



تريندز للبحوث والاســتشــارات TRENDS RESEARCH & ADVISORY



Foreign, Commonwealth & Development Office

## **GC REAIM Expert Policy Note Series** The Nuts and Bolts of the Governance of Military Artificial Intelligence: A Balancing Act Between Technical Depth and Practicality

Yasmin Afina and Giacomo Persi Paoli

May 2025





# **GC REAIM Expert Policy Note Series**

The Nuts and Bolts of the Governance of Military Artificial Intelligence: A Balancing Act Between Technical Depth and Practicality

Authors: Yasmin Afina and Giacomo Persi Paoli

May 2025

Cover photo: unsplash

The Global Commission on Responsible Artificial Intelligence in the Military Domain (GC REAIM) is an initiative of the Government of the Netherlands that was launched during the 2023 REAIM Summit on Responsible Artificial Intelligence in the Military Domain in The Hague. Upon request of the Dutch Ministry of Foreign Affairs, the Hague Centre for Strategic Studies acts as the Secretariat of the Commission.

The GC REAIM Expert Policy Note Series was funded by the Foreign, Commonwealth and Development Office (FCDO) of the United Kingdom. GC REAIM Experts maintained full discretion over the topics covered by the Policy Notes. The contents of the GC REAIM Expert Policy Note series do not represent the views of the Global Commission as a whole. The Policy Notes are intended to highlight some key issues around the governance of AI in the military domain and provide policy recommendations. This Policy Note was released during the Commission Meeting in Abu Dhabi, hosted and graciously supported by Trends Research and Advisory.

© The Hague Centre for Strategic Studies. All rights reserved. No part of this report may be reproduced and/ or published in any form by print, photo print, microfilm or any other means without prior written permission from HCSS. All images are subject to the licenses of their respective owners

HCSS Lange Voorhout 1 2514 EA The Hague

Follow us on social media: @hcssnl

The Hague Centre for Strategic Studies Email: info@hcss.nl Website: www.hcss.nl





### 1. Introduction

Efforts and initiatives related to the governance of artificial intelligence (AI) in the military domain are mushrooming at the international, regional and national levels. Whether through a national strategy document or through a platform for dialogue, such as the Responsible AI in the Military Domain (REAIM) Summit, states around the globe and the wider multi-stakeholder community are grappling with the potentially far-reaching and transformative impact of these technologies. Within the United Nations, just a few months ago, the General Assembly adopted resolution 79/239 on AI in the military domain and its implications for international peace and security.<sup>1</sup> The resolution recognizes both the opportunities and the risks and challenges that these technologies may bring for international peace and security, with on-going work seeking to capture states' views on this issue.<sup>2</sup>

Underlying all these efforts is a widespread recognition that, fundamentally, AI in the military domain is an inherently technical and complex field. As a general-purpose technology, the notion of artificial intelligence extends far beyond a "widget" with a wide range of possible applications. Its nature thus requires governance approaches, initiatives and solutions that balance technical depth with practicality for stakeholders.<sup>3</sup> This echoes the findings of a nascent yet growing body of research in the civilian domain.<sup>4</sup> Yet, despite such acknowledgment even in the military domain, the specifics on *why* regulation and on *how* this will be implemented are consistently missing from the picture.

<sup>&</sup>lt;sup>1</sup> United Nations General Assembly, *Artificial Intelligence in the Military Domain and Its Implications for International Peace and Security*, A/RES/79/239 (31 December 2024), https://unidir.org/wp-content/uploads/2025/03/UN\_General\_Assembly\_A\_RES\_79\_239-EN.pdf.

<sup>&</sup>lt;sup>2</sup> The resolution calls, among others, for states and the wider multi-stakeholder community to submit their views on the opportunities and risks stemming from Al in the military domain for international peace and security. The deadline for submissions was 11 April 2025. The submitted views will be used as a basis to inform a report to be drafted and released by the United Nations Secretary-General in the summer of 2025. UNIDIR drafted a briefing note to support states formulate their views. See Giacomo Persi Paoli and Yasmin Afina, *Al in the Military Domain: A Briefing Note for States* (Geneva: United Nations Institute for Disarmament Research, 2025), https://unidir.org/publication/ai-military-domain-briefing-note-states/. <sup>3</sup> For UNIDIR's work on this, see Yasmin Afina, *The Global Kaleidoscope of Military Al Governance: Decoding the 2024 Regional Consultations on Responsible Al in the Military Domain* (Geneva: United Nations Institute for Disarmament Research, 2024), https://unidir.org/publication/the-global-kaleidoscope-of-military-ai-governance/; Sarah Grand-Clément, *Artificial Intelligence Beyond Weapons: Application and Impact of Al in the Military Domain* (Geneva: United Nations Institute for Disarmament Research, 2024), https://unidir.org/publication/the-global-kaleidoscope.of-military-ai-governance/; Sarah Grand-Clément, *Artificial Intelligence Beyond Weapons: Application and Impact of Al in the Military Domain* (Geneva: United Nations Institute for Disarmament Research, 2023),

