

The authors' views are their own and do not represent the official position of the Institute of Defence and Strategic Studies of the S. Rajaratnam School of International Studies, NTU. These commentaries may be reproduced with prior permission from RSIS and due recognition to the authors and RSIS. Please email to Editor IDSS Paper at <u>RSISPublications@ntu.edu.sg</u>.

No. 006/2025 dated 16 January 2025

The REAIM "Blueprint for Action" Needs Skin in the Game

Manoj Harjani

SYNOPSIS

More than 60 countries endorsed a "Blueprint for Action" at the second Responsible AI in the Military Domain (REAIM) Summit held in September 2024. While this document represents a step forward from the inaugural REAIM Summit's "Call to Action", it will not further the cause of military AI governance without countries committing resources to implement its recommendations.

COMMENTARY

First held in <u>2023</u> at The Hague, the Responsible AI in the Military Domain (REAIM) Summit was jointly initiated by the Netherlands and South Korea to broaden multilateral dialogue on military AI governance beyond the narrower discussions on lethal autonomous weapons ongoing since 2016 under the framework of the Convention on Certain Conventional Weapons.

A major outcome of the inaugural REAIM Summit was a "<u>Call to Action</u>" endorsed by more than 50 countries. While an important first step towards developing norms regarding military AI governance, the Call to Action was framed very broadly and focused on encouraging further dialogue and forming an inclusive community of stakeholders.

Last year, a <u>second REAIM Summit</u> was held in September in Seoul, where more than 60 countries endorsed a "<u>Blueprint for Action</u>" (BFA) that signalled a step up from the Call to Action. The BFA is organised around three issue areas — the impact of AI on international peace and security, implementing responsible AI in the military domain, and envisaging future governance of AI in the military domain.

Three points addressed in the BFA stand out for their potential importance to further develop norms regarding military AI governance: (1) human control should be

maintained over the use of nuclear weapons; (2) legal review procedures under international law are an important tool for military AI governance; and (3) data governance is an important element in military AI governance.

What was unfortunately missing in the BFA was guidance on implementation and on how the required governance capacity might be built, given the wide range of stakeholders involved. Without resource commitments, the BFA cannot advance military Al governance.

As the dates and host country for a third REAIM Summit this year have yet to be announced, there is considerable uncertainty regarding the future of the REAIM process. However, the Netherlands and South Korea led a <u>resolution</u> incorporating key ideas from the REAIM summit that was passed by the First Committee of the UN General Assembly (UNGA) in November 2024. This resolution calls for a report to be submitted by the UN secretary general to the 80th UNGA session later this year, effectively embedding the REAIM process within a larger multilateral forum.

All of this points to potential alternative pathways beyond organising more REAIM summits. Nevertheless, it does not take away from the fact that countries still need to commit resources to implement military AI governance. This need is made more pressing by the rapid adoption of AI by militaries across the globe, seen most visibly in the ongoing conflicts in Gaza and Ukraine.



The Responsible AI in the Military Domain (REAIM) summits held annually since 2023 have produced two documents that are a starting point for developing military AI governance norms. *Image by the Dutch Ministry of Foreign Affairs via <u>Flickr</u>.*

Building Norms

The BFA represents an important step forward in building norms for military Al governance. Though not legally binding, it signals agreement on many key issues. Given that a legally binding agreement may not materialise in the short term, if at all, managing the impact of military Al on international peace and security may depend on the ability of processes like the REAIM summits to entrench norms that deter adverse behaviours.

Three points within the BFA stand out as potentially crucial areas where norms can be developed further. First, by highlighting that human control should be maintained over the use of nuclear weapons, the BFA reflects a growing consensus on the risks to global strategic stability posed by AI's intersection with nuclear weapons.

Such is the level of concern that even the leaders of China and the United States <u>agreed</u> on the sidelines of the 2024 Asia-Pacific Economic Cooperation (APEC) Summit in Lima that human beings and not AI should make decisions on the use of nuclear weapons. Although no formal agreement was signed to this effect, their public statement still carries weight.

