Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy

An increasing number of States are developing military AI capabilities, which may include using AI to enable autonomous functions and systems. Military use of AI can and should be ethical, responsible, and enhance international security. Military use of AI must be in compliance with applicable international law. In particular, use of AI in armed conflict must be in accord with States' obligations under international humanitarian law, including its fundamental principles. Military use of AI capabilities needs to be accountable, including through such use during military operations within a responsible human chain of command and control. A principled approach to the military use of AI should include careful consideration of risks and benefits, and it should also minimize unintended bias and accidents. States should take appropriate measures to ensure the responsible development, deployment, and use of their military AI capabilities, including those enabling autonomous functions and systems. These measures should be implemented at relevant stages throughout the life cycle of military AI capabilities.

The endorsing States believe that the following measures should be implemented in the development, deployment, or use of military AI capabilities, including those enabling autonomous functions and systems:

A. States should ensure their military organizations adopt and implement these principles for the responsible development, deployment, and use of AI capabilities.

B. States should take appropriate steps, such as legal reviews, to ensure that their military AI capabilities will be used consistent with their respective obligations under international law, in particular international humanitarian law. States should also consider how to use military AI capabilities to enhance their implementation of international humanitarian law and to improve the protection of civilians and civilian objects in armed conflict.

C. States should ensure that senior officials effectively and appropriately oversee the development and deployment of military AI capabilities with high-consequence applications, including, but not limited to, such weapon systems.

D. States should take proactive steps to minimize unintended bias in military AI capabilities

E. States should ensure that relevant personnel exercise appropriate care in the development, deployment, and use of military AI capabilities, including weapon systems incorporating such capabilities.

F. States should ensure that military AI capabilities are developed with methodologies, data sources, design procedures, and documentation that are transparent to and auditable by their relevant defense personnel.

G. States should ensure that personnel who use or approve the use of military AI capabilities are trained so they sufficiently understand the capabilities and limitations of those systems in order to make appropriate context-informed judgments on the use of those systems and to mitigate the risk of automation bias.

H. States should ensure that military AI capabilities have explicit, well-defined uses and that they are designed and engineered to fulfill those intended functions.

I. States should ensure that the safety, security, and effectiveness of military AI capabilities are subject to appropriate and rigorous testing and assurance within their well-defined uses and across their entire life-cycles. For self-learning or continuously updating military AI capabilities, States should ensure that critical safety features have not been degraded, through processes such as monitoring.

J. States should implement appropriate safeguards to mitigate risks of failures in military AI capabilities, such as the ability to detect and avoid unintended consequences and the ability to respond, for example by disengaging or deactivating deployed systems, when such systems demonstrate unintended behavior.

In order to further the objectives of this Declaration, the endorsing States will:

- implement these measures when developing, deploying, or using military AI capabilities, including those enabling autonomous functions and systems;
- make public their commitment to this Declaration and release appropriate information regarding their implementation of these measures:
- support other appropriate efforts to ensure that military AI capabilities are used responsibly and lawfully;
- pursue continued discussions among the endorsing States on how military AI capabilities are developed, deployed, and used responsibly and lawfully;
- promote the effective implementation of these measures and refine these measures or establish additional measures that the endorsing States find appropriate; and
- further engage the rest of the international community to promote these measures, including in other fora on related subjects, and without prejudice to ongoing discussions on related subjects in other fora.

The endorsing States recognize that concepts of artificial intelligence and autonomy are subject to a range of interpretations. For the purpose of this Declaration, artificial intelligence may be understood to refer to the ability of machines to perform tasks that would otherwise require human intelligence. This could include recognizing patterns, learning from experience, drawing conclusions, making predictions, or generating recommendations. An AI application could guide or change the behavior of an autonomous physical system or perform tasks that remain purely in the digital realm. Autonomy may be understood as a spectrum and to involve a system operating without further human intervention after activation.