https://unidir.org/publication/artificial-intelligence-beyond-weapons-application-and-impact-of-ai-in-themilitary-domain/. See also Michael C. Horowitz, 'Artificial Intelligence, International Competition, and the Balance of Power (May 2018)', May 2018, https://doi.org/10.15781/T2639KP49.

<sup>&</sup>lt;sup>4</sup> For example, see Anka Reuel et al., 'Position Paper: Technical Research and Talent Is Needed for Effective Al Governance', *CoRR* abs/2406.06987 (2024), https://doi.org/10.48550/ARXIV.2406.06987.

This policy note thus addresses how the technical aspect of military AI can be taken account of in the governance of its development, deployment and use. It first takes stock of existing approaches to the technical dimension of AI governance in the military domain, specifically focusing on international and regional initiatives (in Section 2). It also explores a non-exhaustive series of good practices, in addition to their respective boundaries and limitations (in Sections 3 and 4). Select initiatives on the governance of civilian AI are also dissected (in Section 5) in order to identify possible lessons for the military domain. Finally, the note closes (in Section 6) by offering further food-for-thought as states and the wider international and multi-stakeholder community open their discussions on the governance of these technologies.

# The Technical Dimension of International and Regional AI Governance in the Military Domain: Taking Stock

Efforts, initiatives and processes for governance of AI in the military domain generally recognize the technical dimension of AI and the subsequent need for governance to be embedded into the discussions and deliberations and into the eventual policy products. Without prejudice to what states may negotiate and establish in the future, the following subsections provide a non-exhaustive overview of initiatives of relevance for AI in the military domain and the extent to which the technical dimension has been taken into account.

#### 2.1 Within the United Nations

#### United Nations General Assembly Resolution 79/239

General Assembly resolution 79/239 on "Artificial Intelligence in the Military Domain and Its Implications for International Peace and Security" factors in the technological component throughout its text, in both the preamble and the operative paragraphs.<sup>5</sup> This is particularly evident with:

- The resolution's whole-of-life-cycle approach (preamble paragraph 3)
- The acknowledgment of the contribution made by the technical community and the private sector, in addition to the wider multi-stakeholder community, to addressing the peace and security implications of AI in the military domain (preamble paragraph 13)
- The inclusion of a technological perspective, among others, in the recognition of the opportunities and challenges arising from the application of AI in the military domain (operative paragraph 2)
- The encouragement of states to convene exchanges on responsible application of AI in the military domain with the wider multi-stakeholder community (operative paragraph 5)
- The call for the Secretary-General to invite the views of the scientific community and industry (as part of the wider multi-stakeholder community) on the opportunities and challenges posed to international peace and security by these technologies (operative paragraph 8)

<sup>&</sup>lt;sup>5</sup> United Nations General Assembly, *Artificial Intelligence in the Military Domain and Its Implications for International Peace and Security*, A/RES/79/239 (31 December 2024), https://docs.un.org/en/a/res/79/239.

*Convention on Certain Conventional Weapons and Lethal Autonomous Weapon Systems* When thinking of military AI, lethal autonomous weapon systems (LAWS) often come to mind. However, the caveat that not all military AI constitute LAWS, and not all LAWS necessarily constitute military AI should be noted. In late 2016, a Group of Governmental Experts (GGE) on emerging technologies in the area of lethal autonomous weapons systems (LAWS) was established under the Certain Conventional Weapons (CCW) Convention. The expert-centric nature of the GGE implicitly recognizes the technological dimension of this issue. This is in fact reflected in the historical approach of the CCW framework to the issue of LAWS: prior to the GGE's establishment in late 2016, the CCW regime held "informal meetings of experts" yearly in 2014, 2015 and 2016.

The GGE drafted 11 guiding principles that were adopted by the 2019 CCW Meeting of the High Contracting Parties. These include, among others, a whole-of-life-cycle approach, as well as mentions of these systems' physical security and non-physical safeguards (e.g., cybersecurity). The current work of the CCW GGE on LAWS focuses, inter alia, on a Rolling Text to capture provisional consensus found in the Group. A key component to this text contains a dedicated section on formulating the characterization of LAWS, thus outlining, among other things, their technical nature.

