WORKING TOWARDS ETHICAL AUTONOMOUS WEAPON SYSTEMS: A COMPENDIUM OF ARGUMENTS

Janar Pekarev

■

ABSTRACT. The weaponisation of artificial intelligence (AI) is upending traditional warfare as the development of autonomous weapon systems (AWS) increases in pace and sophistication. As a result, the concept of a weaponised AI where a weapon system that, once activated, can select and engage targets without human intervention, creates serious complications and challenges for international humanitarian law (IHL) enforcement. It also raises fundamental ethical questions as to whether people can delegate life and death decisions and accountability to artificial agents. Many authors have voiced concerns on this subject but thus far a comprehensive presentation of a deeper ethical considerations related to AWS is lacking. This article uses a systematic literature review to provide an overview of the most common arguments for and against using AWS. The results from the review indicate that compliance or non-compliance with the core principles of IHL are the most prevalent concerns in the discourse. The lack of accountability also makes for a strong deontological argument against using AWS, although normative ethics represents only half of the theoretical argumentation.

Keywords: autonomous weapon systems, international humanitarian law, military ethics, criminal liability, artificial agency

Võtmesõnad: autonoomsed relvasüsteemid, relvakonfliktiõigus, militaareetika, kriminaalvastutus, tehislik agentsus

1. Introduction

Ethics as a structuring discourse is well suited to provide a framework for discussions over AWS in a cultural and strategic context[1], although not all scholars necessarily accept this. Gómez de Ágreda argues that ethics does not

---

apply to machines since free will is a necessary prerequisite for the development of ethical standards. According to Gómez de Ágreda, only the users and not the machines are accountable because only human autonomy is relevant in an ethical debate[2]. Bauer contradicts this by asserting that machine ethics very much applies to artificial moral agents[3]. Liao reinforces this stance by submitting that AWS are a critical concern in the discussion of the ethics of AI in general[4], and moreover, ethical considerations should always be incorporated into all phases of their development, along with a worst-case scenario mindset[5]. Umbrello et al. suggest that AWS are theoretically capable of becoming moral actors and, therefore, qualified to make life and death decisions without human intervention[6]. In short, broader research into the ethics of the design and use of these systems is imperative[7].

The purpose of this paper is to identify the most significant problems related to AWS from an ethical perspective. The literature review is used to systematise the current research regarding different perceptions of AWS and identify the aspects that scholars deem the most problematic with AWS. Thus far, a comprehensive approach identifying the main concerns associated with the development of AWS is lacking. Therefore, a good starting point for a more inclusive discussion over the ethical considerations of AWS begins with distinguishing the most common ethical concerns and finding a common thread to guide further research. The article concludes with a discussion of interrelationships between these implications and offers a perspective for developing and using AWS.

[2]   **Gómez de Ágreda, A**. 2020. Ethics of autonomous weapons systems and its applicability to any AI systems. – Telecommunications Policy, Vol. 44, No. 6, 101953. https://doi.org/10.1016/j.telpol.2020.101953. [**Gómez de Ágreda** 2020]

[3]   **Bauer, W. A**. 2020. Virtuous vs. Utilitarian Artificial Moral Agents. – AI & Society, Vol. 35, No. 1, pp. 263–271. https://doi.org/10.1007/s00146-018-0871-3.

[4]   **Liao, S. M**. (ed.). 2020. Ethics of Artificial Intelligence. 1st edition. New York, NY, United States of America: Oxford University Press. [**Liao** 2020]

[5]   **Gómez de Ágreda** 2020, p. 101953.

[6]   **Umbrello, S.; Torres, P.; De Bellis, A. F**. 2020. The future of war: Could lethal autonomous weapons make conflict more ethical? – AI & Society, Vol. 35, No. 1, pp. 273–282. https://doi.org/10.1007/s00146-019-00879-x. [**Umbrello et al**. 2020]

[7]   **Verdiesen, I.; de Sio, F. S.; Dignum, V**. 2019. Moral Values Related to Autonomous Weapon Systems: An Empirical Survey that Reveals Common Ground for the Ethical Debate. – IEEE Technology and Society Magazine, Vol. 38, No. 4, pp. 34–44, 2019, https://doi.org/10.1109/MTS.2019.2948439.

The ethical debate focuses on the terminology of autonomy regarding weapon systems. Although there is no internationally agreed definition of AWS, there is a general understanding that this term refers to an AI-driven machine that can select and engage targets autonomously without human intervention[8]. Autonomous weapon systems may seem relatively intelligent but their functionality should be approached with a degree of caution. They are designed to apply kinetic or non-kinetic force to achieve the desired lethal or nonlethal effects on targets. However, the use of terms such as *fully* or *lethal* can be misleading and ought to be avoided. AWS can also be used in the cyber realm where the application of non-kinetic force by AWS (cyber attacks) against hospitals, electrical grids or a financial system can have devastating consequences and result in deaths or serious physical injuries. This article also accepts that AWS are still in a development phase and a preemptive ban on emerging systems is unrealistic.

