Ethics and Information Technology (2023) 25:25
https://doi.org/10.1007/s10676-023-09685-y

ORIGINAL PAPER

# Why a treaty on autonomous weapons is necessary and feasible

**Daan Kayser**[1]

Accepted: 27 January 2023 / Published online: 22 March 2023


**Abstract**
Militairy technology is developing at a rapid pace and we are seeing a growing number of weapons with increasing levels of autonomy being developed and deployed. This raises various legal, ethical, and security concerns. The absence of clear international rules setting limits and governing the use of autonomous weapons is extremely concerning. There is an urgent need for the international community to work together towards a treaty not only to safeguard ethical and legal norms, but also for our shared security. This article explains why a treaty on autonomous weapons is needed and achievable. It goes into what a treaty could consist of to establish an international norm and set rules and limits on autonomy in weapon systems.

**Keywords** Autonomous weapons · Military applications of AI · Meaningful human control · Humanitarian

## Introduction

A new generation of weapons systems with increasing levels of autonomy is being developed and deployed (PAX, 2021). An example that received a lot of international media attention was the use of the Kargu in Libya. This multi-rotor unmanned aerial vehicle can loiter in a designated geographical area searching for a predefined target type. According to a UN report these weapons systems "were programmed to attack targets without requiring data connectivity between the operator and the munition" (UN Security Council, 2021). In recent years there has been a large increase in the number of companies, from a growing number of countries, that are adding autonomous functions to a wide variety of platforms, from battle tanks, to navy vessels and unmanned aerial vehicles. Until recently it was mainly the USA, Western Europe, South Korea and Israel leading in this field. Now producers from many countries including China, Russia, Turkey, and Eastern Europe are developing weapons systems with autonomous capabilities. While for many of these weapons there is currently still a human operator approving attacks, technically this human approval can easily be removed.

It is deeply concerning that there is a new generation of weapons systems with increasingly autonomous capabilities (using automatic target recognition, facial/object recognition, swarming etc.), without a clear regulatory framework on how these weapons should be used in line with legal and ethical norms. Also, there are no clear legal rules stating which technologies would be fundamentally unacceptable and need to be prohibited. Therefore there is an urgent need for a treaty that addresses this and ensures meaningful human control over the use of force and prevents digital dehumanisation.[1] This article outlines why a treaty is necessary and what it could look like.

## What are autonomous weapons?

There is not yet one agreed definition of autonomous weapons. It is standard practice that a final legal definition is agreed on during treaty negotiations.[2] Various stakeholders use different definitions, but most will agree that the definition of the International Committee for the Red Cross

✉ Daan Kayser
kayser@paxforpeace.nl

[1] PAX, Utrecht, The Netherlands

[1] In the international debate various terms are used to describe the role of the human operator in the deployment of autonomous weapon systems. The general UN working language for this is 'the human element in the use of lethal force', which is intentionally broad and vague. Another example of a term that is used is 'appropriate levels of human judgment'. While human judgement must be retained, the main goal of the human role should be that they can control the effects of an attack on the target and its surroundings. Therefor in this article the term 'meaningful human control' is used.

[2] There are also disarmament treaties that do not include a definition, such as CCW Protocol on Blinding Laser Weapons (Protocol IV).

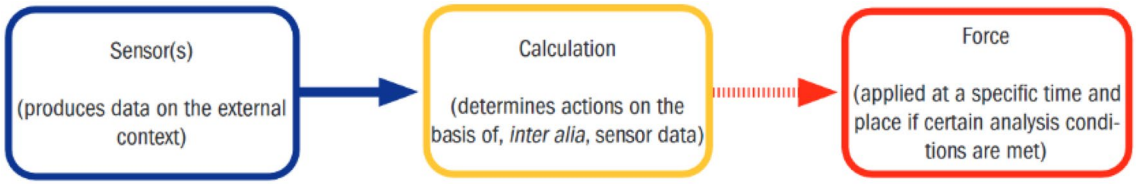


**Fig. 1** Sensor processing (image by Article 36, 2020)

(ICRC) contains the central elements of autonomy in weapons systems and forms a good basis for further work (Amoroso, 2020). The ICRC defines autonomous weapons as "any weapon system with autonomy in its critical functions. That is, a weapon system that can select and attack targets without human intervention" (ICRC, 2016). In other words, a weapons systems that can select and apply force to a target based on sensor inputs, rather than direct human inputs. Once activated by a human user, the weapons system can apply force to a target without direct human approval for a certain period of time. This means that the human user does not necessarily know the specific object to be attacked, and the time and place of the attack. This sensor-based targeting is a function of a weapon system. So a weapon system is functioning autonomously when the processing of sensor data can automatically trigger an application of force (Article 36, 2020 & Stop Killer Robots, 2022). This functionality can be added to numerous platforms including uncrewed aerial vehicles, tanks and naval vessels (Fig. 1).