#### Convention on Certain Conventional Weapons and Lethal Autonomous Weapon Systems

In addition to the CCW GGE on LAWS, the United Nations General Assembly adopted two resolutions on LAWS in 2023 and 2024 (resolutions 78/241 and 79/62, respectively).<sup>6</sup> The resolutions raise, among other things, the technological challenges and concerns that may arise from the "new technological applications in the military domain" (preamble paragraph 3 of resolution 78/241). They also highlight the need for a comprehensive and inclusive approach to address the "full range of challenges and concerns posed by autonomous weapons systems", which includes the technological perspectives to safeguard international peace and security.

#### Human Rights Council Advisory Committee Study

In October 2022, the Human Rights Council requested (through its resolution 51/22) that the Council's Advisory Committee prepare "a study examining the human rights implications of new and emerging technologies in the military domain, while taking into account ongoing discussions within the United Nations system".<sup>7</sup> While the resolution covers more than AI in the military domain, it recognizes that some of the new and

<sup>&</sup>lt;sup>6</sup> United Nations General Assembly, *Lethal Autonomous Weapons Systems*, A/RES/78/241 (22 December 2023), https://docs.un.org/en/A/RES/78/241; United Nations General Assembly, *Lethal Autonomous Weapons Systems*, A/RES/79/62 (2 December 2024), https://docs.un.org/en/A/RES/79/62.

<sup>&</sup>lt;sup>7</sup> United Nations Human Rights Council, *Human Rights Implications of New and Emerging Technologies in the Military Domain*, A/HRC/RES/51/22 (7 October 2022), https://docs.un.org/A/HRC/RES/51/22.

emerging technologies in the military domain "rely on, inter alia, data sets, algorithmbased programming and machine-learning processes" (preamble paragraph 8) and the need to ensure respect for international human rights law across the life cycle of these technologies (preamble paragraph 10).

The Advisory Committee is due to present its study in September 2025 at the Council's 60th session. Following a call for input issued in the autumn of 2023, the Committee has received a number of submissions from Member States and observers, national human rights institutions, non-governmental organizations, civil society organizations and academia.<sup>8</sup>

#### 2.2 Outside the United Nations

#### Responsible AI in the Military Domain

The Netherlands launched the REAIM programme of work by co-hosting, with the Republic of Korea, the inaugural REAIM Summit in February 2023. Part of the summit itself and its agenda was inherently multi-stakeholder. For example, breakout sessions were led by states and non-state stakeholders, thus providing space for technical input notably from the latter but also governmental bodies. These sessions complemented the intergovernmental segment of the summit, which resulted in the adoption of the 2023 Call to Action with its endorsement by a number of states.<sup>9</sup> A similar format was adopted at the second REAIM Summit, held in Seoul in September 2024, which issued the 2024 Blueprint for Action.<sup>10</sup>

In addition, in the lead-up to the Seoul summit and as the second outcome of the 2023 Summit, the Netherlands and the Republic of Korea organized a series of regional consultations in Singapore, Istanbul, Nairobi, Santiago and online, involving both state representatives and select experts. Participants came from a diverse range of backgrounds, such as policymakers, lawyers and technical experts, including from the armed forces.

<sup>&</sup>lt;sup>8</sup> United Nations Human Rights Council, *Human Rights Implications of New and Emerging Technologies in the Military Domain*, n.d., https://www.ohchr.org/en/hr-bodies/hrc/advisory-committee/human-rights-implications.

<sup>&</sup>lt;sup>9</sup> Kingdom of the Netherlands, Ministry of Foreign Affairs and Ministry of Defence, *REAIM 2023 Call to Action* (The Hague: Government of the Netherlands, 16 February 2023),

https://www.government.nl/ministries/ministry-of-foreign-

affairs/documents/publications/2023/02/16/reaim-2023-call-to-action.

<sup>&</sup>lt;sup>10</sup> Republic of Korea, Ministry of Foreign Affairs and Ministry of National Defense, *REAIM 2024 Blueprint for Action* (Seoul: Government of the Republic of Korea, 10 September 2024),

 $https://reaim2024.kr/home/reaimeng/board/bbsDetail.do?encMenuId=4e57325766362f626e5179454e6d6\\e4d4a4d33507a773d3d&encBbsMngNo=366e794c7a644d756342425668444f393053755142673d3d&encBbsNo=6f784e4542386f7735767465766a6531556f4b6149413d3d&ctlPageNow=1&schKind=bbsTtlCn&schWord=#this.$ 

The inaugural REAIM Summit also had a third outcome: the establishment of the Global Commission on Responsible AI in the Military Domain (GC-REAIM).<sup>11</sup> This body is to formulate a series of recommendations for the international community grounded in, among other things, technical evidence. The composition of the GC-REAIM makes it multi-stakeholder by nature.

