



Permanent Mission of the Republic of Armenia to the United Nations MISSION PERMANENTE DE LA FRANCE AUPRÈS DES NATIONS UNIES À NEW YORK Liberté Egalité Fraternité



Permanent Mission of Italy UN - New York





Concept Note

Arria Formula Meeting of the Security Council organized by

Greece, France and the Republic of Korea

Cosponsored by Armenia, Italy, Kingdom of the Netherlands

On

Harnessing safe, inclusive, trustworthy AI for the maintenance of international peace and security

Friday 4 April 2025, 15.00 – 18.00 EST

UN Headquarters, Conference Room 2

Introduction

The field of Artificial Intelligence (AI) has recently seen breakthrough advances, enabling computational technology to engage with many kinds of information at a capacity that would have seemed unlikely a couple of decades earlier. This progress suggests vast opportunities for transforming all aspects of human activity. At the same time, AI poses significant risks and raises questions for the maintenance of international peace and security.

For instance, AI creates significant challenges and opportunities for peacekeeping operations and special political missions mandated by the Security Council. AI-related technologies could be harnessed for peacekeepers' training, logistic support, landmines detection, mapping and clearance, surveillance or monitoring tasks, but this has yet to materialize in practice. Reversely, these technologies can also be misused to carry our malicious cyberattacks, to disseminate hate speech, or to undermine populations' trust in peacekeepers through disinformation campaigns.

Initiatives supporting open, inclusive and trustworthy AI

Artificial Intelligence can play a crucial role in promoting progress and development in our societies. There is a growing international consenus on the need to pursue an inclusive, human-centered, digital transformation that underpins economic growth and sustainable development, maximizes benefits, and manages risks, in line with our shared democratic values and respect for human rights.

Greece has established a High-Level Advisory Committee on Artificial Intelligence (AI) to outline a comprehensive national AI strategy. The Committee's landmark study, <u>Blueprint GREECES AI TRANSFORMATION.pdf</u>, provides guiding principles and flagship projects to drive Greece's AI advancements, with priorities, among others, safeguarding and enhancing democracy, climate mitigation and adaptation, supporting security.

There have been already some landmark international initiatives on AI in the context of peace and security, especially on the military domain. Two REAIM Summits, held in The Hague, in 2023, and in Seoul, in 2024, resulted in the endorsement of the "Call to Action" and the "Blueprint for Action," respectively, calling for and guiding the international community's efforts towards the responsible use of AI in the military domain. The Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy, launched in February 2023, as well as the ongoing discussions of the GGE on LAWS (Group of Governmental Experts on Lethal Autonomous Weapons Systems) within the framework of the Convention on Certain Conventional Weapons/CCW , have contributed to the building of an international consensus around responsible state behavior in the development, deployment, and use of AI applications in the defence sector.

Additionally, the first AI Safety Summit in November 2023, which adopted the Bletchley Declaration, and the AI Seoul Summit in May 2024, which adopted the Seoul Declaration for Safe, Innovative and Inclusive AI and its Annex: the Seoul Statement of Intent toward International Cooperation on AI Safety Science, have paved the way for international cooperation in advancing AI safety.

In February 2025, participants from over 100 countries gathered in Paris to hold the AI Action Summit, where they adopted a joint statement in Inclusive and Sustainable Artificial Intelligence

for People and the Planet, committing to advance Trust and Safety, progress international cooperation on AI safety, address the risks of AI to information integrity and continue the work on AI transparency. During this summit, 27 States endorsed the « Paris Declaration on Maintaining Human Control in AI enabled Weapon Systems », by which they committed to develop, deploy and use AI capabilities in the military domain in a responsible manner, in accordance with international law, to engage in multilateral dialogue and to encourage international cooperation in this field.