## Debate within the UN

Since 2014, autonomous weapon systems have been discussed within the UN Convention on Certain Conventional Weapons (CCW). After years of debate it is now clear that a majority of states agree that certain autonomous weapons should be prohibited and that other autonomous weapons should be regulated.[3] These states see the need to guarantee human control over the use of force and see the need for concrete rules and measures to safeguard this. Unfortunately, this shared ambition has not yet translated into significant progress, as the CCW decides by consensus and a small minority of states has obstructed meaningful progress. After nine years of debate it is time to look for an alternative forum to develop a treaty.

Illustrating growing momentum is the joint statement at the 2022 Human Rights Council that stresses "the importance of human decision-making over the use of force and the imperative that the human element remains central in the use of force". The resolution calls for a study examining the human rights implications of new and emerging technologies in the military domain (Human Rights Council resolution, 2022). Also at the 2022 UN General Assembly seventy states underlined in a joint statement that autonomy in weapons systems "raise serious concerns from humanitarian, legal, security, technological and ethical perspectives". These states recognize the importance of "elaborating the normative and operational framework regulating, where appropriate and necessary, autonomous weapons including through internationally agreed rules and limits." This group includes Austria, Brazil, Chile, Germany, Japan, Mexico, Namibia, the United States and New Zealand (UN General Assembly, 2022). This growing consensus in various forums shows that a new treaty is not only necessary, but also feasible.

Given the great speed at which military technology is developing, it is vitally important to develop a new treaty as soon as possible. This could also happen in another UN forum or a stand-alone process, as was the case with the treaties banning landmines and cluster munitions.

## Concerns

The development of weapon systems with increasing levels of autonomy raises various concerns. Specifically these concerns are related to autonomy in the use of force (critical functions) and not necessarily to other functions like autonomous navigation or take-off and landing. The overarching concern is the loss of human control over the use of force. This has legal, ethical and security implications. First of all, their use without meaningful human control would not comply with International Humanitarian Law (IHL).[4] The law is addressed to humans and they are responsible for applying it. This requires combatants

[3] In the debate at the UN CCW Group Of Governmental Experts on LAWS this is often referred to as the two-tier approach.

[4] The use of weapons during armed conflict is governed by the IHL rules on the conduct of hostilities, including the rules of distinction, proportionality and precautions in attack.

who are carrying out an attack to make context-dependent decisions. If they do not know where and when an attack will take place, or against which specific target, it will be difficult if not impossible to make this legal judgement. Also there is the question of who would be responsible for violations of international law related to the deployment of autonomous weapons.

Furthermore targeting humans with autonomous systems raises fundamental concerns. It is ethically unacceptable to delegate decisions about life and death to algorithms. It goes against human dignity to reduce or remove human moral agency in the decision to kill. It is dehumanising if people are reduced to data points. Also legally, targeting humans with autonomous weapons raises concerns. In IHL humans can have a different legal status depending on the context and their behaviour, from being a protected person to a legitimate target. This changing legal status of humans increases the risk of protected persons being targeted when autonomous weapons are used.

There are also several security concerns. If autonomous weapons are developed and deployed without any kind of regulation, it will lead to more conflict and instability in the world. These weapons will lower the threshold to waging war and make countries more likely to resort to military force rather than to seek political solutions. Their unchecked development would likely lead to an arms race, which will have a destabilising effect. Also, it is unclear how autonomous weapons react and interact with each other, increasing the chance of accidental start or escalation of conflict. The proliferation of these weapons will make them easier to obtain for a wide range of actors. Any military advantage they offer will be temporary and limited in nature. Therefore developing rules and limits are in the interest of countries national security.