2. Methods

In order to gather the most salient views on the ethical considerations of AWS, the author of this study used a systematic literature review based on the guidelines proposed by Petticrew et al.[9]. An implementation of this method made it possible to generate a good overview of the topic and find ways to direct future research efforts. The literature search was conducted using The Web of Science Core Collection, which is a well-regarded and authoritative collection of indexed journals. An initial search through the database revealed many scientific articles on this topic (from various fields of studies) with the number of publications increasing exponentially by year.

2.1. The review protocol

A basic topic search of the core collection database resulted in n = 274 articles (incl. title, abstract, author keywords, and keyword Plus). The refined search

---

8   **US Department of Defense** 2012. Directive 3000.09: Autonomy in Weapon Systems. November 21. Homeland Security Digital Library. Department of Defense. https://www.esd. whs.mil/portals/54/documents/dd/issuances/dodd/300009p.pdf (April 01, 2021).

9   **Petticrew, M.; Roberts, H**. 2006. Systematic Reviews in the Social Sciences: A Practical Guide. Blackwell Publishing. https://doi.org/10.1002/9780470754887. [**Petticrew, Roberts** 2006]

resulted in n = 103 (publications from 2015–2020). This narrowed search resulted only in articles formatted in document types and written in English, i.e., 87% of searched articles. A preliminary analysis showed that many of the studies were not actually focused on the autonomy of weaponry or on the ethical aspects, neither did they define AWS as a research object. Therefore, the search criteria were more narrowly refined using the title keyword – *autonomous weapon systems* (n = 28) – and the topic keywords – *ethical autonomous weapon systems* (n = 21). Table 1 details the identification outcome as well as the eligibility and exclusion criteria. Annex 1 presents a list of 41 publications that corresponded to the criteria.

Table 1. The review protocol

Identi-fication	1. Topic search (autonomous weapon systems) n = 274, timespan: 1980–2020 2. Refined topic search n = 104	
Eligibility	Published from 2015 to 2020, English language (87%), only articles	
	Inclusion criteria	Exclusion criteria
	1. Studies identified in the context of the ethics of warfare and use of force in military affairs 2. Full text published articles indexed in the Core Collection database 3. Focusing on the aspects of AWS as a weaponised AI.	1. Full text from Publisher was not accessible 2. AWS are in the list as an illustrative example of autonomous systems, and the study does not focus on the aspects of weaponry 3. A duplicate article between Title and Topic search
Included	n = 41 (articles)	

A proportional distribution score (see Figure 1) showed that one-fourth of the 41 studies identified eight or more areas of concern in relation to AWS. A review of these studies indicated that two of the articles analysed over 16 different concerns, and six of the studies discussed a minimum of 2 concerns: a total average of 6.3 concerns per article. Although the sample size is relatively small, it does show that there is a pattern in relation to the concerns about AWS and there is sufficient information to allow for research conclusions to be determined. Data saturation refers to
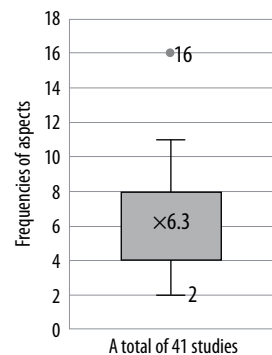

Figure 1. The proportional score

the point where plausibly increasing samples do not provide any new information in relation to the most frequently discussed aspects. It is, therefore, reasonable to assume that further data collection would have yielded similar results and confirmed the preliminary outcome.

2.2. An analytical procedure

A text extraction form[10] was applied to each of the studies in order to narrow the focus to only two categories – the problematic aspects of AWS and their connection to an ethical theory. A problematic aspect or ethical theory needed to be explicitly described and connected to the development, deployment, and use of AWS. When an ethical aspect was detected, the list was supplemented, and that particular aspect was cross-referenced with previous and following samples. The occurrence of areas of concern and ethical theories was measured on a scale of 0–1 to determine their presence or absence and whether or not to include it with the numerical values. A mention of a problematic aspect or an ethical theory needed to appear at least once in the text. Examples 1 and 2 highlight a specific concern about AWS with regard to IHL principles (underlined text). An extract of the results is presented in Table 2.

> *Example 1: Since today's AWSs would likely <u>fail to discriminate</u>, they <u>would not comply with international humanitarian law</u>, hence they should be considered both illegal and ethically impermissible*[11].

> *Example 2: The larger legal concern about these systems is whether a machine would ever be able to <u>uphold international humanitarian law</u>; humans already make mistakes and break IHL, and there is very little doubt that machines would as well*[12].

---

[10] **Petticrew, Roberts** 2006.

[11] **Guersenzvaig, A**. 2018. Autonomous Weapon Systems Failing the Principle of Discrimination. – IEEE Technology and Society Magazine, Vol. 37, Issue 1, pp. 55–61. https://doi.org/10.1109/MTS.2018.2795119.

[12] **Jones, E**. 2018. A Posthuman-Xenofeminist Analysis of the Discourse on Autonomous Weapons Systems and Other Killing Machines. – Australian Feminist Law Journal, Vol. 44, Issue 1, pp. 93–118. https://doi.org/10.1080/13200968.2018.1465333.

