Centre for International
Governance Innovation

Policy Brief No. 193 — January 2025

Human-Machine Interaction and Human Agency in the Military Domain

Ingvild Bode

Key Points

- Militaries increasingly use artificial intelligence (AI) technologies for decision support and combat operations. AI does not replace humans, but personnel interact with AI technologies more frequently.
- Practices of human-machine interaction have the potential to profoundly alter the quality of human agency, understood as the ability to make choices and act, in warfare. Specifically, they introduce distributed agency between humans and machines.
- Forms of distributed agency will be shaped along a spectrum, preserving more room for either human or machine agency. Such practices happen in multiple locations and with multiple, networked systems.
- Accounting for the phenomenon of distributed agency requires going beyond perceiving challenges of human-machine interaction as straightforward problems to solve. Rather, distributed agency needs to be recognized as raising foundational operational, ethical-normative and legal challenges.

Introduction

The proliferation of AI technologies in military decision-making processes around targeting seems to be increasing. At first, the incorporation of AI in the military domain was predominantly examined in relation to weapon systems, frequently referred to as autonomous weapon systems (AWS) that can identify, track and attack targets without further human intervention (International Committee of the Red Cross [ICRC] 2021). Militaries worldwide already employ weapon systems, including some loitering munitions, which incorporate AI technologies to facilitate target recognition, generally depending on computer vision techniques (Boulanin and Verbruggen 2017; Bode and Watts 2023). Although usually operated with human approval, loitering munitions appear to have the potential to dynamically apply force without human intervention. Indeed, various reports from Russia's war in Ukraine have indicated that the Ukrainian army uses loitering munitions that release force without human approval in the terminal stage of operation (Hambling 2023, 2024). These developments firmly underline long-standing and growing concerns about the extent to which the role that humans play in use-of-force decision making when using AI-based systems is diminishing.

However, weapon systems are just one of numerous areas of application where AI is used in the military

About the Author

Ingvild Bode is professor of international relations at the University of Southern Denmark (SDU). She is also director of the Center for War Studies at SDU. Her research focuses on processes of normative change, especially with regard to the use of force. She is the principal investigator of three externally funded research projects: AutoNorms: Weaponised Artificial Intelligence, Norms, and Order (funded by the European Research Council); the AutoPractices project on sustaining and strengthening human agency in military decision-making processes involving AI technologies (funded by the European Research Council); and Human-Machine Interaction: The Distributed Agency of Humans and Machines in Military AI (HuMach, funded by the Independent Research Fund Denmark). Further, Ingvild served as the co-chair of the Institute of Electrical and Electronics Engineers Standards Association Research Group on Issues of AI and Autonomy for Defence Systems. She is also an expert member of the Global Commission on Responsible AI in the Military Domain. Her work has been published in various international journals such as the *European Journal of International Relations*, *Ethics and Information Technology*, *Review of International Studies* and *International Studies Review*. Ingvild's most recent book, *Autonomous Weapons Systems and International Norms* (co-authored with Hendrik Huelss), was published by McGill-Queen's University Press in 2022. Previously, Ingvild was senior lecturer in international relations at the University of Kent, Canterbury (2015–2020) and a Japan Society for the Promotion of Science international research fellow with joined affiliation at United Nations University and the University of Tokyo (2013–2015).

setting. AI technologies are typically considered to improve the effective and rapid analysis of vast quantities of data, making them an appealing choice for a range of military decision-making tasks pertaining to varying levels of risk, such as logistics, recruitment, intelligence and targeting (Grand-Clément 2023). In the military domain, such systems are commonly referred to as AI-based decision support systems (DSS) that "assist decision-makers situated at different levels in the chain of command to solve semi-structured and unstructured decision tasks" (Susnea 2012, 132–33; Nadibaidze, Bode and Zhang 2024).

Militaries are moving toward the incorporation of AI technologies for several reasons, including personnel issues. Human capabilities alone are frequently not considered suited to reviewing the large amounts of data required for gaining battlefield awareness. Further, many militaries, for example, Japan and the United Kingdom, struggle with recruitment and retention, and see an increasing reliance on AI as a potential way of addressing shortages (Jesse Johnson 2024). Integrating AI technologies into military decision making does not simply replace military personnel. Nonetheless, as these technologies gain traction, they have the potential to alter the ways in which many kinds of military personnel do their jobs by incorporating the use of various AI technologies for descriptive, predictive and prescriptive purposes into their routines. This policy brief offers an initial reflection on and assessment of how interacting with AI technologies might alter the exercise of human agency in the military domain, understood as the capacity to decide and act. Instances of human-machine interaction shape a distributed agency between humans and AI. This distributed agency cannot be understood by considering humans and machines as separate entities in isolation because it emerges out of situations of interaction.

The policy brief studies this phenomenon in five parts: First, it examines the notions of human control and human agency in relation to each other, arguing that reflecting on agency instead of control allows for a more comprehensive examination of what it means to use AI technologies in the military domain. Second, it presents a brief empirical overview of how militaries speak of human-machine teaming and human-system integration. Third, it sketches what it means to consider agency as distributed. Fourth, it considers the potential operational, ethical-normative and legal consequences of such a development. Fifth, it suggests practical ways forward for stakeholders involved in the debate about AI in the military domain.

