http://dx.doi.org/10.12775/CLR.2025.010

Igor Vuletić*

ACCOUNTABILITY IN AUTONOMOUS COMBAT: ADDRESSING NEGLIGENCE IN AI WEAPON DEPLOYMENT

Abstract

This chapter advocates the creation of a new criminal offence, Negligent Deployment of Autonomous Weapon Systems, to address the legal and ethical challenges presented by the use of fully AI weapons. This proposed offence seeks to establish clear accountability for individuals and entities responsible for the design, deployment, and operational control of fully AI weapons where their negligence results in unlawful harm or poses a substantial risk of such harm. The existing body of work has primarily focused on the issue of command responsibility and mens rea in relation to military personnel who utilize such systems. The literature has emphasized the need to further explore the question of liability for negligence on the part of manufacturers and developers. This paper seeks to contribute to addressing this legal gap by proposing the introduction of a new criminal offence. Given the importance of this topic for the international community and its potentially far-reaching consequences, the author advocates the harmonization of national criminal laws on this matter and the universal adoption of this or a comparable criminal offence.

Keywords

Intent; negligence; causality; omissions; liability; commander; weapons

* Full Professor of Criminal Law at Josip Juraj Strossmayer University of Osijek, Faculty of Law, Croatia, https://orcid.org/0000-0001-5472-5478, email: vuleticigor600@gmail.com; ivuletic@pravos.hr.

Introduction

The twenty-first century has been marked by a technological revolution characterized by the increasingly advanced development of systems based on autonomous artificial intelligence (AI). The presence of such technologies is evident in nearly all aspects of modern life and, as such, intersects with a vast majority of legal fields. This dynamic is also apparent within criminal justice systems, affecting all major legal frameworks, including those of EU member states, which have, in recent years, intensified their efforts to design appropriate legal frameworks.[1]

The rapid advancement of AI in military applications, particularly in the development of autonomous weapon systems (AWS), has significantly outpaced the evolution of the legal frameworks required to regulate their use. Owing to the substantial tactical advantages these systems offer on the battlefield, leading global military powers are increasingly adopting such weaponry, often with little regard for initiatives like that of the United Nations (UN), which advocates for the limitation or even the outright cessation of further development of autonomous weapons. Under the auspices of the UN, it is emphasized that such weaponry, even if it can be programmed in accordance with international humanitarian law, inherently lacks an ethical component in decision-making. Therefore, it is stressed that an algorithm should never be allowed to have full control over decisions that could result in human casualties. [2]

Given the available information, it remains difficult to accurately assess the extent of human oversight or control over these systems. The details concerning both existing and emerging technologies are largely cloaked in military secrecy, rendering them inaccessible to the public. Experts in this field concur that the involvement of a human operator does not inherently ensure the safe management of such systems.[3] Indeed, human involvement may even heighten the risk, particularly if

---

[1] A. Alqatawna, "Utilizing Artificial Intelligence (AI) in Criminal Justice and Policing", *Comparative Law Review*, Vol. 30, 3 December 2024, p. 10.

[2] https://press.un.org/en/2023/gadis3731.doc.htm [last accessed 31.08.2024].

[3] B. Dresp-Langley, "The Weaponization of Artificial Intelligence: What the Public Needs to Be Aware Of", *Frontiers in Artificial Intelligence*, Vol. 6, 2023, Article 1154184.

the operator lacks sufficient training or if the information provided to the system is unclear or overly complex.

The issue of accountability for breaches of international humanitarian law committed through the use of autonomous weapons is becoming an increasingly prominent topic of debate. The critical question is: who bears responsibility when serious violations of the laws of war are committed by autonomous systems?[4] While the management of military operations remains predominantly a human task, it is noteworthy that at least thirty military forces worldwide are already employing so-called supervised autonomous weapons.[5] These systems enable an autonomous platform to undertake functions such as searching, identifying, tracking, and prioritizing targets, with the human operator making the final decision based on the information the system provides.

However, if the final decision made by the operator leads to unintended consequences, such as civilian casualties or the destruction of civilian objects, this raises a host of complex legal questions. The issue of accountability in such situations is often tied to the "many hands" problem, where it becomes difficult to identify who is ultimately responsible when multiple actors, including both humans and autonomous systems, are involved in the decision-making process. The existing legal frameworks are not yet equipped to adequately address this challenge, as they lack a solid foundation for determining culpability in scenarios involving autonomous systems.

