Submission pursuant to A/RES/79/239

"Artificial Intelligence in the Military Domain"

The following is the AutoNorms project's submission pursuant to Resolution 79/239 on "Artificial intelligence in the military domain and its implications for international peace and security" adopted by the United Nations General Assembly on 24 December 2024. The resolution requests the UN Secretary-General to seek views, including those of Member States, civil society, the scientific community and industry, on "opportunities and challenges posed by the application of artificial intelligence in the military domain, **with specific focus on areas other than lethal autonomous weapons systems**". The AutoNorms team welcomes the opportunity for representatives of academia to submit their views on this important and timely topic.

The AutoNorms project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 852123). Led by Professor Ingvild Bode and hosted by the Center for War Studies at the University of Southern Denmark, the project examines how the integration of artificial intelligence (AI) technologies into weapon systems and military targeting shapes international norms governing the use of force.[1]

Introduction

Over the past 2-3 years, the international debate about applications of AI in the military domain has been characterized by two significant, near-simultaneous changes. First, there has been a **move away from its predominant focus on autonomous or AI technologies in weapon systems** towards considering AI technologies across a wider range of military decision-making tasks, especially in relation to targeting. To reflect his move, this submission focuses on the employment of **AI-based decision support systems (AI DSS),** or systems that are meant to be used as tools to directly or indirectly inform the complex process of use-of-force decision-making, for example, by analyzing large volumes of data, recognizing patterns within the data, predicting scenarios, or recommending potential courses of action to human decision-makers.

Second, there has been a **growing emphasis on human-machine interaction** in the context of using AI in the military domain.[2] This emphasis results from the broad recognition that, even when humans are 'in' or 'on' the loop of targeting decision-making, they need to exercise a sufficient level of oversight, control, and agency over the targeting process. Human oversight is a governance principle featuring prominently across various international initiatives, including A/RES/79/239. However, **dynamics of human-machine interaction as part of the use of AI DSS both introduce new issues and solidify existing sets of challenges that require governance attention.** Our submission highlights these challenges and the need to ensure the exercise of human oversight and agency throughout the full targeting decision-making spectrum. It is structured in three parts, starting with explicating challenges of human-machine interaction, then commenting on the relative under-development of the international debate about AI DSS, and finally, sketching a way forward.

---

[1] The members of the AutoNorms team are Professor Ingvild Bode, Dr Hendrik Huelss, Dr Anna Nadibaidze, Dr Guangyu Qiao-Franco, and Dr Qiaochu Zhang. The AutoNorms project is based at the Center for War Studies, University of Southern Denmark, Odense, Denmark. For more information, please visit our website: www.autonorms.eu.

[2] Ingvild Bode and Anna Nadibaidze, "Symposium on Military AI and the Law of Armed Conflict: Human-Machine Interaction in the Military Domain and the Responsible AI Framework," *Opinio Juris*, April 4, 2024, 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/.

Challenges of human-machine interaction in the use of AI DSS

The use of AI DSS involves various dynamics of human-machine interaction because military personnel such as operators and intelligence analysts routinely and increasingly interact with a network of AI systems throughout the targeting process. These interactions involve multiple challenges **which have the potential to affect the exercise of human agency**, or humans' capacity to understand a system's functions and its effects in a relevant context; deliberate and decide upon suitable actions in a timely manner; and act in a way where responsibility is guaranteed.[3]

Dynamics of human-machine interaction result in **distributed agency between humans and AI systems, where they are not separated into two distinct entities but rather form part of a socio-technical system.**[4] As part of this system, both sides may influence each other in different ways, which then translate into various forms of distributed agency located along a spectrum. In some instances, dynamics of human-machine interaction will offer more opportunities for exercising human agency in targeting decisions. In other instances, however, the humans involved in use-of-force decision-making will be more constrained in their ability to exercise agency.

For example, **humans' ability to exercise agency might be limited by cognitive biases such as automation bias or anchoring bias**. Humans could over-trust AI DSS even when knowing that there might be malfunctions or unintended errors involved, risking an overreliance on algorithmic outputs without engaging in the critical deliberations and assessments that are needed to exercise human agency, especially in critical targeting decisions that might inflict death, destruction, and severe harm. Such biases are typically exacerbated by the increased speed of AI-assisted military decision-making, especially in contexts where there are high levels of pressure to act rapidly. They can also be exacerbated by AI DSS that are used for prescription or recommendations, because such systems restrict the options or courses of action available to human decision-makers.

Moreover, given that AI DSS are likely to be employed not individually but rather as part of a network of systems, the increased complexity of interactions can result in situations where humans act upon some outputs suggested by AI DSS, but do not overall exercise a high quality of agency. Due to these and many other concerns related to interactions between humans and AI DSS, **there is a need to further investigate challenges of human-machine interaction that result in AI DSS not positively 'supporting' humans but rather undermining humans' ability to exercise agency**.[5]

The risks of not addressing challenges of distributed agency are substantial. First, situations where humans are restricted in their exercise of agency **raise questions about compliance with international humanitarian law**, which requires that humans be held accountable and legally responsible for violations of legal principles. Although humans remain officially in control of the selection and engagement of targets, there are concerns about the exact role played by humans in context of using AI DSS in practice.

