Reinventing the wheel? Three lessons that the AWS debate can learn from existing arms control agreements

June 13, 2024, Analysis / Autonomous Weapons / IHL / New Technologies / Special Themes

🕑 12 mins read

Laura Bruun
Researcher, Stockholm
International Peace
Research Institute
(SIPRI)

For more than a decade, states have met at the UN in Geneva to discuss the governance of autonomous weapon systems (AWS). One pandemic, several real-world cases of artificial intelligence (AI) being used in targeting decisions, and numerous meetings later, there is a growing consensus among states that the challenges posed by AWS should be addressed through both prohibitions and restrictions, a so-called 'two-tier' approach. But while there is progress on the basic structure (i.e. two tiers), the actual content of these tiers is debated.

To help states elaborate on possible elements of a two-tiered approach to the governance of AWS, Laura Bruun from the Stockholm International Peace Research Institute (SIPRI) points to three lessons from past arms control negotiations that can be applied to the AWS debate: First, a prohibition does not need to be grounded in a clearly defined class of weapons, second, restrictions can be used to clarify what international humanitarian law (IHL) requires in the specific context of AWS, and third, if there is will (and a need), two-tiered instruments can be grounded in concerns beyond IHL.

Advances of autonomy in weapons systems – which in large part are enabled by the progress of artificial intelligence – are forcing governments to think deeply about how humans and technology not only can, but should, interact in targeting decisions. For more than ten years, *states have met in Geneva* to discuss the big question of how to govern AWS: *weapon systems that, once activated, can select and apply force without further human intervention.*

While states continue to have radically different views on how to address the challenges posed by AWS, there is now a *growing consensus* among states that AWS should, at least be governed through 'two tiers'. A 'two-tiered approach' is intended to, on the one hand, prohibit certain types and uses of AWS, and, on the other hand, place limits and requirements on the development and use of all other AWS.

Now, a key task is to agree on the actual content of these two tiers. What, if any, technical characteristics would make the design and use of AWS off-limits? What restrictions would be necessary to ensure that the use of AWS is permissible and to ensure that human accountability is retained?

While agreeing on what the two tiers should entail is indeed a complex endeavour, it is not the first time states have dealt with such a task. Many existing arms control agreements follow a two-tiered structure – the question is what lessons we can apply to the AWS debate.

AWS may be new to the arms control community, but a two-tiered approach is not

Throughout the history of arms control, diplomats have resorted to two-tiered approaches to address humanitarian and security challenges posed by specific means and methods of warfare. Examples include the *Biological Weapons Convention*, *the Chemical Weapons Convention*, three of the five protocols under the Convention on Certain Conventional Weapons (*Amended Protocol II on Mines, Booby Traps and Other Devices*, *Protocol III on Incendiary Weapons*, and *Protocol IV on Blinding Laser Weapons*), the *Convention on Cluster Munitions* and the *Anti-Personnel Mine Ban Treaty*.

Though all these instruments differ in legal status, content and purpose, they all constitute relatively recent examples of two-tiered arms control agreements, encompassing a mix of outright prohibitions and restrictions (the latter usually consisting of both specific limits on use and/or positive requirements).

To advance already complex discussions on AWS, it is helpful for states to take a step back and situate the AWS debate within the history of arms control; specifically, how existing two-tiered instruments were set up provides at least three useful lessons for the policy process on AWS.

Three things the AWS debate can learn from past approaches to two-tiered regulations

Lesson 1: A prohibition does not need to be grounded in a clearly defined class of weapons.

A central issue in the debate is the issue of how to define AWS, and whether agreeing on a definition is a prerequisite for developing an instrument. Some states, *such as India*, argue that this is necessary because otherwise, states cannot know what they are regulating. *Others*, such as *Pakistan* and *Palestine*, argue that it should not, notably because AWS are considered a *capability* that can be added to existing weapons rather than a distinct category of weapons, and also because AWS represent a developing set of technologies.

To advance the discussion about whether AWS should be treated as a class of clearly defined weapons, it is helpful to consider *what* has been made the subject of two-tiered regulations in the past. Here we see that states have established prohibitions around the following four categories:

1. **Physical descriptions of the weapon**, e.g. chemical and biological weapons, blinding laser weapons, non-detectable, non-self-destructive or non-deactivating mines, cluster munitions and anti-personnel landmines;
2. **Certain functions**, e.g. weapons whose detonation is triggered 'by the presence, proximity or contact of a person or vehicle' or the production of 'a chemical reaction of a substance delivered on the target';
3. **Certain effects**, e.g. weapons designed to 'ignite objects or burning persons' or weapons causing permanent blindness to unenhanced vision; and
4. **Certain uses**, e.g. attacks by air-delivered incendiary weapons within a concentration of civilians or the deployment of anti-personnel landmines directed against the civilian population.

