation driven by socio-technical interactions rather than deliberate choice.

These findings also highlight the limitations of prevailing governance models centered on loop-based human control. Mechanisms such as human-in-the-loop or human-on-the-loop implicitly assume that control can be exercised through technical intervention at discrete decision points. The analysis suggests that such models underestimate the organizational and institutional dimensions of control. They reduce control to procedural oversight instead of treating it as a continuous institutional process. Meaningful control requires not merely human presence, but organizational structures that enable continuous reflection, flexible interpretation of system outputs, and clear lines of responsibility within socio-technical systems.

From a resilience-oriented perspective, the implications are even more far-reaching. If AI-enabled military systems generate uncertainty and systemic risk, then governance strategies focused on optimization and prediction are inherently inadequate. Resilience requires institutional capacities for adaptation, learning, and recovery under conditions of surprise and disruption. In this context, human judgment helps stabilize accelerated decision-making by enabling deeper interpretation, awareness of institutional and strategic implications, and ethical restraint.

The broader relevance of these findings extends beyond military operations to disaster risk management and crisis governance more generally. AI-enabled systems are increasingly deployed in emergency response, critical infrastructure, and public sector decision-making, where similar patterns of acceleration, opacity, and distributed control are emerging (Beriša, Cvetković, & Pavić, 2024, pp. 280–281). The analytical framework developed in this article therefore contributes to a wider understanding of how algorithmic systems reshape risk governance in complex socio-technical environments. Overall, the discussion reinforces the central claim of the article: AI-enabled automation does not eliminate the need for human judgment, but intensifies its strategic importance by transforming the structural conditions under which risk, control, and responsibility are produced and governed.

4.1. Military Judgment as a Strategic Concept

In strategic and crisis contexts, judgment refers to the capacity of decision-makers to interpret ambiguous situations, integrate multiple sources of information, and assess the consequences of action under conditions of uncertainty and time pressure. Unlike procedural compliance or algorithmic optimization, judgment operates in environments characterized by incomplete knowledge, adversarial interaction, and the possibility of irreversible outcomes. For this reason, judgment constitutes a strategic resource embedded within socio-technical systems of command and control. Decision-making failures in high-risk systems rarely result from isolated human error but emerge from systemic interaction patterns among human actors, technological infrastructures, and institutional routines. As Reason (1990) argues, human error often reflects deeper structural conditions and latent organizational vulnerabilities.

This systemic understanding of judgment is reinforced by research on organizational failure and disaster. Vaughan's (1996) concept of normalization of deviance shows how repeated exposure to anomalies can gradually redefine risky conditions as acceptable, embedding danger within routine decision practices. Similarly, Weick's (1995) theory of sensemaking emphasizes that breakdowns in crisis environments arise not primarily from insufficient information, but from erosion of collective interpretive processes under ambiguity. Dekker's (2011) notion of drift into failure further illustrates how locally rational adaptations to performance pressures may progressively erode safety margins without being recognized as risky at the time. In such contexts, judgment is exercised within evolving socio-technical landscapes shaped by incremental organizational change.

Taken together, these perspectives suggest that military judgment cannot be conceptualized solely as an individual cognitive faculty operating in isolation from organizational and technological contexts. Rather, it is a distributed and systemic phenomenon emerging from interactions among institutional practices, technological systems, and political constraints. In crisis situations, its effectiveness depends less on the mere availability of data than on the capacity of organizations to sustain interpretive flexibility, recognize emerging anomalies, and resist normalization of risk.

This systemic view carries important implications for risk and crisis governance in military systems. Improving decision-making cannot be achieved solely through enhanced training or individual accountability; it requires institutional arrangements that preserve diversity of perspectives, enable critical reflection, and maintain sensitivity to early indicators. In strategic contexts, judgment also plays a crucial role in escalation management. Escalation risks often arise not from deliberate aggression, but from misinterpretation, misaligned expectations, and unintended system interaction. By situating technical assessments within broader institutional and strategic contexts and considering second- and third-order effects, judgment helps calibrate responses in ways that preserve stability. Preserving judgment as a central element of command is therefore not resistance to technological innovation, but a condition for responsible crisis governance in complex socio-technical environments.

