

Elizabeth Minor, Advisor, Article 36 – Written Evidence (AIW0017)

About us

Article 36 is a specialist non-profit organisation, focused on reducing harm from weapons. We are part of the steering committees of the global Stop Killer Robots campaign and the UK Campaign to Stop Killer Robots. Our Managing Director, Richard Moyes, is currently the Coordinator of Stop Killer Robots, and is a member of the UK MoD AI Ethics Advisory Panel.

Recommendations

- If the UK government wishes to be at the forefront of shaping new rules on emerging technology,¹ and to set clear international norms² in this area, it must recognise that additional international law on AWS is necessary. The UK should acknowledge that new law would be beneficial if it strengthens the application of existing law and if it protects and upholds key moral and legal principles.
- Whilst it should recognise the valuable work that states have undertaken in the Convention on Conventional Weapons (CCW) to advance the policy conversation on AWS, the UK government should not be dogmatic about the forum in which AWS are discussed. Given the political deadlock in the CCW, such rigidity could give the appearance of simply attempting to prevent progress. Rather, the government should support and participate in discussions in which all states' concerns can be raised - including but not limited to those with military interests in this area. AWS have implications for all countries. Following its endorsement of a broadly supported joint statement in the UN General Assembly (UNGA) last year,³ the government should continue to support engagement on AWS at the UNGA and in other relevant forums.
- Within the UK, the government should convene inclusive, multi-stakeholder discussions to bring together and consult with a range of

¹ As stated in the Integrated Review (2021) <https://www.gov.uk/government/collections/the-integrated-review-2021>

² As stated in parliamentary answer UIN HL2031 <https://questions-statements.parliament.uk/written-questions/detail/2022-07-21/HL2031>

³ See <https://article36.org/updates/ground-breaking-un-joint-statement-on-autonomous-weapons/>

voices on this issue, including civil society, academia and the tech sector.

- To clarify its position on what “context appropriate human involvement” should entail, the UK government should answer the specific policy questions posed in our answer to question 5 below.

1. What do you understand by the term autonomous weapons system (AWS)? Should the UK adopt an operative definition of AWS?

AWS are systems that use sensors to determine where and when force will occur, without this being set specifically by a person. AWS apply a process of functioning of: gathering information via sensors; calculating whether this matches a predetermined ‘target profile’; and, applying force if so. ‘Target profiles’ are encoded representations that serve to approximate human concepts for the purpose of applying force. They will encompass both objects that are intended targets and other objects.⁴ The challenges and risks posed by AWS flow from these characteristics.

The UK government should focus on the broad scope of systems that apply this process of functioning in its policymaking on AWS. The UK government should orientate to broad terminology - such as AWS - rather than forum-specific language (such as Lethal Autonomous Weapons Systems, the terminology used in the CCW) in policymaking.

A joint paper submitted to the CCW by the US and supported by the UK refers to “autonomous weapons systems” as those “weapon systems that, once activated, can identify, select, and engage targets with lethal force without further intervention by an operator”.⁵ The reference to ‘lethal’ force is not useful to include as a qualifier, as this is a possible outcome rather than a characteristic of any weapons system. Otherwise, this working

⁴ The concept of “target profiles” is elaborated in Richard Moyes. (2019), “Target profiles,” Article 36 <https://article36.org/wp-content/uploads/2020/12/Target-profiles.pdf>

⁵ Australia, Canada, Japan, the Republic of Korea, the United Kingdom, and the United States (2023), ‘Draft articles on autonomous weapon systems – prohibitions and other regulatory measures on the basis of international humanitarian law (“IHL”)’ UN document number CCW/GGE.1/2023/WP.4/Rev.1 available at [https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-_Group_of_Governmental_Experts_on_Lethal_Autonomous_Weapons_Systems_\(2023\)/CCW_GGE1_2023_WP.4_Rev1.pdf](https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-_Group_of_Governmental_Experts_on_Lethal_Autonomous_Weapons_Systems_(2023)/CCW_GGE1_2023_WP.4_Rev1.pdf)

definition (which the UK has endorsed) broadly reflects the scope of AWS that we consider should be subject to regulation.

2. What are the possible challenges, risks, benefits and ethical concerns of AWS? How would AWS change the makeup of defence forces and the nature of combat?