Table 2. The text extraction form of aspects

Study	Year	IHL principles	Respon-sibility	Account-ability	Human control	Threshold of armed conflict
Bauer	2020	0	1	1	0	0
Guersenzvaig	2018	1	1	1	0	0
Rosert	2019	1	0	1	1	1
Jones	2018	1	1	1	1	1
Taylor	2020	1	0	1	1	0
Score		**4**	**3**	**5**	**3**	**2**

Ethical theories were measured using a similar method. A criterium of whether the problematic aspects under discussion were related to an ethical theory had to be met. Examples 3 and 4 highlight the occurrence of ethical theories (underlined text). An extract of the results is presented in Table 3.

> *Example 3: the existence of a <u>responsibility gap to pose a strong deontological reason</u> against the deployment of AWS*[13].

> *Example 4: propose an artificial moral agent (AMA) modeled on <u>two-level utilitarianism</u> that exemplifies the strengths of the virtuous AMA while accruing additional benefits*[14].

Table 3. The text extraction form of ethical theories

Study	Year	Utilita-rianism	Deonto-logy	Virtue ethics	Applied ethics	Other
Bauer	2020	1	0	0	0	0
Guersenzvaig	2018	0	0	0	0	1
Rosert	2019	0	0	1	0	0
Jones	2018	0	0	0	1	0
Taylor	2020	0	1	0	0	0
Score		**1**	**1**	**1**	**1**	**1**

---

[13] **Taylor, I**. 2020. Who Is Responsible for Killer Robots? Autonomous Weapons, Group Agency, and the Military-Industrial Complex. – Journal of Applied Philosophy, Vol. 38, Issue 2, pp. 320–334. https://doi.org/10.1111/japp.12469. [**Taylor** 2020]

[14] **Bauer, W. A**. 2020. Virtuous vs. Utilitarian Artificial Moral Agents. – AI & Society, Vol. 35, Issue 1, pp. 263–271. https://doi.org/10.1007/s00146-018-0871-3.

3. Results

A total of 22 areas of concern from a total of 41 studies met the eligibility criteria for the review. Figure 2 shows 19 areas of concern, while three concerns (*i.e., confidence, intervention, and verification*) appeared only once and are not reflected in the figure. Annex 2 presents 22 areas of concern that corresponded to the criteria of a systematic literature review.
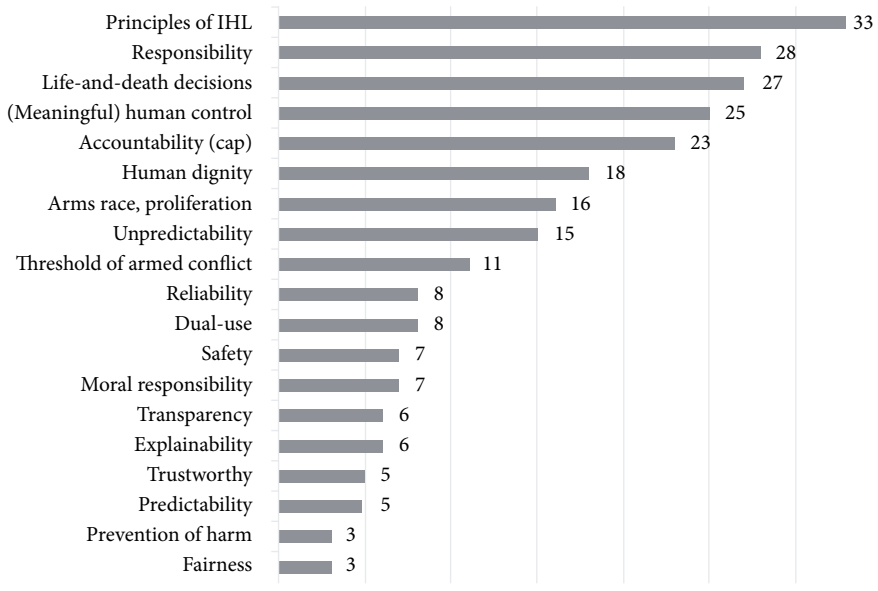

Figure 2. The occurrence of aspects

3.1. Aspects identified in the literature review

The identified areas of concern are very closely interlinked and presenting them one by one in this article would be too lengthy. For this reason, they have been grouped into categories. The following grouping assists in organising and rationalising the presentation of preliminary results: Figure 3 highlights the interlinked areas of concern.

Figure 3. The interlinked areas of concern

Group	Areas of concern
Principles of IHL	compliance with IHL
Criminal liability	accountability, responsibility, moral responsibility
Humanity	human dignity, life and death decisions
Human control	human control (meaningful), intervention, verification
Arms race	arms race, proliferation, threshold of armed conflict
Values	unpredictability, safety, explainability, transparency, predictability, trustworthy, confidence, reliability, fairness, dual-use, prevention of harm

3.1.1. The principles of IHL

As seen in Figure 2, the majority of arguments focused on the fundamental principles of IHL. The principle of distinction was the most frequently mentioned. But for the most part, compliance or non-compliance with the principles of IHL were most often credited as the compelling ethical reasons for not using AWS. The foremost concern among critics was that AWS would be unable to adhere to the IHL principles because the judgment required for these calculations goes beyond the programming and learning abilities of these systems and, in this regard, it is nearly impossible for machines to exceed the capacities of their makers[15]. There are, however, several dissenting authors who believe that AWS could more efficiently adhere to the IHL principles because emotions like fear or hysteria would not cloud the judgment of AWS. Moreover, AI weapons are able to process a lot more sensory information than humans without discarding or distorting it to fit preconceived notions[16].

