Conference on Disarmament

Date .....

Original: English

Note Verbale dated 21 July 2023 from the Permanent Mission of the Islamic Republic of Pakistan transmitting the Working Paper entitled "Addressing the Security and Stability Implications of Military Applications of Artificial Intelligence (AI), and Autonomy in Weapon Systems"

The Permanent Mission of Islamic Republic of Pakistan presents its compliments to the Secretary-General of the Conference on Disarmament and has the honour to request that the attached Working Paper entitled "Addressing the Security and Stability Implications of Military Applications of Artificial Intelligence (AI), and Autonomy in Weapon Systems" be published as an official document of the Conference on Disarmament and circulated to its membership.

The Permanent Mission of Islamic Republic of Pakistan avails itself of this opportunity to renew to the Secretary-General of the Conference on Disarmament the assurances of its highest consideration.

Annex

Working Paper by Pakistan for the Conference on Disarmament, addressing the Security and Stability Implications of Military Applications of Artificial Intelligence (AI), and Autonomy in Weapon Systems

Introduction

1. The development, deployment, integration and use of Artificial Intelligence (AI) for military purposes, as well as autonomous weapon systems continues to gather pace.

2. Simultaneously, there are growing concerns over the absence of global normative and legal guardrails to regulate the weaponization of these new technologies that carry far-reaching security and stability ramifications at the regional as well as international levels.

3. For almost a decade, structured discussions have taken place on Lethal Autonomous Weapon Systems (LAWS) primarily from the prism of International Humanitarian Law (IHL) as well as technical and ethical considerations under the Convention on Certain Conventional Weapons (CCW).

4. Discussions at the CCW on security and stability aspects of LAWS have largely been incidental.

5. In 2016, the informal meeting of experts, convened under CCW, recommended that further consideration should be given to the effects and risks of LAWS on: regional and global security and stability; threshold for armed conflicts; arms race; military value; proliferation to and by non-state actors; and cyber operations in relation to LAWS.

6. A Group of Government Experts (GGE) convened under the CCW, in its 2019 report, noted that consideration should be given to the possible effects of such weapons systems "on regional and global security and stability, including thresholds for armed conflict".

7. While GGE deliberations have helped develop a reasonable level of understanding on various dimensions of LAWS, including on their destabilizing impacts, they have yet to yield a legal and normative framework required to address the grave security and stability ramifications linked to the development, deployment, integration, and use of such weapon systems.

8. Discussions in CCW format have mainly focused on IHL, with insufficient and declining attention to security and stability dimensions of LAWS. Moreover, the GGE discussions have not touched specifically on the military application, integration and use of AI.

A. Security and Stability related Implications

9. The development, deployment, integration and use of AI for military purposes, as well as autonomous weapon systems entail serious repercussions for security and stability at the international and regional levels.

10. These capabilities, when developed and employed without any guardrails, can exacerbate nuclear risks. And when these capabilities are integrated with policies relating to nuclear force posture as well as employment, they can lead to miscalculations, accidents and thereby catastrophic consequences.

11. Concepts of deterrence escalation and risk reduction heavily rely on human rationality, caution, perception and management of the situation politically. Perceived or real absence of human factor renders these concepts without their traditional meaning with risk of automating escalation.

12. The use of AI for generating data and information can engender a false sense of confidence for states considering pre-emptive counterforce destabilizing strikes or targeting

second strike capabilities, thereby posing serious risks and dangers for regional and international stability. Generating disinformation using AI can also be used to manipulate critical decision-making processes.

13. Recognizing the concerns regarding intersection of AI capabilities with Weapons of Mass Destruction (WMDs), a few states have undertaken commitment to maintain human control and involvement in all actions relating to nuclear weapons employment. However, the magnitude of concerns and possible risks, which warrant human control and oversight, are not just limited to WMDs.

14. Threats to peace, security and stability at the regional and global levels resulting from the use of unregulated AI in critical military decision making processes and conventional weapon systems are comparable to any category of WMDs.

15. While the autonomous weapon systems may eliminate the danger of human casualties for the user states, they nonetheless lend themselves to an increased propensity of use and enhanced prospects of symmetric and asymmetric responses, thus lowering the threshold for application of force and armed conflict.

16. In this context and in times of crisis, a low threshold for use of force would be highly destabilizing as it would increase risks of miscalculation and trigger conflict escalation. Completely eliminating human control in such critical times could result in unpredictable consequences and rapid escalation, making it difficult to control the magnitude and duration of conflicts. Introduction of non-linear pathways to escalation with a risk of nuclear war and conventional entanglement will adversely impact strategic stability.

17. Increasing the speed of warfare and compressing decision-making timeframe, particularly in regions with high tensions and history of accidental launch incidents, represents a recipe for an unmitigated disaster.

18. Furthermore, the integration of AI with cyber and outer space domains adds additional layers of complexity and can play a force multiplying role.

19. Some of the more specific but non-exhaustive aspects of security and stability arising from the military use of AI and autonomous weapon systems are given below:

(i) Unpredictability

20. The reliability and predictability of AI-based systems are contingent on algorithms and training data, which can be biased or context-specific. In complex and dynamic real-life situations, the decision-making processes of AI-enabled systems and machine learning may produce unpredictable cascading effects.

