

Joint Working Paper

Submitted by the Republic of Costa Rica, the Republic of Panama, the Republic of Peru, the Republic of the Philippines, the Republic of Sierra Leone and the Eastern Republic of Uruguay

The Convention on the Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed To Be Excessively Injurious Or To Have Indiscriminate Effects (CCW) is a pillar of global efforts to limit warfare in order to protect humanity from indiscriminate destruction and superfluous suffering. At its heart is the shared recognition of our common humanity and of the inherent dignity of all human lives, and of their protection as a legal, ethical, and moral imperative.

Over the years, building on the principles of customary international law and International Humanitarian Law, the Convention has addressed or mitigated threats from technological advances in weaponry. This has led to the adoption and implementation of its various Protocols that prohibit or restrict various conventional weapons.

Today, we are seeing rapid advances in technology, including in the area of artificial intelligence. Used correctly, these advances could propel us towards sustainable and peaceful development, and therefore improve the lives of many. At the same time, they could also be used to develop autonomous weapon systems that threaten to displace humans from their traditional role in decision-making with regard to the use of force. Sufficiently addressing this predicament will uphold the spirit that brought the High Contracting Parties (HCPs) to come together and enact the Convention and render the conventional disarmament architecture future-proof.

Within the framework of the Convention, the Group of Government Experts (GGE) on Emerging Technologies in the Area of Lethal Autonomous Weapons System (LAWS) has been working for the past eight years to explore and agree on possible recommendations on options related to emerging technologies in the area of LAWS, in the context of the objectives and purposes of the Convention, taking into account all proposals – past, present and future – as stated in the mandate given by the Convention’s Fifth Review Conference and the work developed prior to 2016.

The work performed by the GGE thus far, including the national commentaries on the eleven Guiding Principles, have provided a useful foundation for the HCPs to identify potential gaps in existing international law, as it applies to emerging technologies in the area of autonomous weapon systems. There is already a wealth of knowledge – developed and compiled by the GGE and by other stakeholders – to determine specific proposals for a future regulatory framework. Nevertheless, the Guiding Principles shall not be seen as an end in themselves. They also do not constitute an exhaustive response to the ethical, legal and moral concerns presented by autonomous weapon systems.

In 2019, Meeting of High Contracting Parties (MHCP) to the CCW provided the GGE with the mandate to submit to the Sixth Review Conference consensus recommendations in relation to the clarification, consideration and development of aspects of the normative and operational framework on emerging technologies in the area of LAWS.

The GGE should consider the dual nature of this mandate. Its normative aspect includes discussing further granularities with regard to general principles concerning, *inter alia*, autonomy, human control, human-machine interaction, and the applicability of international law, including International Humanitarian Law, International Human Rights Law, and International Criminal Law, as well as ethical and moral considerations. The operative aspect, on the other hand, includes discussions on practical measures to strengthen these principles, and to determine and fulfill gaps *vis-à-vis* new technological developments with military applications.

Our delegations believe that the GGE already has sufficient material to fulfill this mandate, and to therefore produce substantive recommendations for a normative and operational framework. In this context, and without prejudice to our respective national positions, we are submitting this working paper to support the GGE's work and to offer key elements for its consideration.

I. Elements of a normative framework

With less than a year to present recommendations to the Sixth Review Conference, our delegations are of the view that stronger focus should be given to the following principles, which should be the elements of a normative framework on LAWS:

1. Real or hypothetical weapon systems or configurations within the scope of our discussions that are **fully autonomous** are unacceptable and must be prohibited under international law. Such systems target, engage, and apply force, *inter alia* in deciding on the life or death of human beings, without meaningful human control.
2. **Human control** shall be meaningfully maintained over the use of all other autonomous weapons systems in order to uphold legal obligations and to address ethical and moral considerations.
3. Unlimited **autonomy** in weapon systems shall not be pursued or permitted, as this possibility will be incompatible with international law, including International Humanitarian Law, International Human Rights Law, International Criminal Law, and with the respect for human dignity.
4. The use of **target profiles** to identify and use force against humans should also be prohibited because it is legally and morally unacceptable.
5. Weapon systems are not neutral. Algorithm-based programming relies on data sets that can perpetuate or amplify **social biases**, including gender and racial bias, and thus have implications for compliance with international law.
6. Prohibitions and regulations (**positive obligations**) shall be designed to ensure meaningful human control over the use of force (**human-machine interaction**), as only humans can be held legally accountable and responsible for the use of force under international law. Only humans have agency and legal and moral obligations. The use of autonomous weapon systems to target human beings should be ruled out through a prohibition on autonomous weapon systems that are designed or used to apply force against persons. The design and use of autonomous weapon systems that would not be prohibited should be regulated and limited.
7. The **scope of systems under consideration should be broad** and include systems that rely on sensor inputs to identify and engage targets, following a system activation,

emplacement or deployment. In this case, the exact time, place and object to which force will be applied will not be known in advance, thereby generating uncertainty.

8. Sufficient levels of **predictability, foreseeability, reliability, oversight, and explainability** of weapon systems as well as **spatial and temporal constraints** are needed to enable operators to exercise meaningful human control, ensure **legal compliance** and avoid technical vulnerabilities. Further work is needed to determine the type and extent of human involvement and control necessary to ensure compliance with international law and to respond to ethical and moral concerns. Unpredictable autonomous weapon systems should be expressly ruled out, notably because of their indiscriminate effects. In other words, a prohibition on autonomous weapon systems that are designed or used in a manner such that their effects cannot be sufficiently understood, predicted and explained should be prohibited under international law.
9. States have recognized that **meaningful human control** shall be applied throughout the life cycle of the weapon system and to all of its **critical functions**. It would be important to consider the inclusion of recording systems and other technology that would permit effective investigation of, and determination of accountability for alleged wrongful uses of force, and/or misuse of these weapon systems.
10. It is a matter of concern that there is uncertainty about which States are developing and/or acquiring autonomous weapons systems. States should be transparent regarding all aspects of the development of their weapons systems, as well as their **processes for reviewing new weapons**. States should be called upon by the international community and domestic actors to exercise **transparency** in this regard.
11. **Pre-emptive national measures** are needed to prevent developments that may be difficult to reverse later. In parallel, an open-ended and inclusive process to further address this issue should take place at the international level.
12. States already endeavor to exercise responsible behavior in terms of governing the development of technologies, including advanced weapon systems, but such responsible behavior is absent in **armed non-State actors**. Therefore, efforts must be undertaken to prevent the acquisition and proliferation of advanced weaponries by armed non-State actors.
13. States, especially developing countries, retain the **right to pursue development through advanced technologies**.

II. Elements of an operational framework

Further unpacking these normative elements would benefit from a more focused and streamlined discussion in the GGE, building towards a comprehensive and robust normative and operational framework in the area of emerging technologies in LAWS.

Our delegations identify the following elements of an operational framework on LAWS:

1. **Characterization and limitations.** There should be a recognition of acceptable and non-acceptable weapons and weapons systems. In this regard, HCPs should characterize what constitutes LAWS, making it possible to affirm that these weapon systems are by nature unacceptable – and differentiate them from semi-autonomous, automatic, and automated systems. Such differentiation should take into account the weapon or weapon system's life

cycle and the degree of human-machine interaction involved in them. On the basis of these discussions, the determination of the characteristics of current or future weapons that should be prohibited or regulated could be made. For instance, the International Committee of the Red Cross (ICRC) has characterized these weapon systems as those that **select and apply force against targets without human control**. The autonomous weapon 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” (technical indicators function as a generalized proxy for a target). The weapon system fires itself when triggered by an object or person, at a time and place that is not specifically known, nor chosen, by the user.

