Comments on GGE on LAWS rolling text dated 26 July 2024

Contact: Elizabeth Minor elizabeth@article36.org and Richard Moyes richard@article36.org

In this document, prepared by Article 36, we outline our initial thinking on elements of the 'rolling text' issued by the Chair for discussion at the GGE. Our comments are in red on the original text reproduced here.

We are circulating this document informally, with the aim of assisting delegations with the development of their talking points.

The rolling text, as of the start of the GGE, has useful elements, bringing together points that many states have agreed on for some time now. States must now move to negotiate and adopt new law without delay, and these common understandings provide a helpful basis for negotiation on several necessary rules. Unfortunately, based on our past engagement with the CCW, and states' interpretation of its consensus procedures, we expect this text to be weakened over the course of this session; that agreement on it will be unlikely; and that the legally binding instrument the international community needs to agree on autonomous weapons systems will not be concluded in the CCW. Moreover, the current rolling text has critical gaps when it comes to ethical, human rights, and other issues that must be included in a legally binding instrument.

We therefore urge states to use all available opportunities to move forward to negotiate a legally binding instrument in an inclusive forum, which can consider **all** the issues with autonomous weapons systems holistically and comprehensively, and where progress cannot be blocked - noting the possibilities provided by interest in this issue at the UN General Assembly.

Rolling text begins:

With the understanding that nothing has reached consensus until everything has reached consensus, the Group provisionally found consensus on the following formulations for the purpose of advancing its work on a set of elements of an instrument, without prejudging its nature, and other possible measures, in accordance with the Group's mandate.

A working characterization of LAWS
Without prejudging any other options for measures, a lethal autonomous weapon system can be characterized as a weapon system that, after activation, can select and engage a target without further intervention by a human operator.

- We welcome this working characterisation. It reflects the convergence that has been building on what 'autonomous weapons systems' are, broadly aligning with the scope proposed by the International Committee of the Red Cross (ICRC)[1] as well as with the policy position of the Stop Killer Robots campaign,[2] and Article 36's own analysis. This common understanding is a useful basis for proceeding with treaty negotiations.
- Along with many delegations, including the ICRC, we consider the qualifier 'lethal' unnecessary in characterizing and discussing autonomous weapons systems - lethality is an effect rather a characteristic of weapons systems.
- As reflected in this characterization, the key feature of autonomous weapons systems is that the exact object to be struck, and the time and place of the application of force, is determined by sensor processing following human decision-making - rather than directly by a system's users. As the ICRC has put this, an autonomous weapons system: "self-initiates or triggers a strike in response to information from the environment received through sensors and on the basis of a generalized 'target profile'. This means that the user does not choose, or even know, the specific target(s) and the precise timing and/or location of the resulting application(s) of force."[3] It is from this characteristic that legal, ethical, and other challenges arise.

The above description also applies to those weapon systems that require only nominal human input after activation, thus operating without context-appropriate human involvement and judgement.

- We support this inclusion, which highlights the need to avoid bad-faith interpretations or 'loopholes' in our common understanding of autonomous weapons systems. We also note that any weapons system or practice in the use of force that lacks meaningful human control (or "context-appropriate human involvement and judgement") must not be permitted.

It does not apply to autonomous systems that are not weapon systems, manually guided /remote controlled munitions, mines or unexploded explosive ordnance.

- We agree that systems under direct human control, that are not weapons systems, or UXO, do not fall within the scope of autonomous weapons systems. We note that mines do fall within the characterization of autonomous weapon systems above. However, we would emphasize that comprehensive prohibitions on Anti-Personnel Mines are already agreed in the AP Mine Ban Convention, and

[1] See https://www.icrc.org/en/document/icrc-position-autonomous-weapon-systems
[2] See https://www.stopkillerrobots.org/our-policies/
[3] See https://www.icrc.org/en/document/icrc-position-autonomous-weapon-systems

recognise that other regulations and further discussions already exist with respect to this category of weapons.

The above description also does not affect any future understanding and the potential improvement or refinement of the characterization when formulating an international Instrument.