A note on terminology: the brief recognizes AI as an umbrella term to describe the overall effort "to create machines or things that can do more than what is programmed into them" (Gebru 2023). The chosen phrase "AI technologies" emphasizes the intricate, contingent and changeable integration of AI into societies and thereby also seeks to create distance from problematic hype attached to AI (Carmel and Paul 2022). AI technologies rely on various techniques such as computer vision, machine learning, speech recognition and natural language processing. In the military domain, the notion of autonomy precedes, but also connects to, AI technologies. Autonomy typically refers to functions that a system can perform without requiring human intervention, such as intelligence analysis, mobility or targeting (Boulanin and Verbruggen 2017, 19–35). The brief focuses on AI technologies, which can be usefully understood on a spectrum of increasing technological complexity. Further, human-machine interaction in the military space is often thought to involve both robotic and software-based systems.

Human Control Versus Human Agency

At first, almost without exception, stakeholders in the debate about AI in weapon systems and AWS used the term *human control* when discussing the parts humans play in use-of-force decision making (Ekelhof 2017; Taddeo and Blanchard 2022; Amoroso and Tamburrini 2020), frequently adding the qualifier "meaningful" (Roff and Moyes 2016). This concept of human control was, initially, chiefly concerned with the last stage of the targeting process when kinetic force is used.

The exercise of human control has often been described in relation to a decision-making loop, differentiating between in-the-loop, on-the-loop or out-of-the-loop human involvement in weapon systems using AI technologies for targeting (Sharkey 2016). In-the-loop human operators have to authorize the release of force and manage the attack process. Operators either consider specific targets before initiating an attack or choose from a list of targets. Put simply, humans in the loop play a direct and explicit role in decision-making processes that use AI algorithms. *On the loop* describes a set-up where systems can automatically sense and detect targets and fire on them while human operators may only be allocated a time-restricted veto in the targeting process. In other words, human operators only supervise or oversee a process of algorithmic decision making. Finally, *off the loop* describes systems where the process of identifying, tracking and attacking targets happens without human involvement or oversight. These would therefore be characterized as AWS. Keeping a human in or on the loop has transpired as a common regulatory approach to the public policy challenge of algorithmic decision making in and beyond the military domain (Crootof, Kaminski and Price 2023).

But merely including a human at the tail end of decision-making processes involving AI-based systems only renders the nominal appearance of meaningful human control. In other words, humans in or on the loop do not necessarily ensure appropriate or high levels of human engagement and oversight because the human in the loop as a regulatory notion typically does not "account for the humans' needs, skills, or frailties, or anticipate the ways in which working in tandem with a machine will channel and influence that human's behaviour" (ibid., 437; see also Tsamados, Floridi and Taddeo 2024). The human in the loop therefore runs the risk of being ornamental rather than meaningful, only rendering the reassuring appearance of human control rather than having an actual effect on decision making (Brennan-Marquez, Levy and Susser 2019, 754; Bode and Watts 2021; Bode 2023). Further, human-in-the-loop thinking tends to focus on the tail end of decision making, which diverts attention from earlier stages, such as data preparation and processing.

While in-the-loop thinking remains prominent as a regulatory solution, the notion of human control has also expanded beyond the narrow confines of the final stages of a targeting process. Life-cycle models of human control therefore consider multiple human actors involved across different stages of a system, from research and development to testing, use, evaluation after use and reuse/retirement (Blanchard, Thomas and Taddeo 2024; Ekelhof and Persi Paoli 2020; Institute of Electrical and Electronics Engineers Standards Association [IEEE SA] Research Group on Issues of AI and Autonomy in Defence Systems 2024). A life-cycle approach to human control recognizes how the precise shape of AI-based systems is influenced by human decisions, choices, inquiries and behaviour, thereby considering them to be socio-technical in character (Orlikowski 1992).

This perspective can also be conveyed by a shift in terminology from *human control* to *human agency*. The concept of human agency encompasses a broader and more thorough understanding of what decision making and the ability to act entails, rather than solely focusing on the degree of — seemingly unilateral — control exerted by individual humans. Of course, human agency is a multi-faceted, philosophical concept that is open to being understood in many, diverging ways. For the purpose of this brief, rather than delving into a detailed conceptual examination, human agency is understood as "the socioculturally mediated capacity to act" (Ahearn 2001, 112) and as making choices that are oriented around the capacity to practise reason (Hildebrandt 2011, 6). Examining effects on human agency that may result from interacting with AI technologies recognizes that incorporating AI technologies into various processes of decision making involves more than just "delegating" motor capabilities to systems; it also involves people and machines, to some extent, sharing cognitive functions and abilities (Bellanova et al. 2021; Bode and Nadibaidze 2024).

Human-Machine Interaction in Military Thinking

Emphasizing human-system integration or human-machine teaming appears to have become a common trope in relation to both robotic and AI technologies in military decision making. This is particularly visible in policies and strategies pursued by the United States and the United Kingdom.

The US Army's Project Convergence, for example, focuses on simulating experimentation with human-robot-AI interaction on the battlefield and, in line with US aspirations toward Joint All-Domain Command and Control, features exercises across multiple domains and with allied forces, including Australia and the United Kingdom (Demarest and Judson 2023). Project Convergence has been characterized as "a US Army effort to understand how it can use humans and machines together on the battlefield" (Judson 2024), not least through military formations that feature humans and machines. The United Kingdom, meanwhile, has characterized "the effective integration of humans, artificial intelligence (AI) and robotics into warfighting systems" as "at the core of future military advantage" (Development, Concepts and Doctrine Centre 2018, 39).