In the current context, while humans continue to play a pivotal role in military operations, it is becoming increasingly evident that the significance of autonomous systems is on the rise. This growing reliance on autonomous technologies inevitably prompts critical questions about the boundaries of human control and responsibility. Although autonomous weapons provide certain tactical advantages, such as enhanced

---

[4] P. Gaeta, "Who Acts When Autonomous Weapons Strike? The Act Requirement for Individual Criminal Responsibility and State Responsibility", *Journal of International Criminal Justice*, Vol. 21, Issue 5, 2023, pp. 1033–1055.

[5] P. Scharre, *Autonomous Weapons and the Future of War. Army of None*, W. W. Northon & Company, 2016, pp. 51–52, available at: https://ftp.idu.ac.id/wp-content/uploads/ebook/tdg/MILITARY%20PLATFORM%20DESIGN/Army%20of%20None%20Autonomous%20Weapons%20and%20the%20Future%20of%20War.pdf [last accessed 18.03.2025].

reaction speed and precision, their deployment also introduces a range of intricate technical and legal challenges.

These challenges are becoming more acute as autonomous systems are progressively integrated into military operations across the globe. It is clear that further technological advancements will only exacerbate these issues, requiring the international community to address the complexities posed by autonomous weaponry. This will necessitate, not only technological regulation, but also a reevaluation of how fundamental human rights and ethical principles are upheld in the context of modern warfare. Thus, it is imperative that the ongoing debate surrounding autonomous weapons continues to evolve, with a focus on developing solutions that ensure robust control mechanisms and clear accountability for their use. The future of warfare may increasingly depend on our ability to reconcile the advantages of autonomous systems with the need to maintain human oversight and ethical standards.

The majority of existing scientific papers has primarily focused on the issue of command responsibility and *mens rea* in relation to military personnel who utilize such systems. Some legal scholars have also considered the issue of omission, specifically whether it can constitute the *actus reus* of a war crime.[6] However, the literature has emphasized the need to further explore the question of liability for negligence on the part of manufacturers and developers.[7] This paper seeks to contribute to addressing this legal gap by proposing the introduction of a new criminal offence titled "The Negligent Deployment of Autonomous Weapon Systems (NDAWS)".

---

[6] M. Bo, "Criminal Responsibility by Omission for Failures to Stop Autonomous Weapon Systems", *Journal of International Criminal Justice*, Vol. 21, Issue 5, 2023, pp. 1057–1075. See also Y. Gunawan, M.H. Aulawi, R. Anggriawan, and T.A. Putro, "Command responsibility of autonomous weapons under international humanitarian law", *Cogent Social Sciences* 8(1), 2022, pp. 1–16; see also V. Sehrawat, "Autonomous weapon system and command responsibility", *Florida Journal of International Law*, 31(3), 2019, p. 316–337. See also J. Kraska, "Command Accountability for AI Weapon Systems in the Law of Armed Conflict", *International Law Studies*, 97(1), 2021, p. 408–445; see also G. Acquaviva, "Autonomous weapons systems controlled by Artificial Intelligence: a conceptual roadmap for international criminal responsibility", *The Military Law and the Law of War Review*, 60(1), 2022, pp. 89–121; etc.

[7] A. Matthias, "The Responsibility Gap: Ascribing Responsibility for the Actions of Learning Automata", *Ethics and Information Technology*, Vol. 6, Issue 3, 2004, pp. 175–183.

This paper employs a predominantly legal and analytical methodology to examine liability gaps in the negligent deployment of autonomous weapon systems (AWS). Doctrinal legal research is conducted by analysing international and national legal frameworks, including the Rome Statute and various legal traditions. A comparative legal analysis contrasts common law and civil law approaches to negligence and command responsibility, while case law analysis draws on judicial decisions from international and national courts. Theoretical frameworks, such as *mens rea* and duty of care, are explored to establish a legal basis for a new criminal offence. Additionally, a hypothetical case study illustrates a foreseeability-based *mens rea* test, and a policy and regulatory analysis assesses the broader implications of criminal liability on technological development and international legal harmonization.

I. Problem Setting: the Liability Gap of Programmers, Manufacturers and Distributors

The deployment of AWS – machines that can select and engage targets without direct human intervention[8] – complicates the attribution of responsibility when these systems cause harm. The criminal liability of those who create and deploy such systems, specifically programmers and manufacturers, remains a gray area. This gap arises from the difficulty in applying traditional legal principles to autonomous technologies, the complexity of causation in AI, and the ambiguous role of intent in systems that make decisions independently.