- States should take this characterization as a starting point for negotiations on a legally binding instrument, which they must mandate this year. During negotiations, common understandings can be further refined, and a legal definition negotiated if necessary.

Preliminary considerations

Control with regard to weapon systems is needed to uphold compliance with international law, in particular IHL, including the principles and rules of distinction, precautions and proportionality.

Human judgement is essential in order to ensure that the potential use of weapon systems is in compliance with international law, and in particular IHL.

- We welcome the recognition that human control and judgment are key to ensuring compliance with international law in the use of weapons systems,[4] and that this is a point of convergence amongst states. This is a crucial starting point from which states must negotiate legal regulations to address some of the central concerns with autonomous weapons systems.
- It is not apparent to us that control and judgment are more important to compliance in IHL than other bodies of international law, such as IHRL. We suggest that "in particular" should be replaced with "including," and moreover that states should discuss the importance of control and judgment in complying with other bodies of law in more depth.

[4] Not just their "potential" use

Application of international humanitarian law (IHL) to LAWS
IHL applies to all forms of warfare and all kinds of weapons, those of the past, those of the present and those of the future, and is, consequently, independent from the military technology used. IHL, therefore, continues to apply fully to the development, deployment and use of LAWS. - We agree that IHL applies to autonomous weapons systems in armed conflict, but encourage states to consider in more detail the application of other bodies of law during conflict, as well as the need to address the risks and legal implications of the deployment of autonomous weapons systems outside of armed conflict.

Prohibitions and restrictions
- The first four statements below reflect existing IHL. These are useful to re-state as a basis for formulating the legal prohibitions and regulations of autonomous weapons systems that are necessary to ensure compliance with IHL. LAWS that cannot be used in compliance with IHL, including the principles and rules of distinction, precautions and proportionality, are prohibited. It is prohibited in all circumstances to use LAWS which are designed or of a nature to cause superfluous injury or unnecessary suffering, or which are inherently indiscriminate. It is prohibited in all circumstances to make the civilian population as such, individual civilians or civilian objects the object of attack by LAWS. - We stress that ensuring civilians are not the object of attack by autonomous weapons systems would be highly challenging in the event of the use of autonomous weapons systems that target people. The ICRC has noted that it is "difficult to envisage how anti-personnel AWS could be used lawfully under IHL"[5] because determining combatant status is highly contextual and can change,

[5] See https://www.icrc.org/sites/default/files/document_new/file_list/icrc_position_on_aws_and_background_paper.pdf

and is not amenable to encoding as a 'target profile' in any system. States must consider the concerns relating to autonomous weapons systems targeting people in greater depth, including not only IHL but IHRL implications - and the fact that reducing people to objects to be sensed by a weapons system is a fundamental ethical concern when it comes to autonomous weapons systems, and an affront to human dignity. For us, these challenges can only be met by a categorical prohibition on autonomous weapons systems targeting people.

LAWS must not be deployed or used if their effects in attacks cannot be anticipated and controlled, as required by international humanitarian law in the circumstances of their use.

- We welcome the inclusion from Australia et al's 'Draft articles' of the need to ensure that effects of attacks must be capable of being anticipated and controlled[6]: this provides a rationale for more concrete requirements to limit the duration, geographical scope, and scale of attacks - an important point of convergence included in the text below, which should be enshrined in a legal instrument.

LAWS that operate without appropriate control or human judgement are prohibited.

- For us, it is a red line that systems that do not allow for meaningful human control should be prohibited, and that legal regulations are agreed so that other autonomous weapon systems are only used with meaningful human control in practice.

To ensure that LAWS can be used in compliance with IHL, their effects must be adequately predictable, reliable, traceable and explainable.

- These important characteristics are requirements for the characteristics of systems, in order to ensure that the effects of attacks can be anticipated and controlled

For this purpose, States must:

Ensure that LAWS operate with appropriate control and human judgment across the entire life cycle of the weapon systems.