Taking US and UK references as the basis, it appears to be common in Western military thinking to presume and represent the relationship between humans and machines as being hierarchical in nature. Whether this thinking extends to militaries in other geographies is unclear. Assuming a hierarchical relationship between humans and machines allows humans to benefit in a unidirectional manner from the incorporation of AI technologies into military decision making. The primary objective of human-machine integration, for example, as part of joined formations, appears to enable militaries to amalgamate the most advantageous aspects of both humans and machines, considered as distinct entities; as one US Army official described it, "Human-machine integration is all about bringing the right mix of robotic elements to our formations to enable a total formation that is more capable than the sum of either the human or robotic elements" (Michael Cadieux, quoted in Aliotta 2023). The UK Ministry of Defence's joint concept note states that "what humans are best and what machines are best at" (Development, Concepts and Doctrine Centre 2018, 41–42) is the guiding principle for designing tasks to be automated, as well as those that should be undertaken by humans only.

This reasoning holds that human-system integration can address human weaknesses and frailties, allowing military decision makers to benefit from AI's promise of rational efficiency (James Johnson 2024, 17). According to a publication by the US Air Force, for example, human-machine teams "provide decision-makers with a timely and precise intelligence advantage" (Jamieson 2024, 4). Similarly, Project Convergence describes human-machine teams as allowing soldiers to "offload risk" and providing them "with additional information for decision-making" (quoted in Judson 2024). AI technologies, according to this point of view, are essentially tools that humans utilize to further their goals.

These military viewpoints appear to rest on the unfortunately named "men are better at–machines are better at" (MABA–MABA) lists that have been influential in system engineering since at least the 1970s (Dekker and Woods 2002, 240; Parasuraman, Sheridan and Wickens 2000), though the military domain is certainly not unique in taking such an approach, as multiple practical and regulatory examples demonstrate (Crootof, Kaminski and Price 2023; Roth 2016; Brennan-Marquez, Levy and Susser 2019). MABA-MABA lists advocate a manageable division of labour alongside system functions that "relies on a presumption of fixed human and machine strengths and weaknesses" (Dekker and Woods 2002, 240). Such models have received their fair share of disciplinary critique around underappreciating the actual qualitative effects of the straightforward substitution these advocate: rather than machine strengths substituting for human weaknesses in a positive-sum way, this approach also creates new, often unanticipated weaknesses and failures, for example, around long-documented problems with human vigilance (Bainbridge 1983; Bradshaw et al. 2013). UK military thinking around human-machine teaming notably reflects explicitly on human mental capacity, yet still frames the limits of this in interacting with machines in positive-sum, complementary terms: "the ability to dynamically vary the level of active control that operators exercise over systems becomes a fundamental enabler to tempo and team effectiveness" (Development, Concepts and Doctrine Centre 2018, 45).

What Is Distributed Agency?

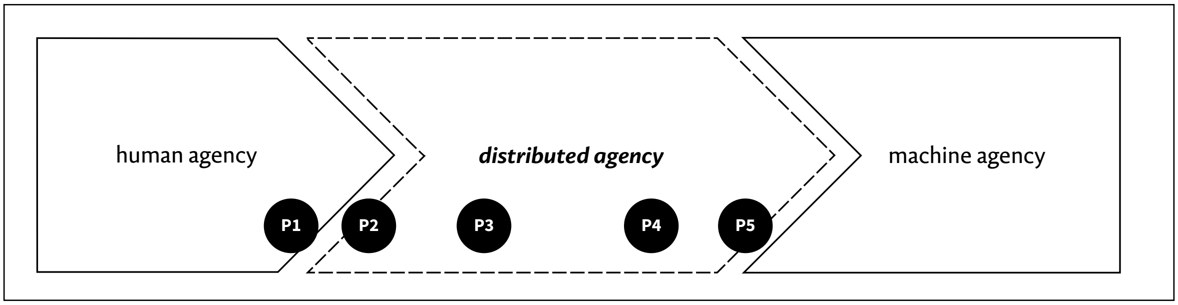
Understanding human-system integration as characterized by a neat, top-down, hierarchical separation between humans and AI technologies underestimates the significance of such interaction patterns. Beyond any conversation about AI, technologies have increasingly been associated with active, agent-like qualities in socio-technical systems by a variety of research programs that span disciplines such as science and technology studies. The basic idea behind this is well articulated by the actor-network theory concept of "actant," which is defined as "whoever or whatever takes part in the process" (Asdal and Ween 2014, 6). This definition therefore encompasses both human and non-human actors as potential actants. It is also important to note that this perspective does not consider agency to rest in the human or the non-human actors themselves, but rather in the relations that exist between them, as well as in the practices performed (Barad 2003, 818). This way of thinking therefore purports conceptualizing agency as being distributed and as dynamically evolving out of links between the technologies or the technological objects and the humans that make up socio-technical systems.

To start comprehending distributed agency and its effects, it remains useful to start from known, ideal-typical distinctions that exist between the ways in which humans tend to make decisions and how AI technologies tend to process outputs, as well as the implications that these differences have for the potential dynamics of interaction. For human agency, this could include the ability to exercise common sense, the ability to make context-dependent judgments, or the tendency to project agency into and anthropomorphize non-human actants (see, for example, James Johnson 2024, 9–16). Machine agency can, by contrast, be associated with a focus on performance rather than competence, the inability to say "I don't know," the tendency to look at details rather than the big picture, and the lack of understanding of "the world" and therefore also of the ability to distinguish between fact and fiction (Bunz 2012; Bender et al. 2021; see also Logan 2024).

Yet, as a result of the fact that humans and AI technologies are components of a relational, complex, socio-technical system, distributed agency needs to be understood beyond these distinctions. The notion of a machine agency that is somehow separate from human processes builds on assumptions that are problematic in the first place. AI technologies have, after all, been designed by various humans, and their design features represent particular, often normative and value-laden choices taken by those who programmed them. Even in the instance of machine learning, any capacities prior to use result from multiple interactions with humans and the subsequent refinement of AI technologies, often subsumed as iterative labelling of data (Jaton 2020; Crawford 2021; Penn 2021). A distributed model of agency should therefore also influence our thinking about what AI technologies are in the first place and the kind of data processing and preparation these involve.