Traditional criminal liability requires both *actus reus* (a guilty act) and *mens rea* (a guilty mind). For someone to be criminally liable, they must not only commit the act, but also do so with the requisite mental state – whether it be intent, knowledge, recklessness, or negligence. Undoubtedly, there are significant differences in the interpretation of the

[8] US Department of Defense, Directive No. 3000.09, 2017, pp. 13–14, https://www.esd.whs.mil/por-tals/54/documents/dd/issuances/dodd/300009p.pdf [last accessed 25.09.2024].

aforementioned categories of guilt between European continental law and common law, specifically Anglo-American law.[9]

The statutes of International Criminal Tribunal for the former Yugoslavia (ICTY) and International Criminal Tribunal for Rwanda (ICTR), as well as the Rome Statute of the International Criminal Court (ICC), uphold this standard, but include the requirement that a commander possesses a certain degree of *mens rea*, meaning they knew or should have known about their subordinates' actions. The Rome Statute sets a narrower scope of liability for civil commanders compared to military commanders. Military commanders can be held accountable for unconscious negligence ("should have known"), while civil commanders must be aware of all relevant circumstances and wilfully neglect their responsibilities.[10] Some national legal systems provide more lenient punishments for negligent command responsibility owing to the fundamental distinction between intentional and negligent offences. For instance, German and Croatian criminal laws treat negligence as a less severe form of command responsibility, aligning with the principles of continental European criminal law. This approach differs significantly from international criminal law, often sparking criticism.[11]

From the perspective of *mens rea*, a negligent commander is expected to know that his/her subordinates are preparing to commit crimes. However, this level of awareness often creates theoretical and practical challenges and is sometimes difficult to prove. Generally, negligence occurs when the individual is unaware of the circumstances underlying a criminal act, but could or should have been aware under the expected standard of care. Negligence represents a breach of due care, combining both objective (standard for any reasonable person) and subjective (specific to the individual's role) elements.[12]

Determining this type of *mens rea* in cases of command responsibility is particularly challenging. Civil law and common law systems differ

---

[9] For more details *see* e.g. T. Fleiner, "Common Law and Continental Law: Two Legal Systems", *Institute of Federalism*, 2005, pp. 5–7.

[10] I. Vuletić, "Rethinking Command Responsibility in the Context of Emerging AI Weapons", *EU and Comparative Law Issues and Challenges Series (ECLIC)*, Vol. 7, 2023, p. 170.

[11] Ibid.

[12] H.-H. Jescheck, T. Weigend, *Lehrbuch des Strafrechts. Allgemeiner Teil*, Duncker & Humblot, Berlin, 1996, pp. 577–582.

in how they conceptualize negligence. In some common law jurisdictions, negligence is further divided into ordinary, gross, and criminal negligence, while other jurisdictions do not make these distinctions. Civil law systems, on the other hand, use distinct terminology, separating *dolus* (intent) from *culpa* (negligence), both of which include subcategories. Terminology differences are especially notable with terms like *dolus eventualis*, which can overlap with both recklessness and intent.[13] Legal scholars have sought to harmonize liability standards to bridge these gaps across systems.[14]

Case studies in national courts reveal how liability for negligence can sometimes be excluded, with some interpretations limiting command responsibility to *dolus eventualis*. Broader interpretations also consider whether commanders accounted for the characteristics of their subordinates during recruitment, such as education, experience, and potential motivations for revenge (e.g., personal losses during the war). This creates a standard akin to the breach of due care, forming a basis for negligence in civil law.[15]

International criminal law, which merges common law and civil law principles (with a stronger influence from common law), has diverse interpretations of these standards. *Ad hoc* tribunals for Yugoslavia and Rwanda use the phrase "had reason to know," while the Rome Statute employs "should have known," a slightly different formulation. Some rulings from *ad hoc* tribunals, like the ICTR's judgment in Bagilishema, rejected negligence as a basis for command responsibility, arguing that it could create conceptual confusion.[16] The ICTY endorsed this stance in the Blaškić case.[17] Conversely, the ICC explicitly recognises that the "should have known" standard involves negligence and imposes a duty on commanders to remain informed about their subordinates' activi-

---

[13] Vuletić, *ibid.*, p. 171.

[14] J. Blomsma, "Fault Elements in EU Criminal Law: The Case for Recklessness", in A. Klip (ed.), *Substantive Criminal Law of the European Union*, Maklu, 2011, pp. 139–159.

[15] Vuletić, *ibid.*, p. 172.