#### Political Declaration

At the REAIM Summit in February 2023, the United States launched the Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy, which was subsequently amended in November 2023.<sup>12</sup> The declaration has been endorsed by over 50 states. Three thematic working groups have been established under this Political Declaration, dedicated respectively to assurance, accountability and oversight.

While largely state-driven and state-oriented, to a large extent the content of the Political Declaration factors in the inherently technical nature of the issue. This is most notable in the paragraphs that put an emphasis on the importance of design procedures, data sources and transparent documentation (paragraph F); user-oriented design and engineering (paragraph H); the safety, security and effectiveness of military Al capabilities and the need for appropriate and rigorous testing and assurance across their life cycle (paragraph I); and safeguards to mitigate risks of failure (paragraph J).

<sup>&</sup>lt;sup>11</sup> Hague Centre for Strategic Studies, *Global Commission on Responsible Artificial Intelligence in the Military Domain (GC REAIM)*, accessed 8 May 2025, https://hcss.nl/gcreaim/.

<sup>&</sup>lt;sup>12</sup> U.S. Department of State, *Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy* (Washington, D.C.: U.S. Department of State, 2023), https://www.state.gov/bureau-of-arms-control-deterrence-and-stability/political-declaration-on-responsible-military-use-of-artificial-intelligence-and-autonomy.

# 3. International Good Practices, Limitations and Boundaries

Whether within or outside the United Nations, a number of good practices can be extracted from the existing governance approaches, initiatives and efforts, some of which are listed above. Conversely, a series of limitations and boundaries can also be highlighted. Identifying these limitations and boundaries is not meant as criticism of these processes, efforts and initiatives; rather, it is an invitation for reflection as to what could complement and further reinforce the processes, efforts and initiatives, and how this could be achieved.

#### 3.1 International Calls for Multi-Stakeholder Input

#### **Good Practices**

A number of these initiatives and efforts call for input from the multi-stakeholder community. At the highest level, General Assembly resolution 79/239 on AI in the military domain and resolution 78/241 on LAWS both operate on similar models: they call for Member States, observers and the wider multi-stakeholder community to submit views that will feed into a report to be drafted by the United Nations Secretary-General. The above-mentioned call for submissions to the Human Rights Council Advisory Committee is similar.

These open calls are a positive approach to collecting input from the wider multistakeholder community. For issues that are as technical as AI in the military domain, such an approach not only provides space for actors such as businesses, research laboratories, academic experts and technologists to share their in-depth perspectives and provide extra food-for-thought on opportunities, challenges and the operationalization of key recommendations and principles for responsible practices. These open calls also incentivize these actors to reflect on how their everyday work could feed into governance, and more generally to foster dialogue with the policymaking community. Furthermore, the public release of submissions (e.g., on the relevant secretariat's webpage) provides a directory for states and for groups of technologists and technical experts in a specific area, including those from the Global South. This reference can then be used to provide a platform for these voices to consolidate future convenings, dialogues and initiatives.

#### Limitations and Boundaries

While most of these open calls clearly state that the submitted reports will somehow feature as part of the final report into which it feeds (e.g., as an annex, as is the case for resolution 79/239), there is not enough concrete transparency on the extent to which these inputs from the multi-stakeholder community influence and shape the final output.

A further issue is outreach and dissemination: parts of the technical community are not even aware of these calls; when they are, the capacity to prepare such submissions is often limited, particularly for organizations from the Global South. Additionally, some of these calls for submissions (e.g., resolution 79/239) require that the language of submissions from non-state entities must be English. While this may be due to limited resources of the relevant secretariat, this restriction would constitute a significant barrier for non-English-speaking organizations, particularly in the Global South. In other cases, the original language in which the submission was received will be maintained (e.g., resolution 78/241). This may open the door to more submissions, but leaving the submission in its original language may limit subsequent access.

#### 3.2 Inclusive Participation

#### **Good Practices**

There are a number of initiatives, platforms and even processes where the multistakeholder community and multidisciplinary representatives can participate in the discussions.

