Do Not Ban "Killer Robots"!

Didier Danet*

Military Academy of Saint-Cyr, didier.danet@st-cyr.terre-net.defense.gouv.fr

Abstract—The ban of so called "Killer Robots" is strongly prescribed by certains States and numerous influential NGOs. Following the report of the Informal Group of Experts created by the Convention on Certain Conventional Weapons, a Governmental Group of Experts is due to examine this issue and formulate recommendations by the end of 2017. In this article, the current state of discussions in Geneva is summarized and the reasons not to ban R&D programs related to hypothetical "Lethal Autonomous Weapon Systems" are put forward.

Robotics and Automation; Autonomous Systems; Military Equipment; Military Computing; Weapons; Artificial intelligence; Legal Factors

I. Introduction

At the end of the Conference on Certain Conventional Weapons (CCW), in December 2016, an important decision was taken concerning the future of military robotics. Following the conclusions of the Informal Group of experts established in 2014, the Meeting of States Parties decided to implement a Group of Governmental Experts (GGE) responsible for analyzing key questions raised by a yet unknown type of weapon system: Lethal Autonomous Weapon System or LAWS. According to the French presentation during the meeting of the Informal Group of Experts in april 2015, LAWS can be defined by the combination of three cumulative conditions [1]:

- A carrier/vector able to move freely on land, in the air or in marine environments not totally known in advance;
- Conducting the targeting and firing of a lethal effector (bullets, missiles, bombs);
- Operating in total autonomy without any human intervention (Human out-of-the-loop).

Even if dramatic improvements are noticeable in military robotics R&D programs, no existing weapon today meets these conditions. However, many institutions, [2] experts [3-5] and NGOs are concerned by their possible emergence mostly due to the spectacular progress achieved in artificial intelligence.

As of today, should we pre-emptively ban LAWS as was the case with the banning of blinding laser systems [6] and other types of weapons not meeting the requirements of International Humanitarian Law and those of military ethics? This is the question posed to the GGE which it must answer by the end of 2017. In this paper, we present the status of negotiations, before submitting two arguments compelling us

not to endorse such a ban because it would rapidly turn out to be both inadequate and ineffective.

II. STATE OF DISCUSSIONS ON LAWS

Discussions on Lethal Autonomous Weapons Systems were dispersedly propagated by protagonists of civil society, more especially traditional human rights organisations (*Human Rights Watch*¹ or *International Committee of the Red Cross*² in particular) and other ad hoc coalitions (notably the *Campaign to stop Killer robots*³ or *Article 36*⁴ among others).

A. First Stage of Intensive Lobbying.

Debates and discussions on "Killer Robots" really started with the report entitled "Losing Humanity: The Case Against Killer Robots".[7] In this solid report, key arguments are outlined for a pre-emptive ban of LAWS. Briefly summarized. the authors argue that the proliferation of robots on the battlefield makes them think that the current trend of replacing humans by machines will, in the next 20 to 30 years, lead us to delegate the selection and engagement of targets to machines without human intervention during an engagement. Such weapons systems would not comply with the regulatory demands of International Humanitarian Law. They would develop added risks of causing greater civilian losses during armed conflicts. Because these military robots would be unable to meet the requirements of fundamental principles such as discrimination or proportionnality, the development and implementation of these weapons systems (including dedicated or related R&D Programs) should be pre-emptively banned.

Similar legal arguments were repeated and further developed by certain non-governmental organisations as was the case with the *Campaign to Stop Killer Robots*. The logic behind such opinions was very similar except, perhaps, for the greater focus put on moral and ethical issues: the lack of human intervention in decision-making would strip it from an ethical dimension and would thereby not comply with the requirements of International Humanitarian Law. The following extract from an article of the *Campaign to Stop Killer Robots* illustrates this: "Allowing life or death decisions to be made by machines crosses a fundamental moral line.