While interacting with AI technologies, human groups that are part of such socio-technical systems are subject to affordances (also known as action potentials), but they are also subject to limitations. AI DSS, for example, afford humans the otherwise inaccessible ability to parse enormous amounts of data in the search for patterns. But, simultaneously, the data analyzed and the outputs presented by such systems are devoid of contextual information that is typically crucial for understanding dynamic patterns of human behaviour (Klonowska 2022, 137–38). The

Figure 1: Spectrum of Distributed Agency



Source: Author.

case of AI technologies and how these are integrated into military decision making also represents particular, and varied, action potentials. The different tasks that AI DSS can be used to fulfil in military decision making underline this (ICRC and Geneva Academy 2024, 9; Klonowska 2022, 135). DSS may be used descriptively, that is collecting, organizing and presenting data pulled from multiple silos and sensors. However, DSS could also be used for predictive tasks that involve the identification of novel patterns and trends in historical data, in conjunction with the capability to generate probabilistic predictions about the future. Additionally, DSS could be used for prescriptive purposes, such as providing advice on preferred courses of action.

Distributed agency will be constituted by and in particular situations of human-machine interaction. Such practices of human-machine interaction may therefore shape different forms of distributed agency along a spectrum (see Figure 1). Some practices of human-machine interaction, visualized illustratively as P1–P5 in Figure 1, may therefore retain more scope for human agency while others may privilege the exercise of forms of machine agency. Tracking the performance of such practices over time can conceptualize how human agency changes through interacting with AI technologies and the practices constituting such changes.

Operational and Normative Consequences and Challenges of Distributed Agency

The reality of distributed agency that arises out of increased human-system integration in military decision making will have profound impacts on operational, ethical-normative and legal aspects. The extent of the possible impacts depends on where the balance of distributed agency will lie, that is, whether it will tilt toward forms of human, or toward forms of machine, decision making.

Operationally, there is, for example, evidence of how humans are psychologically predisposed toward a number of cognitive biases that may, in practice, lead to more leeway for machine forms of agency (Bo and Dorsey 2024; James Johnson 2024, 35–43). One of the best-documented is automation bias, where humans place more confidence and trust in the outputs produced by AI technologies or other forms of automated systems than in their own critical deliberative skills. As a result, humans often defer, more or less uncritically, to system outputs. This tendency increases in the kind of high-pressure and high-stress situations that are representative of warfare (Parasuraman and Manzey 2010, 397). There are also other, often mundane, reasons for why a situation of distributed agency may privilege machine-produced outputs over human exercises of judgment, such as data overload and time constraints. The increasing amount of data that is available to monitor battlefield conditions, coming from, for example, satellites, various sensors and geolocation points, creates an operational push toward integrating more AI processing functionalities — and "a widening gap between human and machine analytical capabilities" (Klonowska 2022, 142). Further, the oft-cited military aspiration to fight at machine speed can further tilt the balance of distributed agency toward machine components, because, in such thinking, the communication links and the delay that humans bring to military decision-making processes only serve to slow down the decision-making loop. This tendency is visible in existing human-in-the-loop or on-the-loop systems, such as various types of air defence systems, where human operators may only have seconds to make a decision about authorizing the release of force (Peterson 2020; Leung 2004; Rogoway 2018). Such a reduced time span makes any critical deliberation or assessment of targets impossible (Bo and Dorsey 2024).

Ethically-normatively, the stakes of distributed agency are high because of how this phenomenon affects the exercise of moral agency in warfare. *Moral agency* refers to the capacity to make ethical judgments about right and wrong while retaining responsibility for the actions and choices pursued. If the balance in distributed agency in practical terms tilts toward more machine agency, the moral agency both of those affected by the use of force and of those using force in warfare will be reduced or perhaps even lost (Renic and Schwarz 2023a). This effect on moral agency results from the systematicity, routinization and technical dehumanization that interacting with AI technologies in warfare produces (Renic and Schwarz 2023b). In other words, something more fundamental changes when humans interact with AI technologies in military decision making. Such perturbations to the exercise of moral agency also shed further light on the human in the loop as a problematic regulatory option. While humans in the loop may only appear to be in control, they are still likely to remain identified as the key moral decision maker for decision-making processes that they actually have limited agency in — a "moral crumple zone" (Elish 2019). Further, such unregulated practices of human-machine interaction as sources of distributed agency also shape — potentially undesirable — social norms, that is, understandings of appropriateness, from the bottom up. In the absence of top-down forms of regulation, such social norms may therefore lead to practically accepting forms of distributed agency with all of its challenges as "appropriate" and "normal" (Bode 2023).

Legally, distributed agency also poses significant challenges, as the system of international law is structured firmly around human accountability and responsibility. Concepts and institutions of justice are human-centric and require human addressees (Bryson, Diamantis and Grant 2017; Bryson and Theodorou 2019). Legal scholarship has long grappled with the "responsibility gaps" that arise out of integrating autonomous and AI technologies into military decision making, sometimes addressed in acknowledging a shared responsibility attributed to different human actors across the design and use processes (Matthias 2004; Crootof 2016; Strasser 2022). The reality of distributed agency reinforces these concerns through acknowledging the substantive agency of AI technologies and their impact on the rationales and outcomes of military decision making (Sienknecht 2024, 194). Assuming distributed agency clearly goes beyond AI technologies merely performing assistance functions to the interaction with AI technologies actually *changing* both the character of military decision making and its outcomes. At present, there are risks of tilting in the direction of machine agency: "the role of algorithms in the underlying calculations could lead states to unwittingly make war-related decisions almost entirely based on machine calculations and recommendations" (Deeks, Lubell, and Murray 2019, 17). Fundamentally, then, distributed agency raises questions about legal compliance for use, *inter alia,* in relation to fundamental principles of international humanitarian law such as precaution, distinction and proportionality, as well as post-use, namely in terms of prosecution (Bo and Dorsey 2024).