These concerns are shared by a wide variety of actors. UN Secretary-General António Guterres has repeatedly stated that lethal autonomous weapons are politically unacceptable and morally repugnant and that they should be prohibited internationally (UN Secretary-General António Guterres, 2019). The ICRC has stated that "loss of human control and judgement in the use of force raises serious humanitarian, legal, and ethical concerns" and has called for a new treaty (ICRC, 2021). In 2018 and 2021, the European Parliament passed resolutions calling for an international treaty (European parliament, 2018 & 2021). But there are also call from industry. For example the German Federation of Industry and numerous individual tech companies have called for a treaty (BDI, 2019). Last but not least, thousands of scientists in the area of computer science have warned for these weapons and called for a treaty (Future of Life, 2016).

## What a treaty could comprise

In the international debate various proposals have been put forward on how to address the issue of increasing autonomy in weapons systems (UNODA, 2022). In this section we put forward our thinking on what a treaty on autonomous weapons should look like.[5] A treaty should contain two main elements:

- A **prohibition** on autonomous weapons systems that:
  - cannot be used with meaningful human control;
  - have a human as the target profile.
- The **regulation** of autonomous weapons through the use of **positive obligations** that guarantee meaningful human control over their use.

**Meaningful human control** should be a central element of a regulatory framework. The implementation of this concept should ensure compliance with legal and ethical norms. Therefore the human user(s) must be able to predict, control and explain the effects of an attack on the target and its surroundings. This does not mean a human user controls the weapon system directly, but the implementation of certain positive obligations should ensure the human user(s) can predict, control and explain the effects. These positive obligations are described below under 'regulation'.

## Prohibitions

The new instrument should prohibit autonomous weapons that *cannot* be used with meaningful human control (fully autonomous weapons). In other words, where the human user(s) cannot predict, control and explain the effects of an attack. For example, this prohibition would include weapons systems that are able to independently modify critical mission parameters during deployment with the aid of machine learning. The new instrument should also prohibit autonomous weapons that have humans as targets. It is unethical to reduce humans to binary data. Furthermore, this does not respect human dignity given the lack of active involvement of humans in the decision-making process.

## Regulation

The use of autonomous weapons systems that not fall under the prohibitions, should be regulated to ensure their use meets legal and ethical requirements.

[5] This section is based on the policy position of the Stop Killer Robots campaign, which PAX co-developed as a member of the Steering Committee

IHL requires that those who plan, decide and carry out an attack must fulfil certain legal requirements (ICRC & SIPRI, 2020). This means the human user(s) must be able to make a context-based decision on the legality of an intended attack. Positive obligations that ensure meaningful human control are necessary to allow a human user to fulfil their obligation to apply the law, make a moral judgement, set limits on use, and also to ensure there is a human who can be held morally and legally accountable. Therefore user(s) must be able to predict, control and explain the effects of an attack on the target and its surroundings, and make an informed decision whether these effects meet legal and ethical standards.

Therefore a treaty should include positive obligations in the design and use of autonomous weapons systems. For example:

- The human user should understand the way the weapons system works and the context where the attack takes place.
- The weapons system should be reliable and predictable.
- The use of the weapons system should be limited in time and geographical area of operation, as well as the type of targets.

The level and form of human control that is required may depend on such factors such as where the system is deployed, for example the ocean versus an urban area.[6] However the implementation of the positive obligations should ensure the human user(s) can predict, control and explain the effects of an attack and ensure these effects are in line with legal and ethical norms.

## How to move forward?

As mentioned above the debate on autonomous weapons has mainly taken place at the CCW, where progress has been impeded by the need for consensus. Of course it is preferable to have as many states involved as possible in the process towards an international treaty. But if that means no treaty can be achieved or one that sets a very low standard, alternative options should be considered as the absence of legal rules on autonomy in weapons systems carry great risks for international peace and security.

The experience of other disarmament treaties such as those on landmines and cluster munitions shows that it can be more effective for a group of countries that want to set a high standard to take the lead. Other countries then often join in the course of the process. A treaty also creates an international norm that influences the conduct of countries that have not signed the treaty. A strict norm that is supported by the majority of countries is better for international peace and security than a situation in which no new norm is set for autonomous weapons, leaving them unregulated and allowing them to be deployed broadly.

The military advantage derived from autonomy can also be achieved with autonomous weapons that *have* meaningful human control, combining the strengths of humans and machines. There are also various ways of defending a country against possible attacks by fully autonomous weapons that do not involve the country in question crossing moral and legal boundaries itself. A treaty that enjoys broad support will considerably reduce the risk of such an attack anyway as it will be able to prevent the unrestricted development and proliferation of these weapons, as well as stigmatize any use. Accordingly, a treaty in line with ethical and legal norms is in states' own security interests.