Inside the United Nations, this is notably the case for the CCW GGE on LAWS due to its unique rules of procedure. Unless stated otherwise (e.g., due to objections expressed by states), the Group's format allows active participation from the multi-stakeholder ecosystem, including civil society and the technical community (e.g., research laboratories). These representatives are even generally provided with dedicated time for interventions and to share inputs to the Group. In addition, they have the potential to collaborate with either states or a United Nations entity to organize side-events, thus helping disseminate research products and perspectives that reach beyond the policymaking realm (e.g., the technical dimension). Moreover, a number of state delegations also include technical experts, which helps to further consolidate expert engagement in the CCW GGE on LAWS and ensure the feasibility and practicality of their respective positions and recommendations. These deliberations also generally have simultaneous interpretation available, thus enabling the participation of both state and non-state representatives in any of the six official languages of the United Nations and ultimately allowing wider participation.

Outside the United Nations, there is also a growing recognition of the importance of multi-stakeholder participation in discussions surrounding AI in the military domain. The REAIM Summits in 2023 and 2024 were largely structured around multi-stakeholder participation through the organization of breakout sessions led by states, civil society, international and regional organizations, industry representatives, and academic experts. In the area of LAWS, Austria convened the 2024 Vienna Conference on Autonomous Weapons Systems. Organized with the objective of addressing and discussing the "profound questions and the various challenges related to the international regulation of autonomous weapons systems", the conference was explicitly open for the participation of "all States, the United Nations, the ICRC, international and regional organisations, as well as representatives of academia, think tanks, industry and civil society".<sup>13</sup>

#### Limitations and Boundaries

Both state and non-state representatives may have limited resources that restrict their ability to make the most of the multi-stakeholder and multidisciplinary nature of discussions, processes and deliberations.

In the case of the CCW GGE on LAWS, its location in Geneva provides opportunities due to its central position in Europe and the presence in the city of a permanent mission of most Member States. But it also creates limitations due to barriers to travel, notably for representatives from the Global South (e.g., due to the need to apply for visas, travel expenses, etc.).

Human resources and capacity may also be limited, restricting the participation of technical representatives in state delegations – notably from the Global South. In fact, some delegations are exclusively represented by Geneva-based representatives who may be handling a series of other portfolios at the same time.

Additionally, unlike the CCW GGE on LAWS, the rules of procedures of such forums do not always guarantee participation by the multi-stakeholder community. While these rules on the exclusive participation of states are in place to enable, under certain circumstances, progress in the deliberations, they may also limit the perspectives represented in the room. In addition to the rules of procedure, limitations on resources and timing may limit states to high-level discussions that do not always allow for technical depth (e.g., the consensus-based rules of procedure within the CCW GGE on LAWS). While the latter limitations may guarantee a more inclusive approach to

<sup>&</sup>lt;sup>13</sup> Republic of Austria, Federal Ministry for European and International Affairs, *Humanity at the Crossroads: Autonomous Weapons Systems and the Challenge of Regulation*, Practical Information Note, Vienna Conference on Autonomous Weapons Systems, 29–30 April 2024,

https://reachingcriticalwill.org/images/documents/Disarmament-fora/other/aws/2024-vienna/documents/practical-info.pdf.

deliberations by ensuring that the voice of every participating state counts, it does not necessarily translate into an ability to reach technical depth.

#### 3.3 Regional Platforms for Discussion

#### **Good Practices**

States are increasingly exploring the provision of regional platforms for discussion on issues surrounding AI in the military domain. Whether more general (e.g., through the REAIM regional consultations in 2024 and upcoming in 2025) or on a specific topic (e.g., on emerging technologies in the area of LAWS), regional dialogues provide a unique opportunity to take stock of and capture specific details of the local security landscape, the socio-cultural context, the state of affairs with regards to digitization and access to technologies, as well as political and historical realities.<sup>14</sup>

Regional conferences and initiatives – such as those hosted by Costa Rica in early 2023, the Philippines in late 2023 and Sierra Leone in early 2024 on autonomous weapon systems – also provide an opportunity to not only raise awareness on these issues, but also to bring ownership to the regions. They also provide a means for information exchange in informal settings, in addition to enabling participants to identify counterparts from other states in the region as well as local experts (e.g., in the 2024 REAIM regional consultations).

#### Limitations and Boundaries

These initiatives are generally resource intensive, from the perspective of human resources, financial resources but also technical resources. While there is a desire to intensify efforts at the regional level, there is also the question of financial sustainability and the scalability of such initiatives.

