Tadeusz ZIELIŃSKI
War Studies University
Military Faculty
t-zielinski@akademia.mil.pl
https://orcid.org/0000-0003-0605-7684
https://doi.org/10.34739/dsd.2025.02.09

BEYOND AUTOMATION: THE ETHICAL, DOCTRINAL, AND OPERATIONAL CHALLENGES OF HUMAN-AI COLLABORATION IN MILITARY DECISION-MAKING

ABSTRACT: The accelerating integration of artificial intelligence (AI) into military systems presents unprecedented opportunities and complex dilemmas, particularly in the domain of decision-making under pressure. This article critically explores the ethical, doctrinal, and operational challenges that arise when human-AI collaboration is employed in high-stakes military contexts. At the heart of the study is the question of how military organizations can construct human-AI teams that simultaneously enhance operational effectiveness, uphold legal and ethical standards, and maintain meaningful human control over critical functions. To address this challenge, the study adopts a qualitative and interdisciplinary methodology that synthesizes doctrinal analysis, particularly of NATO and U.S. Department of Defense frameworks, with in-depth case studies of AI-enabled systems used in military operations. It investigates how AI systems affect traditional decision-making processes by accelerating data synthesis, improving situational awareness, and enabling faster reaction times. However, it also highlights emerging risks, such as automation bias, the erosion of human moral agency, loss of interpretability of algorithmic decisions, and the increasing difficulty in assigning responsibility for outcomes. The findings underscore that without appropriate safeguards, AI could undermine ethical accountability and legal clarity. To mitigate these risks, the article proposes a comprehensive framework for designing effective human-AI teams, which includes the implementation of transparent system architectures, explainable AI (XAI) models, trust calibration strategies, adaptive training modules, and multi-layered oversight mechanisms. Special attention is given to the necessity of doctrinal adaptation, the cultural and institutional readiness of military organizations, and the role of normative principles in regulating machine autonomy. Ultimately, the article argues that responsible innovation in military AI must be grounded in ethical rigor, legal certainty, and strategic prudence. Only by embedding these values into the design, deployment, and governance of human-AI teaming can military institutions ensure that AI enhances, rather than compromises, the legitimacy and effectiveness of defense operations.

KEYWORDS: human-AI teaming, military decision-making, autonomous systems, ethical accountability, artificial intelligence in warfare

POZA AUTOMATYZACJĄ: ETYCZNE, DOKTRYNALNE I OPERACYJNE WYZWANIA WSPÓŁPRACY CZŁOWIEKA I SZTUCZNEJ INTELIGENCJI W PROCESIE PODEJMOWANIA DECYZJI WOJSKOWYCH

ABSTRAKT: Przyspieszająca integracja sztucznej inteligencji (AI) z systemami wojskowymi stwarza bezprecedensowe możliwości operacyjne, ale jednocześnie generuje złożone dylematy, zwłaszcza w zakresie podejmowania decyzji pod presją czasu. Celem artykułu jest identyfikacja wyzwań etycznych, doktrynalnych i operacyjnych, które pojawiają się w sytuacjach wysokiego ryzyka, gdy stosowana jest współpraca człowieka z systemami AI. Problem badawczy sformułowano w postaci pytania: w jaki sposób organizacje wojskowe mogą konstruować zespoły człowiek–AI w taki sposób, aby jednocześnie zwiększać skuteczność operacyjną, zapewniać zgodność z normami prawnymi i etycznymi

oraz utrzymywać realną i znaczącą kontrolę człowieka nad kluczowymi decyzjami? W celu rozwiązania tego problemu badawczego autor zastosował analizę jakościową, łączącą aspekty doktrynalne – w szczególności dokumentów NATO i Departamentu Obrony Stanów Zjednoczonych – z pogłębionymi studiami przypadków zastosowań wojskowych systemów opartych na AI. Analiza koncentruje się na wpływie tych technologii na tradycyjne procesy decyzyjne, w tym na przyspieszenie syntezy danych, poprawę świadomości sytuacyjnej i zwiększenie szybkości reakcji. Równocześnie wskazuje jednak na nowe zagrożenia, takie jak uprzedzenia automatyzacji, erozja sprawczości moralnej człowieka, ograniczona interpretowalność decyzji algorytmicznych oraz narastające trudności w przypisywaniu odpowiedzialności za ich skutki. Wyniki badań podkreślają, że brak odpowiednich zabezpieczeń może prowadzić do osłabienia przejrzystości prawnej i rozliczalności etycznej. W odpowiedzi na te zagrożenia artykuł proponuje kompleksowe ramy projektowania skutecznych zespołów człowiek–AI, obejmujące m.in. przejrzystą architekturę systemową, modele sztucznej inteligencji objaśnialnej (XAI), mechanizmy kalibracji zaufania, adaptacyjne moduły szkoleniowe oraz wielowarstwowe mechanizmy nadzoru. Szczególną uwagę poświęcono potrzebie dostosowania doktryn wojskowych, gotowości kulturowej i instytucjonalnej organizacji wojskowych oraz roli zasad normatywnych w regulowaniu autonomii maszyn. Ostatecznie autor argumentuje, że odpowiedzialna innowacja w zakresie zastosowań AI w środowisku wojskowym musi być zakorzeniona w rygorze etycznym, pewności prawnej i strategicznej dalekowzroczności. Tylko poprzez włączenie tych wartości do projektowania, wdrażania i regulowania współpracy człowieka i AI można zagwarantować, że sztuczna inteligencja wzmocni, a nie osłabi, legitymację i skuteczność operacyjną działań obronnych.

SŁOWA KLUCZOWE: współpraca człowieka i sztucznej inteligencji, proces podejmowania decyzji wojskowych, systemy autonomiczne, odpowiedzialność etyczna, sztuczna inteligencja w działaniach bojowych

INTRODUCTION

The rapid development of AI technologies is transforming military operations, fostering new collaborations between human decision-makers and intelligent systems. As warfare becomes increasingly complex, fast-paced, and data-rich, traditional command structures and decision-making processes are strained in managing risks, analyzing data, and adapting to emerging threats. Consequently, militaries worldwide are integrating AI as a cognitive partner, not just for data processing or automation, but as an active participant capable of influencing, recommending, or initiating actions across various functions.

Central to this shift is the concept of human-AI teaming: a collaborative partnership where humans and intelligent agents work together in decision-making, tasks, and planning. Unlike earlier automation, which followed strict rules under human control, modern AI systems feature adaptive learning, probabilistic reasoning, and autonomous operation in changing environments. This evolution raises essential questions about the distribution of authority, team design, ethical standards, and accountability, especially as the boundaries between human control and machine autonomy become increasingly blurred.

At the core of this paper is the question: How can military organizations design human-AI teams that enhance operational performance while ensuring accountability and compliance with ethical standards? This question addresses both strategic needs and ethical concerns. Strategically, there is increasing acknowledgment that AI provides significant advantages in speed, precise targeting, and predictive analysis. However, these capabilities must also align with commitments to lawful conduct, ethical limits, and democratic control over the use of force. To

address this tension, we need a comprehensive approach – one that considers not just the technical benefits of AI but also the moral and legal responsibilities of military leadership.