Further, accounting for the challenges of distributed agency also means acknowledging that interaction patterns between humans and machines will not happen in the singular but rather in the plural. Multi-layered military decision making involves various humans, systems and AI technologies interacting with each other simultaneously and for different purposes. US military projects such as Project Convergence illustrate this networked model of warfare across various domains and how it brings together multiple human-machine integration situations and practices (Center for Strategic and International Studies 2022). Here, not only may military personnel be required "to control multiple platforms simultaneously" (Cadieux, quoted in Aliotta 2023), but also these platforms may be networked with each other, as well as potentially using data outputs produced by, once again, networked AI DSS. Thinking about distributed agency in this context — and with the backdrop of how even specific, one-to-one human-machine interaction situations may already tilt the overall balance in favour of machine agency — illuminates how the ever-greater reliance on AI technologies in military decision making risks diluting and diminishing spaces for exercising human agency (Arvidsson 2018, 22–25).

Ways Forward

Distributed agency must be recognized as the consequential result of embedding AI technologies into (military) decision making. This brief offers four preliminary ways forward for stakeholders involved in the debate about AI in the military domain:

→ **Raise awareness:** Situations of human-machine interaction resulting from integrating AI in the military domain do not easily remain under human direction and do not necessarily benefit militaries in positive zero-sum form. Considering how distributed agency results from these interactions recognizes this. Thinking on the basis of distributed agency should become part of any governance conversation and initiatives around AI technologies in the military domain. When using terms such as human-system integration and human-machine teaming in their planning documents, militaries should explicitly acknowledge that these concepts constitute independent systems and therefore new entities (Tsamados, Floridi and Taddeo 2024).

→ **Map distributed agency:** This brief offered thinking about distributed agency as a spectrum. Practices of human-machine interaction shape different forms of distributed agency that may privilege spaces for human or machine agency. To understand the evolving trajectory of distributed agency and to manage its consequences, militaries could use this spectrum to plan and map their practices of human-machine interaction across functions and domains. This would allow them to better understand where the balance between spaces for human and machine agency is moving and, on that basis, help assess whether that direction is desirable.

→ **Analyze critically:** Forms of distributed agency that privilege machine over human decision-making spaces can result from all practices of human-machine interaction. To address the challenges arising from this development, political decision makers and militaries will therefore need to follow the mapping, fact-finding exercise outlined in the preceding step with critically analyzing where they consider such forms of distributed agency more or less problematic. This will likely be contingent and context dependent. For example, when militaries consider using AI DSS, forms of distributed agency that privilege machine agency may be less problematic for tasks relating to data processing but more problematic for tasks relating to target nomination. As part of the critical analysis, differentiating between thresholds of risk could be a useful measure for militaries to address the adverse consequences of distributed agency. This stage will require making explicit, ultimately ethical-normative, choices.

→ **Develop principles and practical guidelines:** The critical analysis should serve as the basis for formulating principles and associated practical guidelines to address distributed agency. This ensures that the results of the critical deliberation conducted by political decision makers and militaries actually serve to guide practices. Distributed agency implies radically rethinking what it means to be a human decision maker in warfare and how that decision-making space is mediated by practices of interacting with machines. This will require renewed efforts in the direction of awareness-raising and training for all military personnel encountering situations of human-machine interaction.

Conclusion

Accepting human-machine teaming or human-machine interaction as common occurrences in military decision making also means accepting modifications to the exercise of human agency. These modifications will not be limited to intentional effects. Likewise, even if considered in terms of offering strategic offices, these modifications will amount to more than strategic benefits for human decision makers in the military domain. Interacting with AI technologies in the military domain also looks likely to change, delimit or diminish human agency. It is high time that the international debate about AI in the military context addresses more fundamental problems if we are to fully comprehend how significant the integration of AI technologies is — as well as how to address the operational, ethical-normative and legal ramifications that this decision will have.

Recognizing distributed agency resulting from human-machine interaction does not imply abandoning the requisite need for a high-quality exercise of human agency over decision making in the military domain. Maintaining a human-centric approach to the design and use of AI technologies is critical for meeting that demand. However, such an approach must make sure that humans are and remain capable of exercising their agency. In developing governance rules, it is therefore necessary to think about and find a realistic balance between what military personnel may gain by working with AI-based systems and what they may find adversely affected or lost.

Works Cited

Ahearn, Laura M. 2001. "Language and Agency." *Annual Review of Anthropology* 30 (1): 109–37. https://doi.org/10.1146/annurev.anthro.30.1.109.

Aliotta, Jerome. 2023. "Future of human-machine integration must bring right mix of robotic elements to formations." US Army, October 5. www.army.mil/article/270628/future_of_human_machine_integration_must_bring_right_mix_of_robotic_elements_to_formations.

Amoroso, Daniele and Guglielmo Tamburrini. 2020. "Autonomous Weapons Systems and Meaningful Human Control: Ethical and Legal Issues." *Current Robotics Reports* 1 (4): 187–94. https://doi.org/10.1007/s43154-020-00024-3.

