NOTE FOR NATIONAL DEFENCE:
Ethics of Lethal Autonomous Weapons Systems

Authors: Kari Zaccharias[1,2] and Ketra Schmitt[3]

[1] Research Associate, Gina Cody School of Engineering, Concordia University, Montreal, Canada
[2] Assistant Professor, Centre for Engineering Professional Practice and Engineering Education, University of Manitoba, Winnipeg, Canada
[3] Associate Professor, Centre for Engineering in Society, Concordia University, Montreal, Canada

SUMMARY

- Lethal autonomous weapons systems (LAWS) have the potential to escalate conflicts, blur legal and ethical chains of accountability, and violate international humanitarian law.
- Although ethical arguments are made in favor of LAWS, the ethical drawbacks of these systems are serious and unlikely to be remedied by technological improvements.
- Canada has rejected fully autonomous LAWS, but continues to consider "human in the loop" systems.
- Important questions exist for policymakers about how to build international consensus on LAWS, and how to codify existing consensus in law.

Definitions and Degrees of Autonomy

Lethal autonomous weapons systems (LAWS) rely on sensors and algorithmic processes for decision making, in some cases eliminating the requirement for real-time human intervention in setting objectives, targeting, and executing attacks. There are few fully autonomous weapons systems currently in use; however, the ongoing development of lethal autonomous and semi-autonomous weapons systems, and fully autonomous systems in other fields, demands consideration of the ethical issues that such technologies present.

No universal definition of "autonomous" exists. Autonomous and semi-autonomous systems are usually categorized with respect to the degree of human intervention required for their operation [1]. "Human in the loop," or semi-autonomous weapons systems, require a human operator to identify targets or target groups, and cannot function without human intervention. Autonomous, "human on the loop" weapons systems do not require direct human intervention, but do allow a human operator to monitor the weapon's target identification and engagement, and to stop it if necessary. Fully autonomous weapons systems - "human out of the loop" systems - function entirely independently of human intervention, once activated.

Ethical Issues

(In)stability

Supporters of LAWS frequently point to the military advantages that could be gained by using the capacities of LAWS as a force multiplier [2]. Use of autonomous weapons systems could allow for greater efficacy on the battlefield, significant cost savings, access to previously inaccessible terrain, and reduction of human suffering and loss of life through an increased reliance on non-human warfighters. In turn, critics of LAWS cite the potential for autonomous systems to introduce instabilities that may outweigh these advantages and render the use of LAWS unethical. Two (connected) dimensions of potential instability exist: military instability, through increased proliferation and arms races, and crisis instability, or the accelerated escalation from peace to war, or from lower to higher levels of violence during an existing conflict [3], [4].

In the first case, the dual-use nature of autonomous systems - that is, their development in both military and civilian environments - renders the evolution of LAWS difficult to regulate. The fact that universities and commercial labs are engaged in developing autonomous systems differentiates LAWS from other types of weapons, and may make destabilizing arms race more likely [3]. The potential for a destabilizing arms race was the main concern raised by the international community of AI and robotics researchers in their 2015 open letter supporting a ban on "offensive autonomous weapons beyond meaningful human control" [5]. In the second case, the risk comes from the actions, and especially the interactions, of fully autonomous systems. In some other applications where automated systems are used, e.g. price-setting algorithms and high-frequency trading, the rapid, uncontrolled interactions between different parts of the system have resulted in unpredictable and sometimes disastrous escalation [3]. Fully autonomous weapons systems, with no opportunity for human oversight, have the potential to rapidly escalate and destabilize conflicts. Across both dimensions of instability, critics of LAWS argue that the systems' destabilizing potential outweighs the more narrow stabilizing advantages that might occur within the bounds of specific engagements [6].

Accountability

LAWS blur the chain of accountability that traditionally accompanies military decision making. In fully autonomous, human out of the loop systems, the responsibility for selecting a particular target cannot be traced back to a single human being. Instead, choices are made by a complex network of human and technological actors, including the technological system itself, the human engineers and programmers who created it, the humans who chose to deploy the system, other technologies or systems with which the weapons system interacts, etc. [7], [8]. Similar questions about liability and decision-making play out on a smaller scale with other autonomous and semi-autonomous systems: when a driverless car violates traffic laws, it is unclear who bears the responsibility. LAWS amplify these ethical concerns due to their heightened risk [2]. A flawed decision taken by a fully autonomous weapons system, which may result in the unjustified loss of human life, would have no clear path to accountability.