- We again welcome the recognition of the necessity of control and judgment (which must be meaningful). Use is the part of the 'lifecycle' at which weapons systems will be operated, and meaningful human control and judgement is most critical. At other stages, for example in design, it must be ensured that it will be possible to operate systems with this type of control.

[6] See UN Document CCW/GGE.1/2023/WP.4, Article 1

Limit the types of targets that the system can engage.

- We highlight here again the need to prohibit systems targeting people as part of this. We also note that limiting the types of targets a particular system can engage could reduce the range of unintended objects that fall within its 'target profile' (for example, civilian vehicles that may have similar characteristics to the military vehicles that are the intended target) - which can contribute to more effective control over a system's effects.

Limit the duration, geographical scope, and scale of the operation of LAWS, including through ensuring that LAWS can be deactivated by a human operator after activation and/or that they incorporate self-destruct, self-deactivation or self-neutralization mechanisms.

Limit the number of engagements that LAWS can undertake.

- We welcome the recognition of convergence on the factors outlined in the two points above, which should be key aspects of legal regulation to ensure meaningful human control. We also welcome that possibilities for deactivation etc are recognised as mechanisms that could contribute to limiting the duration, geographical scope and scale of operation of autonomous weapons systems - and caution against such mechanisms being seen as a solution in themselves.

Ensure that LAWS' mission parameters cannot be modified by the system without appropriate control and human judgement.

- This would be one aspect of ensuring that systems that do not allow for meaningful human control are prohibited.

Ensure appropriate training and instructions for human operators of LAWS.

- Ensuring appropriate training and instruction should be considered as one aspect of a general requirement for users to have a sufficient functional understanding of the systems they are using and the context in which they are deploying them (and so what will trigger an application of force by a system). That systems are "adequately predictable, reliable, traceable and explainable," as outlined above, is also needed for functional understanding by users. States should consider this understanding to be one of the overarching requirements in negotiating legal rules to ensure meaningful human control.

Other measures to ensure compliance with IHL
Rigorous testing and evaluation should be conducted to enable a human operator to have a reliable expectation of how the weapon system will perform in the anticipated circumstances of its use. - As above, this should be considered an aspect that can contribute to a general requirement for users to have a sufficient functional understanding of their systems - and this should relate to concrete rather than abstract circumstances of use States should conduct legal reviews of LAWS to understand the weapon's capabilities and limitations, expected circumstances of use, and its anticipated effects in different circumstances. States should conduct reviews to detect possible unwanted bias in data sets. - When it comes to eliminating data bias that could result in discriminatory outcomes in the use of force - for example because of the use of data sets that inevitably reflect the racism, sexism, ableism and other forms of discrimination present in our societies, which are then reproduced in target detection and the application of force, resulting in disproportionate harm to marginalized people - we stress that the key measure must be a categorical prohibition on autonomous weapons systems targeting people. This would go a significant way towards effectively addressing some of the central ethical and human rights problems with autonomous weapons systems. Attempting to detect and mitigate bias in data sets would not be a strong or effective enough response to these challenges: we cannot allow people to be killed or injured through data bias. States should implement measures to reduce unwanted automation bias.

Accountability

Humans must at all times remain accountable in accordance with applicable international law for decisions on the use of force.

- States should be clear that rules on human control and judgment facilitate accountability. In order for accountability to be applied fairly, it must be possible for users to be held meaningfully responsible for the effects of the systems they are deploying - ensuring meaningful human control is essential for this.

Human responsibility for decisions on the use of weapon systems must be retained since accountability cannot be transferred to machines. This should be considered across the entire life cycle of the weapon system.

Accountability for the design, development and use of autonomous weapon systems must be ensured in accordance with applicable international law, including through the operation of such systems within a responsible chain of human command and control.

States must ensure internal mechanisms for the reporting of incidents that may involve violations of IHL. These internal mechanisms should also address incidents related to the operation of LAWS.

States must take all appropriate steps, including legislative and other measures, to prevent and suppress violations [of the agreed measures/international law] by persons or on territory under its jurisdiction or control.

- When negotiating a legally binding instrument, it would be expected to include an article on national implementation and the suppression of violations.