Commentary Paper
N°89
December 2022

Killer Robots: The Ethical Challenges of Lethal Autonomous Weapon Systems (LAWS) and the Third Revolution in Warfare

Manon Grafé

Former CECRI intern, she holds a Master's degree in International Relations from the ULB and a complementary diploma in Management from the ICHEC Brussels School of Management.

Academics call it the third revolution in warfare. Over time, revolutions in warfare have always represented a big step forward in the efficiency and the speed of killing. However, while the first and the second revolution were related to the invention of gunpowder and nuclear weapons, the third revolution in warfare is about lethal autonomous weapons systems (LAWS) [1]. Although this invention may remind you of Terminator, it is not science fiction. These weapons already exist and are developed by several states.

Nevertheless, what are autonomous weapons? As the nickname suggests, killer robots are weapons that can identify targets and open fire on them with the help of artificial intelligence in a fully autonomous manner. However, the international community commonly accepts no precise definition of this concept. This ambiguity is at the core of divergent interests and represents a big stake in significant global economic, political, and military issues.

Lethal Autonomous Weapons Systems

Innovation regarding Artificial Intelligence (AI) technology is a fast innovation process. The idea of a deadly AI system is not new, according to the pace of today's technological advances in the field of military robotization [3]. Massive investments in AI are already made by major powers such as China, Russia, and the United States (US) [4]. The fact that AI is now firmly integrated into military strategies reveals a pressing need to develop appropriate governance structures for managing the risks of its military uses [5]. The quickness of this development has raised various theoretical concerns regarding the morality of LAWS and has led the issue to an international debate.

A dynamic of forces rules the debates. On the one hand, supporters of the development of autonomous weapons, such as arms industries and military powers, argue that the military benefits of these new weapons will dominate security risks. Their argumentation is mainly led by the fact that autonomous weapons are intelligence learning machines that can 'improve the goals and tasks of (...) military operations at almost all levels' [6]. Indeed, the operational and economic benefits of LAWS – such as the reduction of costs and staff and the increasing speed of decision-making in military operations – are regularly claimed by LAWS proponents.

On the other hand, some organizations oppose the very idea of LAWS and call for a preventive ban on them because of the serious ethical concerns they raise. Since 2012, the campaign *Stop the Killer Robots* driven by a coalition of non-governmental organizations, represents the leading actor of the opponent

camp [7]. In this respect, this paper discusses the ethical challenges of this new technology by identifying the arguments of each party.

The autonomy and the human judgement

The first issue at the debate's core is about the weapons' autonomy. There is no shared definition of autonomous weapons because states cannot agree on the level of human control required by the weapon (human in/on/off the loop) [8]. The United Nations supports the idea that these weapons should be subjected to meaningful human control, meaning that a human should consider information regarding the target, the context, and the likely effects of the strike before opening fire [9]. However, the concept is still very vague while the stake remains significant. If the category of LAWS is broadly delimited or defined comprehensively, the definition could also cover some conventional weapons. It would restrict the military freedom of powers such as China, Russia, or the United States. By contrast, if the definition is lightly delimited, it opens the pandora box to all risks and abuses. However, the challenge of defining it is complex because of the speculative nature of the technological evolution of those weapons [10]. Indeed, the particularity of LAWS is that there is no shared experience or understanding. Not or imprecisely defining LAWS in advance carries the risk of confusing them with existing weapons [11].

Moreover, the opponents of LAWS claim the Martens clause, according to which being killed by a machine violates moral values such as human dignity. In this sense, delegating the opening fire to a robot would be unacceptable. Using LAWS could decrease the value of human life in the way a machine would decide to kill itself. There would be a physical and an emotional distance between the human programmer and the robot, which could 'generate an indifference or even a'Gameboy Mentality' on the side of the former' [12]. However, supporters of LAWS argue that human beings are not always more respectful toward human dignity than machines. They often take the example of the bombing of Hiroshima and Nagasaki in this respect [13].