Complexity and Uncertainty

Many of the ethical debates around LAWS are predicated on the assumption that fully autonomous systems will eventually out-perform humans and human-in-the-loop systems with respect to targeting and engagement, and could theoretically be deployed in ways that would correspond to

international humanitarian law. However, this belief is controversial. Determining whether or not an object is an appropriate target is highly complex and context-dependent. For LAWS, such a determination demands not only reliable sensor inputs and high processing speeds, but also knowledge of the situational, geographical, relational, and temporal context of the scenario [7]. Numerous analyses point out the technical difficulties that autonomous systems would face in distinguishing between civilians and combatants, and the danger presented by a technological system's lack of situational awareness [9]. The complexity of the decisions that would be delegated to LAWS call the feasibility of such systems into question.

The required complexity of any fully autonomous system also increases the likelihood of accidents. Horowitz points to normal accidents theory and discussions of AI safety in other contexts to argue that LAWS may be more prone to operational risk than semi-autonomous weapon systems [4]. In addition, multiple fully autonomous systems deployed in the same arena have the potential to generate conflicting goals. Both the complexity of the tasks and the complexity of LAWS themselves pose potential threats to safety and operability.

Asymmetry and Mutual Risk
Moral questions relating to the principles of just war are also relevant to the potential development and deployment of LAWS. Proponents of LAWS have argued that hypothetical automated systems may perform more ethically on the battlefield than human soldiers. Based on these assumptions, some posit there may be a moral responsibility to use LAWS to reduce instances of unethical battlefield conduct [10]. Similar arguments point to the potential asymmetry of combat involving LAWS as a possible moral benefit, positing that when one combatant is "ethically justified in their effort, and the other is not," a technological imbalance between combatants is not an ethical concern [11]. However, critics of such arguments highlight the necessity of considering multiple aspects of just war principles, notably the ethical considerations that apply to the decision to engage (jus ad bellum) in addition to those that apply to conduct during war (jus in bello) [12]. Critics of LAWS argue that the degree of asymmetry posed by LAWS is unprecedented, and that the resulting lack of mutual risk and reciprocity between combatants would undercut the existing ethics of war. The extension of a battlefield into new spheres, sometimes evoked as a benefit of LAWS, also has the potential to change jus ad bellum considerations that would normally play a role in any decision to engage. If asymmetry between combatants leads to disruptions in the moral judgements that lead to engagement, then the potential for destruction and the risk to non-combatants outweigh the possible ethical benefits of LAWS [13], [14].

Human Dignity
Many current efforts opposing LAWS, including the international coalition behind the Campaign to Stop Killer Robots, base their arguments on the issues outlined above, most notably questions about feasibility and accountability. Recent research critiques this approach as excluding basic moral arguments about human dignity [15]. Arguments that focus on human dignity propose that, even if future automated systems could be made to abide by international humanitarian law and rules of engagement, developing and using LAWS would still be morally unacceptable. Such arguments include declarations that machines cannot understand the value of a human life, nor the significance of its loss, and therefore that automated decisions to take a life are an affront to dignity, and concerns that LAWS would diminish our fundamental duty not to harm others [9].

Debate about the relative importance of such fundamental moral arguments to the public and policy discourse on LAWS is ongoing; however, arguments that focus on human dignity have featured prominently in at least one UNIDIR report and are increasingly a part of policy discussions [16].

Existing Policies and Policy Discussions

The United Nations Convention on Certain Conventional Weapons (CCW), which Canada signed in 1981, exists to "ban or restrict the use of specific types of weapons that are considered to cause unnecessary or unjustifiable suffering to combatants or to affect civilians indiscriminately." The CCW has convened regular meetings since 2014 on LAWS, with States Parties agreeing in 2019 to a set of eleven "guiding principles" for the development of LAWS [17]. These guiding principles include an affirmation that international humanitarian law applies to the potential development and use of LAWS, as well as a statement that "Human responsibility for decisions on the use of weapons systems must be retained since accountability cannot be transferred to machines." However, the guiding principles allow for continued development and use.

The CCW has considered proposals to ban LAWS, which have received support from numerous states and non-governmental organizations. However, pre-emptive bans have been opposed by other states, including the United States, Russia, and China [18]. In the December 2019 mandate letter sent by the Prime Minister to former Minister of Foreign Affairs François-Philippe Champagne, Prime Minister Trudeau specified a priority to "advance international efforts to ban the development and use of fully autonomous weapons systems." [19] This priority was not retracted or otherwise contradicted in the Prime Minister's subsequent supplementary mandate letter to incoming Minister of Foreign Affairs Marc Garneau in January 2021 [20].

In advance of the first session of the 2021 CCW's Group of Governmental Experts on Lethal Autonomous Weapons Systems - to be held in August 2021 - Canada offered a commentary on the existing guiding principles. The commentary distinguishes between "lethal" autonomous weapons systems and "fully" autonomous weapons systems, stipulating that these are "related but distinct" concepts [21]. Canada's view is that fully autonomous weapons systems are unacceptable due to their lack of "appropriate human involvement," and that both human control and human judgement must be exercised over any use of force.

