REAIM

Responsible AI in the Military domain **Summit**

September 9-10, 2024
Seoul, Republic of Korea

Co-hosted by The Netherlands | Singapore | Kenya | The United Kingdom

REAIM Blueprint for Action

Artificial Intelligence (AI), as an enabling technology, holds extraordinary potential to transform every aspect of military affairs, including military operations, command and control, intelligence, surveillance and reconnaissance (ISR) activities, training, information management and logistical support.

With the rapid advancement and progress in AI, there is a growing interest by states to leverage AI technology in the military domain. At the same time, AI applications in the military domain could be linked to a range of challenges and risks from humanitarian, legal, security, technological, societal or ethical perspectives that need to be identified, assessed and addressed.

To harness the benefits and opportunities of AI while adequately addressing the risks and challenges involved, AI capabilities in the military domain, including systems enabled by AI, should be applied in a responsible manner throughout their entire life cycle and in compliance with applicable international law, in particular, international humanitarian law.

Building on the Call to Action laid out at the REAIM Summit 2023, we invite all stakeholders including states, industry, academia, civil society, regional and international organizations to:

(The impact of AI on international peace and security)

1. Affirm that AI applications in the military domain should be developed, deployed and used in a way that maintains and does not undermine international peace, security and stability;
2. Recognize that AI applications in the military domain may bring benefits such as increased situational awareness and understanding, precision, accuracy and efficiency, which can enhance the implementation of international humanitarian law and assist in efforts to protect civilians as well as civilian objects in armed conflicts; and AI applications in the military domain may increase effectiveness of and support for peacebuilding and peacekeeping activities, and enhance verification and monitoring capabilities for arms control and other compliance regimes;
3. Recognize also that AI applications can present both foreseeable and unforeseeable risks across various facets of the military domain, which may, inter alia, originate from design flaws, unintended consequences, including from data, algorithmic and other biases, potential misuse or malicious use of the technology and the interaction of AI applications with the complex dynamics of global and regional conflicts and stability, including risks of an arms race, miscalculation, escalation and lowering threshold of conflict;
4. Further recognize that possible high impact applications in the military domain that deserve particular policy attention could include AI-enabled weapons, AI-enabled decision-support systems for combat operations, AI in cyber operations, AI in electronic warfare and AI in information operations;
5. Stress the need to prevent AI technologies from being used to contribute to the proliferation of weapons of mass destruction (WMDs) by state and non-state actors including terrorist groups, and emphasize that AI technologies support and do not hinder disarmament, arms control and non-proliferation efforts; and it is especially crucial to maintain human control and involvement for all actions critical to informing and executing sovereign decisions concerning nuclear weapons employment, without prejudice to the ultimate goal of a world free of nuclear weapons;
6. Underscore the importance of establishing robust control and security measures to prevent irresponsible actors from acquiring and misusing potentially harmful AI capabilities in the military domain, including systems enabled by AI, while bearing in mind that these measures should not undermine equitable access to the benefits of AI capabilities in other non-military domains;

(Implementing responsible AI in the military domain)

7. Affirm that AI applications must be developed, deployed and used in accordance with international law, including, as applicable, the UN Charter, international humanitarian law, international human rights law; and, as appropriate, other relevant legal frameworks, including regional instruments;

8. Stress the importance of establishing national strategies, principles, standards and norms, policies and frameworks and legislation as appropriate to ensure responsible AI applications in the military domain;

9. Acknowledge the following, which are not exhaustive, to ensure responsible AI in the military domain;

   (a) AI applications should be ethical and human-centric.

   (b) AI capabilities in the military domain must be applied in accordance with applicable national and international law.

   (c) Humans remain responsible and accountable for their use and effects of AI applications in the military domain, and responsibility and accountability can never be transferred to machines.

   (d) The reliability and trustworthiness of AI applications need to be ensured by establishing appropriate safeguards to reduce the risks of malfunctions or unintended consequences, including from data, algorithmic and other biases.

   (e) Appropriate human involvement needs to be maintained in the development, deployment and use of AI in the military domain, including appropriate measures that relate to human judgement and control over the use of force.

   (f) Relevant personnel should be able to adequately understand, explain, trace and trust the outputs produced by AI capabilities in the military domain, including systems enabled by AI. Efforts to improve the explainability and traceability of AI in the military domain need to continue.

10. Commit to engaging in further discussions and to promoting dialogue on developing measures to ensure responsible AI in the military domain at the national, regional and international level, including through international normative frameworks, rigorous testing and evaluation (T&E) protocols, comprehensive verification, validation and accreditation (VV&A) processes, robust national oversight mechanisms, continuous monitoring processes, comprehensive training programs, exercises, enhanced cyber security and clear accountability frameworks;

11. Encourage the development of effective legal review procedures, trust and confidence building measures and appropriate risk reduction measures, as well as the exchange of information and consultations on good practices and lessons learned among states; and invite other stakeholders, including industry, academia, civil society and regional and international organizations to actively engage in these efforts, as appropriate, including through regular multi-stakeholder exchanges, dissemination of case studies and other relevant documentation and active participation in collaborative initiatives;

12. Stress that efforts on responsible AI in the military domain can be taken in parallel and do not hamper the efforts on research, development, experimentation and innovation with AI technology;

(Envisaging future governance of AI in the military domain)

13. Recognize that the discussion on the governance of AI in the military domain should include fostering a common understanding of AI technology and its capabilities and limitations, and a shared understanding on the possible applications of AI in the military domain and their potential impacts, including both benefits and risks;

14. Emphasize that such a discussion should take place in an open and inclusive manner to fully reflect wide-ranging views, bearing in mind that different states and regions are at varying stages of integrating AI capabilities in the military domain, come from different security environments and have varying security concerns;

15. Stress the importance of capacity-building, especially in developing countries, to promote full participation of those countries in the discussions on the governance of AI in the military domain, and to facilitate the responsible approach in the development, deployment and use of military AI capabilities;

16. Commit to strengthening international cooperation on capacity-building aimed at reducing the knowledge gap on responsible development, deployment and use of AI in the military domain;

17. Note that data plays a crucial part in AI applications in the military domain, and acknowledge that states and other relevant stakeholders need to engage in further discussions on adequate data governance mechanisms, including clear policies and procedures for data collection, storage, processing, exchange and deletion as well as data protection;

18. Recognize the need for a flexible, balanced, and realistic approach to the governance of AI in the military domain to keep pace with the rapid development and advancement of technologies;

19. Acknowledge developments across multiple initiatives related to the AI applications in the military domain, including the REAIM Summit with its relevant regional events and the establishment of the REAIM Global Commission, the Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy, as well as the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems (LAWS GGE) established under the Convention on Certain Conventional Weapons (CCW), and the discussions in the UN Disarmament Commission and the Conference on Disarmament; take note also of the UN General Assembly Resolution 78/241 on Lethal autonomous weapons systems and relevant regional and international conferences; and stress that these initiatives should be synergistic and complementary, without prejudice to ongoing discussions on related subjects in other fora;

20. Commit to continuing global and regional dialogue on responsible AI in the military domain in an open and inclusive manner with active involvement from and exchange among stakeholders, as appropriate, acknowledging that efforts on responsible AI in the military domain is a task of generations requiring meaningful engagement with the youth.

We invite states to join this Blueprint for Action and also welcome other relevant stakeholders including industry, academia, civil society, regional and international organizations to support and associate with the Blueprint for Action as we continue our efforts to establish responsible AI for the future of humanity.