4.2. Algorithmic Decision-Making in Military Contexts

The increasing deployment of artificial intelligence in military organizations has expanded algorithmic decision-making across intelligence analysis, target identification, logistics optimization, and operational planning. Although typically described as decision-support tools, these systems increasingly shape how information is processed, prioritized, and presented, thereby influencing the substance and timing of military decisions. Algorithmic systems thus participate in structuring decision environments rather than merely assisting human judgment.

From a systems-theoretic perspective, algorithmic decision-making must be understood in terms of control structures rather than component performance. Leveson's (2011) systems-theoretic model emphasizes that risk in complex socio-technical systems arises primarily from flawed system-level control and feedback structures rather than isolated technical malfunction. Applied to military AI, automation does not inherently enhance safety; instead, it redistributes decision authority across human operators, models, and organizational processes.

In algorithmically mediated systems, control emerges from interactions among data inputs, model outputs, human interpretation, and institutional routines. AI systems rely on historical data, yet military environments are characterized by novelty, adversarial adaptation, and rapid contextual change. Under such conditions, embedded model assumptions may diverge from operational realities, producing outputs that appear precise while resting on uncertain foundations. The governance challenge therefore lies less in ideal performance accuracy than in institutional capacity to detect and respond to system drift.

Organizational research further underscores that algorithmic systems are inseparable from the institutional contexts in which they operate. Herrmann and Pfeiffer's (2023) concept of "keeping the organization in the loop" highlights that effective control depends not only on individual oversight but on organizational practices and decision structures. Control is thus not a discrete human intervention but a distributed institutional process.

This embeddedness shapes automation bias and informal delegation. Under time pressure and informational overload, operators may defer to algorithmic recommendations, particularly when systems are perceived as highly reliable. Over time, such deference can produce de facto delegation: formal responsibility remains human, while effective influence shifts toward socio-technical processes that are difficult to scrutinize.

From a systemic risk perspective, the central concern is not occasional technical error but how algorithmic mediation transforms decision conditions. As authority becomes increasingly distributed across human–machine networks, institutional oversight and strategic coherence become more fragile. Failures may emerge not as sudden breakdowns but as gradual misalignments between system behavior and political objectives.

Algorithmic decision-making in military contexts should therefore be analyzed as a governance problem rather than as technical optimization. Integrating AI reshapes control structures in ways that affect crisis stability and escalation management. Preserving meaningful human judgment depends not on inserting humans into workflows, but on designing institutions capable of sustaining accountability and adaptive control under systemic uncertainty.

4.3. Speed, Time Compression, and Strategic Risk

One of the most frequently cited advantages of AI-enabled automation in military operations is speed. By processing large volumes of data and generating recommendations in near real time, algorithmic systems accelerate decision cycles and enhance operational responsiveness. In competitive security environments, speed is often equated with superiority, as the ability to act faster than an adversary is assumed to confer strategic advantage. From a systemic risk and crisis governance perspective, however, acceleration alters the conditions under which judgment, control, and escalation management are exercised.

Time compression reduces the temporal space available for interpretation and deliberation. As algorithmic systems generate outputs at increasing speed, decision-makers are pressured to respond within shorter timeframes, shifting judgment from substantive evaluation toward procedural validation of algorithmic recommendations. This transformation is particularly consequential in crisis situations, where ambiguity and incomplete information require careful sensemaking rather than rapid reaction.

Empirical research supports these concerns. Phillips-Wren and Adya (2020) show that time pressure, information overload, and uncertainty degrade decision quality by increasing cognitive errors, reliance on heuristics, and premature convergence on seemingly optimal solutions. In high-stakes environments such as warfare or disaster response, actors are therefore more likely to simplify complex situations and defer to automated outputs when speed becomes the primary performance criterion.

From a crisis governance perspective, time compression also reshapes escalation dynamics. Boin, ’t Hart, Stern, and Sundelius (2017) emphasize that effective crisis leadership depends not only on rapid response but on institutional capacities for sensemaking, coordination, and legitimacy management under uncertainty. Military environments are inherently adversarial and shaped by deception and incomplete information. Under severe time pressure, algorithmic systems may generate misclassifications or spurious correlations that are difficult to detect in real time. When such outputs are rapidly translated into action, minor anomalies can trigger disproportionate responses before human reasoning can intervene.