3.1.2. Liability issues

Even if AWS were able to follow the principles of IHL in a similar manner to humans, some violations of IHL would most likely still occur. This introduces concerns about the liability of operational AWS. For the purposes of

---

[15] **Sehrawat, V**. 2017. Autonomous Weapon System: Law of Armed Conflict (LOAC) and Other Legal Challenges. – Computer Law & Security Review, Vol. 33, Issue 1, pp. 38–56. https://doi.org/10.1016/j.clsr.2016.11.001.

[16] **Etzioni, A**. 2018. Pros and Cons of Autonomous Weapons Systems (with Oren Etzioni). – Happiness Is the Wrong Metric: A Liberal Communitarian Response to Populism, Vol. 11, pp. 253–263. Dordrecht: Springer. https://doi.org/10.1007/978-3-319-69623-2_16.

this paper, it will be referred to as legal responsibility and accountability. Normally, accountability would refer to the norms that hold people legally responsible for their actions. Legal debates on a UN level constantly echo legal accountability in relation with the IHL principles[17]. The accountability gap is currently one of the most compelling reasons for maintaining human responsibility when deploying AWS[18]. Fortunately, some existing frameworks are already suitable for addressing the accountability gap. Taylor cites the examples of group agency and group responsibility as templates that could provide a way for addressing the essential ethical challenges with using AWS. For instance, holding the military-industrial complex morally accountable for some of the consequences caused by AWS could be a possible means for addressing accountability[19]. It is for this reason that liability aspects can be grouped into a primary concern that is provisionally separate from the IHL principles.

3.1.3. Humanity

Delegating life and death decisions to AWS was another frequently cited concern in the sample studies. Concerns about the abdication of very sensitive decisions to non-human agents were frequent and a central part of the discussions, since the subject is closely intertwined with human dignity[20]. A classification of human beings as mere military objects or lawful targets by an artificial agent also raises very important questions about human rights. In short, the use of AWS in a conflict zone could be a violation of the right to dignity for those affected by the act of force[21]. According to Skerker et al., artificial agents cannot understand the value of human life because they lack the experience of having personal projects or sensing their own mortality. Skerker et al. also assert that since a duty-bearer cannot transfer duties and privileges

[17] **Riebe et al**. 2020.

[18] **Park, S**. 2020. Analysis of the Positions Held by Countries on Legal Issues of Lethal Autonomous Weapons Systems and Proper Domestic Policy Direction of South Korea. – Korean Journal of Defense Analysis, Vol. 32, Issue 3, pp. 393–418. https://doi.org/10.22883/kjda.2020.32.3.004.

[19] **Taylor** 2020.

[20] **Gómez de Ágreda** 2020.

[21] **Heyns, Chr**. 2016. Human Rights and the Use of Autonomous Weapons Systems (AWS) During Domestic Law Enforcement. – Human Rights Quarterly, Vol. 38, Issue 2, pp. 350–378. https://doi.org/10.1353/hrq.2016.0034.

to an entity incapable of bearing those duties and privileges, a human combatant cannot and should not transfer the responsibility of targeting enemy combatants to a robot. Therefore, a human duty-bearer who deploys AWS breaches the martial contract between human combatants and attacks the dignity of the targeted combatants[22].

3.1.4. Human control

It is not possible to discuss weaponised AI without reference to human control. The concept of meaningful human control (MHC) means that situational awareness should apply for timely decision-making, an actual and practical option for intervention, and verification at every stage of a conflict (i.e., from the development phase to the use of lethal force). According to Santoni de Sio et al., several conditions need to be met to satisfy the MHC requirements. The first condition is contingent upon the existence of relevant human (moral) reasons for engaging in a conflict in the first place. The second condition states that the behavior of AWS must be traceable to a proper moral understanding on the part of the humans who designed and deployed the system[23]. Even if the behavior of AWS is traceable, the question then arises as to whether people can understand why an AWS acted in a certain way in a given scenario. This complexity of decision-making by AWS in terms of speed and quantity is not meaningfully controllable by humans in operational scenarios and, therefore, makes regulation and control extremely difficult[24].

3.1.5. Arms race

Although no direct link can be made between arms races and arms proliferation, they are both related to this particular topic. Altmann et al. argue that AWS are susceptible to proliferation and will inevitably create an arms

[22] **Skerker, M.; Purves, D.; Jenkins, R**. 2020. Autonomous Weapons Systems and the Moral Equality of Combatants. – Ethics and Information Technology, Vol. 22, Issue 3, pp. 197–209. https://doi.org/10.1007/s10676-020-09528-0.