This suggests that prohibitions can indeed be flexible and tailored to address different types of concerns, beyond those posed by specific classes of weapons.

Coming back to the AWS debate, the question that needs to be settled is what mix of these would be appropriate. This requires deeper discussions about what concerns states when it comes to AWS: is it certain technical features, specific use cases, certain effects, or something else?

Considering the concerns expressed in the policy debate, it seems particularly relevant to address concerns related to autonomous *functions* in targeting decisions. Namely, concerns around whether certain autonomous functions undermine the ability of humans to perform their obligations under IHL and the ability to investigate and attribute responsibility in case of an unlawful act.

This suggests that the type of category that is relevant to explore further in the context of AWS relates to certain functions. States could, for example, explore the viability of formulating prohibitions on types of autonomous functions that create impermissible configurations of human-machine interactions and do not allow users to exercise the necessary control over targeting decisions.

Lesson 2: Restrictions can be used to clarify what IHL requires in the specific context of AWS.

It is *undisputed* among states that AWS must be developed and used to comply with IHL. However, how IHL applies in the context of AWS is unsettled. A lot of this debate stems from the open-textured nature of many IHL obligations, leaving room for different interpretations. *The obligation to take feasible precautions* in attack (to spare the civilian population, civilians and civilian objects, etc.) is a good example of this debate. While compliance with this rule is *considered particularly critical* in the context of AWS, it is unclear what 'feasible precautions' require from users of AWS and what limits it places on where, against what and how AWS can be used.

It is, however, far from the first time states have debated how IHL applies in the context of specific means or methods of warfare – this is where restrictions (usually forming part of a second tier) in existing instruments provide some useful lessons. In the past, states have used restrictions as an opportunity to provide more specific guidance in terms of how existing obligations under IHL apply in the specific context of the weapon. Existing instruments, for example, place limits on the geographical scope of use ('clearly marked areas' in Amended Protocol II) or through certain means ('air-delivered' in Protocol III). Other examples include specifying what IHL obligations require from users in relation to a certain context (obligations on 'marking, fencing and clearance' in Amended Protocol II) and specifying what the obligations to take feasible precautions entail ('safeguard the marking of areas' in Amended Protocol II and 'training of armed forces' in Protocol IV).

In light of the unclarity about how IHL applies in the context of AWS, states could follow the examples from existing agreements and use the second tier to specify what limits and requirements IHL places on the development and use of AWS. In fact, states involved in the AWS debate, notably within the GGE, have already taken the first steps to this end.

While the specific elements of both tiers are far from settled, the discussion already reflects an agreement that the second tier could be formulated around operational restrictions (notably on types of targets, duration, geographical scope, the scale of use) and specific requirements (e.g. around the training of users and national measures undertaken by states). To further advance efforts to identify elements of a second tier, the approaches and details reflected in existing instruments provide some useful, common reference points.

Lesson 3: If there is will (and a need), two-tiered instruments can be grounded in concerns beyond IHL.

In addition to discussing the content of the two tiers, states are also debating whether a two-tiered approach to the regulation of AWS should reflect only concerns about IHL compliance, or whether broader concerns related to *international human rights law*, *ethics* or *international security* should also be factored in. Tensions around what concerns should – or should not – shape the content of an instrument are far from unique to AWS. How has this been solved in the past?

First of all, most existing arms control have been motivated by widely shared concerns about IHL compliance: prohibitions against blinding laser weapons, anti-personnel landmines and incendiary weapons are all examples of such, grounded in concerns about being inherently indiscriminate or causing unnecessary suffering and superfluous injury. In addition, there are examples of regulations that are motivated by more general humanitarian concerns. For example, the use of anti-personnel landmines was restricted not only because it would contravene IHL, but also because they were considered inhumane.

However, nothing seems to have influenced the final instruments more than national interests and strategic concerns: *Seemingly*, states' perceptions concerning the military utility of a weapon have historically been one of the most important factors in the determination of how a regulation may be structured and defined. This is especially reflected in the protocols to the CCW, where outright prohibitions mainly apply to weapons that states have shown no military interest in anyway, such as permanently blinding weapons.