The potential of AI in military settings is significant. Intelligent systems can analyze large amounts of sensor data, identify patterns that are invisible to humans, simulate outcomes with complex variables, and deliver actionable insights within tight decision-making timeframes. In areas such as missile defense, cyber warfare, and intelligence, surveillance, and reconnaissance (ISR), these capabilities can enhance situational awareness, reduce operator workload, and improve mission success. Additionally, AI's capacity for continuous operation, unaffected by fatigue or emotional states, makes it a valuable partner in high-pressure situations.

Yet delegating decision-making authority to machines, especially in lethal or politically sensitive situations, introduces significant risks. One of the most critical is the potential erosion of human moral agency. AI systems, regardless of their complexity, lack the ability for ethical reasoning, empathy, or legal judgment. When they are given the power to recommend or make decisions involving life-and-death outcomes, they make accountability more complex. Who is responsible if an autonomous system misidentifies a target, breaks a rule of engagement, or escalates a conflict based on flawed data? Traditional models of command responsibility, based on hierarchical authority and human intent, are not suitable to handle these new uncertainties.

This paper examines the historical, operational, ethical, and policy aspects of human-AI teaming. It begins by tracing the development of AI integration in military settings, from early automation tools to today's adaptive, semi-autonomous systems. The analysis then shifts to the real-world challenges of human-AI cooperation in high-pressure situations, where data uncertainty, time constraints, trust calibration, and interface design are crucial in shaping outcomes. In these scenarios, decisions made under stress can have ripple effects across different theaters of operation and political arenas, making the stakes both tactical and strategic.

Equally important are the ethical and doctrinal implications of shared decision-making authority. As AI systems become more integrated into decision processes, military doctrines must adapt to reflect the realities of distributed cognition and machine-assisted judgment. This adaptation involves reevaluating the principle of meaningful human control, redefining the scope of legal responsibility, and updating ethical training to prepare officers for the moral complexities of working with non-human agents. At the same time, policy frameworks must address the interoperability challenges in multinational operations, where different standards, trust levels, and doctrinal assumptions can complicate the integration of AI.

The final section of this inquiry presents design principles and policy recommendations for creating effective human-AI teams. These include keeping human oversight over essential decisions, improving the transparency and interpretability of AI outputs, integrating ethical reasoning into system structures and training programs, and establishing layered oversight mechanisms to ensure traceability and institutional accountability. These recommendations are based on normative frameworks that emphasize the protection of human dignity, adherence to international humanitarian law, and the preservation of democratic oversight in deploying lethal force.

Ultimately, this study aims to add to the ongoing discussion about AI in defense by providing a thorough, multidisciplinary analysis of the challenges and opportunities associated with human-AI teaming. It recognizes that integrating AI into military decision-making is not just a technical task, but a profoundly political, ethical, and strategic one. The future of armed conflict will depend not only on the capabilities of machines but also on the decisions humans make when designing, deploying, and regulating those machines. By clarifying these choices and suggesting ways for responsible innovation, this research intends to help military leaders, policymakers, engineers, and scholars navigate the complex landscape of war in the age of intelligent systems.

THE EVOLUTION OF HUMAN-AI TEAMING IN MILITARY OPERATIONS

The concept of human-AI teaming in military operations has undergone significant evolution over the past twenty years, driven by advances in artificial intelligence technology and the changing strategic needs of modern warfare. Early in this development, automated systems were created to support human operators by performing clearly defined, deterministic tasks. These systems, including autopilots, radar-controlled weaponry, and early command-and-control tools, worked on a strict input-output basis and were limited to specific functions. Although they enhanced operational efficiency and eased the cognitive workload for human operators, these technologies were seen as tools rather than true teammates.

The shift from automation to autonomy marks a fundamental change in military technology and strategy. While automation involves carrying out predefined instructions, autonomy refers to the ability for systems to make independent, adaptable decisions. As Lyons et al. highlight, the psychological and practical differences between these two ideas are significant in human-machine interaction. The primary difference lies in the expectation of agency: automated systems are overseen and controlled, whereas autonomous systems operate, learn, and adjust within the confines of mission goals and operational settings. This progress has been driven by advances in machine learning, sensor fusion, natural language processing, and robotics, technologies that collectively enable systems to operate with minimal human input[1].

Military interest in human-autonomy teaming (HAT) has primarily been driven by the need to enhance mission effectiveness amid increasing complexity and uncertainty. The integration of AI agents into operational units aims to boost response times, minimize human error, and support continuous mission execution in degraded or denied environments. Unlike traditional man-machine relationships, HAT focuses on interdependence and shared intent. Autonomous agents are no longer passive recipients of commands; instead, they actively participate, capable of executing tasks, making recommendations, and even initiating actions based on

[1] J.B. Lyons, K. Sycara, M. Lewis, A. Capiola, *Human–Autonomy Teaming: Definitions, Debates, and Directions*, "Frontiers in Psychology", 2021, 12, p. 1-15.

situational assessments. As a result, the human role shifts from direct control to supervision, guidance, and higher-level decision-making.

A key challenge in building effective human-AI teams involves defining the division of labor between humans and machines. Sycara and Lewis outlined three possible roles for AI systems in teams: tools that assist individual operators, agents that serve as team members, and systems that promote coordination within the team. Initially, military AI applications were primarily limited to the first category, decision-support systems that enhanced situational awareness by providing information[2]. However, as O'Neill et al. describe, the research increasingly shows a shift toward the second and third categories, where AI agents are considered vital team members and facilitators of team coordination[3].

The increasing complexity of military operations has sped up this shift. Modern battlefields feature quick information exchanges, multi-domain interactions, and high-stakes decisions under tight time constraints. In these settings, the cognitive limits of human operators become more apparent. AI systems are particularly well-suited for data-intensive tasks, such as analyzing satellite images, identifying anomalies in communication patterns, and simulating enemy behavior. These abilities enable AI to not only offer options to commanders but also to create, assess, and modify operational plans in real-time. This broader functionality raises key questions about trust, responsibility, and control in human-AI teams.

A key factor in the effectiveness of HAT is human trust in AI agents. As Mayer notes, trust is not a fixed state but a dynamic one influenced by system transparency, reliability, and past performance. Too little trust results in underusing AI, while too much trust can lead to overdependence and operator complacency[4]. The concept of automation bias, as discussed by French and Lindsay, highlights how operators may rely on algorithmic outputs even when these conflict with their judgment. In combat, where mistakes can be deadly, such biases pose serious operational and ethical risks[5].

Case studies from recent conflicts and military experiments highlight both the opportunities and risks of human-AI teaming. The U.S. Department of Defense's Project Maven, for example, utilized AI systems to assist analysts in identifying objects of interest in drone footage. The system significantly reduced workload and boosted detection rates, but it also showed the ongoing need for strong human oversight to interpret ambiguous or context-dependent

[2] K. Sycara, M. Lewis, *Integrating intelligent agents into human teams*, [in:] E. Salas and S.M. Fiore (eds), *Team cognition: Understanding the factors that drive process and performance*, Washington 2004, pp. 203–231, https://content.apa.org/books/10690-010 (20.06.2025).
[3] T.A. O'Neill, Ch. Flathmann, N.J. McNeese, E. Salas, *21st Century teaming and beyond: Advances in human-autonomy teamwork*, "Computers in Human Behavior", 2023, 147.
[4] M. Mayer, *Trusting machine intelligence: artificial intelligence and human-autonomy teaming in military operations*, "Defense & Security Analysis", 2023, 39(4), pp. 521–538,.
[5] S.E. French, L.N. Lindsay, *Artificial Intelligence in Military Decision- Making. Avoiding Ethical and Strategic Perils with an Option-Generator Model*, [in:] B. Koch, R. Schoonhoven (eds), *Emerging Military Technologies: Ethical and Legal Perspectives*, Boston 2022, pp. 53–74, https://brill.com/view/book/edcoll/9789004507951/front-1.xml (24.06.2025).

findings[6]. Similarly, the British Army's experiments with AI-enabled logistics systems have demonstrated how predictive maintenance can enhance operational readiness, although integrating these systems into legacy command structures remains challenging[7].

