Views of members of the scientific community and civil society; specifically, we are a group of academics with expertise in ethical, legal and political dimensions of military Artificial Intelligence and herewith put forward our shared views pursuant to resolution 79/239 "Artificial intelligence in the military domain and its implications for international peace and security" adopted by the General Assembly on 24 December 2024, in accordance with the request of the UN Secretary-General contained in Note Verbale ODA/2025-00029/AIMD

Introduction:

The rapid advancement and integration of AI technologies into targeting operations have sparked ongoing debates surrounding their ethical, legal, and operational implications. Over the past decade, the discourse on AI in warfare has largely centered on autonomous weapon systems (AWS),¹ driven in part by the initiation of discussions in 2013 and the formalization of a regulatory process under the UN Convention on Certain Conventional Weapons (CCW) and the Group of Governmental Experts on Lethal Autonomous Weapons Systems (GGE LAWS), which exclusively focuses on lethal AWS.² However, the increasing integration of AI-based decision-support systems (AI-DSS) into targeting practices³ introduces new layers of complexity that demand closer attention from a broad range of stakeholders. This submission responds to that need, structured around three key components: (1) a brief overview of how AI-DSS are currently used in targeting decisions; (2) an analysis of key concerns, including how these systems shape the potential exercise of human judgement and control and underline fundamental gaps in global governance; and (3) a concluding set of recommendations.

1. Overview of AI-DSS and the joint targeting cycle

Defined as "the process of selecting and prioritizing targets and matching the appropriate response to them, considering operational requirements and capabilities,"⁴ targeting is a core military function at the very heart of warfare. While the potential range of use cases for AI-DSS in military decision-making is broad, in targeting, AI-DSS can be understood to serve as **tools** that use AI techniques to collect and analyze data, provide information about the operational environment as well as actionable recommendations, with the aim of aiding military decision-makers in evaluating factors relevant to legal compliance such as taking precautions and ensuring proportionality in attacks.

More specifically, AI-DSS are increasingly integrated across multiple phases of the joint targeting cycle (JTC), including within target development and prioritization, capabilities analysis, and mission execution. The JTC is a reflective example of a structured process used by military forces to identify,

¹ The latest definition of AWS from the CCW GGE LAWS Rolling Text (26 November 2024): "A lethal autonomous weapon system can be characterized as an integrated combination of one or more weapons and technological components that enable the system to identify and/or select, and engage a target, without intervention by a human user in the execution of these tasks." On file with authors.

² For a brief overview of some of the latest developments of the GGE LAWS see Jeroen van den Boogaard, *Warning! Obstacles Ahead! The Regulation of Autonomous Weapons Systems in the GGE LAWS, Opinio* Juris, 4 March 2024 found at: <u>https://opiniojuris.org/2024/03/04/warning-obstacles-ahead-the-</u> regulation-of-autonomous-weapons-systems-in-the-gge-laws/.

³ There have been several reported uses of AI-DSS by Israel in Gaza and potentially in Lebanon, by both Ukraine and Russia in the ongoing conflict, and by the United States in its actions against Houthi rebels in the Red Sea and in Yemen, to name a few. For a comprehensive overview of literature in this space, see e.g., Anna Nadibaidze, Ingvild Bode, and Qiaochu Zhang, *AI in Military Decision Support Systems, A Review of Developments and Debates,* Centre for War Studies, University of Southern Denmark, November 2024. Found here: <u>https://www.autonorms.eu/ai-in-military-decision-support-systems-a-review-of-developments-and-debates/</u>.

⁴ United States Department of Defense, *Dictionary of Military and Associated Terms*, March 2017, found at: https://www.tradoc.army.mil/wp-content/uploads/2020/10/AD1029823-DOD-Dictionary-of-Military-and-Associated-Terms-2017.pdf.

evaluate, and engage targets while ensuring compliance with operational, legal, and ethical standards,⁵ generally consisting of six (non-linear) phases:

1. End-State and Commander's Objectives: Defining strategic military goals and desired outcomes.

2. **Target Development and Prioritization**: Identifying, verifying/validating, and prioritizing targets based on intelligence and mission goals.

3. Capabilities Analysis: Assessing the available strike options and their effectiveness.

4. **Force Assignment**: Allocating specific military assets (e.g., airstrikes, artillery, cyber operations) to engage the target.

5. **Mission Execution**: Carrying out the targeting operation while ensuring compliance with relevant laws and the rules of engagement.

6. **Assessment**: Evaluating the effectiveness of the operation and adjusting for future operations, if necessary.