States are pursuing AWS due to the perceived military utility that, for example, the increased speed of automated systems and their ability to operate in contested environments might bring, as well as the possibility that developing a new generation of weapons systems could bring strategic advantages.

However, systems operating with the process of functioning described above raise fundamental concerns:

The first relates to control in the use of force. Because an AWS's sensors determine when, where and to what force will be applied - and this will not be chosen specifically by a person - meaningful human control over weapons systems risks being eroded. This is problematic legally and ethically. Where complex systems are built on opaque processes, our ability to understand these tools, or to explain the results that they produce, is reduced. The speed of interaction between complex AWS could also reduce the space for human values and judgement, eroding control and risking unacceptable and harmful outcomes. Increasing remoteness and autonomy could also further displace violence from militaries onto civilians.

Secondly, using AWS to automatically target people would involve treating human beings as objects, a digital dehumanisation that would undermine human dignity and human rights. Biases in systems would, further, reproduce discrimination. Killing with such AWS could also be based on encoded indicators of gender, race or other identities.

AWS also raise wider peace and security concerns. Remoteness and automation risks lowering political thresholds against the use of force. Competition in the development and use of AWS risks producing dangerous arms-racing dynamics between countries. And, crises could escalate through

the use of high-speed systems and competing understandings of what the use of certain systems signifies and how legal principles apply.

Our fundamental concerns with AWS are with the risks to human dignity, and the risks of undermining existing law as a structure through which human moral agency is preserved and collective values are upheld.

3. What safeguards (technological, legal, procedural or otherwise) would be needed to ensure safe, reliable and accountable AWS?

A clear legal structure is needed to ensure the sufficient control of weapons systems. This must both prevent the design and development of systems that cannot be effectively controlled, and the use of other systems in such a way that they are not meaningfully controlled by their users.

Meaningful human control requires the users of AWS to have a sufficient understanding of the effects that the system will have in the area of use (including, what intended and unintended objects will fall within its target profile). It also requires limits on the duration and area of a system's use, such that legal rules can be applied.

4. Is existing International Humanitarian Law (IHL) sufficient to --ensure any AWS act safely and appropriately? What oversight or accountability measures are necessary to ensure compliance with IHL? If IHL is insufficient, what other mechanisms should be introduced to regulate AWS?

No matter how sophisticated, AWS are tools that people use, and it is people that must apply the law - care should be taken not to imply that AWS can hold any agency or responsibility. States at the CCW have collectively recognised that "accountability cannot be transferred to machines."⁶

Increasing automation - whereby the area and period of time over which an AWS can sense, process and apply force to objects without further human

⁶ CCW (2019), Final Report, 'Annex III: Guiding Principles affirmed by the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons System' UN document number: CCW/MSP/2019/9 <http://undocs.org/CCW/MSP/2019/9>

intervention, and whereby the target profiles used in AWS become ever more complex - risks people making legal judgements based on more and more diluted understandings of where, when and to what force will be applied. This presents challenges to the effective application of key principles in existing IHL such as distinction and proportionality. It also risks human responsibility and the human role in legal decision making being eroded, undermining the possibility for meaningful accountability.

Existing IHL does not provide guidance regarding the duration or area of individual attacks: AWS can stretch these parameters, and the development and clarification of the law is needed to respond to this. The adequacy of human legal judgement is dependent upon it being sufficiently granular and contextually informed. Without creating stronger rules to preserve this, the wider fabric of the law will suffer.

The challenges to human dignity raised by the automated processing of people by weapons systems are also not addressed by existing IHL. New laws developing in the civilian sphere, such as the European Union's AI Act, and the General Data Protection Regulation (GDPR), have started to respond to the need to protect individuals from harms related to automated decision making.⁷ This indicates the imperative for new law when it comes to the military sphere, where the consequences of such decision making are most severe for individuals.

So far, states at the CCW have agreed on 11 "Guiding Principles" for their discussions. Though these contain some useful material, these were intended to guide states' future work, rather than representing any sort of end point of agreement.⁸ There remains no international agreement or consensus on how IHL should be best upheld when it comes to the use of AWS.