Empirical studies reviewed by O'Neill et al. show that successful human-AI teams rely on several mediating factors, including team makeup, task setup, training, and interface design. Interpersonal dynamics, which are typically absent in traditional human-automation interactions, become more significant as AI agents assume roles that require coordination, adaptation, and joint problem-solving. For instance, in simulations involving autonomous aerial swarms, human operators who viewed AI agents as competent, communicative, and aligned with their goals were more likely to work effectively with them. This suggests that, under certain circumstances, anthropomorphizing AI agents, although debated, may improve team performance[8].

The doctrinal implications of these developments are extensive. Traditional command-and-control systems rely on hierarchical structures and clear lines of authority. In contrast, successful human-AI teams often need flatter organizational structures where decision-making is shared among human and non-human participants. This change necessitates new training methods, revised engagement rules, and updated legal frameworks to maintain accountability as decision-making roles become more distributed and collaborative.

Equally important are the ethical aspects of human-AI teaming, which require thorough analysis. As Nalin and Tripodi point out, granting decision-making power to AI systems raises concerns about the erosion of human moral responsibility. Autonomous agents, without the ability for ethical reasoning and empathy, may still be assigned tasks, such as targeting decisions, that have life-or-death outcomes. This highlights the ongoing need to preserve meaningful human control, a principle widely supported in international legal and policy discussions[9].

The development of human-AI teaming in military operations reflects a broader shift in how military power is understood, organized, and wielded. As AI systems become increasingly capable, they will not only support human decision-making but also transform the core of military structure and strategy. However, this shift must be managed carefully to ensure that technological advancements support human judgment rather than replace it. Only through intentional design, thorough evaluation, and ongoing ethical reflection can human-AI teams reach their full potential, boosting operational effectiveness while maintaining accountability and legitimacy in warfare.

---

[6] R. Choudhury, *Project Maven: The epicenter of US' AI military efforts*, 2.03.2024, https://interestingengineering.com/military/project-maven-the-epicenter-of-us-ai-military-efforts (24.06.2025).

[7] P. Hinton, *Ones and Zeros: The British Army Unveils its Digital and Data Plan*, 16.05.2023, https://www.rusi.orghttps://www.rusi.org, 24.06.2025 (24.06.2025).

[8] T. O'Neill, N. McNeese, A. Barron, B. Schelble, *Human–Autonomy Teaming: A Review and Analysis of the Empirical Literature*, „Human Factors: The Journal of the Human Factors and Ergonomics Society", 2022, vol. 64, no 5, pp. 904–938.

[9] A. Nalin, P. Tripodi, *Future Warfare and Responsibility Management in the AI-based Military Decision-making Process*, "Journal of Advanced Military Studies", 2023, 14(1), pp. 83–97.

OPERATIONAL CHALLENGES IN HIGH-PRESSURE DECISION-MAKING ENVIRONMENTS

The integration of artificial intelligence into military decision-making has led to significant changes in how operations are conducted, especially in high-pressure situations. As armed forces worldwide implement AI-supported systems in both tactical and strategic settings, the potential for quicker and more precise decisions comes with significant operational, cognitive, and ethical challenges. These issues become even more serious in combat environments characterized by uncertainty, limited time, and incomplete information[10].

Proponents of AI integration argue that intelligent systems provide unmatched capabilities in improving situational awareness, synthesizing complex battlefield data, and lightening the cognitive burden on commanders. AI systems can analyze large, diverse datasets in real-time, detect patterns, and produce probabilistic predictions that may escape even the most skilled human analysts. As demonstrated in recent operational studies and conflict simulations, AI decision-support systems (AI-DSS), such as ALPHA, R-Plan, and JADE, can significantly accelerate the Observe-Orient-Decide-Act (OODA) loop, particularly in air and joint operations[11]. In theory, this speedup provides a vital strategic edge in conflicts involving swift maneuvers, electronic warfare, and sensor overload.

Yet, the speed enabled by AI systems can paradoxically introduce new risks. Critics argue that dependence on AI in high-pressure situations may weaken human judgment, reduce accountability, and increase the chance of unintended escalation. As Johnson highlights, automating decision-making, particularly during the orientation and decision phases of the OODA loop, risks replacing human abilities in interpretation, ethical reasoning, and improvisation in response to circumstances. Human cognition remains uniquely capable of understanding the nuances of intent, detecting deception, and evaluating the political consequences of actions in ways that AI, limited by its training data and structural constraints, cannot match[12].

This problem is worsened by the often opaque nature of many AI systems. Specifically, machine learning applications usually conceal the reasoning behind decisions within complex model structures, making it challenging for human operators to interpret or explain system outputs. Such opacity challenges the idea of meaningful human control, creating two opposite risks: over-reliance, known as automation bias, and under-use, resulting from skepticism about the algorithm. Wischnewski et al. demonstrate that trust calibration, aligning user trust with

[10] A. Choudhary, A. Surbhi, *AI arbitration – Charting the ethical and legal course*, presented at the ETLTC2024 International Conference Series On ICT, Entertainment Technologies, And Intelligent Information Management In Education And Industry Aizuwakamatsu, Japan 2024, p. 020031, https://pubs.aip.org/aip/acp/article-lookup/doi/10.1063/5.0234731 (25.06.2024).

[11] A.O. Morozov, V.O. Yashchenko, *Decision-making technologies in military systems. Challenges and prospects*, „Mathematical machines and systems", 2023, vol .4, pp. 3–10, http://www.immsp.kiev.ua/publications/articles/2023/2023_4/04_23_Yashchenko.pdf (25.06.2025).

[12] J. Johnson, *Automating the OODA loop in the age of intelligent machines: reaffirming the role of humans in command-and-control decision-making in the digital age*, "Defence Studies", 2023, 23(1), pp. 43–67.

a system's true reliability, is crucial for effective human-machine collaboration. Yet, in high-pressure situations, trust is usually fragile and easily affected by stress, time pressure, and organizational culture[13].

Data quality is another critical weakness in AI-enabled decision-making. AI systems are only as dependable as the data on which they are trained and rely during deployment. In fast-changing combat situations, data can be incomplete, outdated, or manipulated by enemies. As the CSET report warns, the risk that AI systems will generate misleading results because of skewed, limited, or adversarial data is high. In such cases, commanders may be misled into trusting AI recommendations too heavily, leading to decisions that could be incorrect or hazardous. Additionally, enemies might exploit AI flaws through methods such as data poisoning or sensor spoofing, thereby manipulating outputs without being detected[14].

