

RESEARCH BRIEF

DEFENDING THE BOUNDARY: CONSTRAINTS AND REQUIREMENTS ON THE USE OF AUTONOMOUS WEAPON SYSTEMS UNDER INTERNATIONAL HUMANITARIAN AND HUMAN RIGHTS LAW

KEY MESSAGES

- Autonomous weapon systems (AWS) tend to be portrayed as ‘weapons of war’, but international humanitarian law (IHL) would never be the sole, and in many instances, it would not be the primary legal frame of reference to assess the legality of their use. Consideration of international human rights law (IHRL) requirements and constraints on the use of AWS must be a part of the debate on AWS, including in the framework of the 1980 Convention on Certain Conventional Weapons (CCW).
- Where IHL permits the ‘categorical’ targeting of security measures, including the use of force, there is scope for the lawful use of an AWS. However, due to procedural requirements and the need to individuate the use of force, this scope is extremely limited under IHRL. IHRL requirements and constraints apply to the use of an AWS in an armed conflict in so far as they are not displaced by IHL.
- To safeguard human dignity and human rights, human agents must:
 - exercise the control necessary to determine, in a timely manner, what legal rules govern applications of force by means of an AWS, and adapt operations as required
 - remain involved in algorithmic targeting processes in a manner that enables them to explain the reasoning underlying algorithmic decisions in concrete circumstances
 - be continuously and actively (personally) engaged in every instance of force application outside of the conduct of hostilities
 - exercise active and constant (continuous or at least frequent, periodic) human control over every individual attack in the conduct of hostilities
 - appropriately bound every attack in spatio-temporal terms to be able to recognize changing circumstances and adjust operations in a timely manner

MAY 2017 | MAYA BREHM

INTRODUCTION

Over recent years, there has been growing debate about the ethical, humanitarian, legal and security implications of AWS. There is, as yet, no agreed definition of an AWS, but the basic idea is that once activated, such a weapon system would, with the help of sensors and computationally intensive algorithms, detect, select and attack targets without further human intervention. According to leading researchers in the field of artificial intelligence (AI) and robotics, the deployment of such systems will be practically, if not legally, feasible within years.

Would the use of an AWS comply with international legal standards for the protection of the human person? This is one of the questions that occupy centre stage in discussions within the framework of the CCW, which has mandated a Group of Governmental Experts to consider the issue (UN doc CCW/CONFV/10). There is general agreement that the use of an AWS would need to comply with IHL and IHRL standards on the use of force, but views diverge on the circumstances in which it would be lawful to use an AWS and on whether additional law is required to ensure respect for the norms that safeguard humanity.

The focus of scholarly inquiry into the legality of AWS has been on compliance with IHL rules on the conduct of hostilities. Comparably little attention has been given to the impact of AWS on human rights protection. With a view to closing this gap and supporting multilateral policy discussions, the Geneva Academy carried out research in 2016 on the requirements and constraints that IHRL places

on the use of force by means of an AWS, both in relation to the conduct of hostilities and for law enforcement purposes, in times of peace as well as during armed conflicts. The use of a 'sentry-AWS' to control a boundary, secure a perimeter or deny access to an area, for example along an international

border – a possible application envisaged by proponents of AWS – forms the backdrop to the legal discussion.

This brief aims to provide policy makers and advocacy groups with a summary of key findings.

To ensure that human beings continue to exercise control in a meaningful manner, the independent operation of an AWS must be adequately bounded.

ENVISIONING AWS, HUMAN CONTROL AND THE APPLICATION OF LAW

It is a long-standing requirement for the moral acceptability, political legitimacy and legality of organized violence that weapons and their consequences must be controllable and controlled. Legal obligations are addressed to human beings, whereas AWS are artefacts, objects made for a certain purpose, devoid of intentionality. AWS would change how human beings exercise control in the use of weapons, and thereby affect their ability to perform legal duties and be accountable for the consequences. An AWS cannot meaningfully be treated as an entity accountable for harm done or a violation of the law. Compliance with the law, as well as ethical and other imperatives, presupposes a measure of human agency in the use of force. This places limitations on permissible 'human-machine configurations' (Suchman and Weber, 2016). To ensure that human beings continue to exercise control in a meaningful manner, the independent operation of an AWS must be adequately bounded. What

'SENTRY-AWS': AUTONOMOUS BORDER CONTROL AND PERIMETER SECURITY

Sentry systems with autonomous capabilities are already deployed, but none of them is currently selecting and attacking targets without direct human intervention. A well-known system of this type is DoDaam's Super aEgis II deployed in the Demilitarized Zone (DMZ) between North and South Korea. According to the manufacturer, it can be equipped with a 12.7mm machine gun, a 40mm grenade launcher or a surface-to-air missile launcher. The system reportedly has the capability to identify, track and destroy a moving target. The original version had an 'auto-firing system', but in present practice a human operator unlocks the system's firing capability.