¹ https://www.hrw.org/topic/arms/killer-robots

² <u>https://www.icrc.org/en/document/statement-icrc-lethal-autonomous-weapons-systems</u>

³ https://www.stopkillerrobots.org/

⁴ http://www.article36.org/issue/autonomous-weapons/

Autonomous robots would lack human judgment and the ability to understand context. These qualities are necessary to make complex ethical choices on a dynamic battlefield, to distinguish adequately between soldiers and civilians, and to evaluate the proportionality of an attack. As a result, fully autonomous weapons would not meet the requirements of the laws of war." [4, 8]

Beyond the specific case of LAWS, many researchers or NGOs raised further concerns on the issue of robotised battlefields in general, including all kinds of military robots: conjectural autonomous systems as well as existing teleoperated drones or automated artillery batteries. [9] According to these experts, the development of any of these military robots has given rise to the same concerns: (i) lowering the triggering threshold of conflicts. Replacing soldiers by robots would lead States possessing them to obviate the political costs linked with human casualties on the battlefield. [10] This could, in turn, encourage them to resort to a regular use of military force in resolving disputes; (ii) a dilution of responsibilities when civilian populations are harmed. In machine-led operations, it would be more difficult to determine who is liable for collateral damages. In fact, who should be incriminated? The designer or the manufacturer of the robot, the officer deploying it in the field or the operator controlling it? Victims would hence find themselves in serious difficulties to seek compensation for the damages suffered. [11]

Such claims received an overwhelming support during the summer of 2015. Famous researchers (Stuart Russell, Stephen Hawking or Yan LeCun) and major industrialists (Steve Wozniak or Elon Musk) all signed an Open Letter calling for a ban of LAWS. [12] Their underlying message is for the greatest part the same as that of HRW or the Campaign to Stop Killer Robots. In the Open Letter lies the idea that LAWS will emerge soon (less than 30 years), that such systems will lower the triggering threshold of conflicts and, what's even more alarming, their costs will fall dramatically offering many belligerents a way to access this frightening technology and triggering an arms race and an uncontrolled proliferation of LAWS. These autonomous "Killer Robots" will become the "Kalashnikovs" of the XXI century, strengthening dictators and fostering terrorist organisations. The conclusion of the Letter is as follows: "Starting a military AI arms race is a bad idea, and should be prevented by a ban on offensive autonomous weapons beyond meaningful human control."6

B. Progressive Institutionalisation

On 30th May 2013, the United Nations Human Rights Council convened in Geneva to examine the report prepared by Christof Heyns, UN Special Rapporteur on extrajudicial, summary or arbitrary executions for the Office of the High Commissioner for Human Rights. [13] The latter raises concerns on the potential development of what he names

⁵ http://www.stopkillerrobots.org/the-problem/

Lethal Autonomous Robotics (LARs) recommending a moratorium on the ongoing research in this field. He states further: "Beyond this, their deployment may be unacceptable because no adequate system of legal accountability can be devised, and because robots should not have the power of life and death over human beings. The Special Rapporteur recommends that States establish national moratoria on aspects of LARs".

Among the large number of nations attending the conference, three concerns were raised. The first one was raised by Pakistan, recommending a ban on the development of LAWS. By contrast, the United Kingdom opposed any attempt on banning or establishing moratoria on such weapons. Between the two extremes, most countries suggested to engage into further discussions: France and Brazil recommending to do so within the framework of the Convention on Certain Conventional Weapons. It was the solution adopted as it had the advantage of offering a space for open debates in a non binding, recognised and favourable setting for all the parties concerned, including multiple stakeholders, especially States and NGOs.

At the CCW, several informal meetings of experts were held in 2014, 2015 and 2016.⁷ Attendants at such conferences were then able to analyse in greater detail the technological developments that have given rise to LAWS. Several panels of experts seeked a definition of what LAWS represented. They addressed legal and ethical issues and evaluated how the legal framework could provide appropriate answers on such systems. All in all, ongoing discussions led to an agreement on some common principles:

- As with all weapon systems, the development and implementation of weapons defined as autonomous systems must in all cases be conducted in compliance with principles of the International Humanitarian Law. Each State should therefore make sure that such systems fully comply with the principles of distinction and discrimination or the proportionate use of force among others;
- The human presence in decision-making and resorting to the use of force is a crucial issue that needs to be tackled in the future;
- The intervention of civil society organisations (such as NGOs) should be maintained for future debates and discussions;
- The choice of the CCW is an appropriate setting for further discussions.