In fact, as put by the civil society organization *Article 36*, the CCW is yet to demonstrate the ability to address humanitarian concerns associated with weapons of 'actual military relevance'. In cases where states have not been able to reach an agreement on how to address specific humanitarian concerns, negotiations have moved outside consensus-based forums like the CCW. The Anti-personnel Mine Ban Treaty and Cluster Convention are examples of this.

This is more of a reminder than a lesson to states involved in the AWS debate: a reminder that arms control is not an objective exercise, it is an inherently politically driven process. Agreeing on how to balance military necessity with humanitarian concerns has historically been a challenge for the arms control community, and as long as AWS are seen (by some, at least) as having military utility, the political will to address parallel humanitarian concerns may be constrained.

States involved in the ongoing AWS debate have fundamentally different perceptions of the legal, ethical, operational, and security challenges posed by AWS, and a key starting point to advance the two-tier discussions is to establish what concerns they collectively want, or do not want, to address in an instrument.

Regulating a process rather than a technology?

Situating the AWS debate within the history of arms control reminds us that the states do not need to reinvent the two-tiered wheel entirely. Diplomats have before faced – and solved – similar challenges.

With that said, some challenges are also new. The dilemmas posed by AWS appear more complex than those posed by previously regulated means and methods of warfare. For example, AWS do not produce characteristic effects that are deemed inherently unlawful, nor can they necessarily be

characterized as an inherently unlawful class of weapons like chemical or biological weapons.

The unique characteristics of AWS pose critical legal and ethical questions about how to ensure a (meaningful) human role in targeting. With that, the arms control community is forced to consider, perhaps for the first time, whether it wants to expand arms control to regulate *processes* of decision-making rather than, let's say, a specific technology; that is, whether it is seeking to establish an instrument that, through prohibitions and restrictions, aims to codify what is required by humans to comply with IHL and, in turn, defines what limits such requirements place on the use of autonomous functions.

See also:

- Alexander Blanchard, **The road less travelled: ethics in the international regulatory debate on autonomous weapon systems**, April 25, 2024
- Ingvild Bode, **Tom F.A. Watts, Loitering munitions: flagging an urgent need for legally binding rules for autonomy in weapon systems**, June 29, 2023
- Vincent Boulanin, Marta Bo, **Three lessons on the regulation of autonomous weapons systems to ensure accountability for violations of IHL**, March 2, 2023
- Laura Bruun, **Autonomous weapon systems: what the law says – and does not say – about the human role in the use of force**, November 11,2021

Tags: AI, artificial intelligence, autonomous weapons, AWS, Geneva Conventions, IHL, international humanitarian law, modern warfare

## *You may also be interested in:*

Internment pursuant to the Fourth Geneva Convention during an international armed conflict: practice and reflections from Norway

🕐 10 mins read

 Analysis / Autonomous Weapons / IHL / New Technologies / Special Themes
Camilla G. Cooper

The Fourth Geneva Convention was the first humanitarian law convention dedicated to protections for civilians ...
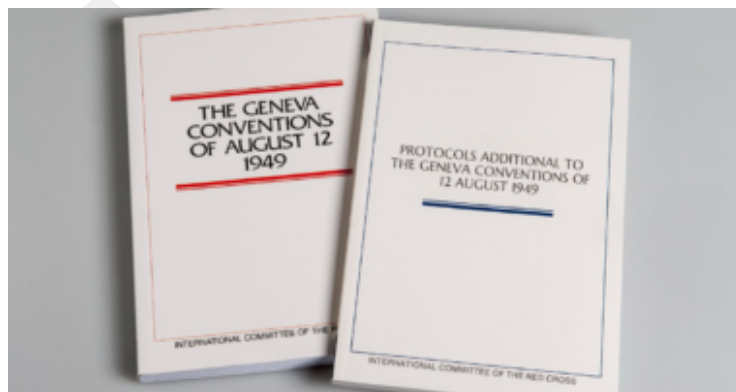
Civilian internment in international armed conflict: when does it begin?

🕐 12 mins read

 Analysis / Autonomous Weapons / IHL / New Technologies / Special Themes
Jelena Pejic

The legal and practical issues related to the exceptional practice of internment of protected persons ...