Autonomous systems of this type are advertised for use in diverse operational environments, ranging from 'military force protection' to 'homeland security' and the 'interception' of trespassers around power plants or airports. Such systems can comprise mobile components, allowing them to patrol a boundary or area; some can be equipped with weapons branded as 'non-lethal'. Partisans of sentry-AWS expect that their deployment will obviate the need to expose human security personnel to danger. They also hope to 'close the kill chain' by combining target detection, identification and the capability to fire.

‘meaningful human control’ involves, concretely, remains to be determined.

THE APPLICABLE LAW: A QUESTION OF HUMAN INTENT AND CONTROL

The scope for the lawful use of an AWS depends to a significant degree on the legal rules governing the use of force in a concrete instance, and on the interaction between IHRL and IHL.

In times of peace, the use of an AWS for law enforcement purposes would be governed by IHRL standards on the use of force. In times of armed conflict, including in situations of hostilities governed by IHL, human rights protection does not cease. In such situations, the rules of IHL and IHRL apply concurrently and need to be reconciled. Whereas IHL is the primary reference point for assessing the use of an AWS as a means of warfare, any other exercise by states of their authority or power involving the use of an AWS continues to be assessed within a law enforcement framework (taking account of IHL).

In the use of an AWS, the spatial, temporal and causal remoteness of human intervention from the locus of force application affects intent- and control-related determinants of the applicable law:

- A will to ‘wage war’ against another state (*animus belligerendi*), for instance, cannot be presumed. Due to the lack of proximate human involvement in the use of force, an AWS cannot trigger an international armed conflict ‘on its own’. Its applications of force thus remain governed by IHRL standards on the use of force.
- In order to use an AWS to conduct hostilities during an armed conflict, human agents of a party to the conflict must exercise sufficiently proximate control over the AWS to use it as a means of warfare. If an AWS is not activated by a human agent with the intent to conduct hostilities, the ‘belligerent nexus’ between the AWS’ applications of force and the armed conflict cannot be presumed. Such applications of force remain governed by law enforcement standards (taking account of IHL).

In spite of the predominant portrayal of AWS as weapons of war, IHL is never the sole nor in many instances,

the primary legal frame of reference to assess the legality of their use. To comply with all applicable law, human agents must exercise the control necessary to determine, in a timely manner, what legal rules govern any specific use of force, and adapt operations as required.

THE DUTY TO INDIVIDUATE THE USE OF FORCE UNDER IHRL

The use of force by means of an AWS, in pursuit of a legitimate law enforcement objective, would expose anyone falling within the parameters of a valid target to a real and immediate risk to life. To safeguard life, a state deploying an AWS has a duty to take all measures necessary to effectively prevent anyone potentially falling within the system’s target parameters, but who may not be legally killed, from entering the system’s sensor and weapon range.

As the right to life is inherent in every person, it is not enough to stipulate that potentially lethal force may be used to target a certain category of people presumed to pose a threat (e.g. persons convicted of a violent crime attempting to flee a detention facility). To be legal, the use of force must also be justified in the concrete circumstances prevailing at the time. The use of lethal force is only justifiable if the particular person that force is directed at poses an imminent threat of death or serious injury.

To comply with the requirement that lethal force be used only as a last resort whilst minimizing the risk of deprivation of life or bodily harm, human agents must be continuously, actively and, arguably, personally engaged in every instance of force application. Due to the need to individuate the use of force under IHRL, the scope for the lawful use of an AWS is extremely limited.

THE LIMITED SCOPE FOR CATEGORICAL TARGETING UNDER IHL

In contrast, in times of armed conflict the necessity to use force may in certain circumstances be presumed. IHL permits the ‘categorical targeting’ of persons based on their ‘status’ or (imputed) membership in a group (e.g. combatants), or of objects due to their ‘nature’. This allows the broadening of the context within which the legality of AWS use is assessed to that of an attack (as a whole) as defined in IHL, rather than individual applications of force or acts of violence, and opens up

Due to the need to individuate the use of force under IHRL, the scope for the lawful use of an AWS is extremely limited.

limited scope for the lawful use of an AWS.

However, to ensure that targeting rules can be applied in a manner that effectively protects the victims of war, even though the number and context of specific acts of violence is not known when an attack is launched, human agents must bound every attack appropriately in spatio-temporal terms and retain sufficient control to recognize changing circumstances and to adjust operations in a timely manner. This calls for active and constant, in the sense of continuous or at least frequent, periodic, human control over every individual attack. Human control during the conduct of hostilities must also safeguard the opportunity to shift to a law enforcement model of force application when this becomes factually possible and, thus, legally mandated.

HUMAN CONTROL AS A PROCEDURAL SAFEGUARD

AWS take shape against the backdrop of practices of automated killing, mass surveillance, large-scale interception of personal data and algorithm-based profiling that are attracting strong criticism from human rights actors. Developments enabling the algorithmic construction of targets of security measures, including the use of force, can be expected to sustain and promote these practices.