During the Fifth Review Conference of the High Contracting Parties to the Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons which convened in 2016, it was decided to accept the proposals made by the Informal Meetings of Experts to establish a Group of Governmental Experts (GGE). The Group will be responsible for preparing recommendations that will be

http://www.unog.ch/80256EE600585943/(httpPages)/37D51189AC4FB6E1C1257F4D004CAFB2?OpenDocument

⁶ Refer to the following webpages: http://futureoflife.org/open-letter-autonomous-weapons/

submitted to the Meeting of the High Contracting Parties, planned to convene at the end of 2017. The Work Plan assigned to the CGE is wide-ranging as it covers the following: a definition of LAWS complying with International Humanitarian Law, the political and legal responsibilities of States, moral and ethical issues, the impact on security and on regional and global stability, the risks of an arms race and arms proliferation risks...

The work accomplished by the group of experts and the discussions between stakeholders of the CCW have led to some clarifications on their respective concerns.

The initial promoters of a pre-emptive ban⁸ have now been re-joined by several African 9 and South American countries. 10 More importantly, China, a permanent member of the U.N. Security Council, raised some uncertainties for International Humanitarian Law to address the challenges posed by LAWS. The Chinese declaration raised three major issues: respect for the principles of discrimination, respect for the principles of proportionality and determining the responsibility of actors violating International Humanitarian Law. China is consequently in favour of elaborating legally binding instruments based on the CCW protocol of 1995 prohibiting the use and transfer of blinding laser weapons. The following statement illustrates this point: "China supports the development of a legally binding protocol on issues related to the use of LAWS, similar to the Protocol on Blinding Laser Weapons, to fill the legal gap in this regard." 11 For the supporters of the ban, China's stance in this respect is all that more important in that it is the only permanent member of the Security Council to have adopted it.

A renewed activism by NGOs can hence be expected this year, before the annual meeting of High Contracting Parties is held in November 2017.

III. WHY A PRE-EMPTIVE BAN ON LAWS WOULD BE WRONG

A. Categorisation and Terminological Confusion on LAWS

As the general aim of opponents to "Killer Robots" is to obtain the prohibition of certain types of weapon, the minimal requirement expected from them is to precisely know what is the definite object of the prohibition. In other terms: which kind of weapon systems should be on the list of prohibited ones and which ones should not? What are the distinctive characteristics of an "Autonomous" Weapon System? How

⁸ Pakistan, Ecuador, Cuba, Egypt, Holy See

to determine the border between "Autonomous" and "Non autonomous" Weapons Systems? This is both a matter of legality (the rule of law must be clear) and effectiveness of the law (a confusing rule gives rise to interpretations and litigation) Considering the multiple definitions of the term "autonomous" given by the opponents to "Killer Robots", a satisfying and legally usable definition is obviously not on the agenda.

It's firstly important to stress that LAWS do not exist at this time; they are purely conjectural. This point is not discussed among States and NGOs and it is even the basic assumption presented in the report through which the Informal Group of Experts was created in 2013. With the work progressing, it is now clear that a sharp distinction has to be made between, on the one hand, teleoperated systems (directly managed by a team of usually numerous operators) and automated systems (where the operator relies on the machine to perform more or less basic data processing tasks and/or boot to full capacity) and, on the other hand, autonomous systems (where the operator is out-of-the-loop for any decision-making). The problems raised by these two kinds of machines are completely different and should not be considered within the same framework. Unlike the two first categories (teleoperated and automated systems), autonomous systems such as LAWS do not exist and many obstacles stand in their way. One of the most famous companies, Microsoft, had to withdraw its "chatbot" Tay just one day after its implementation because malicious individuals had perverted the deep learning processes in order to transform Tay in a nazi, racist and machist speaker. The coming of an autonomous machine which would be able to decide and act like a soldier on the battlefield is yet hypothetical and, in spite of the scientists' progresses, can not be extrapolated. This means that a pre-emptive ban on weapon systems would be decided whereas their nature, components, and characteristics are still unknown. There is hence a striking difference with these systems and blinding laser systems, the example given by China to support its claim. It can therefore be admitted that one cannot compare the formal prohibition of current weapon systems where their nature, characteristics or dangerous effects have been fully explained with an object that does not exist and which major characteristics cannot be described.