As recommended by the ICRC, the use of autonomous weapon systems to target human beings should be ruled out through a prohibition on autonomous weapon systems that are designed or used to apply force against persons. On the other hand, the design and use of autonomous weapon systems that would not be prohibited should be regulated, including through a combination of (1) **limits on the types of target**, such as constraining them to objects that are military objectives by nature; (2) **limits on the duration, geographical scope and scale of use**, including to enable human judgement and control in relation to a specific attack; (3) **limits on situations of use**, such as constraining them to situations where civilians or civilian objects are not present; and (4) **requirements for human-machine interaction**, notably to ensure effective human supervision, and timely intervention and deactivation.

2. **Accountability.** This element shall summarize current convergences with regard to ensuring that accountability with regard to the use of weapons or weapon systems remain firmly resting on persons, including individuals and parties to conflict. This element should also include stipulations on the accountability and culpability of individuals, states and corporations in a potentially criminal act using LAWS and in regard to relevant prohibitions and regulations.
3. **Transparency.** This element shall address current uncertainties regarding development or possible acquisition by State actors of advanced weapons or weapon systems that could fall under certain characterizations that may require either prohibition or regulation. Transparency should include all aspects of development of these weapons or weapon systems across their entire life cycles, as well as national processes for reviewing them in accordance with the Marten’s clause.
4. **Exchange of good practices.** This element shall explore possible mechanisms for HCPs to exchange views, information, guidance, and insights on their domestic operationalization of these normative elements, including best practices in undertaking weapons review under Marten’s clause, among others.
5. **Public safeguards.** This element shall explore possible positive obligations with regard to ensuring responsible State behavior in the aspects of developing, utilizing, and deploying advanced weapons and weapon systems. These include means to uphold obligations under International Humanitarian Law, including mainstreaming them in doctrines and trainings, and measures on risk assessment and mitigation in the design, development, testing, and deployment cycles of emerging technologies in weapons system.
6. **Industry safeguards.** This element shall explore means to enhance physical and non-physical security of advanced technologies that could potentially be weaponized, including cyber-security against hacking or data spoofing, with the view to reducing the risk of acquisition and proliferation of advanced weapons and/or weapon systems by armed non-

State actors. This could also enhance peaceful uses of advanced technologies, including artificial intelligence. The current Group of Experts mechanism under the Second Amended Protocol (AP II) and the Fifth Protocol (PV) of the CCW could be a useful model in this regard.

III. Way Forward: Beyond the Sixth Review Conference

Our delegations welcome the continued call of the Secretary-General of the United Nations and other world leaders and advocates for a global prohibition and internationally agreed limits on autonomous weapon systems, and to move expeditiously to address humanitarian, legal, ethical, moral, human security and other concerns over the implications of developments in this area.

A **legally-binding instrument** would strengthen the existing framework of international law. Therefore, anything short of this, including a political declaration or voluntary applicable guidance, can only be acceptable as an intermediary and/or complementary step towards a legally-binding instrument.

Civilian protection is eroded by **systems that target people**, or those that cannot be effectively controlled. Remoteness and autonomy risk lowering thresholds against the use of force. **Autonomy** invites more autonomy in response, which could result in arms race, proliferation, as well as illegal transfers and diversion of these systems. Moreover, autonomy in weapons and weapon systems may increase existing inequalities and further erode international peace and security. These threats may be exacerbated by cyberattacks and/or acquisition of these weapons and weapon systems by non-State actors.

To prevent non-compliance with international law, unpredictable technologies should not be authorized to make legal decisions, including those concerning the use of force, nor should a human operator make these decisions with no real understanding of, and proximity to the context of an attack.

The Sixth Review Conference will provide States with the space and opportunity to consider the substantive work of the Group in the past years, and undertake a decision for a negotiating mandate.

Our delegations are of the view that HCPs in the GGE need to:

1. Recommend a new mandate for the GGE to initiate immediately an open-ended process to negotiate a legally binding instrument.
2. Propose that the elaboration of this new legal norm includes specific prohibitions and regulations (i.e. positive obligations) for the design, development, production, deployment, and use of autonomous weapon systems.
3. Indicate that States will determine the key aspects, including scope, core obligations, framework for regular meetings, and other aspects, over the course of the formal negotiations.