Intrusive mass surveillance systems put in place in recent years generate large amounts of digitized data to which AI techniques can be applied, with minimum human intervention. Using an AWS capable of detecting individuals or objects matching certain criteria and tracking them would entail surveillance of people's habits of everyday life and social relationships. It would involve the automatic processing of personal data with the potential to undermine key data protection principles and the right to privacy.

The mining of massive datasets offers the possibility of categorizing individuals on the basis of some observable characteristics so as to infer other characteristics with a view to taking individual decisions relating to them or predicting their attitude or behaviour. This constitutes profiling – an approach criticized for its dehumanizing quality in that it tends to reduce the person to the profile generated by automated processes. When used as a basis for decision making, this can have unfair, stigmatizing and

discriminatory impacts.

Autonomous targeting also subjects individuals to measures of an automatic nature, treating all people indiscriminately like objects. It precludes deliberative human intervention (Asaro, 2012) and the exercise of discretion (Lieblich and Benvenisti, 2016). The calculated blindness to individual circumstances involved in the use of an AWS is an affront to human dignity and can amount to cruel, inhuman or degrading treatment.

To safeguard human dignity and human rights, human agents must remain involved in algorithmic targeting processes in a manner that enables them to explain the reasoning underlying algorithmic decisions in concrete circumstances. This is essential to ensuring the availability of an effective remedy, accountability for the use of force and for maintaining public confidence in states' adherence to the rule of law, in times of peace as well as war.

CONCLUDING REMARKS

Legal norms already regulate and limit algorithmic decision making and automated killing but new technologies and evolving security practices challenge the categories and disrupt the 'human-machine configurations' around which the legal regulation of force is articulated. This generates controversies and uncertainties about the applicability and meaning of existing norms, thus diminishing existing law's capacity to serve as a guidepost.

In addition, accommodating new practices within an existing legal framework bears the risk that existing rules are preserved formally, but filled with a radically different meaning. In light of this, an explicit, formal, legal requirement to exercise meaningful human control in the use of force would help to safeguard human dignity and human rights.

To safeguard human dignity and human rights, human agents must remain involved in algorithmic targeting processes in a manner that enables them to explain the reasoning underlying algorithmic decisions in concrete circumstances.

SOURCES

P. Asaro, 'On Banning Autonomous Weapon Systems: Human Rights, Automation, and the Dehumanization of Lethal Decision-Making', 94 *International Review of the Red Cross* 886 (2012) 687–709.

E. Liebllich and E. Benvenisti, 'The Obligation to Exercise Discretion: Why Autonomous Weapons Systems are Unlawful', in N. Bhuta, S. Beck, R. Geiß, H.-Y. Liu and C. Kreß (eds), *Autonomous Weapons Systems: Law, Ethics, Policy*, Cambridge University Press, 2016, pp 245-283.

L. Suchman and J. Weber, 'Human-Machine Autonomies', in N. Bhuta et al (eds), *Autonomous Weapons Systems: Law, Ethics, Policy*, Cambridge University Press, 2016, pp 75–102.

Final Document (Advance Version), Fifth Review Conference of the High Contracting Parties to the Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to Be Excessively Injurious or to Have Indiscriminate Effects, UN doc CCW/CONFV/10, 23 December 2016.

SELECTED REFERENCES

Campaign to Stop Killer Robots

CCW, Group of Governmental Experts on Lethal Autonomous Weapons Systems

International Committee for Robot Arms Control (ICRAC)

International Committee of the Red Cross, 'Views of the ICRC on Autonomous Weapon System', 11 April 2016

UNIDIR, 'The Weaponization of Autonomous Technologies', Research Project



© Nobordernetwork

THE GENEVA ACADEMY

The Geneva Academy provides post-graduate education, conducts academic legal research and policy studies, and organizes training courses and expert meetings. We concentrate on branches of international law that relate to situations of armed conflict, protracted violence, and protection of human rights.

WEAPONS LAW

Our research on the regulation of weapons under international law focuses on challenges posed by certain weapon technologies and practices of armed violence and on the development and updating of the Weapons Law Encyclopedia – a pioneering online compilation of information on the regulation of weapons under international law. Research on the legality of AWS carried out in 2016 was funded by the Swiss Federal Department of Foreign Affairs.

**The Geneva Academy
of International Humanitarian Law
and Human Rights**

Villa Moynier
Rue de Lausanne 120B
CP 1063 - 1211 Geneva 1 - Switzerland
Phone: +41 (22) 908 44 83
Email: info@geneva-academy.ch
www.geneva-academy.ch

**© The Geneva Academy
of International Humanitarian Law
and Human Rights**

This work is licensed for use under a Creative Commons Attribution-Non-Commercial-Share Alike 4.0 International License (CC BY-NC-ND 4.0).