21. The absence of transparency, common understandings, parameters for predictability, reliability, understandability, and explainability further amplifies the risks to regional and international stability posed by the military AI capabilities and autonomous weapon systems.

(ii) Asymmetrical Warfare

22. Autonomous weapon systems could also entail additional consequences such as anonymous and clandestine operations, including targeted killings in foreign territories. This can lead to reprisals by the targeted state on different fronts, bypassing the protections offered by IHL.

23. Such asymmetrical means not only undermine compliance with existing IHL but also contribute to the erosion of norms, potentially fueling further conflict escalation.

(iii) Challenges for Accountability

24. The reduction or removal of human decision-making in targeting and engagement also raises concerns over state responsibility, accountability, and attribution for the actions taken by autonomous weapon systems. The challenge of accountability also makes it difficult to operationalize the principles of humanity, necessity and proportionality.

25. Ambiguity in assigning responsibility and accountability may arise, potentially shifting it to operators or developers. Establishing effective oversight and redressal

mechanisms therefore are a critical requirement to address suspected, reported, or documented violations related to the development, deployment, or use of such weapon systems.

(iv) Risks of Arms Race and Proliferation

26. AI based weapon systems are already being factored into strategic and security doctrines of states. Absent any meaningful restraints, these developments are bound to trigger costly arms races, as nations may feel compelled to acquire such technology to maintain strategic balance or advantage. These developments would also increase the risks of proliferation of such weapon systems to non-state actors.

(v) Cyber Attacks

27. Autonomous weapon systems and military AI capabilities are also vulnerable to cyber-attacks. Malicious actors can exploit vulnerabilities in these systems, potentially manipulating their intended functions and possible system takeover by a third party or an adversary, which may lead to unintended consequences or misuse.

B. The Case for Multilateral Responses including at the CD

28. The scale of legal, humanitarian, technical and security challenges resulting from the use of AI for military purposes including in weapon systems necessitates a multifaceted multilateral response at various fora.

29. The ongoing work within the ambit of the CCW should continue with an aim to develop international rules through a new Protocol. Such a Protocol would need to clearly spell out prohibitions and regulations governing autonomous weapon systems to ensure compliance with the IHL consistent with the objectives and purposes of the CCW. In this context, attention is invited to the proposal submitted by Pakistan on an "International Legal Instrument on Lethal Autonomous Weapons Systems (LAWS)" as contained in CCW/GGE.1/2023/WP.3/Rev.1.

30. Without prejudice to the discussions in the CCW that pivot on IHL, there are other aspects of military AI capabilities and autonomous weapon systems that need to be addressed.

31. The security dimensions of AI for military purposes and autonomous weapon systems highlight the need for further elaborating their impacts on regional and global stability, understanding threat perceptions, pursuing arms control and restraints in the security doctrines, eschewing measures that propel arms races, and crafting of trust and confidence-building among States.

32. In considering control and regulatory measures, the inherent right of states to access and use AI for peaceful purposes and for achievement of Sustainable Development Goals (SDGs) shall be upheld.

33. The Conference on Disarmament (CD) is well-suited to evolve effective measures in these realms, as it brings together military significant states on equal footing to discuss and negotiate arms control issues, while protecting their vital security interests.

34. The security implications emanating from the use of AI for military purposes and autonomous weapon systems have direct relevance to several agenda items of the CD.

35. The AI for military purposes and autonomous weapon systems are a cross-cutting theme to the areas of nuclear disarmament, prevention of war, prevention of an arms race in outer space, and new types of weapons of mass destruction.

36. Similarly, the discourse on risk reduction needs to factor in the risks posed by the development, deployment, integration and use of military AI capabilities and autonomous weapon systems.

37. For this purpose, following course of action is proposed for consideration by CD member states:

(i) Continued deliberations on the implications of regional and global security as well as stability resulting from the development, deployment, integration and use of AI for military purposes and autonomous weapon systems under existing agenda item 5, titled, "New types of weapons of mass destruction and new systems of such weapons; radiological weapons."

(ii) Introduction of new agenda item in the 2024 session of the CD titled, "`New technologies: the development, deployment, integration and use of AI for military purposes, and autonomous weapon systems"

(iii) Subsequently, establishment of a subsidiary body under this new agenda item to advance substantive work on this subject and to pursue negotiations on arms control measures pertaining to AI for military purposes and autonomous weapon systems to address security and stability related concerns.

38. Failure to address these serious risk and potential threats to peace and security would not only entail dire humanitarian consequences, but also oblige states to defend themselves with the capabilities at their disposal vis-à-vis their perceived adversaries without any guardrails. Continued absence of legal and normative framework will engender the undesirable but likely consequences of a spiraling arms race, weakening of the already stressed arms control architecture, and a more precarious regional and international security environment.

39. Conversely, commencing deliberations in the CD on the security aspects of the development, deployment, integration and use of AI for military purposes and autonomous weapon systems will have a positive synergetic effect on the ongoing discussions in the CCW and elsewhere.

40. Such discussions will also be in tune with the calls made by several leaders including the UN Secretary General, international organizations and civil society to give serious attention to this subject, which entails transformational impacts on international peace and security.

41. These deliberations may also help unlock the decades long impasse in the CD; contribute to revitalization of the work of the CD; and send a clear signal to the international community that CD is responsive to the risks arising from application and use of AI for military purposes and autonomous weapon systems.