Another operational concern focuses on how AI systems are integrated organizationally and procedurally. Introducing AI-DSS into current command-and-control structures often clashes with hierarchical military cultures and strict doctrinal frameworks. Military decision-making is not only a matter of cognitive processes; it also involves institutional layers, such as authorization, legal review, and interagency coordination. AI systems that produce options outside accepted norms or offer unfamiliar outputs might be ignored, misunderstood, or misused. This organizational resistance can undermine the very advantages of speed, insight, and foresight that AI systems aim to provide[15].

A deeper philosophical and epistemological critique examines the nature of decision-making itself. As Johnson and Erskine and Miller argue, strategic decisions in warfare often rely on human intuition, moral judgment, and the ability to act in the face of ambiguity and ethical risk. AI, by contrast, performs well in environments where objectives and constraints can be clearly defined[16]. However, warfare is full of contradictory information, emotional stakes, and adversarial actions designed to confuse and provoke. Relying on machines to make decisions in such contexts risks not only producing mistaken outcomes but also undermining the moral agency and political responsibility that should support the use of force[17].

Nevertheless, advocates of carefully scoped AI integration emphasize that human-machine teams, when properly designed, can outperform both unaided humans and fully autonomous systems. This superior performance relies on transparent interfaces, strong oversight, and clear

[13] M. Wischnewski, N. Krämer, E. Müller, *Measuring and Understanding Trust Calibrations for Automated Systems: A Survey of the State-Of-The-Art and Future Directions*, „Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems", presented at the CHI '23: CHI Conference on Human Factors in Computing Systems ACM, Hamburg Germany 2023, pp. 1–16, https://dl.acm.org/doi/10.1145/3544548.3581197 (26.06.2025).
[14] Center for Security and Emerging Technology, *CSET's 2024 Annual Report*, Center for Security and Emerging Technology 2025, https://cset.georgetown.edu/publication/2024-annual-report/ (26.06.2025).
[15] A. Shutov, *Artificial Intelligence and International Security*, „International Affairs", 2024, vol. 70, no 005, pp. 254–256, http://dlib.eastview.com/browse/doc/99985268 (25.06.2025).
[16] J. Johnson, *The AI Commander Problem: Ethical, Political, and Psychological Dilemmas of Human-Machine Interactions in AI-enabled Warfare*, "Journal of Military Ethics", 2022, 21(3–4), pp. 246–271.
[17] T. Erskine, S.E. Miller, *AI and the decision to go to war: future risks and opportunities*, "Australian Journal of International Affairs", 2024, 78(2), pp. 135–147.

doctrines. As Meimandi et al. argue, HAT should be viewed as a socio-technical system that includes cognitive models of stress, trust, and adaptation. In such systems, AI agents serve not as replacements for human decision-makers but as force multipliers, providing hypotheses, detecting anomalies, and managing data complexity, while leaving the final decisions to human operators[18].

ETHICAL AND DOCTRINAL IMPLICATIONS OF SHARED DECISION AUTHORITY

As artificial intelligence systems become more integrated into military operations, the division of decision-making authority between humans and machines raises complex ethical and doctrinal issues. At the heart of this challenge is the question of how much control should be given to autonomous systems, especially in situations involving lethal force, strategic escalation, or quick tactical decisions[19].

Supporters of shared decision-making authority argue that autonomous systems, when properly constrained and guided, can enhance operational accuracy, alleviate cognitive overload for human operators, and mitigate the risk of human error under stress. AI agents can operate at speeds and scales beyond human limits, quickly processing sensor data, anticipating adversary actions, and suggesting optimal courses of action. In areas such as missile defense or cyber intrusion response, delays caused by relying solely on human decision-making could compromise mission success or national security. As a result, military doctrine has started to adopt "human-on-the-loop" rather than "human-in-the-loop" setups, especially in situations where rapid response is critical[20].

Advocates also highlight AI's potential to enhance ethical decision-making by mitigating the impact of emotions, stress, or bias, factors that can compromise human judgment in combat. An AI system programmed to follow the rules of engagement strictly might prevent impulsive or retaliatory actions that could violate the principle of proportionality. For example, algorithms can be designed to detect civilian structures or sensitive targets and reject strike proposals that pose excessive risks. In this way, AI can help achieve the broader strategic goal of reducing collateral damage while upholding military effectiveness[21].

However, critics warn that speeding up decision-making through AI delegation might compromise ethical accountability and legal clarity. In lethal situations, key principles of international humanitarian law, distinction, proportionality, and necessity, require contextual reasoning and moral judgment that current AI systems cannot provide. As Rashid et al. note, even the most

[18] K.J. Meimandi, M.L. Bolton, P.A. Beling, *Human-Agent Teaming: A System-Theoretic Overview*, Preprints 26.01.2024, https://www.techrxiv.org/users/716315/articles/701117-human-agent-teaming-a-system-theoretic-overview?commit=05d58d706b35a50d12dcd7bac3feefe5f0440723 (25.06.2025).
[19] M. Bistron, Z. Piotrowski, *Artificial Intelligence Applications in Military Systems and Their Influence on Sense of Security of Citizens*, "Electronics", 2021, 10(7), p. 871.
[20] N.J. McNeese, Ch. Flathmann, T.A.O'Neill, E. Salas, *Stepping out of the shadow of human-human teaming: Crafting a unique identity for human-autonomy teams*, "Computers in Human Behavior", 2023, 148.
[21] J.E. Márquez Díaz, *Benefits and challenges of military artificial intelligence in the field of defense*, "Computación y Sistemas", 2024, 28(2), pp. 309-323.

advanced autonomous systems lack consciousness and intentionality, raising serious concerns about responsibility in cases of civilian casualties or violations of the law of armed conflict[22].

One of the most urgent ethical concerns is the loss of human moral agency. Giving machines control over life-and-death decisions risks spreading accountability across human-machine networks, where no single person can be held responsible for results. The 2024 CSET report highlights that the lack of transparency in AI decision-making, particularly in black-box neural networks, makes it more challenging to assign blame or explain actions following an incident. Even when humans technically have authority, automation bias can cause operators to unquestioningly trust algorithmic outputs, effectively passing ethical responsibility to machines[23].

Real-world examples demonstrate these dilemmas. During operations in Syria, the use of semi-autonomous loitering munitions revealed a concerning pattern: human operators increasingly accepted machine-generated target designations with minimal oversight, prioritizing speed and operational tempo over verification. Similarly, in DARPA OFFSET program simulations, human supervisors often failed to override AI decisions, even when those decisions conflicted with established rules of engagement, exposing the psychological and procedural barriers to maintaining meaningful human control[24].

Broader concerns about the cultural and institutional impacts of machine delegation also appear in the literature. Military ethics, which are typically grounded in human virtues such as courage, prudence, and responsibility, face risks of erosion when responsibility is delegated through algorithms. These issues become even more significant in multinational operations, where different legal systems and normative frameworks make it harder to maintain shared accountability[25].

Doctrinally, military organizations are grappling with how to define the limits of machine agency. Frameworks such as NATO's doctrine on autonomous systems[26] and the U.S. Department of Defense's Directive 3000.09 specify that humans must maintain ultimate responsibility for decisions involving the use of force[27]. However, the practical integration of AI into targeting, surveillance, and command functions creates ambiguities that these frameworks find difficult to resolve. As McNeese et al. argue, command-and-control models need a fundamental rethinking to address the unique roles, capabilities, and liabilities of AI agents within military structures[28].