Countries benefit from multilateralism and a world order based on norms, not just because these norms form the foundation of our society but also because they contribute to our security and economic prosperity. These norms deserve to be protected and consolidated. That is why like-minded states should join forces and take action to bring about an international treaty that guarantees meaningful human control over the use of force.

## Conclusion

Technology is developing at a rapid pace and we are seeing a growing number of weapons with increasing levels of autonomy being developed and deployed. This raises various legal, ethical, and security concerns. The absence of clear international rules setting limits and governing the use of autonomous weapons is extremely concerning. There is an urgent need for the international community to work together towards a treaty not only to safeguard ethical and legal norms, but also for our security.

## References

Daniele Amoroso. (2020). '*Autonomous Weapons Systems and International Law: A Study on Human-Machine Interactions in Ethically and Legally Sensitive Domains*'. (Edizioni Scientifiche Italiane 2020). p. 17

BDI. (2019). https://english.bdi.eu/article/news/waiting-is-the-biggest-risk/ https://documents-dds-ny.un.org/doc/UNDOC/GEN/N21/037/72/PDF/N2103772.pdf?OpenElement__;!!NLFGqXoFfo8MMQ!osNWB-NwNHbPEl8wFuxHkFJkF1DWeFtIBuvr

[6] Due to the inherent unpredictability of autonomous weapons one could argue that autonomous weapons should not be used in populated areas

YpKYwFKFPLP023BJtVqZgguYezLjpPY1ih0Mzt0feZhLBpgRVI-fwrdT7F80
European Parliament. (2018). resolution www.europarl.europa.eu/doceo/document/TA-8-2018-0341_EN.html; (2021) resolution www.europarl.europa.eu/doceo/document/A-9-2021-0001_EN.pdf
Article 36. (2020). '*Autonomy in weapons systems—considering approaches to regulation',* https://article36.org/updates/short-approaches-to-regulation/; Stop Killer Robots (2022) 'Our policy position', https://www.stopkillerrobots.org/our-policies/
Future of Life. (2016). '*Autonomous Weapons Open Letter: AI & Robotics Researchers',* https://futureoflife.org/open-letter/open-letter-autonomous-weapons-ai-robotics/
Human Rights Council resolution. (October 2022). '*Human rights implications of new and emerging technologies in the military domain',* https://ecnl.org/sites/default/files/2022-10/Res1.pdf
ICRC. (2016). '*Views of the International Committee of the Red Cross (ICRC) on autonomous weapon systems'* https://www.icrc.org/en/document/views-icrc-autonomous-weapon-system
ICRC. (2021). '*Autonomous weapons: The ICRC calls on states to take steps towards treaty negotiations'* https://www.icrc.org/en/document/autonomous-weapons-icrc-calls-states-towards-treaty-negotiations
PAX. (2021). '*Increasing autonomy in weapon systems'*, https://paxforpeace.nl/what-we-do/publications/increasing-autonomy-in-weapons-systems
ICRC & SIPRI. (2020). '*Limits on autonomy in weapon systems'* https://www.sipri.org/media/press-release/2020/new-sipri-and-icrc-report-identifies-necessary-controls-autonomous-weapons
UN Secretary-General António Guterres. (2019). *'Autonomous weapons that kill must be banned, insists UN chief'* https://news.un.org/en/story/2019/03/1035381
UN Security Council. (March 2021). '*Letter dated 8 March 2021 from the Panel of Experts on Libya established pursuant to resolution 1973 (2011) addressed to the President of the Security Council'* https://documents-dds-ny.un.org/doc/UNDOC/GEN/N21/037/72/PDF/N2103772.pdf?OpenElement
United Nations General Assembly. (21 October 2022). *'Joint Statement on Lethal Autonomous Weapons Systems First Committee, 77th United Nations General Assembly Thematic Debate – Conventional Weapons',* https://estatements.unmeetings.org/estatements/11.0010/20221021/A1jJ8bNfWGlL/KLw9WYcSnnAm_en.pdf
UNODA. (2022). 'CCW GGE LAWS 2022', https://meetings.unoda.org/ccw/convention-certain-conventional-weapons-group-governmental-experts-2022