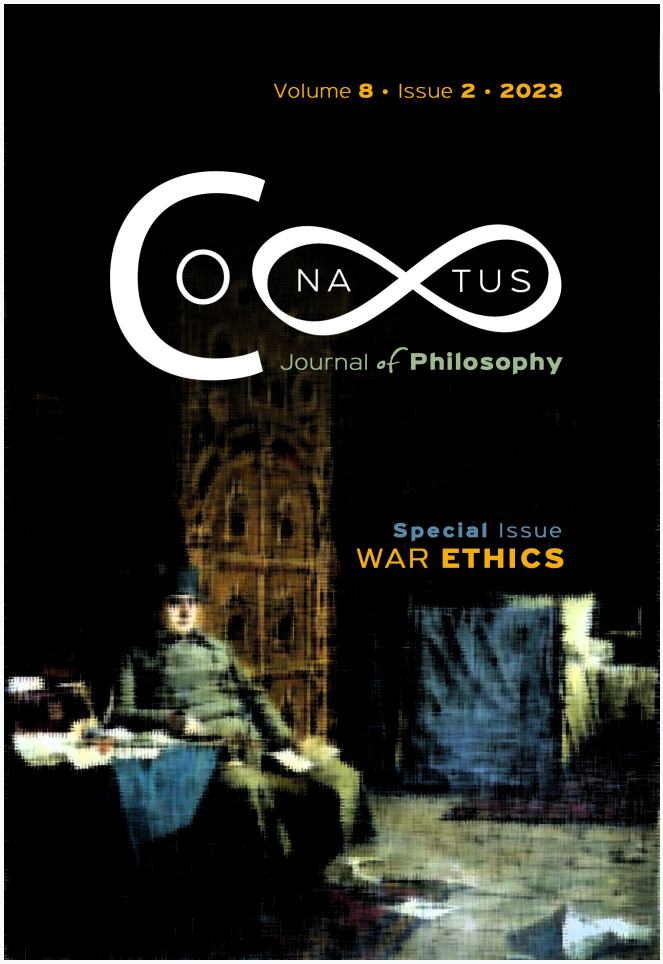
Conatus - Journal of Philosophy

Vol 8, No 2 (2023)

Conatus - Journal of Philosophy SI: War Ethics



Killer Robots and Inauthenticity: A Heideggerian Response to the Ethical Challenge Posed by Lethal Autonomous Weapons Systems

Ashley Roden-Bow

doi: 10.12681/cjp.34864

To cite this article:

Roden-Bow, A. (2023). Killer Robots and Inauthenticity: A Heideggerian Response to the Ethical Challenge Posed by Lethal Autonomous Weapons Systems. *Conatus - Journal of Philosophy*, *8*(2), 477–486. https://doi.org/10.12681/cjp.34864

A. Roden-Bow · *Conatus* 8, no. 2 (2023): 477-486
doi: https://doi.org/10.12681/cjp.34864

Killer Robots and Inauthenticity: A Heideggerian Response to the Ethical Challenge Posed by Lethal Autonomous Weapons Systems

Ashley Roden-Bow
TASIS, Portugal
E-mail address: ashley.roden@tasisportugal.org
ORCID iD: https://orcid.org/0009-0004-9370-1187

Abstract

This paper addresses the ethical challenges raised by the use of lethal autonomous weapons systems. Using aspects of the philosophy of Martin Heidegger, the paper demonstrates that lethal autonomous weapons systems create ethical problems because of the lack of moral agency in an autonomous system, and the inauthentic nature of the deaths caused by such a system. The paper considers potential solutions for these issues before arguing that from a Heideggerian standpoint they cannot be overcome, and thus the development and use of lethal autonomous weapons systems should be resisted and prohibited.