There is also the desire, by states and experts from the regions, for further clarity on the ambitions and objectives for such a dialogue. For example, they may want to spell out whether the discussions should remain at a high level or whether there may be room for a more technical exchange of information. In the latter case, they must decide the modalities for such cooperation while ensuring that these efforts complement one another and do not hamper global alignment, resulting in harmful fragmentation across regions and across communities.<sup>15</sup>

<sup>&</sup>lt;sup>14</sup> For an overview of the 2024 REAIM regional consultations, see Yasmin Afina, *The Global Kaleidoscope of Military Al Governance: Decoding the 2024 Regional Consultations on Responsible Al in the Military Domain* (Geneva: United Nations Institute for Disarmament Research, 2024), https://unidir.org/wp-

content/uploads/2024/09/UNIDIR\_The\_Global\_Kaleidoscope\_of\_Military\_Al\_Governance.pdf.

<sup>&</sup>lt;sup>15</sup> Yasmin Afina, *The Global Kaleidoscope of Military AI Governance: Decoding the 2024 Regional Consultations on Responsible AI in the Military Domain* (Geneva: United Nations Institute for Disarmament Research, 2024), https://unidir.org/wp-

content/uploads/2024/09/UNIDIR\_The\_Global\_Kaleidoscope\_of\_Military\_AI\_Governance.pdf.

# 4. Beyond International Initiatives: National Frameworks and Efforts

Beyond international initiatives, a number of lessons can also be drawn from national approaches to incorporating technological considerations into governance approaches, ultimately ensuring their practicality and effective implementation. Two specific areas can be highlighted: national strategy documents and other tactical governance frameworks.

#### 4.1 National Strategy Documents

A growing number of states are adopting national strategy documents on AI, albeit with limited reach into the military domain, with a few exceptions. These strategy documents not only constitute a means to operationalize international, high-level principles at the national level, they also provide specific mandates to government bodies and national agencies for the governance of these technologies, including in the military domain when applicable.

National strategy documents have been recognized as being of importance even for international governance, as they further complement and reinforce the latter. They thereby provide more room for the technical element in addition to concrete steps for implementation, operationalization and monitoring.<sup>16</sup>

#### 4.2 Other Tactical Governance Frameworks

Existing governance tools within the military domain – such as military doctrines, procurement guidelines, standard operating procedures (SOPs), tactics, techniques and procedures (TTPs), logbooks and after-action reports – may all be applied for the development, deployment and use of AI in the military domain. Whether they are applied as they are or specifically adapted for these technologies, these tools constitute highly specialized and targeted means to strike the balance between, on the one hand, the need to factor in these technologies' inherent technological complexity and, on the other, practicality at the governance level.<sup>17</sup>

<sup>&</sup>lt;sup>16</sup> For more, see Yasmin Afina, *Draft Guidelines for the Development of a National Strategy on Al in Security and Defence* (Geneva: United Nations Institute for Disarmament Research, October 2024), https://unidir.org/wp-

content/uploads/2024/10/guidelines\_for\_the\_development\_of\_national\_strategy\_web-2.pdf.

<sup>&</sup>lt;sup>17</sup> For more information, see Giacomo Persi Paoli and Yasmin Afina, "The Tactical Governance of Artificial Intelligence in the Military Domain," *GC-REAIM Policy Notes* (2025, forthcoming), copy on file with the authors.

# 5. Lessons from Select Initiatives in the Civilian Space

The international governance of civilian applications of AI is often characterized as more advanced than the governance of military applications. This results from a blend of factors that include the sensitivity of the military domain, the (lack of) political will, tensions and distrust, as well as limited resources and capacity. While it is clear that AI in the military domain will bring its own set of considerations due to its unique nature, the following three initiatives are worth further consideration in terms of how a similar approach adapted to the military domain may be explored.

# 5.1 Standards: The International Telecommunication Union and the Institute of Electrical and Electronics Engineers

The International Telecommunication Union (ITU) has an inherently multidisciplinary composition, with both policy experts and technical staff. Leveraging this, the ITU is conducting work on development of AI standards building on its work as part of the AI for Good programme. The standards specifically seek to enable the utilization of AI for "orchestrating 5G and future networks, multimedia innovation, assessing and improving the quality of digital services, and improving energy efficiency".<sup>18</sup> The ITU has established focus groups for each of these specific applications as part of its established processes within the study groups of its dedicated Telecommunication Standardization Sector (ITU-T).

