

**Sixth 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**

20 December 2021

English only

Geneva, 13-17 December 2021

Item 12 of the agenda

Review of the scope and operation of the Convention and its annexed protocols

**Translating Ethical Concerns into a Normative
and Operational Framework for Lethal
Autonomous Weapons Systems**

Submitted by Holy See**I. Introduction**

1. A stark common denominator throughout the history of humankind has been the tendency to research, develop and use technology for offensive and defensive purposes. Regrettably, this has often relegated ethical considerations to a secondary role. The development of regulations to protect from and mitigate the negative consequences of hostilities has always come too late. The grim experiences of chemical, biological and nuclear weapons, as well as antipersonnel landmines and cluster munitions, are just a few dramatic examples.

2. The research on emerging technologies in the area of the so-called Lethal Autonomous Weapons Systems (LAWS), including the weaponization of artificial intelligence, does not seem to escape this logic. The development of LAWS will provide the capacity of altering the nature of warfare, raising legal and ethical implications.

3. In the view of the Holy See, the challenges are not limited only to the realm of international humanitarian law (IHL) and international human rights law. As the development of LAWS would compel all States to reassess their military capabilities, there are also serious implications for peace and stability. For these reasons, after eight years of specific discussions on this issue within the CCW framework, the Holy See deems it urgent that the CCW adopts an ambitious forward-looking approach through the commencement of negotiations of a legally binding instrument to address the issues raised by LAWS.

4. To this end, mindful of the valuable work carried out by the Group of Governmental Experts (GGE) throughout the years, the negotiations will be crucial in agreeing on important elements, including characteristics, the different relevant levels of autonomy encompassed, the scope of operation (defensive/offensive) and specific regulations and prohibitions.

5. The purpose of this contribution is twofold: first, to present some of the ethical and legal concerns that the Holy See has raised since the beginning of the discussion on LAWS; and second, to submit to other High Contracting Parties certain elements for a possible way forward.

6. While appreciating how new technologies could ensure greater adequacy with certain requirements of humanity (e.g. by avoiding or correcting what, during a given action, could prove to be dangerous for human beings), the Holy See deems it of the utmost importance to give such ethical considerations a more prominent role and to retain the fundamental reference to the dignity of the human person at the heart of the discussion on LAWS.



7. It should also be noted that ethical principles do not hinder the research, development and use of technologies. However, ethics can orient technologies towards an ultimate horizon which is not based merely on the criteria of utility or efficiency, but on furthering the common good of humanity. AB Pope Francis has suggested: "we can once more broaden our vision. We have the freedom needed to limit and direct technology; we can put it at the service of another type of progress, one which is healthier, more human, more social, more integral!"¹

II. Ethical Concerns

8. An autonomous weapons system can never be a morally responsible subject. The unique human capacity for moral judgment and ethical decision-making is more than a complex collection of algorithms, and such a capacity cannot be reduced to programming a machine, which as "intelligent" as it may be, remains an object. A machine can execute instructions and rules, but it is a mistake to say that it can "decide" or "judge". Performing an action (or omitting to do so) represents a process intended to be conscious and controlled. In particular, certain aspects of judgment and of decision-making need interpretation and "prudence" (sagacity). Furthermore, the application of rules and principles requires an understanding of the contexts and specific situations which cannot be subsumed under universal formal rules, however articulated they may be, and which cannot be pre-established a priori into algorithms. In this regard, autonomous weapons systems could mistakenly consider normal, in the statistical sense of the term, and thus acceptable, those actions that international law prohibits, or that, albeit not explicitly proscribed by current IHL, remain forbidden by the dictates of public conscience and by ethics.

9. Autonomous weapons systems, equipped with self-learning or self-programmable capabilities, necessarily give way to a certain level of unpredictability, which could, for instance, "deviate" into actions targeting non-combatants in order to maximize their destructive efficiency, thus flouting the principle of distinction. If functioning without any direct human supervision, such systems could also make mistakes in identifying the intended targets due to some unidentified "bias" induced by their "self-learning capabilities". The concept of a swarm of autonomous weapons further aggravates this risk since the stochastic nature of the swarm could lead to excessive injuries and indiscriminate effects, in stark contradiction with IHL.

10. Completely relinquishing to machines the decision over the spatial and temporal application of lethal force removes, or at least obfuscates, the moral burden intrinsically associated with military operations. LAWS could create a qualitatively new problem, by implementing an unintended shortcut that maximally satisfies the goals programmed into it, with disregard for common ethical considerations and without any understanding of the principle that the end does not justify the means. As a consequence, the inherent dignity of the human person is reduced to meaningless and interchangeable data ("digital reductionism"). Indeed, a machine, which is a thing, cannot truly think, feel, decide or be accountable for its "action", because an autonomous machine has no real body or (substantial) history and no real relations; machines do not have the essential capabilities to truly judge, think, will and act, thus precluding their moral agency. Machines can only simulate human behavior. Here, it is vital to stress an important epistemological point: reality, as such, can never be reduced to a "representation", a "simulation" of reality. To blur the distinction between a being and its representation and its modelling is an epistemological error that will inevitably lead to serious moral problems. For these reasons, adequate, meaningful and consistent human supervision will always be necessary.