Keywords: *Heidegger; ethics of war; weaponry; artificial intelligence; authenticity; drone warfare; moral status of robots*

Long a topic of science fiction and philosophical contemplation, lethal autonomous weapons systems – weapons systems that select and apply lethal force to targets without human intervention[1] – have moved from the realm of possibility to actuality. The use of these systems remains controversial and deniable.[2] Whether or not

[1] "What You Need to Know About Autonomous Weapons," *International Committee of the Red Cross*, accessed May 1, 2023, https://www.icrc.org/en/document/what-you-need-know-about-autonomous-weapons.

[2] Robert F. Trager, and Laura M. Luca, "Killer Robots Are Here – and We Need to Regulate

lethal autonomous weapons systems have yet been used in combat, the availability of their use necessitates urgent dialogue about the ethical issues raised by the use of these systems, and what steps, if any, can be taken to mitigate the ethical issues raised. The focus throughout this paper will be on these ethical issues, rather than the technical aspects of lethal autonomous weapons systems. In the paper I will argue that the ethical problems I raise are inherent within the concept of lethal autonomous weapons systems, and thus cannot be overcome with a technical solution.

German philosopher Martin Heidegger is often cited as amongst the twentieth century's most important thinkers. His refusal to provide an explicit ethics in his writings, nor – some would argue – to consistently communicate in a manner that is clear and direct, has provided a barrier for Heidegger's work to be applied in a military ethics context. This challenge is multiplied by Heidegger's much debated collaboration with the National Socialist regime which has led to calls to remove the philosopher's works from the philosophical canon (for example in a book by Emmanuel Faye[3]). In this paper I aim to demonstrate that despite his personal moral and political failings, Heidegger's writings on technology can still provide a useful contribution to the discussion connected to lethal autonomous weapons systems, and that this contribution can be used in a way that avoids obscurantism, whilst providing a different approach from recent research taking a similar focus (e.g., Brayford[4]). To do this, I will utilize Heidegger's own writings on technology, and also those of Hubert Dreyfus whose critique of artificial reason is built on Heideggerian foundations.

From an everyday standpoint, the ethical problems connected to the use of lethal autonomous weapons systems are similar to those arising from self-driving vehicles (see Hansson et al. for an overview of these[5]). Despite the potential for automated systems to conduct decision making more accurately in rapid high-stress situations such as car crashes or on a battlefield, the idea of non-human "intelligence" making life and death decisions without direct human input or confirmation

---

Them," *Foreign Policy*, May 11, 2022, https://foreignpolicy.com/2022/05/11/killer-robots-lethal-autonomous-weapons-systems-ukraine-libya-regulation/.

[3] Emmanuel Faye, *Heidegger: The Introduction of Nazism into Philosophy in Light of the Unpublished Seminars of 1933-1935* (New Haven, CT: Yale University Press, 2011).

[4] Kieran M. Brayford, "Autonomous Weapons Systems and the Necessity of Interpretation: What Heidegger Can Tell Us About Automated Warfare," *AI and Society* (2022): 1-9.

[5] Sven Ove Hansson, Matts-Åke Belin, and Björn Lundgren, "Self-Driving Vehicles – An Ethical Overview," *Philosophy and Technology* 34, no. 4 (2021): 1383-1408.

feels uncomfortable for many. The use of lethal autonomous weapons systems is troublesome from a Heideggerian standpoint from two positions: firstly, because artificial intelligence – despite appearances – is incapable of reaching the status of moral agency, and secondly, because the kind of warfare conducted with lethal autonomous weapons systems would be inauthentic and thus unethical. In what follows, I aim to explain both these objections before making some suggestions for how we might respond to the ethical dilemmas raised.

Philosophical discomfort with the use of lethal autonomous weapons systems may be connected to the role a lethal automated weapon would play in a *jus in bello* context. Under traditional just war theory, the commanders, soldiers, and other actors are human, with all the psychological, biological, and ethical characteristics that come along with this. One could counter that advances in the area of artificial intelligence might work towards overcoming this obstacle as AI comes closer and closer to mimicking human intelligence. Despite their age, Hubert Dreyfus' writings on artificial reason in works such as *What Computers Can't Do* can help us explore this concern. Dreyfus was originally writing at a time of confidence in AI research that aimed to create a system capable of human-like reasoning. In response to this confidence, Dreyfus insisted that such efforts failed to account for the complexities of our everyday Being-in-the-world (to follow Heidegger's terminology), and what we might call "common sense."