[23] **Santoni de Sio, F.; van den Hoven, J**. 2018. Meaningful Human Control over Autonomous Systems: A Philosophical Account. – Frontiers in Robotics and AI, Vol. 5, Article 15. https://doi.org/10.3389/frobt.2018.00015.

[24] **Huelss, H**. 2020. Norms Are What Machines Make of Them: Autonomous Weapons Systems and the Normative Implications of Human-Machine Interactions. – International Political Sociology, Vol. 14, Issue 2, pp. 111–128. https://doi.org/10.1093/ips/olz023.

race that has the potential to result in an increased number of crises, greater instability, and higher risks of escalation[25]. Warren states that, in the digital age, proliferation is possible for any algorithm by creating or transmitting algorithms that are adapted for use by robotic devices[26]. In addition to the aforementioned concerns, the introduction of AWS would potentially reduce the threshold for armed conflicts. The inherent properties of AWS – low cost, military efficiency, expendability, and precision – could encourage proliferation while lowering the threshold for their use and, therefore, undermine the geopolitical security[27]. Barbé et al. stress that resorting to AWS could erode the existing security norms while lowering the threshold for a required force[28]. However, some authors argue that the AWS race could be similar to the nuclear arms race wherein mutual deterrence has actually increased the lethal force threshold. Nevertheless, the urgency for developing AWS could start a new arms race, resulting in a proliferation of artificial agents that could potentially alter the armed conflict threshold.

3.1.6. The values of AWS

Figure 3 reveals eleven values that characterise AWS. These values are complementary and can be categorised into a single group, so it is not necessary to highlight their individual features separately. Riebe et al. argue that predictability and, to varying degrees, reliability are essential values that are inherent to emerging technologies[29]. Therefore, these values can also be considered as aspects of AWS, which means that they can be placed into a single group. Although these values may be related to the aspects discussed earlier, the specific features are more connected with the values of AWS.

---

[25] **Altmann, J.; Sauer, F**. 2017. Autonomous Weapon Systems and Strategic Stability. – Survival. Global Politics and Strategy, Vol. 59, Issue 5, pp. 117–142. https://doi.org/10.1080/00396 338.2017.1375263.

[26] **Warren, A.; Hillas, A**. 2020. Friend or Frenemy? The Role of Trust in Human-Machine Teaming and Lethal Autonomous Weapons Systems. – Small Wars and Insurgencies, Vol. 31, Issue 4, pp. 822–850. https://doi.org/10.1080/09592318.2020.1743485.

[27] **Umbrello et al**. 2020.

[28] **Barbé, E.; Badell, D**. 2020. The European Union and Lethal Autonomous Weapons Systems: United in Diversity? – European Union Contested: Foreign Policy in a New Global Context. Johansson-Nogues, E.; Vlaskamp, M. C.; Barbé, E. (eds.). Cham: Springer International Publishing AG, pp. 133–152. https://doi.org/10.1007/978-3-030-33238-9_8.

[29] **Riebe et al**. 2020.

3.1.7. Results of a theoretical approach to ethics

The aspects of AWS analysed in the review can be categorised according to normative theory traditions. Figure 4 reveals that several of the considerations do not provide a basis for a discussion over the theoretical framework. These are marked as *Other*. Almost half of the addressed aspects did not involve an argumentation based on ethical theories, i.e., the authors did not provide a theoretical framework for a discussion, while in the other half, the discussion was based on an ethical theory.

Figure 4. The theoretical approach

Theories	Utilitarianism	Deontology	Virtue ethics	Applied ethics	Other
Score	2	10	5	6	19

Although many of the authors did not directly link the issue of liability and compliance with IHL to ethics, in many cases, these legal aspects are nevertheless implicit from a deontological point of view. Several authors have argued that life and death decisions occurring outside of human agency or control should be reason enough to impart a solid deontological opposition against the use of AWS. In this study, the position of deontological ethics was also found to be the overarching theme grounding the respective values as well as the dominant discourse in relation to the ethical aspects.

4. Discussion

4.1. Who can decide on matters of life or death?

One of the central issues in using AWS is the substitution of a combatant with an artificial agent. This creates multiple dilemmas that, as the results of the review show, have no easy solutions and may even require us to completely rethink our current ethical values. Appropriate levels of human judgment or sufficient human control are intertwined with the problematic aspects of AWS discussed in the review. It is hard to accept that control over life and death decisions could be delegated to an artificial agent. Heyns argues that control is followed by legal responsibility and without MHC, there cannot be meaningful accountability. Thus, the notion of MHC in the context of AWS seems paradoxical – as long as there is MHC, there cannot be full autonomy[30].

---

[30]   **Bhuta, N.; Beck, S.; Geiss, R.; Liu, H-Y.; Kress, C**. 2016. Autonomous Weapons Systems: Law, Ethics, Policy. Cambridge University Press.

Moreover, life and death decisions include responsibility, which is simply not conceivable without human control. According to Sharkey, some people see the right to abortion and euthanasia as a representation of human worth and dignity, whereas others see them as an affront to human dignity, so the meanings attributed to dignity can vary according to context. This ethical discretion becomes highly acute when an artificial agent is authorised to use lethal force in the context of an armed conflict. However, even if we consider the law and the risk of a minimal liability gap, we must still achieve a consensus on whether an artificial agent's right to kill is morally acceptable. In short, responsibility, human control, and humanity are strongly interlinked: these issues cannot be resolved separately and must be addressed simultaneously.