Standards are often raised as a means to complement the suite of tools for the governance of AI, including in the military domain. While the ITU's standards do not touch upon military applications per se, efforts have been proposed to socialize standards as a means to complement governance initiatives in the military domain with technical solutions that are implementable and practical. In particular, Brazil has been a consistent champion of the first global ontological standard of the Institute of Electrical and Electronics Engineers (IEEE) for ethically driven robotics and automation (IEEE STD 7007-2021). This has been demonstrated by Brazil's tabling of a dedicated working paper to the CCW GGE on LAWS.<sup>19</sup>

<sup>&</sup>lt;sup>18</sup> International Telecommunication Union. *Artificial Intelligence*. Geneva: ITU, n.d. https://www.itu.int/en/action/ai/Pages/default.aspx.

<sup>&</sup>lt;sup>19</sup> CCW Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems, "The Possible Uses of IEEE's First Global Ontological Standard for Ethically Driven

#### 5.2 Oversight and Implementation Measures of Regulatory Frameworks: The European Union's AI Act and the Council of Europe's Framework Convention on Artificial Intelligence

The European Union's Al Act, which entered into force in mid-2024 for member states of the European Union, mandates those states to establish or designate, as part of their national implementation plans, three types of national authority:

- 1. A market surveillance authority, notably entrusted with ensuring the surveillance of the compliance of products (Article 3(26), EU AI Act)
- 2. A notifying authority, which will be responsible for "setting up and carrying out the necessary procedures for the assessment, designation and notification of conformity assessment bodies and for their monitoring" (Article 3(19), EU AI Act)
- 3. A national public authority or body to supervise or enforce the respect of fundamental rights in relation to the use of high-risk AI systems (Article 77(2), EU AI Act)

The establishment or the designation of specific national authorities, each with precise and distinct mandates ensures the practical and effective implementation of policy, factoring in the technical element of Al governance.

Similarly, the Council of Europe's Framework Convention on Artificial Intelligence contains provisions on effective oversight mechanisms. Specifically, Article 26 of the Convention mandates each party to "establish or designate one or more effective mechanisms to oversee compliance with the obligations in this Convention". Article 15 also mandates state parties to ensure "effective procedural guarantees, safeguards and rights" where an AI system "significantly impacts upon the enjoyment of human rights", in addition to the adoption or maintenance of "measures for the identification, assessment, prevention and mitigation of risks posed by artificial intelligence systems" through its Article 16.

What oversight and the need for operationalization mean for the governance of Al in the military domain is a subject to much discussion. While difficulties may arise at the international level, at the national level it could be useful, indeed necessary, to establish or designate a dedicated body or entity responsible for coordinating oversight and implementation of both international principles and national strategy priorities (e.g., as outlined in strategy documents). Challenges may arise not only due to the sensitive nature of such applications, but also due to capacity, resource availability, and the multi-agency and multidisciplinary nature of the development, deployment and use of these technologies.

Robotics and Automation Systems as a Building Block for Confidence Building Measures Regarding the Responsible Development and Use of Autonomous Weapons Systems," Working Paper Submitted by Brazil, CCW/GGE.1/2024/WP.1, 4 March 2024, https://docs.un.org/CCW/GGE.1/2024/WP.1.

# 5.3 Technical Cooperation: Proposals for a "CERN for Al" and Scientific Panels

In the civilian domain, calls for an international AI research laboratory to play a role akin to that played in the field of nuclear physics by the European Organization for Nuclear Research (CERN) are increasingly gaining traction. The original CERN was motivated by a desire to harness the "good" of nuclear physics in the post-Second World War era. Similarly, a CERN for AI would enable states to pave the way for international cooperation for the collective advancement of scientific discoveries and solutions, including to foster AI safety.<sup>20</sup>

Considering the sensitivities of AI in the military domain, it is likely that such an approach would have little influence on military applications of AI. There may, however, be room to explore collaborative opportunities for international cooperation on the development, deployment and use of AI to foster international peace and security, such as AI applications to support humanitarian responses to natural disaster or to enhance civilian protection – two applications that would benefit the wider international cooperation, such an approach would be grounded in technical expertise and scientific evidence, thus ensuring the practicality and effectiveness of its products and outputs. The scientifically grounded nature of the CERN for AI proposal echoes a recommendation made by the United Nations Secretary-General's High-Level Advisory Board on AI to establish an international scientific panel on AI, as well as the commitment in the Global Digital Compact to establish within the United Nations "a multidisciplinary Independent International Scientific Panel on AI" (paragraph 56(a)).<sup>22</sup>

<sup>&</sup>lt;sup>20</sup> Elliot Jones, "A 'CERN for AI' – What Might an International AI Research Organization Address?" in *Artificial Intelligence and the Challenge for Global Governance*, Chatham House, 7 June 2024,

https://www.chathamhouse.org/2024/06/artificial-intelligence-and-challenge-global-governance/02-cern-ai-what-might-international.