Second, the BFA encourages the development of legal review procedures for military AI governance under international law. Under <u>Article 36</u> of the Additional Protocol I to the Geneva Conventions — the foundation for international humanitarian law — countries have an obligation to determine whether the "study, development, acquisition or adoption of a new weapon, means or method of warfare … would, in some or all circumstances, be prohibited" under international law.

However, only a few countries have the capacity and capabilities to conduct legal reviews of new weapons, and Additional Protocol I does not define how countries should determine the legality of new weapons. Another challenge is that military AI is not a countable, physical weapon like a missile. As a general-purpose technology, the use of AI in the military domain will be significantly harder to pin down for assessment in a legal review under Article 36.

Finally, the BFA also emphasises the importance of data governance in the overall governance of military AI. Although this point may seem obvious given AI's reliance on data, governance processes for data are not always integrated with those for AI. Furthermore, given the sensitivity of military activity, it is not a straightforward process to evaluate military data sets, and there is no incentive for countries to open them to external scrutiny.

Making Implementation a Reality

Although the BFA has identified important areas where military AI governance can be advanced, it does not contain any guidance on implementation. It also lacks an accompanying institutional structure that is essential for countries to commit resources to realise its implementation.

One possible modality for implementation is setting up working groups focusing on specific issue areas, each led by a co-host of the REAIM process. Since there are now five co-hosts — Kenya, the Netherlands, Singapore, South Korea, and the United Kingdom — the burden of coordinating and running working groups can be shared.

To support the working groups, the mandate of the existing <u>Global Commission on</u> <u>Responsible Artificial Intelligence in the Military Domain</u> (GC REAIM) could be expanded through financial contributions by countries that have endorsed the BFA. The GC REAIM was set up by the Netherlands after the inaugural REAIM Summit to support the development of norms, so it is ideally placed to support implementation of the BFA. A further opportunity to explore would be whether to institutionalise the regional consultations organised in the run-up to the second REAIM Summit. These workshops were held during the first half of 2024 in Singapore, Istanbul, Nairobi, and Santiago, while a virtual event was organised for Europe and North America. Perhaps the most important feature of the regional consultations was their inclusiveness: smaller and less economically developed countries were given a voice and an opportunity to reflect on their positions regarding military AI governance. Crucially, they also emphasised the importance of capacity building and of regional coordination as a stepping stone to global consensus.

Involving regional organisations in the implementation of the BFA may seem a messy prospect, particularly since each entity functions differently and faces different challenges. However, for the BFA to have any chance of being implemented, different regions and individual countries will need flexibility to decide what is important and practicable. Funding the BFA's implementation at a regional level may also be more palatable for many countries compared to contributing to a global institutional structure that they have less influence over.

Looking Ahead

There is no shortage of challenges ahead for military AI governance. Whatever pathway the REAIM process takes, the BFA will be an important element. It will not be easy to convince countries to have skin in the game, and this is where the REAIM cohosts must demonstrate leadership. Furthermore, even if the REAIM process gets subsumed within the UNGA's First Committee, there is still value in regional organisations implementing the BFA according to their unique circumstances and available resources.

In the case of Southeast Asia and ASEAN, Singapore is well positioned to lead, given its active involvement in the REAIM process as a co-host and convenor of the Asia regional consultation workshop in 2024. The first challenge will be to encourage more countries in the region to endorse the BFA. Other than Singapore, only two countries have done so — Brunei and the Philippines.

The other challenge will be building capacity. The region's wide range of economic development and military capabilities, as well as its complex geopolitical dynamics involving the superpowers, means that this will not be a straightforward task. Furthermore, ASEAN's established modality for regional cooperation and coordination — particularly its preference for non-interference in the internal affairs of member countries — will pose difficulties in implementing initiatives such as legal reviews.

Manoj Harjani is Research Fellow and Coordinator of the Military Transformations Programme at the S. Rajaratnam School of International Studies.

S. Rajaratnam School of International Studies, NTU Singapore Block S4, Level B3, 50 Nanyang Avenue, Singapore 639798