The legality of LAWS and the respect of international humanitarian law (IHL)

The legality and the respect of international humanitarian law (IHL) are at the core of the second moral issue. Indeed, the respect of IHL remains for many an indispensable condition for the implementation and use of autonomous weapons. Again, the critics of LAWS highlight the difficulty of evaluating the future capacities of autonomous weapons. Subsequently, it is hard to consider that machines would be adequately qualified to respect the fundamental principles of IHL. For example, according to the IHL principle of precaution and distinction, weapons that fire indiscriminately are illegal. Indiscriminate weapons are weapons that do not distinguish between military and civilian targets and cause unnecessary damage. For LAWS, the respect of such principles would be converted into algorithms, but there is no assurance that datasets would not bias these algorithms [14]. Besides, contemporary wars show us that it is increasingly difficult to distinguish a civilian form from a combatant form. Indeed, applying the distinction principle is contextual and depends on complex features such as direct participation in the hostilities and the 'hors de combat' function [15]. Therefore, converting this IHL principle into an IT parameter is hard. Furthermore, the proportionality principle – according to which collateral damages should be limited depending on military advantages [16] – is also complex to convert into algorithms because of its contextual features. Overall, these situations need to be seen through a case-by-case analysis, depending on each situation's specific geopolitical and strategic context. In this sense, human analytical skills seem to meet these requirements more than computers [17].

On the other hand, supporters of LAWS highlight the humanitarian benefits of these technologies and argue that machines can respect IHL rather than humans who have already committed various war crimes. They state that if the robot's behavior is indistinguishable from human behavior in a given context, there would be a moral duty to develop them [18]. It refers to the principle of unnecessary risk according to which, in an armed conflict, we

should minimize the risks faced by combatants if we can do so. Developing autonomous systems that have similar or better capabilities than humans would compel us to deploy them. Supporters of autonomous weapons also argue that these machines would better respect IHL because they cannot feel emotions such as fear, revenge, stress, sadness, or hate. Subsequently, they would be more objective in the decision-making [19]. They would commit fewer crimes than humans who violate IHL by following their basic instincts. Those weapons would not have any reason to dissimulate information or violations committed during military operations. They would be more transparent. These machines would not be 'bound to the shifting subjectivities of the mind' [20].

In this sense, the question of legal liability is interesting. The uncertain level of autonomy granted to a LAWS implies that the responsibility for a LAWS attack is difficult to identify. Would the human programmer of the weapon be responsible for the attack? Or would the weapon itself be? Would it be the builder of the LAWS? Or the state conducting the operation? These questions are at the core of the debate.

The adverse effects of LAWS

The third question discussed in the LAWS debate concerns the potential adverse effects of this technology.

Firstly, as technology develops, the autonomy of LAWS will simultaneously increase. In the long run, it could imply an interaction of multiple LAWS and the creation of 'self-organizing swarms' [21]. The risk of swarms represents 'the possible use of a very large number of coordinated LAWS to saturate the opponent' [22]. Swarms of this size would be considered weapons of mass destruction (WMD) regarding the military power they could engage. The risk is that no human or any machine could be able to face this kind of threat. The fact that these weapons are not predictable is dangerous because predictability represents the basis for efficacity and security in the military field [23]. Among the adverse effects of LAWS, there is also the risk of malfunctioning of LAWS.

Secondly, when we talk about technological evolution, it is often related to cybersecurity issues. Indeed, such weapons might be subject to cyber hacking and spoofing [24]. For example, LAWS could be hijacked to carry out terrorist attacks.

Thirdly, the use of LAWS could lower the threshold for conflict: 'reducing human costs, they would provide an incentive to go to war' [25]. Under the pretext that machines could replace humans in the field, governments could decide to send LAWS on the ground without going through a democratic process. Furthermore, the arms race toward LAWS would involve using non-conventional weapons by weak states to defend themselves against such technological weapons, which are illegal under international law.

Fourthly, these powerful weapons could also fall into the hands of non-state actors or small groups as terrorists that would be able 'to unleash massive levels of destruction and kill in great numbers' [26]. These technologies might become propaganda tools by authoritarian regimes or other violent extremists. Indeed, authoritarian governments would be able to target any activist or critic, undermining any democratic effort inside the State. It could lead to a world in 'which civil discourse, liberal values, and the fundamental institutions of democracy could face relentless attacks' [27]. However, advocates of LAWS – echoing Grotius – respond that just because some people abuse this technology does not mean it should be banned. They consider that abusive use of LAWS does not represent a sufficient condition for a preventive ban [28].
Finally, and beyond all these risks, there are still those that the experts have not yet identified. Again, the unpredictability of LAWS remains the main threat.