<sup>&</sup>lt;sup>21</sup> Yasmin Afina, *Remarks for the United Nations Security Council Arria-Formula Meeting "Harnessing Safe, Inclusive, Trustworthy AI for the Maintenance of International Peace and Security"*, 4 April 2025, https://unidir.org/wp-content/uploads/2025/04/UNIDIR\_Yasmin-Afina\_Briefing\_UNSC\_Arria\_AI\_4-April-2025-3.pdf.

<sup>&</sup>lt;sup>22</sup>High-level Advisory Body on Artificial Intelligence, *Governing AI for Humanity: Final Report* (New York: United Nations, 2024),

https://www.un.org/sites/un2.un.org/files/governing\_ai\_for\_humanity\_final\_report\_en.pdf. For the Global Digital Compact, see United Nations, *Pact for the Future: Global Digital Compact and Declaration on Future Generations*, Annex I (New York: United Nations, 2024), https://www.un.org/sites/un2.un.org/files/sotf-pact\_for\_the\_future\_adopted.pdf.

# 6. Food-for-Thought for the Way Ahead

It is clear that, in order for AI governance pathways to plot the right course between technical depth and practicality, they must be multilayered and multidisciplinary by nature. The governance of AI in the military domain is, ultimately, a complex patchwork of international, regional and national legal frameworks and non-binding tools, requiring the participation and contribution of policymakers, lawyers, ethicists, technologists, states, industries, academic experts, research laboratories and civil society organizations. Their engagement must be meaningful and carefully designed to ensure that governance approaches, principles and eventual solutions are indeed implementable, practicable and accessible.

As the international community continues to grapple with this issue, in the light of the points raised above, readers are invited to reflect on the following questions and considerations:

- What governance is needed and at what level? Since balancing technical depth and practicality requires a multilayered approach, what is desirable and achievable will differ at each level. Further reflection is required on what is concretely needed internationally, regionally and nationally, and how to ensure their complementarity.
- What capacity needs to be built or enhanced? Capacity-building is important in order to ensure that governance solutions are implemented effectively and in a practical manner. Acknowledging this, there is a clear need to invest in the required technological, human and financial resources. What such efforts mean in practice across sectors, and how to ensure that they are meaningful, sustainable over time and effectively cascade in the context of Al in the military domain remain yet to be clarified.
- How should dual-use technologies be defined, governed and addressed? Much of the contemporary discussions, including the formulation of governance solutions and pathways, distinguishes the civilian domain from military applications. However, it may be argued that most contemporary technologies, if they are not dual-use by nature, are dual-use in practice. Beyond a technology itself, its components (notably data) can also be dual-use. There may thus be a need to reflect on, and potentially redefine, existing assumptions, understandings and approaches to the very concept of "dual-use" and to re-calibrate the parameters around which these technologies are governed. Ensuring their responsible development, procurement, use and eventual decommissioning and disposal will require careful considerations that not only transcend disciplines and sectors, but may also invite a re-questioning and re-evaluation of existing assumptions that separate the civilian and military domains.

# **About the Authors**

#### Dr. Yasmin Afina

Yasmin Afina is a Researcher for the Security and Technology Programme at UNIDIR, where her research covers the intersection between international security policy, international law and artificial intelligence. Yasmin recently obtained her PhD in law from the University of Essex, where her research focused on the international humanitarian law considerations for the development of AI-enabled technologies for military targeting. Yasmin also holds an LL.M. from the Geneva Academy of International Humanitarian Law and Human Rights, an LL.B. from the University of Essex, as well as a French Bachelor of Laws and Postgraduate degree (Maîtrise) in International Law from the Université Toulouse I Capitole.

#### Dr. Giacomo Persi Paoli

Giacomo is the Head of the Security and Technology Programme at the United Nations Institute for Disarmament Research (UNIDIR). His expertise spans the science and technology domain with emphasis on the implications of emerging technologies for security and defence. His recent work focused on artificial intelligence and cyber security. Before joining UNIDIR, Giacomo was Associate Director at RAND Europe where he led the defence and security science, technology and innovation portfolio as well as RAND's Centre for Futures and Foresight Studies. He also served for 14 years as warfare officer in the Italian Navy. He holds a PhD in Economics from the University of Rome, a master's degree in Political Science from the University of Pisa, and a master's degree in International Affairs and Diplomacy from the University of Trieste.

HCSS Lange Voorhout 1 2514 EA The Hague

Follow us on social media: @hcssnl

The Hague Centre for Strategic Studies