

*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 [RSISPublications@ntu.edu.sg](mailto:RSISPublications@ntu.edu.sg).*

No. 085/2023 dated 29 November 2023

## Barriers to New Arms Control Regulation on AI

*Simon Cleobury*

### SYNOPSIS

*Calls for regulation of Artificial Intelligence (AI) in the military domain are increasing. However, there are significant barriers to adopting new arms control agreements in this area. A comprehensive legally binding instrument seems unlikely, and any new multilateral agreement will take time to negotiate. In the shorter term, major AI-using states should reassure us by making unilateral and joint declarations about retaining human responsibility for decision-making. Such negative AI assurances could then be endorsed by a UN Security Council resolution.*

### COMMENTARY

In July 2023, the UN Security Council discussed the potential impact of AI on international peace and security for the first time. Disarmament diplomats have increasingly recognised the need to discuss the impact of emerging technologies like AI on arms control, but there is no consensus on how and where to have such discussions. There is now significant momentum on the need for AI regulation (in the broadest sense of the term), which necessitates examining some of the barriers to reaching agreement on regulating AI in the arms control context and thinking about what is achievable right now.



The road to a comprehensive agreement governing military use of AI could be a long one. One way to navigate potential barriers and pitfalls for arms control regulation on AI could be through unilateral declarations or moratoria, which could be extended to joint statements or political declarations over time. *Image from Pixabay.*

## **What Are the Barriers?**

One of the biggest barriers to any arms control is the fact that the major military powers do not want to be constrained in the kind of weapons systems they can develop and use. While all the major countries have signed up to principles that state that human responsibility will always be retained for decision-making, they may want to keep their options open as to exactly where the line on autonomy should be drawn. While there are many countries that would sign up to a legally binding agreement on AI regulation, if the major users of AI in the military domain stayed away, then its effectiveness would be limited.

Concluding arms control agreements can be a long process. Even agreeing to set up a negotiation process, which requires agreement on an appropriate forum, its mandate, and the rules of procedure, can take a long time – that’s before you even get to the negotiations. Even in a collaborative environment, coming to an agreement would take time, especially when there are extremely tricky technological issues to grapple with.

Related to the two barriers noted above is the issue of timing. When is the right time to attempt such a negotiation? Some argue that it is too soon to regulate AI from an arms control perspective because we do not know where the technology will end. Those in this “too soon” camp prefer not to attempt regulation whilst the scientific advances are ongoing.

Others say that the genie is out of the bottle, that AI is a thing now and that it is in fact too late to regulate it. Those in this “too late” camp would argue that AI is now too difficult to regulate. The middle ground, and probably the largest constituency, would argue that something needs to be done soon. However, for the reasons given above, getting a new process up and running will not be easy.

For an arms control agreement to be successful, the parties need to trust that each other will comply with its provisions. Assuming that there has been sufficient trust to start negotiations on AI regulation (a big if), then the parties will also need to ensure there is a verification mechanism that enables parties to be reassured that others are complying with the agreement. Verifying the use of AI is different from verifying something tangible like a weapon. To be successful, any verification mechanism will not only need to be carefully drafted, but it will also have to rely on a certain level of transparency and collaboration, something that is hard to achieve when there are low levels of trust.

The consensus rule is an important part of bodies such as the Conference on Disarmament. It also applies in processes such as the one on Lethal Autonomous Weapons Systems. Unfortunately, the consensus rule will make it hard to reach any comprehensive agreement on regulating the military use of AI through such structures. For some it may seem attractive to go outside these structures, to the UN General Assembly, for example, but this would likely result in agreements that none of the key states join because of the loss of protection arising from the consensus rule. This is a common dilemma in multilateral arms control.

A major difference with traditional arms control agreements, which are among states, is that here many of the major players are from the private sector and not part of government. These AI industry actors will need to be included in the discussions. Working out the best way for these two entirely different sets of stakeholders to collaborate will not be straightforward.