In *Being and Time*, the work responsible for Martin Heidegger's renown, Heidegger outlines a rich and compelling explanation of humans' place in the world.[6] According to Heidegger, the mistake made in much previous philosophizing has been to consider the world in which we dwell as being a collection of external objects that we encounter and interact with, but which are wholly separate from us. For Heidegger, this atomistic vision of the self is an error. The individual person does not, has not, and cannot exist except as part of the world. Not only this, but our Being-in-the-world is essentially a Being-with-in-the-world.[7] By this, Heidegger means that our everyday way of Being – what he calls our "everydayness" – consists of existing in and as part of the world alongside others whose mode of Being is also Being-in-the-world and Being-with-others. According to Dreyfus, the complexity of this everydayness is such, that the creation of artificial intelligence has had to work from the basis of treating the background

[6] Martin Heidegger, *Being and Time*, trans. John Macquarrie and Edward Robinson (Oxford: Blackwell, 1962), 78-90.

[7] Ibid., 149-168.

or context in which the AI is functioning "as an object with its own set of preselected descriptive features."[8] All of which is to say that if Heidegger is right about the nature of human existence, and if Dreyfus is right about the insurmountable nature of replicating the complexities of everyday Being-with-in-the-world, we can assume that even a futuristic perfected lethal autonomous weapons system will not be able to achieve self-awareness, and therefore cannot be considered a moral agent.

Ethics – military ethics included – works from the basis that there are agents involved who can be held responsible for their choices and actions. A lethal autonomous weapon now seems to sit in a grey zone, lacking the personality to have agency and yet seemingly too autonomous or self-directed to be solely a tool. If we accept that even the most advanced lethal autonomous weapon system will not be the self-aware beings we encounter in movies like *Terminator* or *Blade Runner*, it is still worth considering alternative, more reduced, forms of agency. One such approach would be to consider the lethal autonomous weapons system within the context of ethical deliberations as a limited agent akin to an animal or child. Animals and children are not tools, and act in a way that is at least partly self-directed. Yet philosophically and legally, we do not hold animals or children responsible for their acts in the same way we would an adult human. By this we recognize the limitations at play in the reasoning processes that exist in an animal or child. We could perhaps see the reasoning of a futuristic lethal autonomous weapon system in the same way. This however would raise additional issues. For one, is the still-developing reasoning power of children not one of the reasons for our revulsion at the idea of using child soldiers? The AI-child/animal equivalency suffers from other weaknesses. Children and animals are usually accepted to have certain rights (albeit more limited rights in the case of animals), which provide barriers to placing them in situations where they might suffer moral or physical injury. Surely this risk of injury is the reason in the modern era we would defer from using child soldiers and animals in combat situations, even if their other limitations could be overcome. The need to protect children and animals from these is amplified by their lack of ability to understand and consent to the combat situation they are being placed in. The equivalency also still overlooks both the complexity of the decision-making processes at hand, and also the lack of self-directedness a lethal autonomous weapons system would

[8] Hubert Dreyfuss, *What Computers Still Can't Do: A Critique of Artificial Reason* (Cambridge, MA: MIT Press, 2009), 56.

have. Animals and children may be guided at least in part by their instinct or other psychological factors outside of their control, but this cannot be satisfactorily substituted with the role of algorithms in the autonomous lethal weapons system. Ultimately, even in the context of a "self-learning" system, the initial algorithms or instructions programmed into the weapon act as a technological "first cause." This first cause is not biological or theological – at least not directly. It is the action of human agents.