Arvidsson, Matilda. 2018. "Targeting, Gender, and International *Posthuman*itarian Law and Practice: Framing The Question of the Human in International Humanitarian Law." *Australian Feminist Law Journal* 44 (1): 9–28. https://doi.org/10.1080/13200968.2018.1465331.

Asdal, Kristin and Gro Birgit Ween. 2014. "Writing Nature." *Nordic Journal of Science and Technology Studies* 2 (1): 4–10. https://doi.org/10.5324/njsts.v2i1.2130.

Bainbridge, Lisanne. 1983. "Ironies of Automation." *Automatica* 19 (6): 775–79. https://doi.org/10.1016/0005-1098(83)90046-8.

Barad, Karen. 2003. "Posthumanist Performativity: Toward an Understanding of How Matter Comes to Matter." *Signs: Journal of Women in Culture and Society* 28 (3): 801–31. https://doi.org/10.1086/345321.

Bellanova, Rocco, Kristina Irion, Katja Lindskov Jacobsen, Francesco Ragazzi, Rune Saugmann and Lucy Suchman. 2021. "Toward a Critique of Algorithmic Violence." *International Political Sociology* 15 (1): 121–50. https://doi.org/10.1093/ips/olab003.

Bender, Emily M., Timnit Gebru, Angelina McMillan-Major and Shmargaret Shmitchell. 2021. "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?" In *FaccT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–23. https://doi.org/10.1145/3442188.3445922.

Blanchard, Alexander, Christopher Thomas and Mariarosaria Taddeo. 2024. "Ethical governance of artificial intelligence for defence: normative tradeoffs for principle to practice guidance." *AI & SOCIETY, February* 19. https://link.springer.com/10.1007/s00146-024-01866-7.

Bo, Marta and Jessica Dorsey. 2024. "Symposium on Military AI and the Law of Armed Conflict: The 'Need' for Speed — The Cost of Unregulated AI Decision-Support Systems to Civilians." *Opinio Juris* (blog), April 4. https://opiniojuris.org/2024/04/04/symposium-on-military-ai-and-the-law-of-armed-conflict-the-need-for-speed-the-cost-of-unregulated-ai-decision-support-systems-to-civilians/.

Bode, Ingvild. 2023. "Practice-based and public-deliberative normativity: retaining human control over the use of force." *European Journal of International Relations* 29 (4): 990–1016. https://doi.org/10.1177/13540661231163392.

Bode, Ingvild, and Anna Nadibaidze. 2024. "Symposium on Military AI and the Law of Armed Conflict: Human-Machine Interaction in the Military Domain and the Responsible AI Framework." *Opinio Juris* (blog), April 4. https://opiniojuris.org/2024/04/04/symposium-on-military-ai-and-the-law-of-armed-conflict-human-machine-interaction-in-the-military-domain-and-the-responsible-ai-framework/.

Bode, Ingvild and Tom Watts. 2021. *Meaning*-Less *Human Control: Lessons from air defence systems on meaningful human control for the debate on AWS*. Drone Wars UK, February 19. https://dronewars.net/2021/02/19/meaning-less-human-control-lessons-from-air-defence-systems-for-lethal-autonomous-weapons/.

———. 2023. "Loitering Munitions and Unpredictability: Autonomy in Weapon Systems and Challenges to Human Control." Odense, Denmark: University of Southern Denmark. www.autonorms.eu/loitering-munitions-and-unpredictability-autonomy-in-weapon-systems-and-challenges-to-human-control/.

Boulanin, Vincent and Maaike Verbruggen. 2017. *Mapping the Development of Autonomy in Weapon Systems*. November. Stockholm, Sweden: Stockholm International Peace Research Institute. www.sipri.org/publications/2017/policy-reports/mapping-development-autonomy-weapon-systems.

Bradshaw, Jeffrey M., Robert R. Hoffman, David D. Woods and Matthew Johnson. 2013. "The Seven Deadly Myths of 'Autonomous Systems.'" *IEEE Intelligent Systems* 28 (3): 54–61. https://doi.org:10.1109/MIS.2013.70.

Brennan-Marquez, Kiel, Karen Levy and Daniel Susser. 2019. "Strange Loops: Apparent Versus Actual Human Involvement in Automated Decision Making." *Berkeley Technology Law Journal* 34 (3): 745–72. https://doi.org/10.15779/Z385X25D2W.

Bryson, Joanna J. and Andreas Theodorou. 2019. "How Society Can Maintain Human-Centric Artificial Intelligence." In *Human-Centered Digitalization and Services*, edited by Marja Toivonen and Evelina Saari, 305–23. Singapore: Springer International Publishing.

Bryson, Joanna J., Mihailis E. Diamantis and Thomas D. Grant. 2017. "Of, for, and by the people: the legal lacuna of synthetic persons." *Artificial Intelligence and Law* 25 (3): 273–91. https://doi.org/10.1007/s10506-017-9214-9.

Bunz, Mercedes. 2012. *Die stille Revolution: wie Algorithmen Wissen, Arbeit, Öffentlichkeit und Politik verändern, ohne dabei Lärm zu machen*. Berlin, Germany: Suhrkamp.

Carmel, Emma and Regine Paul. 2022. "Peace and Prosperity for the Digital Age? The Colonial Political Economy of European AI Governance." *IEEE Technology and Society Magazine* 41 (2): 94–104.

Crawford, Kate. 2021. *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. New Haven, CT: Yale University Press.

Center for Strategic and International Studies. 2022. "Project Convergence: An Experiment for Multidomain Operations." February 16. Washington, DC: Center for Strategic and International Studies. www.csis.org/analysis/project-convergence-experiment-multidomain-operations.