Secondly, the perimeter of "Killer Robots" to be banned, widely differs among those demanding a ban of these systems, especially but not only when it comes to the definition of "Autonomy". The Open Letter, although signed by thousands of specialists, has been the source of much confusion in this respect. Not only does the Letter refer to the notion "Beyond meaningful control" [3] which has been categorically rejected by the Chinese Delegation's declaration of December 2016¹²,

⁹ Algeria

Algeria

Chile, Costa Rica, Mexico, Nicaragua (April 2016)
Argentina, Guatemala, Panama, Peru, Venezuela (December 2016)

Position paper submitted by the Chinese delegation to the CCW 5th Review Conference, Geneva, 12/2016. Webpage: http://www.unog.ch/80256EDD006B8954/(httpAssets)/DD1551E60648CEBBC125808A005954FA/\$file/China's+Position+Paper.pdf

[&]quot;Discussions about definitions should focus on several core issues: 1. Levels of autonomy and criteria for their determination; 2. Relations and distinctions between automation, autonomy and remote control; 3. The mode of human involvement and the human role which requires a strict

but it introduces a supplementary distinguishing criterion which had never been mentioned before: the ban would only concern "offensive" weapons. In other words, this means that a distinction should be made between offensive and defensive weapons, the latter obviating all bans and/or prohibitions. This crucial but undecidable distinction would lead to many divergent interpretations and engender endless debates or litigations concerning forbidden or authorized programs, quiclky jeopardizing the whole legal provisions.

Thirdly, the partisans of the pre-emptive ban regularly confuse teleoperated systems with autonomous systems. This mix of genres obviously appeared in July 2016, when the Dallas Police used a teleoperated land robot to neutralise an armed individual who had already killed five policemen, refused to surrender and continued shooting at the police surrounding him. 13 The web site of the Campaign to Stop Killer Robots¹⁴ made it quite clear that the robot used in this case did not fall into the category of LAWS, as the robot in question was a teleoperated system under human control. A further report stated that the decision taken to detonate an explosive charge by the robot was made by officers in charge of the operation. In other words, the use of this robot in no way illustrates and clarifies the debate on LAWS. Yet the conclusions reached by the NGO are precisely calling for a ban of LAWS: "The Campaign to Stop Killer Robots calls for a pre-emptive ban on fully autonomous weapon systems... The campaign does not seek to prohibit armed drones or remote-controlled vehicles and the use of the term "killer robot" is inappropriate for this incident." Following the same reasoning, it's odd to give so much exposure to an incident which was completely irrelevant with the issue and only created greater confusion on the claims this NGO put forward concerning the ban.

B. Technological Change and the Evolution of Law

Furthermore, beyond the absence of clear definition of the weapon systems which are supposed to be ban, there is a second reason for rejecting a premptive-ban plea: nothing proves this type of weapon would not comply with the requirements of International Humanitarian Law and deprive victims in seeking compensation for damages suffered. [14, 15]

The above claim has been particularly well-researched in the report "Losing Humanity". Having considered the potential progress artificial intelligence can bring, it concludes: "But even with such compliance mechanisms, fully

definition and cannot be replaced by such vague concepts as 'human judgement' or 'meaningful human control'. »

autonomous weapons would lack the human qualities necessary to meet the rules of International Humanitarian Law. These rules can be complex and entail subjective decision making, and their observance often requires human judgment. For example, distinguishing between a fearful civilian and a threatening enemy combatant requires a soldier to understand the intentions behind a human's actions, something a robot could not do." In other words, a robot would by essence be unable to draw a clear distinction in interpreting the intentions of a potentially hostile individual. There is cause for objecting this hypothesis as distinguishing between a fearful individual and an enemy combatant is based more on concrete behaviour than on likely intentions. It would moreover be unrealistic to think that the decision taken by a soldier on the battlefield is based on a deep interpretation to understand the intentions of a supposed enemy. Such a soldier is under intensive stress and is not prepared for some serious psychological analysis as making use of his/her weapon is practically always a reflex action rather than dragging his/her feet in procrastination.