Instead, one might better consider the person or persons who deployed or programmed, or ordered the deployment or programming, of the lethal autonomous weapons system to be the moral agent or agents culpable for breaches of *jus in bello* principles. Though these people do not actively select a particular target and confirm the use of lethal force, it is their earlier decisions and actions which lead to the use of force.

We cannot be confident that even the most advanced lethal autonomous weapons system would be free from error, or even always superior to a human placed in the same position. As Bonnefon highlights in the context of self-driving cars, a car that is "just safer than the average human driver means that it is not as safe as many, many drivers."[9] Similarly, a lethal autonomous weapons system may make fewer mistakes than most humans under pressure, this does not mean that in all situations the system would be better than *all* humans. Yet even if we put this to one side and assume that a perfected lethal autonomous weapons system will be more accurate than even the most skillful human in *all* situations, and only harm those ethically liable to harm under *jus in bello* principles, from a Heideggerian standpoint, this shift from considering the lethal autonomous weapon system to be a limited agent to being a tool would not be sufficient for the death caused by the lethal autonomous weapons system to be considered authentic, and therefore – for our purposes – ethical.

In a chapter addressing the subject, Statman argues that the use of lethal autonomous weapons systems would result in unfair, disrespectful, and riskless warfare.[10] From a Heideggerian perspective it is the inauthenticity of the potential deaths caused by lethal autonomous

[9] Jean-François Bonnefon, "Trusting Self-Driving Cars Is Going to Be a Big Step for People," interview by Jonathan O'Callaghan, *Horizon Magazine*, April 2, 2019, https://ec.europa.eu/research-and-innovation/en/horizon-magazine/trusting-self-driving-cars-going-be-big-step-people.

[10] Daniel Statman, "Drones and Robots: On the Changing Practice of Warfare," in *The Oxford Handbook of Ethics of War*, eds. Seth Lazar and Helen Frowe, 472-487 (Oxford: Oxford University Press, 2020), 475-478.

weapons systems which creates these problems. In order to explain this, we will need to address both the role of death in Heidegger's thought and Heidegger's understanding of technology.

For Heidegger, our entire lives of achieved and missed possibilities are leading up to our deaths after which those possibilities are closed. He characterizes this as our Being-toward-death.[11] Heidegger's approach towards death in *Being and Time* is perhaps the most traditionally existential aspect of this work, finding echoes in the philosophies of Søren Kierkegaard and Jean-Paul Sartre. For Heidegger, by closing off all future possibilities for us, our death offers us the possibility to be whole. We are no longer an incomplete work in progress, but instead have completed our process of self-creation. The question of the authenticity of death is central for Heidegger. Unlike in other aspects of our life in which we can be dissolved into the undifferentiated mass that Heidegger calls "the They," our death is our own. We experience the death of others as being "there alongside"[12] but we have "no way of access to the loss-of-Being"[13] suffered by the one who dies. We will only access this experience in a genuine sense when it is us ourselves who die. All of which is to say that for Heidegger's philosophy, death is a centrally important aspect of our existence. As such, the authenticity of one's death is central to the authenticity of our existence as a whole. We need now to connect this to Heidegger's understanding of technology, which will then be connected back to lethal autonomous weapons systems as an example of a technology which, by design, causes death.

Heidegger's writings on technology, in particular his essay *The Question Concerning Technology*, sets out a skeptical, but not hostile approach to our relationship with technology. In the essay Heidegger argues "the essence of technology is by no means anything technological."[14] By this, he is addressing a tendency to view technological advancement as something inevitable or value neutral. According to Heidegger, there is an inclination to see technology as a means to an end, or a tool[15] to achieve some particular goal. If technology seems to be "threaten[ing] to slip from human control"[16] – a fear that is per-

[11] Heidegger, *Being and Time,* 279-311.

[12] Ibid., 282.

[13] Ibid.

[14] Martin Heidegger, *The Question Concerning Technology and Other Essays*, trans. William Lovitt (New York: Garland, 1977), 4.

[15] Ibid.