Within this framework, AI DSS are assumed to serve primarily as informational and analytical tools which support human decision-making rather than supplant it. However, this assumption and framing obscures how AI-DSS influence human cognitive processes within the JTC. This impact on human decision-making is often underestimated and remains insufficiently examined, leaving critical discussions about the role of AI-DSS largely absent from current policy debates.

2. Analysis of Key Concerns

a. (Meaningful) Human Judgement and Control

AI-DSS are often portrayed as enhancing human decision-making and the quality of decisions therein. The perception of AI-DSS as mere subsidiary tools has led to a narrative that the integration of AI-DSS poses fewer challenges than AWS, given that these systems do not directly "engage" targets (i.e., they do not have an inherent capability to directly carry out the use of force) and are tools that assist human commanders. The outputs are ostensibly ultimately reviewed through (several layers of) human oversight, such as processes of verifying and validating targets using additional intelligence sources.⁶ As a result, errors or inaccuracies in AI-DSS outputs are often seen as non-critical, based on the assumption that robust human oversight and appropriate control will compensate for them. However, closer examination reveals that this control is frequently superficial, offering only the appearance of, rather than actual meaningful, or context-appropriate, human judgement and control.

This is because AI-DSS structure and condition the quality of human control and oversight and limit the ways control and oversight can be exercised. The use of AI-DSS creates a shared decision-making space between human military personnel and AI technologies. States appear to have recognized and focused on many of the advantages of this shared decision-making space for military personnel, i.e., how the use of AI-DSS advances human decision-making through offering data-driven insights. But using AI-DSS also delimits the capacity to exercise human oversight and control because of the

⁵ Michael Schmitt et al, *Joint and Combined Targeting: Structure and Process*, Chapter 13 in Jens David Ohlin (ed) *Weighing Lives in War* (Oxford, 2017). See also, Jessica Dorsey and Marta Bo, *AI-Enabled Decision-Support Systems in the Joint Targeting Cycle: Legal Challenges, Risks, and the Human(e) Dimension*, forthcoming 2025, *International Law Studies.* "Targeting generally involves four key steps: (1) objectives and guidance, (2) planning, (3) execution, and (4) assessment. Encapsulating these four key steps, the United States and NATO outline their targeting processes through similar six-phase cycles [addressed in this submission]. As the reader can discern, different states employ different doctrines for targeting. What is important ... is not necessarily the specific labels for various steps followed by any given state, but rather how and when compliance with the principle[s of IHL are] incorporated into the targeting process."

⁶ Alexander Blanchard and Laura Bruun, *Automating Military Targeting: A Comparison Between Autonomous Weapon Systems and AI-Enabled Decision Support Systems*, Stockholm International Peace Research Institution (SIPRI) forthcoming 2025 (draft on file with authors).

technologies' complexity and the increased speed (and therefore scale) it can bring to decision-making processes. Rather than supporting human oversight, using AI-DSS may risk humans becoming little more than reactive cogs in socio-technical systems.⁷ Moreover, this configuration risks amplifying adverse human biases, such as automation bias, anchoring bias, or cognitive action bias, to the detriment of exercising qualitatively high levels of human control.⁸ Considering AI-DSS as a distinct form of technology therefore reveals significant challenges associated with military AI and human oversight, challenges that extend beyond those that arise when simply integrating the technology in weapon systems.

Recent conflicts have shown the risks associated with AI-DSS being employed in critical functions, such as target selection and even nomination, and their conditioning and constraining of human involvement, affecting the fulfilment of core legal obligations embedded within the JTC. The use of AI-DSS raises fundamental concerns about whether human decision-makers can retain adequate cognitive autonomy over the JTC process or whether humans will become overly reliant on algorithmic outputs for critical judgements in the context of armed conflict.⁹ Consequently, there are significant legal concerns regarding the effects of such systems on decision-making processes and use of force decisions and ability for users to comply with IHL obligations, especially with respect to the obligation to take all feasible precautions to minimize civilian harm to the greatest extent possible in attack and comply with the principles of distinction and proportionality.¹⁰

Importantly, these concerns are not new. There is extensive debate around how to preserve meaningful human judgment and human agency when conducting IHL-evaluative legal assessments, in the context of AWS. These discussions—which include expert analysis on accountability, human-machine interaction, automation bias, and the effect of AI systems on legal and ethical reasoning¹¹—provide valuable lessons that must inform discussions around military AI and specifically the use of AI-DSS in military contexts.

b. AI-DSS: Understudied, Under-Addressed and Unregulated

Framing AI-DSS as mere tools, has led to an underestimation and lack of analysis on the way their use affects the cognitive decision-making process within the JTC. The relative lack of attention paid to AI-

⁷ Ingvild Bode, *Human-Machine Interaction and Human Agency in the Military Domain*, Policy Brief No. 193 (Waterloo, ON: Centre for International Governance Innovation, 2025).