A delegation of life-and-death decisions to AWS raises a fundamental ethical question about humanity and poses new challenges for the law and human control over weaponised AI. With this in mind, Melancon (2020) argues that a debate on accountability is focused on *who* is responsible for the decision to use force, ignoring another key aspect of the use of force: the targeting criteria on which the decision to use force is based[31]. This makes a strong case for argumenting on the use of deadly force and should be the core of the AWS debate for the following reasons.

Positive target identification has a central role in the decision-making process of whether or not to use lethal force. Any life and death decision made by a human is irrelevant if AWS already have the licence to decide whether to legitimise a military target. In other words, a person does not have to press the button if they already gave consent for positive target identification to AWS. Following this distinction, it no longer matters whether a person or AWS launch the attack and acquire the target if the use of force is limited to the degree, intensity, and duration necessary to remove the threat. The great difficulty lies not in the actual decision-making but determining whether an attack on a military target can meet the IHL principle of distinction. If the consequences of an attack are unlawful, then criminal liability must ensue, assessed by identifying whether the target was a legitimate military object. Signature strikes illustrate this complex issue[32].

---

[31]   **Melancon, A-A**. 2020. What's Wrong with Drones? Automatization and Target Selection. – Small Wars and Insurgencies, Vol. 31, Issue 4, pp. 801–821. https://doi.org/10.1080/09592318. 2020.1743486. [**Melancon** 2020]

[32]   **Melancon** 2020.

4.2. Combatants vs. AWS

Autonomy is followed by accountability as legislations seek to systematise and enforce value judgments on human interaction. Humans exercise autonomy by enacting decisions that affect their lives. If humans are deprived of the ability to enact a decision then their autonomy is compromised[33]. Any level of unsupervised autonomy for machines comes at the expense of human autonomy[34]. Soltanzadeh et al. suggest that artificial moral agents with customised settings of ethics might make it possible for users to make morally significant decisions. But these settings of ethics would not be fixed in the algorithms but, instead, be incorporated into the decision-making mechanisms as higher-order commands[35]. De Lucia Dahlback draws from the example of deep learning by arguing that there is no substantial difference between AI and human intelligence. An attribution of a mind to a machine, weapon, or any other device capable of learning subsequently implies the attribution of all of the inherent properties (emotional and rational) of the human mind[36].

This brings up the question of whether AWS can be trained in military ethics just as soldiers are (context-based rules of engagement). If each soldier applied their knowledge of military ethics in an individual manner, then it may be impossible to train AWS to replicate these modes of ethical reasoning. However, AWS do not necessarily have to be seen as the end of humanity because at least as many soldiers lack situational awareness and empathy or compassion. The question comes down to whether or not it is possible to create machines with human-like cognitive capacities. If we are unable to find an answer then the entire discipline of AI ethics may be irrelevant[37]. Still, regardless of ethical training, it is necessary to consider the possibility that AWS could exercise the same level of autonomy as a military commander. It is essential to understand that AWS (when they are applied) are not just

---

[33] **Soltanzadeh, S.; Galliott, J.; Jevglevskaja, N**. 2020. Customizable Ethics Settings for Building Resilience and Narrowing the Responsibility Gap: Case Studies in the Socio-Ethical Engineering of Autonomous Systems. – Science and Engineering Ethics, Vol. 26, Issue 5, pp. 2693–2708. https://doi.org/10.1007/s11948-020-00221-5. [**Soltanzadeh et al**. 2020]

[34] **Gómez de Ágreda** 2020.

[35] **Soltanzadeh et al**. 2020.

[36] **De Lucia Dahlbeck, M**. 2020. AI and Spinoza: A Review of Law's Conceptual Treatment of Lethal Autonomous. – AI & Society, July 10. https://doi.org/10.1007/s00146-020-01014-x.

[37] **Coeckelbergh, M**. 2020. AI Ethics. Cambridge, MA: The MIT Press.

sophisticated robots but rather, hypothetically speaking, a new form of intelligence that might not fit into human categories.

Nevertheless, Liao states that after assembling and examining the leading moral theories, it becomes clear that the use of AWS and the licence for or authorisation of autonomous violence is morally wrong[38]. Even beyond the ethical scope, there is also a collective sense that something is wrong or at least not entirely right about the use of AWS. According to Verdiesen et al., the empirical results of two studies (conducted with military personnel and civilians) examining the moral perceptions of AWS indicated that both groups were more anxious about the deployment of AWS than they were about the use of human-operated drones, and both groups felt that AWS have less respect for human life and dignity[39]. However, since human beings do not possess the analytical capabilities in terms of speed and data acquisition that AWS are able to provide in some extreme situations, they could be used as a last resort for self-defence.

The ethical dilemma here may even be more significant than *the trolley problem*. There might be an *a priori* prejudice that people prefer to be protected by AWS in a state of war rather than by fellow citizens, and in the same situation, they would most likely rather to fight against enemy troops than adversarial AWS. Let us not forget the point raised by Del Monte that, in a state of emergency, people would rather risk a machine than a human life[40].