Conclusion

Recent years have seen significant military power investment in this type of weaponry. The United States has already allocated a budget of 18 billion US$ for autonomous weapons between 2016 and 2020 [29], partnering with significant defense industries such as Lockheed Martin, Boeing, Raytheon, and Northrop Grumman [30]. The opacity and the uncertainty of the technological innovations of military powers have raised many concerns among

world political leaders in recent years.

Several events have also made the news in recent years, bringing the debate to the forefront internationally. Since 2014, sentry robots have been installed in the Korean Demilitarised Zone (DMZ) on the border with North Korea to replace human soldiers in the South Korean army. These robots can autonomously identify and destroy targets, functioning as a 'human on the loop' system (see above) [31]. Moreover, the debate has taken a significant turn since March 2021 when the first autonomous drone attack in Libya may have been identified, according to a report by the U.N. Panel of Experts on Libya [32]. Militia fighters pursued by the Tripoli government forces 'were hunted down and remotely engaged by the unmanned combat aerial vehicles or the lethal autonomous weapons systems', according to the report [33].

An important meeting in Geneva discussing the limits on using lethal autonomous weapons was held in December 2021 within United Nations Convention on Certain Conventional Weapons framework. In the past, the objective of this convention has been to restrict some of the world's cruelest conventional weapons such as land mines, booby traps, and incendiary weapons. However, no agreement has been reached by the UN community, due to the strong opposition from States such as India, Russia, and the US [34]. The meeting was the culmination of several years of research into different approaches to mitigating the threats of LAWS. For the critics of LAWS, the discussion in Geneva was a great disappointment [35].
In this respect, the current unpredictable Ukrainian-Russian conflict is unlikely to improve the debate on this subject.

The risk caused by the inertia of the debate is that technological advances will overtake political discussions. As the autonomous arms race is already underway and is not about to end, the need for legislation on the development and use of LAWS is more pressing than ever.

Notes

[1] Toby Walsh, 'Regulating the Third Revolution in Warfare - AIIA', Australian Institute of International Affairs, 18 October 2018, https://www.internationalaffairs.org.au/australianoutlook/2062-the-world-that-ai-made/.

[2] Nevertheless, France has already given a restrictive definition that could help the reader to understand about what this paper is about: 'LAWS are mobile systems, capable of adapting to their environment on land, sea or in the air, and to the behaviour of the agents around them, and of selecting a target and firing a lethal munition autonomously, i.e. without any human supervision or validation (...) Equipped with a self-learning capacity in an evolving environment, their behaviour is not totally predictable'(free translation from French), quoted in Jean-Baptiste Jeangène Vilmer, 'Diplomatie Des Armes Autonomes: Les Débats de Genève', Politique Étrangère 81, no. 3 (2016): 121.

[3] Eric James Beyer, 'Elusive Ethics: Robotic Warfare and Autonomous Weapons', media, Interesting Engineering (blog), 16 March 2021, https://interestingengineering.com/ethics-of-robotic-warfare-and-autonomous-weapons.

[4] Ibidem.

[5] Darrell M. West and Jack Karsten, 'It's Time to Start Thinking about Governance of Autonomous Weapons', Brookings, 10 May 2019, https://www.brookings.edu/blog/techtank/2019/05/10/its-time-to-start-thinking-about-governance-of-autonomous-weapons/.

[6] Ibidem.

[7] 'The Campaign To Stop Killer Robots', Stop Killer Robots, accessed 26 March 2021, https://www.stopkillerrobots.org/learn/.

[8] For a definition, see Jean-Baptiste Jeangène Vilmer, 'Terminator Ethics: faut-il interdire les «robots tueurs»?', Politique etrangere Hiver, no. 4 (8 December 2014): 153-154.

[9] Ibid., 156–157.

[10] Ibid., 153.

[11] Jean-Baptiste Jeangène Vilmer, 'Diplomatie Des Armes Autonomes: Les Débats de Genève', loc. cit., 121.

[12] Regina Surber, 'Artificial Intelligence: Autonomous Technology (AT), Lethal Autonomous Weapons Systems (LAWS) and Peace Time Threats', ICT for Peace Foundation, 2018, 10.

[13] Jean-Baptiste Jeangène Vilmer, 'Terminator Ethics: faut-il interdire les «robots tueurs»?', loc.cit., 159.

