Revised working paper

Submitted by Austria

1. This paper builds upon the discussions that have taken place in the framework of the Group of Governmental Experts on emerging technologies in the area of lethal autonomous weapons systems established under the Convention on Certain Conventional Weapons. A growing convergence of views has emerged among many participating states on the issue of regulating autonomous weapons systems following a two-tier structure. Such an approach would be technology-neutral and functional, with some weapons systems requiring prohibitions and others regulations.

2. Against this background, the aim of this working paper is to highlight what might be considered as the key elements from which to build meaningful human control over autonomous weapon systems and the use of force as well as to protect human dignity.

   (a) To enable compliance with existing international law, including international humanitarian law, those authorizing any use for a weapon system that integrates autonomy in its critical functions of selecting and applying force to a target, must:

      i. Have an adequate functional understanding of the system under consideration. They must understand what circumstances or conditions will trigger an application of force by the system, including conditions that would trigger an unintended engagement. Moreover, they must be able to trace back the outcome of the use of force to human agency.

      ii. Adequately assess the context in which the system is to be used. This includes evaluating the circumstances or conditions that would trigger an application of force in relation to the given environment and the time and duration of its use.

      iii. Limit the duration and geographical area of the system’s functioning, and the number of engagements that a system can undertake, to the extent necessary to enable them to make informed judgements about the anticipated outcomes of the use of force in accordance with legal obligations. Those authorizing any use must be in a position to readjust, interrupt or deactivate a system if continued functioning would place it outside the context for which such informed legal judgements have been made.
Autonomous weapons systems that cannot meet all three of these conditions would risk causing effects that cannot be adequately explained, predicted or sufficiently controlled. Such systems are unacceptable, and so systems that cannot meet these conditions should be prohibited.

(b) Autonomous weapons systems that select and engage persons as targets in a manner that violates the dignity and worth of the human person as well as the principles of humanity or the dictates of public conscience are unacceptable and must be prohibited.

3. Further international legal regulation is necessary to ensure the above-mentioned conditions that are fundamental to protecting the integrity of international law, in particular international humanitarian law, in the context of autonomy in weapons systems.

4. It is also a requirement throughout research, development, acquisition and use of autonomous weapons systems to constantly review and reassess any possible changes and modifications in the system’s functioning with regard to fulfilment of the conditions listed above. This should include technical aspects such as ‘machine learning’ and any datasets upon which system functions are based. This monitoring process should be embedded in an adequate multi-layered international regulatory framework that would entail regular review of the implementation of prohibitions and positive obligations in order to ensure meaningful human control is preserved over the use of force, and legal and moral rules and ethical principles are protected in the design, development and use of autonomous weapon systems.