Crootof, Rebecca. 2016. "War Torts: Accountability for Autonomous Weapons." *University of Pennsylvania Law Review* 164 (6): 1347–1402.

Crootof, Rebecca, Margot E. Kaminski and W. Nicholson Price II. 2023. "Humans in the Loop." *Vanderbilt Law Review* 76 (2): 429–510. https://scholarship.law.vanderbilt.edu/vlr/vol76/iss2/2.

Deeks, Ashley, Noam Lubell and Daragh Murray. 2019. "Machine Learning, Artificial Intelligence, and the Use of Force by States." *Journal of National Security Law & Policy* 10 (1): 1–25.

Dekker, S. W. A. and D. D. Woods. 2002. "MABA-MABA or Abracadabra? Progress on Human–Automation Co-ordination." *Cognition, Technology & Work* 4: 240–44. https://doi.org/10.1007/s101110200022.

Demarest, Colin and Jen Judson. 2023. "What Project Convergence will look like after bucking its yearly rhythm." *Defense News*, October 11. www.defensenews.com/battlefield-tech/c2-comms/2023/10/11/what-project-convergence-will-look-like-after-bucking-its-yearly-rhythm/.

Development, Concepts and Doctrine Centre. 2018. "Joint Concept Note 1/18: Human-Machine Teaming." Swindon, UK: UK Ministry of Defence. https://assets.publishing.service.gov.uk/media/5b02f398e5274a0d7fa9a7c0/20180517-concepts_uk_human_machine_teaming_jcn_1_18.pdf.

Ekelhof, Merel A. C. 2017. "Complications of a Common Language: Why It Is so Hard to Talk about Autonomous Weapons." *Journal of Conflict and Security Law* 22 (2): 311–31. https://doi.org/10.1093/jcsl/krw029.

Ekelhof, Merel and Giacomo Persi Paoli. 2020. "The Human Element in Decisions about the Use of Force." Geneva, Switzerland: United Nations Institute for Disarmament Research. https://unidir.org/publication/the-human-element-in-decisions-about-the-use-of-force/.

Elish, Madeleine Clare. 2019. "Moral Crumple Zones: Cautionary Tales in Human-Robot Interaction." *Engaging Science, Technology, and Society* 5: 40–60. https://doi.org/10.17351/ests2019.260.

Gebru, Timnit. 2023. "Don't Fall for the AI Hype," January 19, in *Tech Won't Save Us* podcast with Paris Marx, 1:03:32. https://techwontsave.us/episode/151_dont_fall_for_the_ai_hype_w_timnit_gebru.

Grand-Clément, Sarah. 2023. *Artificial Intelligence Beyond Weapons: Application and Impact of AI in the Military Domain*. Geneva, Switzerland: United Nations Institute for Disarmament Research.

Hambling, David. 2023. "Ukraine's AI Drones Seek and Attack Russian Forces Without Human Oversight." *Forbes*, October 17. www.forbes.com/sites/davidhambling/2023/10/17/ukraines-ai-drones-seek-and-attack-russian-forces-without-human-oversight/.

———. 2024. "Ukraine Rolls Out Target-Seeking Terminator Drones." *Forbes*, March 21. www.forbes.com/sites/davidhambling/2024/03/21/ukraine-rolls-out-target-seeking-terminator-drones/.

Hildebrandt, Mireille. 2011. "Introduction: a multifocal view of human agency in the era of autonomic computing." In *Law, Human Agency and Autonomic Computing: The Philosophy of Law Meets the Philosophy of Technology*, edited by Mireille Hildebrandt and Antoinette Rouvroy, 1–11. New York, NY: Routledge.

ICRC. 2021. "ICRC Position and Background Paper on Autonomous Weapon Systems." May 12. Geneva, Switzerland: IRCR. www.icrc.org/sites/default/files/document_new/file_list/icrc_position_on_aws_and_background_paper.pdf.

ICRC and Geneva Academy. 2024. *Artificial Intelligence and Related Technologies in Military Decision-Making on the Use of Force in Armed Conflicts: Current Developments and Potential Implications*. Geneva, Switzerland: ICRC.

IEEE SA Research Group on Issues of AI and Autonomy in Defence Systems. 2024. *A Framework for Human Decision Making Through the Lifecycle of Autonomous and Intelligent Systems in Defense Applications*. New York, NY: IEEE.

Jamieson, Dash. 2024. "Human Machine Teaming: The Intelligence Cycle Reimagined." The Mitchell Forum No. 53, January. Mitchell Institute for Aerospace Studies. https://mitchellaerospacepower.org/wp-content/uploads/2024/01/MI_Forum_53-HMT-FINAL.pdf.

Jaton, Florian. 2020. *The Constitution of Algorithms: Ground-Truthing, Programming, Formulating*. Inside Technology Series. Cambridge, MA: MIT Press.

Johnson, James. 2024. *The AI Commander: Centaur Teaming, Command, and Ethical Dilemmas*. New York, NY: Oxford University Press.

Johnson, Jesse. 2024. "Japan's Defense Ministry unveils first basic policy on use of AI." *The Japan Times*, July 2. www.japantimes.co.jp/news/2024/07/02/japan/sdf-cybersecurity/.

Judson, Jen. 2024. "The robots are coming: US Army experiments with human-machine warfare." *Defense News*, March 25. www.defensenews.com/unmanned/2024/03/25/the-robots-are-coming-us-army-experiments-with-human-machine-warfare/.