One thing is for certain: the progress of AI will raise entirely new ethical challenges. Like many others, Hynek identifies the CCW framework of the UN as an appropriate forum for negotiating the multiple aspects of AWS. Any decision to ban or regulate AWS would mainly have to be based on GGE recommendations[41]. Finally, as was stressed earlier, the areas of concern in relation to AWS are strongly interlinked and can be grouped together in multiple ways. The issues of concern include: accountability, explainability, and the capacity for a timely intervention needed to uphold MHC. MHC

---

[38] **Liao** 2020.

[39] **Verdiesen, I.; Santoni de Sio, F.; Dignum, V**. 2019. Moral Values Related to Autonomous Weapon Systems: An Empirical Survey That Reveals Common Ground for the Ethical Debate. – IEEE Technology and Society Magazine, Vol. 38, Issue 4, pp. 34–44. https://doi.org/10.1109/MTS.2019.2948439.

[40] **Del Monte, L. A**. 2018. Genius Weapons: Artificial Intelligence, Autonomous Weaponry, and the Future of Warfare. Illustrated edition. Amherst, New York: Prometheus Books.

[41] **Hynek, N.; Solovyeva, A**. 2020. Operations of Power in Autonomous Weapon Systems: Ethical Conditions and Socio-Political Prospects. – AI & Society, Vol. 36, Issue 1, pp. 79–99. https://doi.org/10.1007/s00146-020-01048-1.

also denotes responsibility as an essential value of AWS. Although we were able to detect multiple subjects in the literature review that highlighted the problematic aspects of AWS, the analysis was limited to a single database and a small number of samples. Therefore, it remains disputable whether all of the identified problematic aspects of AWS are relevant and can be included.

5. Conclusion

In this study, a literature review was utilised to come to some empirical conclusions that would fill the void on the problematic aspects of AWS and help us to better understand the main concerns associated with introducing AWS in the battlefield. The review revealed that the principles of IHL, responsibility, life-and-death decisions, and meaningful human control are the most frequent concerns. Preliminary results of theoretical reasoning highlighted that almost half of the problematic aspects were not based on an argumentation related with ethical theories. The systematic approach revealed that the deontological ethical position gave a primary foundation for the debate on the ethical aspects of AWS, i.e., their development, use, and moral and ethical perceptions.

References

Altmann, J.; Sauer, F. 2017. Autonomous Weapon Systems and Strategic Stability. – Survival. Global Politics and Strategy, Vol. 59, Issue 5, pp. 117–142.
https://doi.org/10.1080/00396338.2017.1375263.

Barbé, E.; Badell, D. 2020. The European Union and Lethal Autonomous Weapons Systems: United in Diversity? – European Union Contested: Foreign Policy in a New Global Context. Johansson-Nogues, E.; Vlaskamp, M. C.; Barbé, E. (eds.). Cham: Springer International Publishing AG, pp. 133–152.
https://doi.org/10.1007/978-3-030-33238-9_8.

Bauer, W. A. 2020. Virtuous vs. Utilitarian Artificial Moral Agents. – AI & Society, Vol. 35, Issue 1, pp. 263–271.
https://doi.org/10.1007/s00146-018-0871-3.

Bhuta, N.; Beck, S.; Geiss, R.; Liu, H-Y.; Kress, C. 2016. Autonomous Weapons Systems: Law, Ethics, Policy. Cambridge University Press.

Coeckelbergh, M. 2020. AI Ethics. Cambridge, MA: The MIT Press.

De Lucia Dahlbeck, M. 2020. AI and Spinoza: A Review of Law's Conceptual Treatment of Lethal Autonomous. – AI & Society, July 10.
https://doi.org/10.1007/s00146-020-01014-x.

Del Monte, L. A. 2018. Genius Weapons: Artificial Intelligence, Autonomous Weaponry, and the Future of Warfare. Illustrated edition. Amherst, New York: Prometheus Books.

Etzioni, A. 2018. Pros and Cons of Autonomous Weapons Systems (with Oren Etzioni). – Happiness Is the Wrong Metric: A Liberal Communitarian Response to Populism, Vol. 11, pp. 253–263. Dordrecht: Springer.
https://doi.org/10.1007/978-3-319-69623-2_16.

Gómez de Ágreda, A. 2020. Ethics of Autonomous Weapons Systems and Its Applicability to Any AI Systems. – Telecommunications Policy, Vol. 44, Issue 6, 101953.
https://doi.org/10.1016/j.telpol.2020.101953.

Guersenzvaig, A. 2018. Autonomous Weapon Systems Failing the Principle of Discrimination. – IEEE Technology and Society Magazine, Vol. 37, Issue 1, pp. 55–61.
https://doi.org/10.1109/MTS.2018.2795119.

Heyns, Chr. 2016. Human Rights and the Use of Autonomous Weapons Systems (AWS) During Domestic Law Enforcement. – Human Rights Quarterly, Vol. 38, Issue 2, pp. 350–378.
https://doi.org/10.1353/hrq.2016.0034.