⁸ Dorsey and Bo, supra n. 5

⁹ *Ibid*; see also Anna Nadibaidze, Ingvild Bode, and Qiaochu Zhang, *AI in Military Decision Support Systems, A Review of Developments and Debates,* Centre for War Studies, University of Southern Denmark, November 2024. Found at: <u>https://www.autonorms.eu/ai-in-military-decision-support-systems-a-review-of-developments-and-debates/</u>

¹⁰ Article 57 of the First Additional Protocol to the Geneva Conventions. See also Dorsey, Bo supra n. 5 (on AI-DSS and their effects on the principle of precautions); Jessica Dorsey, *Proportionality under Pressure: The Effects of AI-Enabled Decision Support Systems, the Reasonable Commander Standard and Human(e) Judgment in Targeting*, forthcoming *International Review of the Red Cross* (2025) (on AI-DSS and their effects in the context of IHL proportionality assessments).

¹¹ Marta Bo, *Autonomous Weapons and the Responsibility Gap in light of the Mens Rea of the War Crime of Attacking Civilians in the ICC Statute*, 19 Journal of International Criminal Justice 2021; Bo, M., Bruun, L. and Boulanin, V., *Retaining Human Responsibility in the Development and Use of Autonomous Weapon Systems: On Accountability for Violations of International Humanitarian Law Involving AWS* (SIPRI: Stockholm, Oct. 2022), p. 41; Boulanin, V., Bruun, L. and Goussac, N., *Autonomous Weapon Systems and International Humanitarian Law: Identifying Limits and the Required Type and Degree of Human–Machine Interaction* (SIPRI: Stockholm, June 2021), p. 54; and Bruun, L., Bo, M. and Goussac, N., *Compliance with International Humanitarian Law in the Development and Use of Autonomous Weapon Systems: What Does IHL Permit, Prohibit and Require?* (SIPRI: Stockholm, Mar. 2023), p. 24. Elke Schwarz, "The (im)possibility of meaningful human control for lethal autonomous weapons systems," *Humanitarian Law and Policy*, 29 August 2018, found at: <u>https://blogs.icrc.org/law-and-policy/2018/08/29/im-possibility-meaningful-human-control-lethal-</u> <u>autonomous-weapon-systems/</u>

DSS so far can partly be attributed to the fact that such systems are seen to be used with a human *in* or *on* the loop framework, with their outputs ostensibly reviewed by one or more individuals during the targeting process. As a result, current understandings of AI-DSS use appear to align with widely supported principles of human control and oversight. However, this gap in the debate is also caused by a lack of transparency around how specific AI-DSS function, and a consistent failure to comprehensively examine how they are being used in practice.

Additionally, the persistent focus on AWS at the expense of AI-DSS obscures the growing reliance on AI in shaping operational and strategic outcomes. Unlike AWS, which have been debated in the framework of the CCW for the past decade, AI-DSS lack a comparable institutional platform. Attention to AI-DSS remains scattered across various initiatives but these efforts have yet to provide the dedicated regulatory focus or coordination needed.

3. <u>Recommendations:</u>

- i) **Reassert** the central role of human cognitive and legal reasoning in military operations by implementing safeguards that ensure key legal assessments remain grounded in human(e) judgment. Leverage existing insights from debates on AWS and research on human-machine teaming and human-computer interaction to inform discussions on AI-DSS.
- ii) **Recognize** and address the incremental effects of AI-DSS design and use on human cognitive reasoning and critical deliberation. Promote awareness and attentiveness as a crucial part of reasserting and strengthening the exercise of human agency in targeting decision-making.
- iii) Reinforce calls for greater attention to the implications of AI-DSS in armed conflict. Utilize platforms such as the UN General Assembly's First Committee on Disarmament and International Security and the Global Commission on the Responsible Use of AI in the Military Domain to foster inclusive and complementary discussions on the associated risks and systemic changes AI-DSS introduce.

Submitted by:

Jessica Dorsey, Assistant Professor of International Law, Utrecht University; Expert Member, Global Commission on the Responsible Use of AI in the Military Domain (corresponding author, <u>j.l.dorsey@uu.nl</u>)

Dr. Marta Bo, Senior Researcher, T.M.C. Asser Instituut, The Hague

Dr. Ingvild Bode, Professor of International Relations and Director of the Center for War Studies, University of Southern Denmark; Expert Member, Global Commission on the Responsible Use of AI in the Military Domain

Dr. Elke Schwarz, Professor of Political Theory, Queen Mary University; Vice-Chair of the International Committee of Robot Arms Control