---

[3] Anna Nadibaidze, Ingvild Bode, and Qiaochu Zhang, *AI in Military Decision Support Systems: A Review of Developments and Debates* (Odense: Center for War Studies, 2024), https://www.autonorms.eu/ai-in-military-decision-support-systems-a-review-of-developments-and-debates/.

[4] Ingvild Bode, *Human-Machine Interaction and Human Agency in the Military Domain*, Policy Brief No. 193 (Waterloo, ON: Centre for International Governance Innovation, 2025), https://www.cigionline.org/publications/human-machine-interaction-and-human-agency-in-the-military-domain/.

[5] Anna Nadibaidze, "Do AI Decision Support Systems 'Support' Humans in Military Decision-Making on the Use of Force?" *Opinio Juris*, November 29, 2024, https://opiniojuris.org/2024/11/29/do-ai-decision-support-systems-support-humans-in-military-decision-making-on-the-use-of-force/.

Second, these concerns also extend to the risk of **negatively affecting moral agency and responsibility in warfare**. Challenges of human-machine interaction that result in distributed agency would allow humans to feel less morally responsible for decisions that could affect other people's lives. They also risk making the human role a nominal, 'box-checking' exercise which can *de facto* be compared with AI DSS playing an 'autonomous' role because the human role is substantially reduced.

Third, there are **security and operational risks related to distributed agency dynamics**, especially when they give too prominent roles to AI DSS and algorithmic outputs. AI systems often malfunction, are trained on biased sets of data which do not apply beyond the training context or specific contexts of use, as well as integrate assumptions that might not be strategically or operationally beneficial.

Various types of biases, issues of trust, uncertainties, targeting and military doctrines, political and societal contexts in which AI DSS are used—all these aspects **can lead to dynamics of distributed agency which limit the exercise of human agency and prioritize algorithmic outputs**. It is important to investigate these dynamics and ensure that distributed agency provides more opportunities than limitations to human decision-makers in warfare.

Relative under-development of the international debate on AI DSS

Despite increasing reports about the use of AI DSS in recent and ongoing armed conflicts, and the significant challenges and risks they pose to the effective exercise of human agency, **the international debate on human-machine interaction in the use of AI DSS remains insufficiently developed**, particularly within intergovernmental UN settings. Current discussions on AI in the military domain, including those within the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems (GGE on LAWS), have focused on the use of AI at the tail-end of the targeting process, specifically autonomy and AI in weapon systems. This narrow focus risks overlooking or failing to address critical normative, legal, ethical, security, and operational risks that can proliferate and compound throughout the entire targeting decision-making process.

An increasing, albeit still limited, number of stakeholders are raising this issue at international multistakeholder forums, such as the Summits on Responsible Artificial Intelligence in the Military Domain (REAIM). Some international non-governmental organisations and research institutes—such as the International Committee of the Red Cross (ICRC), the Stockholm International Peace Research Institute (SIPRI), the UN Institute for Disarmament Affairs (UNIDIR), and the Asser Institute—have initiated discussions on challenges posed by AI in the military domain, beyond the issue of autonomy in weapon systems. Despite this progress, there remains **a clear need to develop a more comprehensive and inclusive international multistakeholder debate to guide the responsible development and deployment of AI DSS in military contexts.**

Way forward

In closing, we sketch three ways intended to move the international debate about applications of AI in the military domain forward:

1. **Increase awareness for the implications of practices of designing, developing, and using AI DSS.**
   States and other stakeholders across industry, civil society, and academia engaged in the governance, development, and use of AI DSS for military targeting must consider the implications of their practices. These practices influence what counts as 'appropriate' ways of considering and employing AI DSS and thereby shape what becomes the accepted, requisite quality of human oversight and agency exercised over

the whole process of use-of-force decision-making. To increase such awareness, the debate pursuant to A/RES/79/239 at the UNGA First Committee should centrally focus on the issue of AI DSS in the military domain.

2. **Consistently map both 'best' practices and 'problematic' practices associated with the design, development, and use of AI DSS.** To get a better sense of the direction that the design, development, and use of AI DSS take, states and other stakeholders need to closely map their own (and others') practices. While there have been some limited efforts to exchange potential best practices, we also need to be attentive to practices with potentially problematic effects. This should encompass practices exercised across the full lifecycle of AI systems from development to use and post-use review. Mapping such practices would offer stakeholders a better overview of which practices may be beneficial, i.e., provide opportunities for the exercise of human agency, and which practices may be problematic, i.e., limit the exercise of human agency, and therefore assess the desirability of particular practices.

3. **Pursue the debate on AI DSS within a multistakeholder format.** States should work with diverse stakeholders—including academics across social sciences and technical disciplines, civil society representatives, and international organizations—to develop normative guidance and regulation, especially regarding the human role in military decision-making. Moreover, top-down processes towards governing AI DSS should be accompanied by a bottom-up, standard-setting process focused on establishing operational standards. Such an inclusive approach could strike a balance between national security and humanitarian concerns, while reinforcing the need to ensure that humans can exercise agency in use-of-force decisions.