[16] Ibid., 5.

haps even more prescient today than when Heidegger was writing – the focus turns to our will to master this technology.[17] All of this, Heidegger argues, is a mistake caused by the faulty premise that technology is merely a means to achieve a particular end. Instead, Heidegger sets out his theory of technological Enframing – this is the argument that technology is a filter through which we encounter and interact with the world. Technological Enframing falsely appears as a neutral state masquerading as bland everydayness – similar to the way in which Zizek[18] and others view the hidden role of capitalist ideology in the modern world. Technological Enframing, according to Heidegger, causes us to encounter the world as a collection of resources to be utilized for technological processes. He uses the example of a hydroelectric plant on the Rhine reducing this great river into "something at our command."[19] As such, for Heidegger, technological Enframing changes our relationship with the world from one of dwelling to one of domination or utilization.

As part of a lecture entitled "The Danger" given in Bremen in 1949, Heidegger obliquely addresses the evils of the Holocaust through criticisms of the use of technological means to commit mass murder in the following, much debated, quotation:

> Hundreds of thousands die in their masses. Do they die? They perish. They are put down. Do they die? They become pieces of inventory of a standing reserve for the fabrication of corpses.[20]

The quotation has understandably been criticized. Firstly, "millions" would have been a much more accurate scale for the desolation than "hundreds of thousands." Secondly, Heidegger's own role in actively supporting the earlier stages of the Nazi regime which led to murder and destruction on an unprecedented scale also goes unsaid and un-apologized for in that text (or indeed elsewhere in his work). Despite this, Heidegger – flawed as he was – sets out an important point about the importance of dignity in death. He adds that "to die [...] means to

---

[17] Ibid.

[18] The exploration of John Carpenter's movie *They Live* in Zizek's documentary *The Pervert's Guide to Ideology* probably sets this out the clearest.

[19] Heidegger, *The Question Concerning Technology and Other Essays*, 16.

[20] Martin Heidegger, "The Danger," in *Breman and Freiburg Lectures: Insight Into That Which Is and Basic Principles of Thinking*, trans. Andrew J. Mitchell, 44-63 (Bloomington, IN: Indiana University Press, 2012), 53.

carry out death in its essence. To be able to die means to be capable of carrying this out."[21] Following through Heidegger's philosophy on technology, we can conclude that when people are conceived of as a mere resource, we close off both their possibility of their authentic dwelling-in-the-world, and their possibility of an authentic death.

From a Heideggerian standpoint, the use and proliferation of lethal autonomous weapons systems would reduce the targets of these weapons to a mere resource in the world to be dispensed with. Their existence terminated without even a nod from another human being – even one based in an office on another continent. This discomfort surrounding authentic or honorable deaths is not just a Heideggerian concern, being also present in many traditional warrior codes (see the work done by Shannon French[22]). The use of lethal autonomous weapons systems may protect the indirect user from the moral wound of direct involvement in causing death, but surely this carries the risk of making the causing of death less challenging, and perhaps therefore less avoided. In the context of nuclear deterrence, the previous major technological threat to accepted military ethics, Roger Fisher famously made the following suggestion:

> My suggestion was quite simple: Put [the nuclear] code number in a little capsule, and then implant that capsule right next to the heart of a volunteer. The volunteer would carry with him a big, heavy butcher knife as he accompanied the President. If ever the President wanted to fire nuclear weapons, the only way he could do so would be for him first, with his own hands, to kill one human being. The President says, "George, I'm sorry but tens of millions must die." He has to look at someone and realize what death is – what an innocent death is. Blood on the White House carpet. It's reality brought home. When I suggested this to friends in the Pentagon they said, "My God, that's terrible. Having to kill someone would distort the President's judgment. He might never push the button."[23]

Forty-plus years later, we must again consider if we want the causing of death to be easier or more challenging for those who command it.

---

[21] Ibid., 53.

[22] Shannon French, *The Code of the Warrior: Exploring Warrior Values Past and Present* (Lanham, MD: Rowman & Littlefield, 2017).