Time compression further alters the architecture of control in socio-technical systems. As decision cycles accelerate, the practical capacity for human oversight diminishes even when formal authority remains unchanged. Operators may remain nominally “in the loop,” yet lack the temporal and cognitive resources required to critically assess algorithmic recommendations. Intervention becomes increasingly symbolic, preserving the appearance of control while effective decision power migrates toward automated processes. Rapid interactions between automated components can generate feedback loops and cascading effects that outpace human comprehension, creating nonlinear escalation pathways that are difficult to govern.

These dynamics suggest that effective governance of AI-enabled military systems depends on how time is structured in decision processes. Rather than treating acceleration as an inherent advantage, institutions must recognize the strategic value of delay, redundancy, and friction in high-risk environments. Preserving temporal space for judgment and institutional control is therefore a strategic challenge, not merely a technical one.

4.4. Strategic and Operational Risks of AI-Enabled Automation

The integration of artificial intelligence into military operations introduces strategic and operational risks that cannot be fully captured by conventional notions of technical reliability or system performance. While AI-enabled automation is often framed as enhancing efficiency and responsiveness, its deployment in high-stakes socio-technical systems alters the structural conditions under which risk is produced and governed. From a systemic perspective, these risks are not exceptional side effects of malfunction, but emergent properties of tightly coupled decision infrastructures.

Perrow's (1984) theory of normal accidents provides a foundational framework for understanding these dynamics. In complex systems characterized by tight coupling and high interdependence, failures are structurally inevitable and cannot be fully prevented through improved design or oversight. Accidents arise from nonlinear interactions among system components, where small disturbances propagate rapidly and unpredictably, producing cascading effects that exceed localized control. Applied to AI-enabled military infrastructures, this perspective suggests that automation intensifies systemic risk by increasing interdependence, coupling, and speed across decision processes.

These vulnerabilities are reinforced by organizational dynamics. Dekker's (2011) concept of drift into failure captures how adaptation to performance pressures and efficiency demands gradually erodes safety margins. Failures emerge not from sudden breakdowns but from cumulative processes of local optimization that shift systems toward unsafe conditions. In military contexts, AI integration may accelerate such drift by embedding automation within established routines. As algorithmic tools become normalized, institutions may progressively redefine acceptable risk and delegate greater authority to automated systems without fully recognizing long-term implications for control and accountability.

From a strategic standpoint, these dynamics generate escalation risks. When algorithmic systems mediate critical functions such as threat detection and operational planning, errors or misinterpretations can have far-reaching effects. In tightly coupled systems, localized anomalies—such as misclassified data or adversarial manipulation—may trigger rapid responses that propagate across domains and levels of command. Such responses may be perceived as deliberate escalation rather than technical outputs, increasing the likelihood of unintended conflict. Escalation in these scenarios results less from intentional aggression than from interactions among automated processes operating under uncertainty and speed.

Helbing's (2013) analysis of globally networked risks further illuminates the strategic implications of AI-enabled automation. In highly interconnected systems, risk arises primarily from structural interdependence rather than isolated failure. Increasing integration and automation allow disturbances to spread quickly across domains. Local incidents can therefore escalate into systemic crises. Military AI systems are embedded within global information infrastructures and cyber networks exhibiting precisely these characteristics. Their consequences therefore extend beyond discrete operational settings and may reverberate across broader security ecosystems.

Another critical risk concerns the diffusion of responsibility within AI-enabled systems. Traditional command structures presuppose identifiable human decision-makers accountable for outcomes. Automation complicates this assumption by distributing decision-relevant functions across operators, models, and organizational processes. When actions are shaped by opaque systems whose outputs are collectively produced, responsibility becomes more difficult to locate. This diffusion may weaken incentives for caution and erode norms of command accountability.

Risks also emerge from adversarial strategies. Opponents may exploit dependencies on automation by manipulating data, generating deceptive signals, or provoking automated responses. Such strategies are particularly effective when systems operate at speeds that limit human oversight and when organizational cultures privilege technical outputs over judgment. The more tightly decision-making is coupled to algorithmic processes, the greater the incentive to target those processes as leverage points.

Taken together, these dynamics underscore the limits of viewing AI-enabled automation as a neutral enhancement of military capability. Automation reshapes decision architectures by increasing coupling, accelerating interaction, and redistributing authority across socio-technical systems. These transformations amplify cascading failures, unintended escalation, and responsibility diffusion—risks that are systemic rather than accidental. Managing these challenges therefore requires governance arrangements capable of monitoring system-level dynamics, preserving accountability, and sustaining adaptive control under uncertainty.