---

[22] A.B. Rashid, A.K. Kausik, A. Al Hassan Sunny, M.H. Bappyet, *Artificial Intelligence in the Military: An Overview of the Capabilities, Applications, and Challenges*, "International Journal of Intelligent Systems", 2023, 1, pp. 1-31.
[23] Center for Security and Emerging Technology, *CSET's 2024 Annual Report*..., op. cit.
[24] T.H. Chung, R. Daniel, *DARPA OFFSET: A Vision for Advanced Swarm Systems through Agile Technology Development and Experimentation*, "Field Robotics", 2023, 3, pp. 97–124.
[25] G. Stanovsky, R. Keydar, G. Perl, E. Habba, *Beyond Benchmarks: On The False Promise of AI Regulation*, arXiv 2025, https://arxiv.org/abs/2501.15693 (25.06.2025).
[26] NATO, *Summary of the NATO Artificial Intelligence Strategy*, 22.10.2021, https://www.nato.int/cps/en/natohq/official_texts_187617.htm (24.06.2025).
[27] Department of Defense, *DoD Directive 3000.09 Autonomy in Weapon Systems*, U.S. Department of Defense 25.01.2023, https://www.esd.whs.mil/portals/54/documents/dd/issuances/dodd/300009p.pdf (25.06.2025).
[28] N.J. McNeese et al., *Stepping out of the shadow of human-human teaming*..., op. cit.

Disparities in regulatory readiness worsen this challenge. Although initiatives like the European Union's AI Act[29] and UN Resolution A/78/L.49[30] set normative standards for human oversight, their implementation is inconsistent and often aspirational. As highlighted in the editorial of *New Biotechnology*, maintaining effective human oversight becomes increasingly difficult as systems grow more complex, autonomous, and opaque. Human-in-the-loop models may become unfeasible for quick-response situations, while human-on-the-loop oversight remains vulnerable to automation bias, fatigue, and organizational pressures[31].

Opponents of doctrinal adaptation warn that reconfiguring command structures to normalize AI authority might undermine the ethical foundations of warfare. As Johnson notes, the political legitimacy of military action partly comes from the assurance that humans, not machines, exercise judgment in using lethal force. Incorporating AI into strategic decision-making also risks triggering algorithmic escalation, where autonomous systems respond to perceived threats based on unclear rules or flawed data, potentially igniting broader conflicts without human oversight[32].

However, there is a counterargument that strategic deterrence and operational stability might benefit from clearly defined, well-regulated human-AI teamwork. When roles and limitations are transparent, autonomous systems can enhance decision speed without sacrificing legal or ethical standards. Examples from the Israeli Defense Forces and the U.S. Indo-Pacific Command demonstrate that AI tools, when utilized within strict oversight protocols and transparent interfaces, can support rather than replace human authority. In such setups, shared decision-making serves as augmented cognition: AI provides analytical depth, while humans offer moral and contextual judgment[33].

To address these tensions, emerging scholarship suggests a layered approach to ethical oversight and doctrinal adaptation. This approach incorporates explainable AI (XAI), real-time auditing tools, and standardized accountability records that detail decision-making processes and responsible parties. Regulatory frameworks, such as the European Union's AI Act and UN Resolution A/78/L.49, highlight principles of transparency, reliability, and the protection of human dignity, which should form the foundation of any military doctrine involving AI delegation.

Moreover, military oversight systems can learn from other high-risk industries. For example, using reinforcement learning with human feedback (RLHF) in medical diagnostics provides a model for enhancing AI behavior to align with human preferences. Similar techniques

[29] European Commission, *EU Artificial Intelligence Act*, 12.07.2024, https://artificialintelligenceact.eu/ai-act-explorer/ (24.06.2025).
[30] UN General Assembly, *Seizing the opportunities of safe, secure and trustworthy artificial intelligence systems for sustainable development*, UN General Assembly 11.03.2024, https://documents.un.org/doc/undoc/ltd/n24/065/92/pdf/n2406592.pdf (24.06.2025).
[31] A. Holzinger, K. Zatloukal, H. Müller, *Is human oversight to AI systems still possible?*, "New Biotechnology", 2025, 85, pp. 59–62, https://doi.org/10.1016/j.nbt.2024.12.003.
[32] J. Johnson, *The AI Commander Problem*..., op. cit.
[33] N. Polemi, I. Praça, K. Kioskli, A. Bécue, *Challenges and efforts in managing AI trustworthiness risks: a state of knowledge*, "Frontiers in Big Data", 2024, 7, p. 1381163.

could be adapted for military AI by incorporating ethical limits and operational rules through repeated, supervised learning.

Ultimately, the ethical and doctrinal implications of shared decision-making authority in military contexts cannot be solved through technical design alone. They require renewed commitment to normative inquiry, strong legal frameworks, and institutional cultures that support human judgment, accountability, and restraint. As AI systems become more capable and autonomous, military and political leaders have a responsibility to ensure that delegation enhances, rather than weakens, the ethical conduct and strategic control of warfare. This means that doctrine must evolve not just in response to technological advances but in ways that protect the principles of just war theory, democratic accountability, and the rule of law[34].

DESIGNING EFFECTIVE HUMAN-AI TEAMS: PRINCIPLES AND POLICY RECOMMENDATIONS

As artificial intelligence systems assume increasingly critical roles in military operations, designing effective human-AI teams becomes a key strategic objective. Beyond just technology, the success of these teams depends on organizational vision, ethical alignment, clear operations, and consistent doctrine. Human-AI interaction is not just a technical interface; it's a socio-technical relationship marked by mutual adaptation, shared agency, and institutional accountability.

A widely supported principle in the emerging doctrine of human-AI collaboration is maintaining meaningful human control over the system. This concept emphasizes that humans must hold ultimate authority over critical decisions, particularly those involving the use of lethal force. Military organizations seek to implement this principle through systems such as human-in-the-loop and human-on-the-loop, which enable humans to approve or override machine recommendations. For example, systems such as the U.S. Navy's Aegis Combat System provide rapid threat assessment and targeting but require human confirmation before firing[35]. These frameworks aim to strike a balance between ethical and legal accountability while leveraging AI's rapid processing power. However, during rapid-fire scenarios, such as missile interception or cyber defense, the time needed for human confirmation often becomes impractical. In such cases, delays caused by human decision-making could jeopardize missions or endanger forces. Consequently, insisting on human intervention at all times may limit the capabilities of autonomous systems, especially where quick action is crucial and machines outperform humans.

The debate over meaningful human control also raises a deeper philosophical question about responsibility in technologically mediated warfare. As AI systems gain more autonomy, traditional boundaries of individual accountability start to blur. In complex multi-agent

[34] Z. Zahedi, S. Kambhampati, *Human-AI Symbiosis: A Survey of Current Approaches*, arXiv 2021, https://arxiv.org/abs/2103.09990 (26.06.2025).
[35] Lockheed Martin, *Aegis Combat System*, 27.05.2025, https://www.lockheedmartin.com/en-us/products/aegis-combat-system.html (24.06.2025).

operations, decision-making becomes more distributed across human-machine networks. This decentralization challenges classical legal frameworks for assigning liability in cases of misconduct or unintended harm. The idea that a commander or operator can be entirely held responsible for an AI system's actions, especially one governed by probabilistic inference, emergent behavior, or opaque neural architectures, requires new models of shared responsibility and institutional accountability. These legal and doctrinal changes must be supported by cultural shifts within military institutions, emphasizing ethical leadership, mission command, and professional judgment as operational roles continue to evolve[36].