At the UN level, the challenges and opportunities related to AI in the context of international peace and security have been discussed twice in formal Security Council meetings. Firstly, in July 2023 the Council discussed the implications of Artificial Intelligence for international peace and security, especially the risks posed for UN peacekeepers and personnel, but also the potential for AI to be used to mitigate risks of civilian casualties in conflict. Recently, the Security Council convened in December 2024 to follow up with subsequent initiatives of Member States aimed at addressing the implications of AI for international peace and security, with a focus to the preservation of the value of human life and dignity, as well as to the application of AI in the military domain.

Furthermore, in December 2023 the Security Council convened in Arria Formula to discuss on Artificial Intelligence (AI) and its impact on hate speech, disinformation and misinformation in the context of the maintenance of international peace and security.

The General Assembly adopted recently Resolution 79/239, with which it affirmed that international law, including the Charter of the United Nations, international humanitarian law and international human rights law, applies to matters governed by it that occur throughout all stages of the life cycle of artificial intelligence, including systems enabled by artificial intelligence, in the military domain. A report of the Secretary General on the matter is expected in 2025. The General Assembly is also in the process of establishing an independent international scientific panel and a global dialogue on the governance of AI, both grounded in the principles and objectives laid out in the Global Digital Compact.

The way forward

With Action 27 of the Pact for the Future the world leaders expressed their willingness to seize the opportunities associated with new and emerging technologies and address the potential risks posed by their misuse. In this vein, they recognized that rapid technological change presents opportunities and risks to our collective efforts to maintain international peace and security. Also, they proclaimed that International law, including the Charter, will guide our approach to addressing these risks. On the specific domain of AI, they decided to continue to assess the existing and potential risks associated with the military applications of artificial intelligence and the possible opportunities throughout their life cycle, in consultation with relevant stakeholders.

Greece, France and the Republic of Korea decided to convene an Arria Formula meeting of the Security Council, cosponsored by Armenia, Italy and the Kingdom of the Netherlands, so as to build up on its previous meetings and examine the AI realm under an additional spectrum: What

could be practical responses of the UN, including the Security Council, to national, subregional, regional and international initiatives on responsible use of Artificial Intelligence, in a way to contribute to maintenance of international peace and security, especially through regulation, non-proliferation, and the prevention of the diversion of AI capabilities in the military domain, enhancement of rule of law, democratic values, social cohesion and economic development?

The meeting aims at offering to the participants the possibility to assess the relevant initiatives, as well the state of play within UN, including the Security Council, on AI in the context of international peace and security. Moreover, it can serve as a platform so as the Member States have the chance to share national experiences and lessons learned on tangible benefits from the AI revolution, that can contribute to international peace and security.

Topics for Discussion

During the discussion, the participants are encouraged to engage on the following guiding questions:

1. How do initiatives supporting open, inclusive and trustworthy AI contribute to the responsible use of this technology, thus supporting the maintenance of international peace and security?

2. Building on these initiatives, and on the recognition that international law applies to AI, how can the UN, including the Security Council within its mandate, leverage the responsible use of AI to advance international peace and security?

3. Which could be the practical responses to challenges posed to information integrity, through malicious cyberattacks, dissemination of hate speech and disinformation campaigns, especially in the context of UN peace operations?

Format

All Member States of the UN as well as permanent observers and non-governmental organizations accredited to the UN are invited to attend.

Following the panelists' interventions, members of the Security Council will be invited to take the floor.

Other delegations wishing to make an intervention are encouraged to inform the organizers in advance as time will be limited.

All delegations are encouraged to speak no more than 3 minutes.

Interpretation will be provided.

Briefers

1) Mr. Giannis Mastrogeorgiou, Special Secretary for Strategic Foresight at Presidency of the Hellenic Government & Coordinator of the National Advisory Committee on Artificial Intelligence of Greece

2) Dr. Yasmin Afina, Al Researcher, Security and Technology Programme, United Nations Institute for Disarmament Research

3) Ms Charlotte Scaddan, Senior Adviser on Information Integrity, UN Department of Global Communications