This analysis reinforces the central argument of the article: AI-enabled automation intensifies the strategic importance of human judgment by altering the architecture of risk and control within military systems. Rather than reducing uncertainty, automation often transforms and amplifies it by embedding decision-making within complex infrastructures. Preserving stability and legitimacy depends not on eliminating human involvement, but on strengthening institutional capacities for oversight and systemic risk governance.

4.5. Human Control Beyond the Loop

Debates on the governance of artificial intelligence in military systems frequently invoke concepts such as human-in-the-loop, human-on-the-loop, or human-in-command as safeguards against excessive automation. These formulations suggest that the presence of a human operator ensures control, accountability, and legitimacy. In practice, however, such models often reduce control to the ability to approve, override, or terminate algorithmic processes at discrete decision points. In complex socio-technical systems, control cannot be adequately understood as a procedural checkpoint appended to automated workflows.

In this article, control is conceptualized in three interrelated dimensions. First, operational override authority: the capacity to interrupt or modify algorithmic outputs at specific decision points. Second, institutional accountability: the alignment between decision influence and responsibility within command structures. Third, interpretive capacity: the ability of organizations to critically assess, contextualize, and question algorithmic recommendations. Reducing control to loop-based oversight conflates these dimensions and risks obscuring how formal human presence may coexist with diminished substantive influence. Recent scholarship on ethical governance of artificial intelligence in defence further demonstrates that translating high-level principles into practice requires explicit institutional choices regarding life-cycle models, auditing mechanisms, traceability standards, and accountability structures (Blanchard, Thomas, & Taddeo, 2025). These governance design decisions directly shape whether human oversight remains substantive or becomes merely procedural.

The limitations of loop-based control become particularly evident under conditions of time compression and systemic complexity. As AI systems generate recommendations at speeds that exceed human deliberative capacities, opportunities for meaningful intervention narrow. Operators may remain formally “in the loop,” yet lack the temporal, informational, or institutional resources required for critical assessment. Under such conditions, human involvement risks becoming symbolic, preserving the appearance of control without its substantive content.

Herrmann and Pfeiffer’s (2023) concept of “keeping the organization in the loop” underscores the importance of organizational practices, workflows, and decision structures in shaping how AI systems function in practice. Control is exercised through institutional routines and governance frameworks that determine how algorithmic outputs are interpreted and acted upon. The relevant unit of analysis is therefore not the human–machine interface alone, but the socio-technical organization as a whole. This broader conception has significant implications for accountability and legitimacy. As decision-relevant functions become distributed across humans, models, and institutional processes, it becomes increasingly difficult to articulate responsibility for specific outcomes or justify decisions in political and legal terms.

From a systemic risk perspective, the pursuit of perfect technical control is both unrealistic and counterproductive. Efforts to maximize predictability may produce brittle systems that fail catastrophically when confronted with novelty. Resilience-oriented approaches offer an alternative. Rather than eliminating uncertainty, resilience emphasizes institutional capacity to absorb disturbances, respond to anomalies, and maintain functionality under stress (Hollnagel et al., 2006; Linkov & Trump, 2019).

In strategic contexts, this broader conception of control plays a stabilizing role in escalation management. By introducing interpretive depth, political awareness, and ethical restraint into decision processes, human judgment counterbalances pressures of speed and automation. Institutional friction—often viewed as inefficiency—can function as a safeguard against precipitous action and unintended escalation. Effective governance of AI-enabled military systems therefore depends less on the formal presence of human operators at discrete decision points, and more on institutional designs capable of sustaining accountability, interpretive capacity, and adaptive control under systemic uncertainty.

4.6. Implications for Risk Governance and Crisis Management

The preceding analysis shows that AI integration in military decision-making cannot be understood solely through models of optimization or individual oversight. Escalation risk emerges as an institutional governance problem arising from interactions between automated control architectures and political decision structures. From a risk governance perspective, contemporary risks cannot be effectively managed through centralized, technocratic, or purely predictive control models. As Renn (2008) and Renn, Klinke, & van Asselt (2011) argue, modern risk environments are characterized by complexity, uncertainty, and ambiguity that limit purely technical approaches. AI-enabled military systems exemplify these conditions by embedding decision-making within tightly coupled infrastructures that generate nonlinear and emergent behaviors. Effective governance therefore requires institutional capacities for reflexivity, coordination, and continuous learning rather than reliance on algorithmic prediction alone.