Furthermore, the claim made that the use of force by means of these robots runs counter to legal principles is rather puzzling. The partisans supporting the pre-emptive ban are in fact galvanised by two contradictory concerns. On one hand, some of them (the authors of the Open Letter for example) are concerned by the possible emergence of military robots with a frightening potential due to the spectacular developments of artificial intelligence systems. 15 In this veine, the fear is that super-intelligent machines could free themselves from human control, take death or life decisions on their own and engage fire against alleged ennemies. The risk would be, for humanity, to be deprived of control over too sophisticated machines. The moto of the preemptive ban could be "Too intelligent not to be banned". But, for other at the opposite, whatever the development of artificial intelligence, military robots will always be unable to get basic distinguishing ability which a "simple" soldier is always supposed to master and can apply with success despite a somewhat limited knowledge of legal matters. ¹⁶ The moto would then be "Too stupid not to be banned". Too intelligent or not? Either machines of the future will be endowed with sensors and particularly sophisticated processing capabilities, with the potential, no one should doubt, to "learn" the fundamental principles of International Humanitarian Law and even comply with these principles with some discernment. Or, they will not. From all the above, it can hence be admitted that machines of the future will never reach such a level of sophistication and will never become autonomous, the human operator always remaining in-the-loop or on-the-loop for decision-making and for controlling

719

" **A** .

autonomous-weapons/

¹³ For example: Sam Thielman, Use of police robot to kill Dallas shooting suspect believed to be the first in US history, The Guardian, 8 juillet 2016,

https://www.theguardian.com/technology/2016/jul/08/policebomb-robot-explosive-killed-suspect-dallas

https://www.stopkillerrobots.org/2016/07/explosives-used-on-us-police-robot-to-kill/

[&]quot;Autonomous weapons have been described as the third revolution in warfare, after gunpowder and nuclear arms." Webpage: http://futureoflife.org/open-letter-

^{16 &}quot;As a result, fully autonomous weapons would not meet the requirements of the laws of war." Webpage: http://www.stopkillerrobots.org/the-problem/

operations. The preemptive ban is therefore supported by two opposite reasonings which will make difficult to get a general agreement on such matters as dual R&D programs, offensive or defensive weapon systems...

Concerning the dilution of responsibilities, would civilian populations be harmed by LAWS, the concerns of the promoters of the pre-emptive ban are also highly divergent. Arguably, the complex operations led on modern battlefields where there are multiple actors involved, have already haighlighted awareness on the difficulties in defining individual and/or collective responsibilities. The growing sophistication of machines and the decision-making by quasiautomated devices could give the impression that such difficulties will develop further. This may be true. Nonetheless, will this deprive victims from their rights? Nothing is so sure for, even if conjectural LAWS were to emerge and proliferate on battlefields, the fact remains that any alleged errors would be pinned on at least two actors: (i) Either the fault lies with a bad design of the machine (e.g. weaknesses in the learning-process on the rules and regulations of International Humanitarian Law) and it is then the robot's manufacturer - or the designer of its faulty component - who bears the blame; (ii) Or, if the fault was made due to the bad use of the machine, it's the officer in charge of leading military operations who bears the blame for any damages caused. Excepting a far-fetched possibility for "singularity", humans will always spearhead the design and development of machines, which rules out the assumption of the absence of accountability of weapon systems.