11. A classic foundation of legal systems is the recognition of the human person as a morally responsible subject that could be sanctioned for her/his wrongdoing and be obliged to provide redress for the damage caused. This notion of responsibility originates from the profound reality of the human person as a free and rational being. Removing human agency from the moral equation as its fundamental point of reference is problematic not only from the point of view of ethics, but also from the point of view of the foundation of law, including IHL, the application of which entails a fundamental reference to interpretation, good faith

¹ Pope Francis, *Laudato si'* n. 112.

and prudential judgement. This recognition is implied in the use of IHL terminology such as "anticipated", "may be expected", "superfluous injury", "unnecessary suffering" etc.

12. If important decision-making powers over the use of force are delegated to a weapons system whose behavior is unpredictable or whose scope of operation is not well-defined or known (as in the case where the autonomous weapons system is equipped with self-learning capabilities), the crucial action/responsibility nexus would be inevitably jeopardized. This would result in a glaring contradiction and thus undermine the essential legal foundations on which international institutions are based, including the prevention of crimes and the prosecution of those who commit them. It is only humans who are able to appreciate the results of their actions and understand the connections between cause and effect. Ultimately, respecting this legal/ethical starting point is a way to ensure respect for the meaning of political authority and its human content and meaning.

13. The Martens clause - which is at the intersection of IHL and ethics, but more importantly represents a legal obligation enshrined in Additional Protocol I to the Geneva Conventions - could already offer a priori a crucial compass for the regulation and prohibition of LAWS. In this regard, the growing awareness of these issues, also among prominent scientists, engineers, researchers, military, ethicists and the larger civil society community, represents a change in the public perception, which is also a driving force behind the implementation and enforcement of IHL. These concerns go well beyond respect for IHL and the CCW, and they attest to the far-reaching implications and urgency of the discussion.

III. The Way Forward - a normative and operational framework

14. The crucial challenge, then, is how to translate concretely these ethical concerns into a practical solid outcome. In the view of the Holy See, it is imperative to ensure adequate, meaningful and consistent human supervision over weapon systems: it is only humans who are able to see the results of their actions and understand the connections between cause and effect. This would not be the case with LAWS which could never "understand" the meaning of their actions.

15. Adequate human supervision means that the human being introduces sufficient conditions, in the management of weapons systems, to preserve the aforementioned ethical principles and to ensure compliance with IHL.

16. Meaningful human supervision implies that, ultimately, there is always the reference to the human person that must guide the research, development, and use of weapons systems, even in the absence of specific legal regulations, as implied by the "Martens Clause".

17. Consistent human supervision entails that at no time the weapons systems would have the capacity to contradict what the human authority has prescribed as the main purpose or result of its intervention. It would be, in fact, inconsistent that a weapon system deployed to fulfill a particular mission begin to adopt behavior inconsistent with its prescribed purposes.

18. From this point of view, it seems that the mentioned principles imply the requirements of predictability and reliability of autonomous systems. Indeed, even if LAWS appear to manifest degrees of freedom of "behavior", they must be required in all circumstances to achieve the prescribed ends and implement the required intentions and objectives of the responsible authority.

19. If particular behaviors cannot be controlled at all times, all possible behaviors must always be a priori circumscribed. From this point of view, it would therefore be impossible to accept systems capable of "learning" completely new behaviors.

20. As was stressed before, the Holy See deems that the challenges presented by LAWS are not limited only to the realm of IHL and IHRL. They also raise serious implications for peace and stability. In his address to the 75th Session of the UN General Assembly, Pope Francis warned that "we are witnessing an erosion of multilateralism, which is all the more serious in light of the development of new forms of military technology, such as lethal autonomous weapons systems which irreversibly alter the nature of warfare, detaching it further from human agency".²

21. The Holy See believes that to prevent an arms race and the increase of inequalities and instability, there is an imperative and urgent duty to translate the wealth of knowledge produced in the sessions of the GGE throughout the years into a concrete normative and operational framework, rooted in ethical considerations. Now is the time to prevent LAWS from becoming the reality of tomorrow's warfare. The CCW should make a courageous decision that could hopefully lead toward the prohibition of lethal autonomous weapons, like it did in the past concerning other types of weapons.

22. As affirmed in the preamble of the CCW, the High Contracting Parties recognized "the importance of pursuing every effort which may contribute to progress towards general and complete disarmament under strict and effective international control" and reaffirmed "the need to continue the codification and progressive development of the rules of international law applicable in armed conflict".

23. This Review Conference presents an opportunity to take ambitious decisions. It would be appropriate to consider enforcing a moratorium on the development and use of LAWS pending the negotiations of a legally binding instrument to address the challenges raised by LAWS.

24. Besides convening these negotiations, it could be desirable and beneficial for States, in the long term and in a broader scope than LAWS and the CCW per se, to consider establishing an International Organization for Artificial Intelligence, to facilitate, and have the right of all States to participate in, the fullest possible exchange of scientific and technological information for peaceful uses and for the promotion of the common good and integral human development.

25. Amidst the global pandemic, it is important to place emerging technologies at the service of humanity. The research on emerging technologies should be oriented towards combating the real challenges that affect the international community. This fight will not be won by developing autonomous weapons systems, but by placing technology at the service of the human person and by orienting it towards the common good. Ultimately, this will be a choice well beyond the scope of the CCW and the discussion on LAWS. Nevertheless, it is a choice that must be made by, and will have consequences on, humanity as a whole.