A central implication is the need to reconceptualize control as a systemic and adaptive process. Theories of complex systems and normal accidents suggest that tightly coupled socio-technical systems generate irreducible uncertainty and cannot be fully stabilized through design. Governance must shift from exclusive emphasis on failure prevention. Institutions need adaptive capacities to detect anomalies, respond to disturbances, and recover from disruption. This approach aligns with resilience-oriented frameworks focused on system performance under stress (Hollnagel et al., 2006; Linkov & Trump, 2019). In crisis management, these insights challenge assumptions that speed and automation inherently enhance control. As Boin, ’t Hart, Stern, and Sundelius (2017) note, effective crisis leadership depends on sensemaking, coordination, communication, and legitimacy management under uncertainty. AI-enabled automation may undermine these capacities if responsiveness displaces interpretive judgment.

Organizational design represents another critical implication. Governance frameworks must recognize that risk emerges from interactions among technology, routines, and institutional incentives. As socio-technical systems evolve, locally rational adaptations may gradually erode safety margins and normalize risky practices. Effective governance therefore requires continuous monitoring, structured mechanisms for critical evaluation, and institutionalized diversity of perspectives in decision-making.

These implications extend beyond military contexts. AI-enabled decision systems are increasingly deployed in disaster response, critical infrastructure management, and emergency governance, where similar patterns of uncertainty, time pressure, and interdependence are present. The framework developed here thus contributes to a broader understanding of how AI reshapes risk governance across high-stakes domains.

Taken together, effective governance of AI-enabled socio-technical systems depends less on predictive accuracy and technical control than on institutional capacities for adaptation, learning, and resilience. Rather than seeking to eliminate uncertainty through automation, organizations must acknowledge uncertainty as intrinsic to complex systems. Human judgment, clearly defined accountability, and institutionalized mechanisms for critical reflection remain central resources for managing systemic risk and preserving legitimacy in algorithmically mediated decision environments.

Policy-oriented frameworks on military AI governance similarly emphasize multi-stakeholder approaches, life-cycle management, trust-building measures, and safeguards against destabilization effects associated with accelerated and opaque decision architectures (Afina & Persi Paoli, 2024).

These insights suggest several governance design principles for AI-enabled high-risk systems. First, escalation-sensitive contexts require preservation of deliberative space and, where appropriate, the intentional slowing of decision cycles. Second, independent challenge and review functions should be institutionalized to counter automation bias and epistemic convergence. Third, AI systems should incorporate traceability and auditability mechanisms that enable retrospective reconstruction of decision pathways. Fourth, decision-makers should be trained to recognize anomaly patterns and model failure modes rather than treating algorithmic outputs as default baselines. Finally, clear lines of responsibility must be assigned for both deployment decisions and operational outcomes. Together, these principles translate systemic analysis into actionable governance design.

These dynamics are not limited to military systems. In civilian disaster risk management, similar structural patterns are observable. Control redistribution appears in algorithm-supported dispatch systems and AI-assisted triage decisions. Time compression is evident in automated alert infrastructures and rapid-response protocols triggered by real-time data feeds. Systemic risk amplification emerges in tightly coupled critical infrastructures where localized disruptions cascade across energy, communication, and transport networks. Responsibility diffusion can arise when AI-supported public decisions obscure lines of accountability between technical systems and public authorities. Recognizing these parallels strengthens the relevance of the present framework for broader disaster risk governance contexts.

Despite these contributions, the present analysis remains subject to important limitations. First, the study is conceptual and theory-driven and does not include empirical testing, case-based validation, or quantitative assessment. The arguments advanced here should therefore be interpreted as analytical propositions rather than empirically verified claims. Second, the review of AI safety scholarship is selective and oriented toward governance and systemic risk perspectives rather than technical model robustness, verification methods, or engineering safeguards. Third, although the focus on military contexts enables examination of escalation dynamics and high-stakes decision environments, not all findings may fully generalize to civilian disaster governance systems characterized by different institutional structures and accountability regimes. Future research would benefit from empirical case studies, comparative institutional analysis, and interdisciplinary collaboration between governance scholars and technical AI researchers.