[14] Peter Asaro, 'Algorithms of Violence: Critical Social Perspectives on Autonomous Weapons', Social Research: An International Quarterly 86, no. 2 (2019): 538.

[15] 'Rule 47. Attacking persons who are recognized as hors de combat is prohibited. A person hors de combat is:
(a) anyone who is in the power of an adverse party;
(b) anyone who is defenceless because of unconsciousness, shipwreck, wounds or sickness; or
(c) anyone who clearly expresses an intention to surrender;
provided he or she abstains from any hostile act and does not attempt to escape.'
ICRC, Rule 47 of Customary IHL Database, https://ihl-databases.icrc.org/customary-ihl/eng/docs/v1_rul_rule14, accessed on 24 August 2021.

[16] 'Rule 14. Launching an attack which may be expected to cause incidental loss of civilian life, injury to civilians, damage to civilian objects, or a combination thereof, which would be excessive in relation to the concrete and direct military advantage anticipated, is prohibited.' ICRC, Rule 14 of Customary IHL Database, https://ihl-databases.icrc.org/customary-ihl/eng/docs/v1_rul_rule14, accessed on 24 August 2021.

[17] Jean-Baptiste Jeangène Vilmer, 'Terminator Ethics: faut-il interdire les «robots tueurs»?', loc.cit., 161.

[18] Ibidem.

[19] Regina Surber, ibid., 10.

[20] Eric James Beyer, ibidem.

[21] Regina Surber, loc. cit., 10.

[22] Jean-Baptiste Jeangène Vilmer, 'Diplomatie Des Armes Autonomes: Les Débats de Genève', loc. cit., 128 (free translation from French).

[23] Ibid., 127.

[24] Peter Asaro, loc. cit., 552.

[25] Jean-Baptiste Jeangène Vilmer, 'Terminator Ethics: faut-il interdire les «robots tueurs»?', loc. cit., 159.

[26] Peter Asaro, loc. cit., 552.

[27] Ibid., 553.

[28] Jean-Baptiste Jeangène Vilmer, 'Diplomatie Des Armes Autonomes: Les Débats de Genève', loc. cit., 124.

[29] The Editors, 'Don't Let Robots Pull the Trigger', Scientific American, 1 March 2019, sec. Policy, https://doi.org/10.1038/scientificamerican0319-6.

[30] Adam Satariano, Nick Cumming-Bruce, and Rick Gladstone, 'Killer Robots Aren't Science Fiction. A Push to Ban Them Is Growing.', The New York Times, 17 December 2021, sec. World, https://www.nytimes.com/2021/12/17/world/robot-drone-ban.html.

[31] Alexander Velez-Green, 'The Foreign Policy Essay: The South Korean Sentry—A "Killer Robot" to Prevent War', Lawfare, 1 March 2015, https://www.lawfareblog.com/foreign-policy-essay-south-korean-sentry—-killer-robot-prevent-war.

[32] Emma Farge, 'U.N. Talks Adjourn without Deal to Regulate "Killer Robots"', Reuters, 17 December 2021, sec. World News, https://www.reuters.com/article/us-un-disarmament-idAFKBN2IW1UJ.

[33] 'Letter dated 8 March 2021 from the Panel of Experts on Libya Established pursuant to Resolution 1973 (2011) addressed to the President of the Security Council', final report (New York: UN. Panel of Experts Established pursuant to Security Council Resolution 1973 (2011), 8 March 2021), https://digitallibrary.un.org/record/3905159 quoted in, Maria Cramer, 'A.I. Drone May Have Acted on Its Own in Attacking Fighters, U.N. Says', The New York Times, 3 June 2021, sec. World, https://www.nytimes.com/2021/06/03/world/africa/libya-drone.html.

[34] James Dawes, 'An Autonomous Robot May Have Already Killed People – Here's How the Weapons Could Be More Destabilizing than Nukes', The Conversation, accessed 4 January 2022, http://theconversation.com/an-autonomous-robot-may-have-already-killed-people-heres-how-the-weapons-could-be-more-destabilizing-than-nukes-168049.

[35] Adam Satariano, Nick Cumming-Bruce, and Rick Gladstone, 'Killer Robots Aren't Science Fiction. A Push to Ban Them Is Growing.', loc. cit.