Klonowska, Klaudia. 2022. "Article 36: Review of AI Decision-Support Systems and Other Emerging Technologies of Warfare." In *Yearbook of International Humanitarian Law*, volume 23 (2020), edited by Terry D. Gill, Robin Geiß, Heike Krieger and Rebecca Mignot-Mahdavi, 123–53. The Hague, the Netherlands: T. M. C. Asser Press. https://link.springer.com/10.1007/978-94-6265-491-4_6.

Leung, Rebecca. 2004. "The Patriot Flawed?" CBS News, February 19. www.cbsnews.com/news/the-patriot-flawed-19-02-2004/.

Logan, Sarah. 2024. "Tell me what you don't know: large language models and the pathologies of intelligence analysis." *Australian Journal of International Affairs* 78 (2): 220–28. https://doi.org/10.1080/10357718.2024.2331733.

Matthias, Andreas. 2004. "The responsibility gap: Ascribing responsibility for the actions of learning automata." *Ethics and Information Technology* 6 (3): 175–83. https://doi.org/10.1007/s10676-004-3422-1.

Nadibaidze, Anna, Ingvild Bode and Qiaochu Zhang. 2024. *AI in Military Decision Support Systems: A Review of Developments and Debates*. Odense, Denmark: Center for War Studies.

Orlikowski, Wanda J. 1992. "The Duality of Technology: Rethinking the Concept of Technology in Organizations." *Organization Science* 3 (3): 398–427.

Parasuraman, R., T. B. Sheridan and C. D. Wickens. 2000. "A model for types and levels of human interaction with automation." *IEEE Transactions on Systems, Man, and Cybernetics — Part A: Systems and Humans* 30 (3): 286–97.

Parasuraman, Raja and Dietrich H. Manzey. 2010. "Complacency and bias in human use of automation: an attentional integration." *Human Factors: The Journal of the Human Factors and Ergonomics Society* 52 (3): 381–410. https://doi.org/10.1177/0018720810376055.

Penn, Jonnie. 2021. "Algorithmic Silence: A Call to Decomputerize." *Journal of Social Computing* 2 (4): 337–56. https://doi.org/10.23919/JSC.2021.0023.

Peterson, Nolan. 2020. "Iran Admits to Shooting Down Ukrainian Airliner." *The Daily Signal*, January 10. www.dailysignal.com/2020/01/10/why-it-looks-like-iran-shot-down-ukrainian-airliner/.

Renic, Neil C. and Elke Schwarz. 2023a. "Inhuman-in-the-Loop: AI-targeting and the Erosion of Moral Restraint." *Opinio Juris* (blog), December 19. https://opiniojuris.org/2023/12/19/inhuman-in-the-loop-ai-targeting-and-the-erosion-of-moral-restraint/.

———. 2023b. "Crimes of Dispassion: Autonomous Weapons and the Moral Challenge of Systematic Killing." *Ethics & International Affairs* 37 (3): 321–43. https://doi.org/10.1017/S0892679423000291.

Roff, Heather M. and Richard Moyes. 2016. "Meaningful Human Control, Artificial Intelligence and Autonomous Weapons." Briefing paper prepared for the Informal Meeting of Experts on Lethal Autonomous Weapons Systems, UN Convention on Certain Conventional Weapons, April. London, UK: Article 36. www.article36.org/wp-content/uploads/2016/04/MHC-AI-and-AWS-FINAL.pdf.

Rogoway, Tyler. 2018. "Watch Iron Dome Accidentally Launch 11 Interceptors at Machine Gun Fire." The War Zone, March 27. www.thedrive.com/the-war-zone/19651/too-sensitive-watch-iron-dome-accidentally-launch-11-interceptors-at-machine-gun-fire.

Roth, Andrea L. 2016. "Trial by Machine." *Georgetown Law Journal* 104 (5): 1245–1305.

Sharkey, Noel. 2016. "Staying in the Loop: Human Supervisory Control of Weapons." In *Autonomous Weapons Systems: Law, Ethics, Policy*, edited by Nehal Bhuta, Susanne Beck, Robin Geiß, Hin-Yan Liu and Claus Kreß, 23–38. Cambridge UK: Cambridge University Press.

Sienknecht, Mitja. 2024. "Proxy responsibility: addressing responsibility gaps in human-machine decision making on the resort to force." *Australian Journal of International Affairs* 78 (2): 191–99. https://doi.org/10.1080/10357718.2024.2327384.

Strasser, Anna. 2022. "Distributed responsibility in human–machine interactions." *AI and Ethics* 2 (3): 523–32. https://doi.org/10.1007/s43681-021-00109-5.

Susnea, Elena. 2012. "Decision Support Systems in Military Action: Necessity, Possibilities and Constraints." *Journal of Defense Resources Management* 3 (2): 131–40.

Taddeo, Mariarosaria and Alexander Blanchard. 2022. "A Comparative Analysis of the Definitions of Autonomous Weapons Systems." *Science and Engineering Ethics* 28 (5): 37. https://doi.org/10.1007/s11948-022-00392-3.

Tsamados, Andreas, Luciano Floridi and Mariarosaria Taddeo. 2024. "Human control of AI systems: from supervision to teaming." *AI and Ethics*, May. https://doi.org/10.1007/s43681-024-00489-4.

About CIGI

À propos du CIGI

Credits

Managing Director and General Counsel Aaron Shull
Director, Program Management Dianna English
Program Manager and Research Associate Kailee Hilt
Senior Publications Editor Jennifer Goyder
Graphic Designer Sepideh Shomali

67 Erb Street West
Waterloo, ON, Canada N2L 6C2
www.cigionline.org

CIGI