A second key concern in human-AI teaming is trust calibration and system transparency. Without a clear understanding of how reliable and capable the system is, human operators might either rely too much on or too little trust in AI outputs. Both situations pose serious risks: automation bias can cause people to accept flawed suggestions without question, while unwarranted doubt can lead to ignoring valuable insights. To address these issues, military designers are increasingly using explainable AI tools and user-friendly interface designs. Systems that display uncertainty, explain their recommendations, or rank options help humans make more informed decisions. In ISR applications, for example, interfaces that display confidence levels and provide explanations for detections have improved analyst accuracy and situational awareness[37].

The importance of calibrated trust is amplified in joint or coalition operations, where trust must be maintained not only between humans and machines but also among human teammates from different nations or branches of service. AI systems that act unpredictably or give poorly explained recommendations can disrupt team cohesion, cause friction in decision cycles, and undermine overall mission confidence. Although there is increasing agreement on the need for interpretability, technical efforts to develop transparent models often face trade-offs between clarity and performance. Highly interpretable algorithms may not have the predictive power of deep learning systems, while simplifying models can hide the complexity of real-world variables. This balance between performance and explainability is not just a technical challenge; it is a strategic one, affecting risk management, legal compliance, and public trust.

Meanwhile, the development of layered oversight mechanisms has become a key policy focus. These systems are designed as multi-level frameworks that combine real-time supervision with post-event accountability tools, such as logging systems, audit trails, and institutional checks. The goal is to sustain operational flexibility without compromising legal accountability or ethical standards. NATO's Allied Ground Surveillance program, which integrates human review with continuous system monitoring, exemplifies oversight in action[38]. Effective oversight must operate across multiple levels, tactical, operational, and strategic, ensuring that decisions

[36] B. Lou, T. Lu, T.S. Raghu, Y. Zhang, *Unraveling Human-AI Teaming: A Review and Outlook*, arXiv 2025, https://arxiv.org/abs/2504.05755 (26.06.2025).
[37] H.W. Meerveld, R.H.A Lindelauf, E.O. Postma, M. Postma, *The irresponsibility of not using AI in the military*, "Ethics and Information Technology", 2023, 25(1), p. 14, https://doi.org/10.1007/s10676-023-09683-0.
[38] NATO, *NATO Intelligence, Surveillance and Reconnaissance Force (NISRF)*, 25.03.2025, https://www.nato.int/cps/en/natohq/topics_48892.htm (24.06.2025).

made by AI systems can be explained, justified, and, if needed, challenged. Post-mission analysis and legal review should include AI-generated logs and decision pathways. Real-time systems should also allow for human overrides and mission pauses in the event of abnormal behavior[39].

However, institutionalizing oversight across diverse military commands and allied structures presents significant challenges. Bureaucratic delays, overlapping jurisdictions, and inconsistent doctrines can hinder coordination and reduce the responsiveness of the system. Moreover, oversight often competes with the demand for operational tempo, especially in expeditionary or joint operations where quick adaptation is prioritized over procedural rigor. Finding the right balance between oversight and agility remains an unresolved issue in military AI integration. The resource dimension must also be considered: effective oversight requires skilled personnel, technical infrastructure, and organizational capacity, resources that can be stretched thin during prolonged or large-scale operations[40].

Contextual calibration of human and AI roles is another vital factor. Successful human-AI teams need to be designed to respond to task requirements, mission goals, and environmental complexity. In logistical planning, AI might act as the primary decision-maker for inventory control, while in decisions involving rules of engagement, human judgment should take precedence. Programs like DARPA's Gremlins project demonstrate the advantages of role-specific division, where unmanned aerial vehicles autonomously handle navigation while humans maintain authority over target validation. The benefit of such hybrid systems is their complementarity: humans offer interpretive flexibility and ethical awareness, whereas machines provide speed, consistency, and endurance[41].

However, strict role partitioning can create vulnerability. In dynamic and adversarial settings, role expectations may change quickly. Systems that are too specialized or narrowly defined might fail to adapt to new situations, requiring reprogramming or manual intervention. For instance, an AI system built solely for ISR might struggle to prioritize or interpret humanitarian goals if it encounters displaced civilians during a surveillance mission. Therefore, flexibility in role assignment should be a key design focus, allowing teams to reconfigure their command structure as events unfold[42].

The integration of ethical reasoning into system design and operator training is another vital frontier. While technical performance is crucial, AI systems used in warfare must align with normative commitments consistent with international humanitarian law and democratic principles. This includes embedding ethical constraints into the system architecture, such as filters that identify proportionality violations or suspend operations in cases of legal ambiguity,

[39] A. Blanchard, M. Taddeo, *The Ethics of Artificial Intelligence for Intelligence Analysis: a Review of the Key Challenges with Recommendations*, "Digital Society", 2023, 2(1), p. 12.
[40] W. Hoffman, H.M. Kim, *Reducing the Risks of Artificial Intelligence for Military Decision Advantage*, Center for Security and Emerging Technology March 2023, https://cset.georgetown.edu/publication/reducing-the-risks-of-artificial-intelligence-for-military-decision-advantage/ (24.06.2025).
[41] DARPA, *Gremlins*, 5.11.2021, https://www.darpa.mil/news/2021/gremlins-airborne-recovery (24.06.2025).
[42] J. Pöhler, N. Flegel, T. Mentler, K. Van Laerhoven, *Keeping the human in the loop: are autonomous decisions inevitable?*, "i-com", 2025, 24(1), pp. 9–25.

and educating users in ethical literacy. The UK Ministry of Defence's Human Security Advisors initiative demonstrates how ethics can be put into practice within command environments[43]. However, critics question the practicality of encoding ethics into algorithms. Moral judgment often relies on nuanced contextual understanding, cultural norms, and political sensitivities that are difficult to formalize and articulate. Systems with fixed ethical heuristics may perform reliably under typical conditions but could fail in edge cases where ambiguity and uncertainty prevail. There is also the concern of "ethical outsourcing," where operators believe that compliance is built into the machine, thereby relieving them of moral accountability.

Beyond technical and ethical design, human-AI teams must be compatible across different services and allied forces. Multinational missions, such as those led by NATO or the UN, require consistency in doctrine, terminology, and interface logic. In this regard, policy frameworks have recommended standardizing AI-human protocols and interface formats. The NATO Communications and Information Agency has already begun working on interoperability guidelines for AI integration. These measures aim to reduce misunderstandings, enable data sharing, and enhance coordination among nations. However, standardization is not universally accepted. Critics argue that uniform doctrine could hinder innovation, limit national sovereignty, or overlook culturally specific approaches to warfare. In fields such as strategic deterrence or offensive cyber operations, where secrecy and unpredictability are assets, shared protocols may create vulnerabilities or restrict doctrinal flexibility[44].

Ultimately, designing effective human-AI teams requires ongoing empirical testing. Field experiments, red teaming, and iterative design tests are essential for identifying failure points and enhancing interaction protocols. These methods can uncover unexpected bottlenecks, user interface issues, and hidden biases in system behavior. Using simulation environments, synthetic training data, and operational sandboxes provides a low-risk approach to testing human-AI interactions before live deployment. In the long run, feedback mechanisms such as after-action reviews and lessons-learned databases should incorporate AI-related insights, helping to build a growing knowledge base of best practices in human-machine teaming[45].