Huelss, H. 2020. Norms Are What Machines Make of Them: Autonomous Weapons Systems and the Normative Implications of Human-Machine Interactions. – International Political Sociology, Vol. 14, Issue 2, pp. 111–128.
https://doi.org/10.1093/ips/olz023.

Hynek, N.; Solovyeva, A. 2020. Operations of Power in Autonomous Weapon Systems: Ethical Conditions and Socio-Political Prospects. – AI & Society, Vol. 36, Issue 1, pp. 79–99.
https://doi.org/10.1007/s00146-020-01048-1.

Jones, E. 2018. A Posthuman-Xenofeminist Analysis of the Discourse on Autonomous Weapons Systems and Other Killing Machines. – Australian Feminist Law Journal, Vol. 44, Issue 1, pp. 93–118.
https://doi.org/10.1080/13200968.2018.1465333.

Liao, S. M. (ed.) 2020. Ethics of Artificial Intelligence. 1st edition. New York, NY: Oxford University Press.

Melancon, A-A. 2020. What's Wrong with Drones? Automatization and Target Selection. – Small Wars and Insurgencies, Vol. 31, Issue 4, pp. 801–821.
https://doi.org/10.1080/09592318.2020.1743486.

Park, S. 2020. Analysis of the Positions Held by Countries on Legal Issues of Lethal Autonomous Weapons Systems and Proper Domestic Policy Direction of South Korea. – Korean Journal of Defense Analysis, Vol. 32, Issue 3, pp. 393–418.
https://doi.org/10.22883/kjda.2020.32.3.004.

Petticrew, M.; Roberts, H. 2006. Systematic Reviews in the Social Sciences: A Practical Guide. Blackwell Publishing.
https://doi.org/10.1002/9780470754887.

Riebe, T.; Schmid, S.; Reuter, Chr. 2020. Meaningful Human Control of Lethal Autonomous Weapon Systems: The CCW-Debate and Its Implications for VSD. – IEEE Technology and Society Magazine, Vol. 39, Issue 4, pp. 36–51.
https://doi.org/10.1109/MTS.2020.3031846.

Santoni de Sio, F.; van den Hoven, J. 2018. Meaningful Human Control over Autonomous Systems: A Philosophical Account. – Frontiers in Robotics and AI, Vol. 5, Article 15.
https://doi.org/10.3389/frobt.2018.00015.

Sehrawat, V. 2017. Autonomous Weapon System: Law of Armed Conflict (LOAC) and Other Legal Challenges. – Computer Law & Security Review, Vol. 33, Issue 1, pp. 38–56.
https://doi.org/10.1016/j.clsr.2016.11.001.

Skerker, M.; Purves, D.; Jenkins, R. 2020. Autonomous Weapons Systems and the Moral Equality of Combatants. – Ethics and Information Technology, Vol. 22, Issue 3, pp. 197–209.
https://doi.org/10.1007/s10676-020-09528-0.

Soltanzadeh, S.; Galliott, J.; Jevglevskaja, N. 2020. Customizable Ethics Settings for Building Resilience and Narrowing the Responsibility Gap: Case Studies in the Socio-Ethical Engineering of Autonomous Systems. – Science and Engineering Ethics, Vol. 26, Issue 5, pp. 2693–2708.
https://doi.org/10.1007/s11948-020-00221-5.

Taylor, I. 2020. Who Is Responsible for Killer Robots? Autonomous Weapons, Group Agency, and the Military-Industrial Complex. – Journal of Applied Philosophy, Vol. 38, Issue 2, pp. 320–334.
https://doi.org/10.1111/japp.12469.

Umbrello, S.; Torres, P.; De Bellis, A. F. 2020. The Future of War: Could Lethal Autonomous Weapons Make Conflict More Ethical? – AI & Society, Vol. 35, Issue 1, pp. 273–282.
https://doi.org/10.1007/s00146-019-00879-x.

US Department of Defense 2012. Directive 3000.09: Autonomy in Weapon Systems. November 21. Homeland Security Digital Library. Department of Defense.
https://www.esd.whs.mil/portals/54/documents/dd/issuances/dodd/300009p.pdf.

Verdiesen, I.; Santoni de Sio, F.; Dignum, V. 2019. Moral Values Related to Autonomous Weapon Systems: An Empirical Survey That Reveals Common Ground for the Ethical Debate. – IEEE Technology and Society Magazine, Vol. 38, Issue 4, pp. 34–44.
https://doi.org/10.1109/MTS.2019.2948439.

Warren, A.; Hillas, A. 2020. Friend or Frenemy? The Role of Trust in Human-Machine Teaming and Lethal Autonomous Weapons Systems. – Small Wars and Insurgencies, Vol. 31, Issue 4, pp. 822–850.
https://doi.org/10.1080/09592318.2020.1743485.

Annex 1.
The list of publications for systematic literature review (https://datadoi.ee/handle/33/339).

Annex 2.
The areas of concern of systematic literature review (https://datadoi.ee/handle/33/426).

Major **JANAR PEKAREV**, MA
Junior Research Fellow at the Department of Applied Research, Estonian Military Academy