### **The Pitfalls to Manage**

It seems inevitable that regulation of the military use of AI will be fragmented and piecemeal. Given all the barriers discussed above, it is unlikely that there will be a single, comprehensive agreement. Even if there were one, the other parts of the disarmament architecture, particularly the existing treaties, would need to align with it in some way. The key will be to ensure consistency in regulation across the instruments.

In recent years, the disarmament community has had to contend with parallel processes on cyber and outer space. A two-track process, or even multi-track processes, might well be pursued on AI too. If this happened, such parallel processes would need to complement each other to the greatest extent possible.

### **Options for Overcoming these Barriers**

The road to a comprehensive agreement governing military use of AI could be a long one. Many are calling for an international legally binding agreement to be concluded now to ensure that decisions around use of force are always made by a human. Reassuringly, the major military powers have all said that they would never want such decision-making power to rest with AI. However, none has given such an assurance in a legally binding form. Here I set out some options for what can be done relatively quickly to provide the reassurance that many are seeking around the military use of AI.

One method for states to reassure each other on the military use of AI could be through unilateral declarations or moratoria. In the arms control world, such moratoria have been declared by nuclear weapon states on nuclear testing and on the production of fissile material for use in nuclear weapons. Moratoria are not perfect, they do not constitute legally binding agreements with other states, they are hard to verify, and they can be easily revoked at any time. However, they are useful as a first step and confidence-building measure.

An extension to the unilateral declarations would be joint statements or political declarations, whereby a group or bloc of countries agree a set of principles around the use of AI. For example, the G7 Leaders' Statement on the Hiroshima AI Process and the US Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy.

Declarations are important steps towards norm setting. However, they fall short of being legally binding and there is no accountability mechanism attached to them. One possible way to inject a certain level of legal commitment would be through a Security Council resolution. The resolutions on negative security assurances relating to the use of nuclear weapons can provide some inspiration here. Under resolutions 984 and 1887, the permanent members of the Security Council gave security assurances against the use of nuclear weapons to non-nuclear-weapon states that are parties to the Treaty on the Non-Proliferation of Nuclear Weapons (NPT). What if they did something similar for AI?

They could start by giving national assurances that they would never use AI in the command, control, and communications of their nuclear weapons and such assurances could then be endorsed in a resolution. Obtaining similar assurances from the non-NPT nuclear possessor states would be harder, but the resolution could contain a call for them to make similar declarations. The resolution could then call on all states to make declarations that they would never use AI in the decision-making process for the use of force and always retain human responsibility for such decisions.

While not perfect – because Security Council resolutions can be ignored and a call on a state to do a certain thing is not as strong as signing an agreement that obliges that state to do that thing – it would provide some level of legally binding commitment. It could also be agreed relatively quickly.

Such an approach would not go far enough for those arguing for a legally binding instrument, but for the reasons I have set out in [my previous paper](#) and this one, the prospects for agreeing such a treaty anytime soon are not good. If (and it is a big if) the Security Council was able to agree to this, it would be a building block for future negotiations and a welcome confidence-building measure.

An alternative and/or additional way of achieving the same result would be for the negative AI assurance to appear in a statement by the leaders of the permanent five members of the UN Security Council, rather like the Prevention of Nuclear War statement from January 2022.

## Conclusion

Concluding any legally binding arms control agreement on AI will be challenging. Going down the Security Council route might not seem like a good idea right now. However, the debate on AI held under the United Kingdom's presidency of the Security Council in July has opened the door for a Council agenda item on AI. There is an opportunity for it to play an important role in a cutting-edge international security issue.

**Simon CLEOBURY** is Head, Arms Control and Disarmament, Geneva Centre for Security Policy. He submitted this IDSS Paper in collaboration with the Military Transformations Programme at the Institute for Defence and Strategic Studies (IDSS), RSIS.

---

S. Rajaratnam School of International Studies, NTU Singapore  
Block S4, Level B3, 50 Nanyang Avenue, Singapore 639798  
T: +65 6790 6982 | E: [rsispublications@ntu.edu.sg](mailto:rsispublications@ntu.edu.sg) | W: [www.rsis.edu.sg](http://www.rsis.edu.sg)