Questions for Policymakers

- What role should Canada play in international discussions of LAWS? How will Canada advance its commitment to a ban of fully autonomous weapons systems?
- What constitutes "human judgement" and "human control" over the use of force? How should an appropriate level of human-machine interaction be identified and mandated?
- How will Canada respond to the position, advocated by some of our traditional allies, that existing international law is sufficient to address the ethical challenges of LAWS?
- What funding priorities should DND adopt to define and implement appropriate human judgement and control in emerging weapons technology?

References

[1] K.M. Sayler, "Defense primer: U.S. policy on lethal autonomous weapons systems," Congressional Research Service, December 1, 2020. [Online] https://fas.org/sgp/crs/natsec/IF11150.pdf.

[2] A. Etzioni and O. Etzioni, "Pros and cons of autonomous weapons systems," Military Review, May-June 2017, pp. 72-81.

[3] J. Altmann and F. Sauer, "Autonomous weapon systems and strategic stability," Survival - Global Politics and Strategy, vol. 59, no. 5, pp. 117-142, Sept. 2017.

[4] M.C. Horowitz, "When speed kills: Lethal autonomous weapon systems, deterrence and stability," Journal of Strategic Studies, vol. 42, no. 6, pp. 764-788, Sept. 2019.

[5] "Autonomous weapons: An open letter from AI and robotics researchers," Future of Life Institute, https://futureoflife.org/open-letter-autonomous-weapons/ (accessed July 15, 2021).

[6] G. Tamburrini, "On banning autonomous weapons systems: From deontological to wide consequentialist reasons," in Autonomous Weapons Systems: Law, Ethics,Policy, N. Bhuta et al., eds., Cambridge, UK: Cambridge University Press, 2016, pp. 121–141.

[7] H.M. Roff, "The strategic robot problem: Lethal autonomous weapons in war," Journal of Military Ethics vol. 13, no. 3, pp. 211-227, July 2014.

[8] N.E. Sharkey, "The evitability of autonomous robot warfare," International Review of the Red Cross, vol. 94, no. 886, pp. 787-799, Summer 2012.

[9] A. Sharkey, "Autonomous weapons systems, killer robots and human dignity," Ethics and Information Technology, vol. 21, no. 2, pp. 75-87, June 2019.

[10] R.C. Arkin, "The case for ethical autonomy in unmanned systems," Journal of Military Ethics, vol. 9, no. 4, pp. 332-341, Dec. 2010.

[11] B.D. Strawser, "Moral predators: The duty to employ uninhabited aerial vehicles," Journal of Military Ethics, vol. 9, no. 4, pp. 342-368, Dec. 2010.

[12] J.C. Galliot, "Uninhabited aerial vehicles and the asymmetry objection: A response to Strawser," Journal of Military Ethics vol. 11, no. 1, pp. 58-66, Mar. 2012.

[13] C.G. Kels, "Closing argument: At the outer bounds of asymmetry," Journal of Military Ethics, vol. 11, no. 3, pp. 223-244, Sept. 2012.

[14] J.C. Galliot, "Viewpoint article closing with completeness: The asymmetric drone warfare debate," Journal of Military Ethics, vol. 11, no. 4, pp. 353-356, Dec. 2012.

[15] E. Rosert and F. Sauer, "How (not) to stop the killer robots: A comparative analysis of humanitarian disarmament campaign strategies," Contemporary Security Policy, vol. 42, no. 1, pp. 4-29, Jan. 2021.

[16] United Nations Institute for Disarmament Research, "The weaponization of increasingly autonomous technologies," March 28, 2018. [Online] https://unidir.org/publication/weaponization-increasingly-autonomous-technologies-artificial-intelligence.

[17] United Nations, "Final report - Meeting 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," December 13, 2019. [Online] https://undocs.org/CCW/MSP/2019/9.

[18] K.M. Sayler, "International discussions concerning lethal autonomous weapons systems," Congressional Research Service, April 19, 2021. [Online] https://fas.org/sgp/crs/weapons/IF11294.pdf.

[19] Office of the Prime Minister, "Minister of Foreign Affairs Mandate Letter," December 13, 2019. [Online] https://pm.gc.ca/en/mandate-letters/2019/12/13/minister-foreign-affairs-mandate-letter.

[20] Office of the Prime Minister, "Minister of Foreign Affairs Supplementary Mandate Letter," January 15, 2021. [Online] https://pm.gc.ca/en/mandate-letters/2021/01/15/minister-foreign-affairs-supplementary-mandate-letter.

[21] "Commentary by Canada on the operationalization of the Guiding Principles affirmed by the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems," United Nations Office for Disarmament Affairs, https://documents.unoda.org/wp-content/uploads/2021/06/Canada_Commentary-on-the-Guiding-Principles.pdf (accessed July 15, 2021).