Designing effective human-AI teams, therefore, requires more than technological sophistication. It demands ongoing engagement with competing priorities: speed versus safety, precision versus adaptability, autonomy versus accountability. The principles and policy directions outlined in this chapter provide a foundation for managing these trade-offs. However, their implementation must be iterative, context-aware, and validated through practical experience. Human-AI teams are not fixed entities; they are evolving socio-technical systems shaped by doctrine, knowledge, and the demands of war. Future design efforts must focus not only on more innovative algorithms but

[43] Ministry of Defence, *Defence Artificial Intelligence Strategy*, 15.06.2022, https://www.gov.uk/government/publications/defence-artificial-intelligence-strategy/defence-artificial-intelligence-strategy (24.06.2025).
[44] B.O. Adelakun, T.G. Majekodunmi, O.S. Akintoye, *AI and ethical accounting: Navigating challenges and opportunities*, "International Journal of Advanced Economics", 2024, 6(6), pp. 224–241.
[45] G. Bansal, B. Nushi, E. Kamar, E. Horvitz, D.S. Weld, *Is the Most Accurate AI the Best Teammate? Optimizing AI for Teamwork*, arXiv 2020, https://arxiv.org/abs/2004.13102 (26.06.2025).

also on wiser institutions – ones that empower human judgment, reinforce ethical standards, and maintain strategic control amid the uncertainties of twenty-first-century conflict[46].

CONCLUSION

The integration of artificial intelligence into military decision-making marks one of the most significant changes in modern defense policy and practice. As intelligent systems play an increasing role in tactical, operational, and strategic activities, the core aspects of how decisions are made, who makes them, and the moral and legal considerations involved are being reshaped. The potential of AI to enhance speed, accuracy, and data-driven reasoning is substantial, but it also raises serious concerns about ethical oversight, human accountability, and maintaining strategic stability.

At the core of this transformation is the challenge of designing human-AI interactions that enhance rather than replace human agency. AI technologies can handle large amounts of information, recognize patterns in complex datasets, and suggest courses of action at unprecedented speeds. However, integrating such systems into high-stakes military settings raises essential questions about trust, supervision, interpretability, and the limits of algorithmic reasoning. Risks like cognitive overload, automation bias, and mission drift must not be ignored. Human operators must remain prepared and empowered to critically evaluate AI outputs, intervene when necessary, and ultimately take responsibility for decisions with ethical and political implications.

The legitimacy of military operations in democratic societies relies on clear accountability, compliance with international law, and ethical conduct of force. AI systems, while powerful, lack the ability for moral judgment and legal reasoning. As these systems assume more autonomous roles, institutions must develop safeguards to ensure that technological decision-making does not compromise the core principles of armed conflict. This includes implementing transparent system designs, clear audit trails, and strong oversight mechanisms that maintain the traceability and legitimacy of decisions.

Moreover, integrating AI into multinational military operations requires a common understanding of human-machine roles, interoperability standards, and ethical principles across different strategic cultures. The success of coalition missions relies not only on the performance of AI systems but also on the compatibility of doctrines, trust among partners, and mutual acknowledgment of responsibilities. Without coordinated efforts, the spread of AI in military systems could create new fault lines in alliance politics, lead to unequal capabilities, and increase the risk of strategic misalignment.

Ultimately, the future of human-AI teaming in military contexts must be guided by a commitment to responsible innovation. This involves recognizing the limits of automation, reaffirming the vital role of human judgment, and integrating ethical reasoning throughout the entire

[46] Á.A. Cabrera, A. Perer, J.I. Hong, *Improving Human-AI Collaboration With Descriptions of AI Behavior*, "Proceedings of the ACM on Human-Computer Interaction", 2023, 7(1), pp. 1–21.

life cycle of AI systems – from design and development to deployment, use, and review. It requires policymakers, military leaders, engineers, and legal experts to collaborate in shaping AI integration that enhances operational effectiveness while upholding the core values of accountability, humanity, and strategic prudence.

The way forward will require ongoing reflection, thorough empirical assessment, and organizational agility to keep pace with changing technologies and security conditions. It will also require continuous investment in ethical leadership, cross-disciplinary cooperation, and flexible doctrines to ensure AI functions as a responsible tool of statecraft rather than a source of strategic chaos. Only by adopting a principled and multidisciplinary approach can countries effectively manage the profound challenges posed by artificial intelligence for warfare in the twenty-first century, ensuring that these technologies bolster rather than erode the moral and legal foundations of military power.

REFERENCES

Adelakun Beatrice Oyinkansola, Majekodunmi Tomiwa Gabriel, Akintoye Oluwole Stephen. 2024. "AI and ethical accounting: Navigating challenges and opportunities". International Journal of Advanced Economics 6(6): 224–241. https://doi.org/10.51594/ijae.v6i6.1230.

Bansal Gagan, Nushi Besmira, Kamar Ece, Horvitz Eric, Weld Daniel S. 2020. "Is the Most Accurate AI the Best Teammate? Optimizing AI for Teamwork". In https://arxiv.org/abs/2004.13102.

Bistron Marta, Piotrowski Zbigniew. 2021. "Artificial Intelligence Applications in Military Systems and Their Influence on Sense of Security of Citizens". Electronics 10(7): 871. https://doi.org/10.3390/electronics10070871.

Blanchard Alexander, Taddeo Mariarosaria. 2023. "The Ethics of Artificial Intelligence for Intelligence Analysis: a Review of the Key Challenges with Recommendations". Digital Society 2(1): 1-28. https://doi.org/10.1007%2Fs44206-023-00036-4.

Cabrera Ángel Alexander, Perer Adam, Hong Jason I. 2023. Improving Human-AI Collaboration With Descriptions of AI Behavior. In: Proceedings of the ACM on Human-Computer Interaction. https://doi.org/10.1145/3579612.

Center for Security and Emerging Technology. 2025. CSET's 2024 Annual Report. https://cset.georgetown.edu/publication/2024-annual-report/.

Choudhary Ashiv, Surbhi Adya. 2024. AI arbitration – Charting the ethical and legal course. In: The ETLTC2024 International Conference Series On ICT, Entertainment Technologies, And Intelligent Information Management In Education And Industry Aizuwakamatsu. Japan. In https://pubs.aip.org/aip/acp/article-lookup/doi/10.1063/5.0234731.

Choudhury Rizwan. 2024. Project Maven: The epicenter of US' AI military efforts. https://interestingengineering.com/military/project-maven-the-epicenter-of-us-ai-military-efforts.

Chung Timothy H., Daniel Roshan. 2023. "DARPA OFFSET: A Vision for Advanced Swarm Systems through Agile Technology Development and Experimentation". Field Robotics 3: 97–124. https://doi.org/10.55417/fr.2023003.

DARPA. 2021. Gremlins. https://www.darpa.mil/news/2021/gremlins-airborne-recovery.

Department of Defense. 2023. DoD Directive 3000.09 Autonomy in Weapon Systems. In https://www.esd.whs.mil/portals/54/documents/dd/issuances/dodd/300009p.pdf.