A new legally binding instrument on AWS would be the best way to address this, and must be negotiated by states. This will need to contain prohibitions and regulations to prevent the automation of killing and ensure meaningful

⁷ See for example Christina Huchro (2022) 'Foundations of the GDPR: Principles relevant to discussions on autonomous weapons,' Article 36 <https://article36.org/wp-content/uploads/2022/03/AWS-Foundations-of-the-GDPR.pdf>

⁸ See Richard Moyes (2019), 'Critical Commentary on the Guiding Principles' Article 36 <https://article36.org/wp-content/uploads/2019/11/Commentary-on-the-guiding-principles.pdf>

human control over weapons systems. The negotiation of a legally binding instrument is now supported by the majority of countries at the CCW.

5. What are your views on the Government's AI Defence Strategy and the policy statement 'Ambitious, safe, responsible: our approach to the delivery of AI-enabled capability in Defence'? Are these sufficient in guiding the development and application of AWS? How does UK policy compare to that of other countries?

Through the Defence AI Strategy and policy statement, the UK government recognises that lines need to be drawn when it comes to AWS - specifically, that systems which identify, select and attack targets without "context-appropriate human involvement" would be unacceptable. This is valuable. Nevertheless, the position so far lacks sufficient detail.

Towards clarifying where the lines of unacceptability are when it comes to AWS, and the elements that are required for "context-specific human involvement" in the use of weapons systems (which we would consider similar in content to "meaningful human control"), we recommend the UK government answers the following specific policy questions. Article 36 elaborated these in a publication in 2019, and has previously sought answers to them from the government:⁹

1. For systems that process sensor inputs to determine where and when to apply force (without further human involvement) would the following be acceptable or unacceptable?
 - a. Systems that are designed to identify people as targets on the basis of human biometrics?
 - b. Systems that identify different groups of people as targets on the basis of perceived racial, gender or age characteristics?
 - c. Systems where the sensor-identifiable characteristics of possible targets can change or develop, within the system, after it has been activated and without being specifically certified by a human?

⁹ Richard Moyes (2019), 'Autonomy in weapons systems: mapping a structure for regulation through specific policy questions,' Article 36 <https://article36.org/wp-content/uploads/2019/11/regulation-structure.pdf>

- d. Situations where the human users understand what the system is intended to target, but do not know the actual physical/emission characteristics that will be identified as a target – such as where target profiles have been built through current neural network/ machine learning?
 - e. Situations where the human users do not have an understanding of what, other than intended targets, might also be identified as targets by the system?
2. For systems that process sensor inputs to determine where and when to apply force (without further human involvement) are the following assertions reasonable?
- a. Human users should be fully responsible for verifying the risk to civilians from their use of a system;
 - b. Systems that will target both certain civilian objects and certain military objects should not be used in situations where those objects are intermingled;
 - c. The geographic area over which a sensor-targeting function can occur should be controlled such that human users can fulfil their legal obligations;
 - d. The duration over which a sensor-targeting function can occur should be controlled such that human users can fulfil their legal obligations;
 - e. The time at which a sensor-targeting function may occur should be sufficiently proximate to the application of human legal judgement for that legal judgement to be relevant to the circumstances in which the function will occur;
 - f. The number of applications of force that a system can undertake in an individual attack should be set by the human users;
 - g. Human users need to understand the actual weapon effects (type of force) that such systems will create.

The UK's stated ambition to be a normative leader in the area of AWS is not matched in reality.

The UK has made useful contributions to policy discussion at the CCW, particularly recently, including through contributing useful content on broad measures and approaches that would facilitate meaningful human control

and could provide a basis for regulation (such as those outlined in written contributions on the mitigation of the risk of unintended engagements, and good practices related to human-machine interaction¹⁰).

Nevertheless, the UK is in a minority position internationally in opposing the elaboration of new international law on AWS (preferring an approach of political standards or good practices supported by a minority). Around 90 states have now expressed support for the negotiation of a legally binding instrument.¹¹ In February at a regional conference in Costa Rica, Latin American and Caribbean countries committed to work for a new treaty, providing one political foundation for work ahead.¹²

The UK position that existing IHL is sufficient to regulate AWS should be seen as the government maintaining a holding pattern whilst waiting for others to set the terms of progress, rather than providing normative direction. Previous UK governments have maintained similar positions on other weapons issues, before changing tack when others have taken the initiative to start the negotiation of new laws.