[23] Roger Fisher, "Preventing Nuclear War," *Bulletin of the Atomic Scientists* 37, no. 3 (1981): 16.

Perhaps a similar safeguard should be considered for political leaders who request the use of lethal autonomous weapons systems today?

There are no easy answers to the questions raised by lethal autonomous weapons systems. Heidegger's response to these challenges is to turn to art – in particular, poetry – as an alternative to technological Enframing.[24] Though this is understandable in the context of his philosophy, a more practical approach may be more effective for us in the short to medium term. The response to these conclusions, if they are correct, must be to act to prohibit the use of lethal autonomous weapons systems and to treat their use in much the same way as we do chemical weapons and other inhuman acts of war. The development, improvement, and use of lethal autonomous weapons systems is not something that is inevitable or value neutral. Whether we are civilian ethicists or military practitioners, our duty therefore is to highlight this lack of inevitability, and to encourage those in the position to think again (whether in universities, ministries of defense, or private laboratories) to reconsider.

Acknowledgements

An earlier version of this paper was presented at the annual International Society for Military Ethics in Europe (Euro ISME) conference held in Athens, Greece, on May 11-12, 2023. I am grateful for all the comments, suggestions, and queries I received from various attendees. All errors and omissions remain my own.

References

Bonnefon, Jean-François. "Trusting Self-Driving Cars Is Going to Be a Big Step for People." Interview by Jonathan O'Callaghan. *Horizon Magazine*, April 2, 2019. https://ec.europa.eu/research-and-innovation/en/horizon-magazine/trusting-self-driving-cars-going-be-big-step-people.

Brayford, Kieran M. "Autonomous Weapons Systems and the Necessity of Interpretation: What Heidegger Can Tell Us About Automated Warfare." *AI and Society* (2022): 1-9.

Dreyfuss, Hubert. *What Computers Still Can't Do: A Critique of Artificial Reason*. Cambridge, MA: MIT Press, 2009.

Faye, Emmanuel. *Heidegger: The Introduction of Nazism into Philosophy in Light of the Unpublished Seminars of 1933-1935*. New Haven, CT: Yale University Press, 2011.

[24] Martin Heidegger, *The Question Concerning Technology and Other Essays*, 34-35.

Fisher, Roger. "Preventing Nuclear War." *Bulletin of the Atomic Scientists* 37, no. 3 (1981): 11-17.

French, Shannon. *The Code of the Warrior: Exploring Warrior Values Past and Present*. Lanham, MD: Rowman & Littlefield, 2017.

Hansson, Sven Ove, Matts-Åke Belin, and Björn Lundgren. "Self-Driving Vehicles – An Ethical Overview." *Philosophy and Technology* 34, no. 4 (2021): 1383-1408.

Heidegger, Martin. "The Danger." In *Breman and Freiburg Lectures*: *Insight into That Which Is and Basic Principles of Thinking*. Translated by Andrew J. Mitchell, 44-63. Bloomington, IN: Indiana University Press, 2012.

Heidegger, Martin. *Being and Time*. Translated by John Macquarrie and Edward Robinson. Oxford: Blackwell, 1962.

Heidegger, Martin. *The Question Concerning Technology and Other Essays*. Translated by William Lovitt. New York: Garland, 1977.

International Committee of the Red Cross. "What You Need to Know About Autonomous Weapons." https://www.icrc.org/en/document/what-you-need-know-about-autonomous-weapons.

Statman, Daniel. "Drones and Robots: On the Changing Practice of Warfare." In *The Oxford Handbook of Ethics of War*, edited by Seth Lazar and Helen Frowe, 475-478. Oxford: Oxford University Press, 2020.

Trager, Robert F., and Laura M. Luca. "Killer Robots Are Here – and We Need to Regulate Them." *Foreign Policy*, May 11, 2022. https://foreignpolicy.com/2022/05/11/killer-robots-lethal-autonomous-weapons-systems-ukraine-libya-regulation/.