Erskine Toni, Miller Steven E. 2024. "AI and the decision to go to war: future risks and opportunities". Australian Journal of International Affairs 78(2): 135–147. https://doi.org/10.1080/10357718.2024.2349598.

European Commission. 2024. EU Artificial Intelligence Act. In https://artificialintelligence-act.eu/ai-act-explorer/.

French Shannon E., Lindsay Lisa N. 2022. Artificial Intelligence in Military Decision- Making. Avoiding Ethical and Strategic Perils with an Option-Generator Model. In: Bernhard Koch, Richard Schoonhoven (eds.), Emerging Military Technologies: Ethical and Legal Perspectives, 53–74. Brill.

Hinton Patrick. 2023. Ones and Zeros: The British Army Unveils its Digital and Data Plan. https://www.rusi.orghttps://www.rusi.org.

Hoffman Wyatt, Kim Heeu Millie. 2023. Reducing the Risks of Artificial Intelligence for Military Decision Advantage. In https://cset.georgetown.edu/publication/reducing-the-risks-of-artificial-intelligence-for-military-decision-advantage/.

Holzinger Andreas, Zatloukal Kurt, Müller Heimo. 2025. "Is human oversight to AI systems still possible?". New Biotechnology 85: 59–62. https://doi.org/10.1016/j.nbt.2024.12.003.

Johnson James. 2022. "The AI Commander Problem: Ethical, Political, and Psychological Dilemmas of Human-Machine Interactions in AI-enabled Warfare". Journal of Military Ethics 21(3–4): 246–271. https://doi.org/10.1080/15027570.2023.2175887.

Johnson James. 2023. "Automating the OODA loop in the age of intelligent machines: reaffirming the role of humans in command-and-control decision-making in the digital age". Defence Studies 23(1): 43–67. https://doi.org/10.1080/14702436.2022.2102486.

Lockheed Martin. 2025. Aegis Combat System. In https://www.lockheedmartin.com/en-us/products/aegis-combat-system.html.

Lou Bowen, Lu Tian, Raghu T.S., Zhang Yingjie. 2025. "Unraveling Human-AI Teaming: A Review and Outlook". In https://arxiv.org/abs/2504.05755.

Lyons Joseph B., Sycara Katia, Lewis Michael, Capiola August. 2021. "Human–Autonomy Teaming: Definitions, Debates, and Directions". Frontiers in Psychology 12: 589585. https://doi.org/10.3389/fpsyg.2021.589585.

Márquez Díaz Jairo Eduardo. 2024. "Benefits and challenges of military artificial intelligence in the field of defense". Computación y Sistemas 28(2). https://doi.org/10.13053/cys-28-2-4684.

Mayer Michael. 2023. "Trusting machine intelligence: artificial intelligence and human-autonomy teaming in military operations". Defense & Security Analysis 39(4): 521–538. https://doi.org/10.1080/14751798.2023.2264070.

McNeese Nathan J., Flathmann Christopher, O'Neill Thomas A., Salas Eduardo. 2023. "Stepping out of the shadow of human-human teaming: Crafting a unique identity for human-autonomy teams". Computers in Human Behavior 148: 107874. https://doi.org/10.1016/j.chb.2023.107874.

Meerveld H.W., Lindelauf R.H.A., Postma E.O., Postma M. 2023. “The irresponsibility of not using AI in the military”. Ethics and Information Technology 25(1): 14. https://doi.org/10.1007/s10676-023-09683-0.

Meimandi Kiana Jafari, Bolton Matthew L., Beling Peter A. 2024. “Human-Agent Teaming: A System-Theoretic Overview”. Preprints. In https://www.techrxiv.org/users/716315/articles/701117-human-agent-teaming-a-system-theoretic-overview?commit=05d58d706b35a50d12dcd7bac3feefe5f0440723.

Ministry of Defence. 2022. Defence Artificial Intelligence Strategy. In https://www.gov.uk/government/publications/defence-artificial-intelligence-strategy/defence-artificial-intelligence-strategy.

Morozov A.O., Yashchenko V.O. 2023. “Decision-making technologies in military systems. Challenges and prospects”. Mathematical Machines and Systems 4: 3–10.

Nalin Alessandro, Tripodi Paolo. 2023. “Future Warfare and Responsibility Management in the AI-based Military Decision-making Process”. Journal of Advanced Military Studies 14(1): 83–97. https://doi.org/10.21140/mcuj.20231401003.

NATO. 2021. Summary of the NATO Artificial Intelligence Strategy. In https://www.nato.int/cps/en/natohq/official_texts_187617.htm.

NATO. 2025. NATO Intelligence, Surveillance and Reconnaissance Force (NISRF). In https://www.nato.int/cps/en/natohq/topics_48892.htm.

O’Neill Thomas A., Flathmann Christopher, McNeese Nathan J., Salas Eduardo. 2023. “21st Century teaming and beyond: Advances in human-autonomy teamwork”. Computers in Human Behavior 147: 107865. https://doi.org/10.1016/j.chb.2023.107865.

O’Neill Thomas A., McNeese Nathan, Barron Amy, Schelble Beau. 2022. “Human–Autonomy Teaming: A Review and Analysis of the Empirical Literature”. Human Factors: The Journal of the Human Factors and Ergonomics Society 64(5): 904–938. https://doi.org/10.1177/0018720820960865.

Pöhler Jonas, Flegel Nadine, Mentler Tilo, Van Laerhoven Kristof. 2025. “Keeping the human in the loop: are autonomous decisions inevitable?” i-com 24(1): 9–25. https://doi.org/10.1515/icom-2024-0068.

Polemi Nineta, Praça Isabel, Kioskli Kitty, Bécue Adrien 2024. “Challenges and efforts in managing AI trustworthiness risks: a state of knowledge”. Frontiers in Big Data 7: 1381163. https://doi.org/10.3389/fdata.2024.1381163.

Rashid Adib Bin, Kausik Ashfakul Karim, Al Hassan Sunny Ahamed, Bappyet Mehedy Hassan. 2023. Artificial Intelligence in the Military: An Overview of the Capabilities, Applications, and Challenges. In: Yu-an Tan (ed.), International Journal of Intelligent Systems. https://doi.org/10.1155/2023/8676366.

Shutov Alexandrovich. 2024. “Artificial Intelligence and International Security”. International Affairs 70(005): 254–256.

Stanovsky Gabriel, Keydar Renana, Perl Gadi, Habba Eliya. 2025. “Beyond Benchmarks: On The False Promise of AI Regulation”. In https://arxiv.org/abs/2501.15693.

Sycara Katia, Lewis Michael. 2004. Integrating intelligent agents into human teams. In: Eduardo Salas, Stephen M. Fiore (eds.), Team cognition: Understanding the factors that drive process and performance, 203–231. American Psychological Association.

UN General Assembly. 2024. Seizing the opportunities of safe, secure and trustworthy artificial intelligence systems for sustainable development. In https://documents.un.org/doc/undoc/ltd/n24/065/92/pdf/n2406592.pdf.

Wischnewski Magdalena, Krämer Nicole, Müller Emmanuel. 2023. Measuring and Understanding Trust Calibrations for Automated Systems: A Survey of the State-Of-The-Art and Future Directions. In: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems. Hamburg.

Zahedi Zahra, Kambhampati Subbarao. 2021. "Human-AI Symbiosis: A Survey of Current Approaches". In https://arxiv.org/abs/2103.09990.