5. Conclusions

This analysis argues that artificial intelligence–enabled automation fundamentally reshapes military decision-making through the dynamics of systemic risk, risk governance, and socio-technical complexity. Rather than approaching AI as a discrete technological innovation or a set of technical tools, the analysis has conceptualized AI as a structural intervention in high-risk decision infrastructures. From this perspective, the central challenge is not whether algorithmic systems improve operational efficiency, but how their integration reshapes the emergence, distribution, and governance of risk in complex and tightly coupled socio-technical systems.

This analysis demonstrates that AI-enabled automation intensifies rather than diminishes the strategic importance of human judgment. In complex environments characterized by uncertainty, adversarial interaction, and time pressure, judgment cannot be reduced to rule application or algorithmic optimization. It functions as a systemic and institutional capacity for interpretation, sense-making, and responsibility that strengthens organizational abilities to recognize emerging anomalies, manage uncertainty and contain escalation risks. Attempts to replace judgment with automated decision-making risk undermining the conditions that make high-risk systems governable and legitimate.

These findings underscore that many of the risks associated with AI-enabled military systems are not accidental or exceptional, but structural in nature. Failures and unintended consequences emerge not primarily from technical malfunctions or individual errors, but from systemic properties such as tight coupling, acceleration, opacity, and distributed control. In this sense, AI does not simply introduce new vulnerabilities, but amplifies existing patterns of systemic risk by increasing interdependence, speed, and automation across decision processes.

This analysis exposes the limitations of prevailing approaches to human control, particularly loop-based models that frame oversight as a technical intervention at specific decision points. Such models underestimate the organizational and political dimensions of control, reducing it to procedural supervision rather than treating it as a continuous, responsibility-bearing governance process. Meaningful human control, as argued here, must be understood as an institutional capacity embedded in organizational practices, decision cultures, and governance frameworks.

Effective risk governance requires a shift from optimization and failure prevention toward resilience-oriented approaches that emphasize adaptation, learning and recovery. Rather than seeking to eliminate uncertainty through automation, institutions must cultivate capacities to operate under uncertainty by preserving interpretive flexibility, organizational reflexivity and institutional accountability.

This resilience-oriented perspective reconceptualizes the role of AI in high-risk systems. AI should not be understood as a substitute for human judgment or as a means of achieving perfect control, but as one component of broader socio-technical systems that remain inherently uncertain and vulnerable to systemic failure. The strategic challenge lies not in maximizing speed or efficiency, but in designing institutional arrangements that can absorb disturbances, detect emerging risks, and prevent cascading failures. Human judgment, in this sense, emerges not as a source of error to be minimized, but as a critical resource for sustaining stability, legitimacy, and responsible decision-making in complex environments.

While this article has focused on military decision-making, its analytical framework has broader relevance for disaster risk management and crisis governance more generally. AI-enabled systems are increasingly deployed in domains such as emergency response, critical infrastructure and environmental governance, where decisions are made under conditions of deep uncertainty and systemic interdependence. In all these contexts, the integration of algorithmic systems raises similar questions about control, accountability, and resilience. These findings advance a broader understanding of how AI reshapes risk governance in contemporary societies.

In conclusion, the strategic problem posed by AI-enabled automation is not how to replace human judgment with machines, but how to integrate technological capabilities without eroding the institutional foundations of responsible governance. As socio-technical systems become more complex and automated, managing systemic risk will depend less on prediction and optimization. It will depend more on institutions that sustain judgment, learning, and adaptive control. Treating judgment as a strategic resource, rather than as an obstacle to efficiency, offers a more robust and sustainable approach to governing AI in high-risk decision environments and to strengthening resilience in the face of systemic uncertainty.

Author Contributions: The author conceptualized the study, conducted the theoretical analysis and wrote the manuscript.

Funding: This research was supported by the project *"Changes in the Strategic Environment of the Republic of Serbia – Trends and Implications for National Security and Defense"*, funded by the Ministry of Defence of the Republic of Serbia.

Acknowledgements: The author acknowledges the use of Grammarly Premium and ChatGPT in the process of translating and improving the clarity and quality of the English language in this manuscript. The AI tools were used solely for language enhancement and were not involved in developing the scientific content. The author takes full responsibility for the originality, validity, and integrity of the manuscript.

Conflicts of Interest: The author declares no conflict of interest.