¹⁰ Australia, Canada, Japan, the Republic of Korea, the United Kingdom, and the United States (2023), 'Draft articles on autonomous weapon systems – prohibitions and other regulatory measures on the basis of international humanitarian law ("IHL")' UN document number CCW/GGE.1/2023/WP.4/Rev.1 available at [https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-_Group_of_Governmental_Experts_on_Lethal_Autonomous_Weapons_Systems_\(2023\)/CCW_GGE1_2023_WP.4_Rev1.pdf](https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-_Group_of_Governmental_Experts_on_Lethal_Autonomous_Weapons_Systems_(2023)/CCW_GGE1_2023_WP.4_Rev1.pdf)

¹¹ See Automated Decision Research, 'State positions' for monitoring <https://automatedresearch.org/state-positions/>

¹² See <https://article36.org/updates/latin-american-and-caribbean-states-lead-the-way-towards-a-treaty-on-autonomous-weapons/>

6. Are existing legal provisions and regulations which seek to regulate AI and weapons systems sufficient to govern the use of AWS? If not, what reforms are needed nationally and internationally; and what are the barriers to making those reforms?

A new internationally negotiated legally binding instrument on AWS is needed. This should contain a combination of prohibitions and regulations. It must prohibit systems that are triggered by the presence, proximity or contact of a person, as well as those that cannot be adequately understood, their area or duration sufficiently limited, or that otherwise cannot be used with meaningful human control. It must also contain positive obligations to ensure all weapons systems are used with meaningful human control.

At the CCW, there has been significant progress in recent years in the development of shared understandings around AWS, and broad policy convergence among a range of states that currently differ on whether new international law is needed. For example, there is increasing acknowledgement that the regulation of AWS requires a two-tiered approach of recognising the range of systems that are prohibited and regulating the rest, as well as a broad recognition of a link between IHL compliance and human control.¹³

Nevertheless, no legally or politically binding regulatory outcomes will be achieved in the CCW. Russia in particular has taken on the role of blocking any developments in recent years - but other states invested in the development of AWS would also be unwilling to allow progress by consensus in this forum. The UK government and others must find constructive ways forward to address the need to regulate AWS and develop strong international norms given this reality.

In a positive development, the UK endorsed a broadly supported joint statement on AWS at the UNGA during 2022. This recognised that autonomy in weapons systems raises serious concerns from humanitarian, legal, security and ethical perspectives; highlighted that a combination of

¹³ For further detail, see Elizabeth Minor (2023), 'Laws for LAWS: Towards a treaty to regulate lethal autonomous weapons,' Friedrich Ebert Stiftung <https://article36.org/wp-content/uploads/2023/02/Laws-for-LAWS.pdf>

prohibitions and regulations provides a way forward; and committed states to strengthen efforts to address the issue. This joint statement was significant in bringing together states with a range of views on whether new law is necessary (including the UK, US, and other NATO states, alongside a diverse range of states calling for a legally binding instrument) that nevertheless share significant common ground on policy content and the need to find a way forward, despite a lack of progress in the CCW.¹⁴

The government should continue to engage with this broad grouping of states that have identified common ground, and with any work around AWS in the UNGA and in any other forums. The UNGA is a forum that includes all states - including those that have expressed concerns on AWS but are not party to the CCW. The UNGA is also able to make progress based on majority approval rather than consensus (which can be used by certain states to mean veto). It has produced treaties on weapons issues before that the UK has participated in. States with military interests in AWS have been well represented in the discussion at the CCW, and their views would continue to be represented in more inclusive discussions, even if some may insist on boycotting such processes.

The UK government should consider setting international norms on AWS through new international law to be a strategic priority. The government must push to set such standards, and to draw clear lines of principle that will provide an international reference point, even without the participation of some of the 'major military powers' that are pursuing increasing autonomy in weapons systems. If it does not do so, international norms will still develop. However, these will risk being shaped solely by the behaviour and practice of AWS users, including adversaries the UK is concerned about, risking a race to the bottom in the pursuit of military advantage. This is not only likely to have unacceptable humanitarian consequences, but is unlikely to be in the UK government's own perceived strategic interests.

Elizabeth Minor
Article 36
April 2023

¹⁴ See <https://article36.org/updates/ground-breaking-un-joint-statement-on-autonomous-weapons/>