IV. CONCLUSION

From the crossbow to the atomic bomb and nowadays to LAWS, advances in military science and in armaments technologies have historically always instigated various forms of antagonism and fear due to the destructive power of new military inventions, lower thresholds in armed conflicts or the damages caused to civilian populations. Those very same reasons animate the promoters requesting a preemptive ban of LAWS. Contesting the relevance of this preemptive ban, obviously does not equate to an appeal for researchers and defense companies to develop and produce weapon systems that would be free to kill whomever they want, based on some mysterious algorithm that would get rid of any human control. On the contrary, innovative weapon systems need to be regulated effectively. But, this fundamental principle, on which every stakeholders agree, cannot be enforced through poor legal provisions. This would be the case if a preemptive ban were to be adopted for hypothetical weapon systems which precise nature, characteristics and effects are largerly unknown and if a faint or blurred border between authorized and forbidden R&D programs was engendering so many interpretations and litigations that the legal framework would be widely and quickly infringed. Contesting a special preemptive ban for the category of hypothetical weapon

systems not only prevents the implementation of an inefficient legal framework but offers the crucial advantage to maintain the unity of this legal framework which is already applying to each and every weapon systems, even when they are still in the early phases of the R&D programs. Doing so seems to be the best solution to conciliate innovation in weapon systems, protection of combattants and populations, compliance by the rules of IHL and military ethics, accountability of men in charge and, more generally, control of the use of force on the battlefield.

V. REFERENCES

- 2 UNIDIR: 'The Weaponization of Increasingly Autonomous technologies: Addressing Competing Narratives (Phase II)', UNIDIR Research Project, 2017, pp. http://www.unidir.org/programmes/emerging-security-issues/the-weaponization-of-increasingly-autonomous-technologies-addressing-competing-narratives-phase-ii
- 3 Asaro, P.: 'On banning autonomous weapon systems: human rights, automation, and the dehumanization of lethal decision-making."', International Review of the Red Cross 2012, 94, (886), pp. 687-709
- 4 Roff, H.M.: 'The strategic robot problem: Lethal autonomous weapons in war.', Journal of Military Ethics, 2014, 13, (3), pp. 211-227
- 5 Sharkey, N.E.: 'The evitability of autonomous robot warfare..', International Review of the Red Cross 2012, 94, (886), pp. 787-799
- 6 Carnahan, B.M., and Marjorie Robertson: 'The Protocol on' Blinding Laser Weapons': A New Direction for International Humanitarian Law.', The American Journal of International Law 90.3 (1996): 484-490., 1996, 90, (3), pp. 484-490
- 7 Docherty, B.: 'Losing Humanity. The Case Against Killer Robots' (Human Right Watch & IHRC, 2012. 2012)
- 8 Sharkey, N.E.: 'Saying 'no!'to lethal autonomous targeting', Journal of Military Ethics 2010, 9, (4), pp. 369-383
- 9 Chamayou, G.: 'Théorie du drone' (La Fabrique Editions, 2013. 2013)
- 10 Howlader, D., Giordano, and James: 'Advanced Robotics: Changing the Nature of War and Thresholds and Tolerance for Conflict-Implications for Research and Policy.', The Journal of Philosophy, Science & Law,, 2013, 13, (2), pp. 1-19
- Asaro, P.M.: 'Robots and responsibility from a legal perspective.', Proceedings of the IEEE, 2007, pp. 20-24.
- 12 Gibbs, S.: 'Musk, Wozniak and Hawking urge ban on warfare AI and autonomous weapons', The Guardian, 2015, july 27th
- Heyns, C.: 'Report of the Special Rapporteur on extrajudicial, summary or arbitrary executions, Christof Heyns, A/HRC/23/47', UN Human Rights Council 2013, Promotion and protection of all human rights, civil, political, economic, social and cultural rights, including the right to development, 23 session, Agenda Item 3, pp. http://www.unog.ch/80256EDD80006B88954/(httpAssets)/80684AB80253F83935B80255C80242C81257CC200429C200427C/\$file/Report+of+the+Special+Rapporteur+on+extrajudicial.pdf
- Anderson, K.e.W., Matthew C.: 'Law and ethics for autonomous weapon systems: Why a ban won't work and how the laws of war can', in Editor (Ed.)^(Eds.): 'Book Law and ethics for autonomous weapon systems: Why a ban won't work and how the laws of war can' (Columbia University, 2013, edn.), pp. 33
- 15 Schmitt, M.N.: 'Autonomous weapon systems and international humanitarian law: a reply to the critics.', Harvard National Security Journal