6. References

1. Afina, Y., & Persi Paoli, G. (2024). Governance of artificial intelligence in the military domain: A multi-stakeholder Perspective on priority areas. Geneva, Switzerland: United Nations Institute for Disarmament Research.
2. Beriša, H., Cvetković, V., & Pavić, A. (2024). Implications of artificial intelligence and cyberspace on risk management capabilities. International Journal of Disaster Risk Management, 6(2), 279–295. https://doi.org/10.18485/ijdrm.2024.6.2.18.
3. Blanchard, A., Thomas, C., & Taddeo, M. (2025). Ethical governance of artificial intelligence for defence: Normative tradeoffs for principle to practice guidance. AI & Society, 40, 185–198. https://doi.org/10.1007/s00146-024-01866-7.
4. Boin, A., 't Hart, P., Stern, E., & Sundelius, B. (2017). The politics of crisis management: Public leadership under pressure (2nd ed.). Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/9781316339756.
5. Defense Innovation Board. (2019). AI ethical principles for the Department of Defense. Washington, DC: U.S. Department of Defense.
6. Dekker, S. (2011). Drift into failure: From hunting broken components to understanding complex systems. Farnham, UK: Ashgate.
7. Hanspal, M. S., & Behera, B. (2024). Role of emerging technology in disaster management in India: An overview. International Journal of Disaster Risk Management, 6(2), 133–148. https://doi.org/10.18485/ijdrm.2024.6.2.9.
8. Helbing, D. (2013). Globally networked risks and how to respond. Nature, 497(7447), 51–59. https://doi.org/10.1038/nature12047.
9. Herrmann, T., & Pfeiffer, S. (2023). Keeping the organization in the loop: A socio-technical extension of human-centered artificial intelligence. AI & Society, 38, 1523–1542. https://doi.org/10.1007/s00146-022-01391-5.
10. Hollnagel, E., Woods, D. D., & Leveson, N. (2006). Resilience engineering: Concepts and precepts. Aldershot, UK: Ashgate.
11. Jovičić, R., Gostimirović, L., & Milašinović, S. (2024). Use of new technologies in the field of protection and rescue during disasters. International Journal of Disaster Risk Management, 6(1), 111–122. https://doi.org/10.18485/ijdrm.2024.6.1.8
12. Leveson, N. G. (2011). Engineering a safer world: Systems thinking applied to safety. Cambridge, MA: MIT Press.
13. Linkov, I., & Trump, B. D. (2019). The science and practice of resilience. Cham, Switzerland: Springer. https://doi.org/10.1007/978-3-030-04565-4.
14. Milenković, D., Cvetković, V., & Renner, R. (2025). Community resilience indicators based on the BRIC method. International Journal of Disaster Risk Management, 7(1), 79–104. https://doi.org/10.18485/ijdrm.2024.6.2.6.
15. Perrow, C. (1984). Normal accidents: Living with high-risk technologies. Princeton, NJ: Princeton University Press.
16. Phillips-Wren, G., & Adya, M. (2020). Decision making under stress: The role of information overload, time pressure, complexity, and uncertainty. Journal of Decision Systems, 29(2), 99–112. https://doi.org/10.1080/12460125.2020.1768680
17. Reason, J. (1990). Human error. Cambridge, UK: Cambridge University Press.
18. Renn, O. (2008). Risk governance: Coping with uncertainty in a complex world. London, UK: Earthscan.
19. Renn, O., Klinke, A., & van Asselt, M. B. A. (2011). Coping with complexity, uncertainty and ambiguity in risk governance: A synthesis. Ambio, 40(2), 231–246. https://doi.org/10.1007/s13280-010-0134-0.
20. Scharre, P. (2018). Army of none: Autonomous weapons and the future of war. New York, NY: W.W. Norton.

21. Ulal, S., Saha, S., Gupta, S., & Karmakar, D. (2023). Hazard risk evaluation of COVID-19: A case study. International Journal of Disaster Risk Management, 5(2), 81–101. https://doi.org/10.18485/ijdrm.2023.5.2.6.
22. Unal, B., & Richard, U. (2024). Governance of artificial intelligence in the military domain (UNODA Occasional Papers No. 42). New York, NY: United Nations Office for Disarmament Affairs.
23. Vaughan, D. (1996). The Challenger launch decision: Risky technology, culture, and deviance at NASA. Chicago, IL: University of Chicago Press.
24. Weick, K. E. (1995). Sensemaking in organizations. Thousand Oaks, CA: Sage.