[16] *Prosecutor v. Bagilishema*, Appeals Decision Reasons (ICTR), Case No. ICTR-95-1A, 3 July 3 2002, para. 32– 37 For further comments, see J. S. Martinez, "Understanding Mens Rea in Command Responsibility", *Journal of International Criminal Justice,* Vol. 5, No. 3, 2007, pp. 647–660.

[17] *Prosecutor v. Tihomir Blaškić*, Appeals Chamber Judgment (ICTY), Case No. IT – 95 – 14 – A, 29 Jul 2004, para. 63.

ties. Breaching this duty establishes command responsibility, which is supported in academic literature.[18] In summary, the negligent standard for command responsibility remains contentious, ambiguous, and challenging to substantiate in practice.

When it comes to AWS, these concepts become difficult to apply. Programmers and manufacturers do not directly commit the harmful acts caused by AWS. The harmful action is instead executed by an autonomous system, distancing the human actors from the actual harm. AI systems, particularly those designed for military use, operate with varying levels of autonomy.[19] In highly autonomous systems, AI can make decisions without human intervention, raising questions about who is ultimately responsible for the decisions made by the machine. The traditional legal concept of agency, which relies on direct human control, is ill-suited for AI systems that operate independently. This separation between the human creator and the autonomous agent introduces a significant barrier to applying existing legal doctrines to hold programmers and manufacturers criminally liable.

AI systems, especially those used in autonomous weapons, are built through complex layers of decision-making algorithms. Programmers write the code, but the AI system may learn and evolve over time through machine learning, further distancing the end result from the original programming. The causal link between a programmer's code and the AI system's eventual decision to engage a target can become obscured.[20] The problem of causation is compounded by the possibility of emergent behaviour – actions taken by AI systems that were not explicitly programmed or anticipated by the developers. AI systems are often the product of collaboration among many actors, including software developers, hardware engineers, military operators, and commanders. This distribution of responsibilities across multiple actors complicates the attribution of criminal liability. If an AI weapon malfunctions or makes an erroneous decision, it may be difficult to determine who

---

[18] C. Meloni, *Command Responsibility in International Criminal Law*, TMC Asser Press, Den Haag, 2010, pp. 183–184.

[19] V. Boulanin & M. Verbruggen, *Mapping the Development of Autonomy in Weapon Systems*, "SIPRI", Solna, Sweden, 2017, pp. 36–54.

[20] For further details on the functioning of AWS and the types of such systems *see* e.g. Vuletić, *ibid*., pp. 165–169.

among these actors is legally responsible. The involvement of many parties creates a diffusion of responsibility, making it challenging to pin down criminal liability on any one individual or group.

One of the central challenges in holding programmers and manufacturers criminally liable is the question of intent. Criminal law typically requires that the defendant acted with a certain state of mind – intent to cause harm, knowledge of potential harm, recklessness, or negligence. In the context of AI weapons, programmers and manufacturers generally do not intend to cause harm directly; instead, they design systems that are intended to perform tasks autonomously, which may include the use of force. If harm occurs as a result of the AI system's decisions, the programmers and manufacturers may not have had any intent or direct knowledge that their actions would lead to that specific outcome. Negligence and recklessness differ primarily in the level of awareness and intent involved in the wrongful conduct. Negligence occurs when a person fails to exercise the care that a reasonable person would in a similar situation, resulting in harm. It is unintentional, as the person does not, but should have, realized the risk their actions pose. In contrast, recklessness involves a conscious disregard of a substantial and unjustifiable risk, where the person is aware of the danger, but chooses to ignore it. For example, negligence might involve a driver causing an accident because they were distracted and failed to notice a red light, whereas recklessness could involve a driver speeding through a crowded school zone, fully aware of the potential harm, but acting anyway. While both result in harm, recklessness carries a higher degree of culpability owing to the deliberate disregard for safety.[21]

While negligence and recklessness are lower standards of *mens rea* than intent, they still pose challenges in the context of AI development. For example, to prove negligence, one must show that the defendant failed to exercise reasonable care in a situation where harm was foreseeable. However, the unpredictable nature of AI systems makes it difficult to determine what constitutes reasonable care. AI weapons can make decisions in real-time, often in complex and dynamic environments, making it nearly impossible for developers to foresee every po-

---

[21] J. B. Brady, "Recklessness, Negligence, Indifference, and Awareness", *The Modern Law Review*, Vol. 43, Issue 4, 1980, pp. 381–399.

tential harmful outcome.[22] Proving recklessness would require showing that the programmer or manufacturer acted with a conscious disregard for the risks involved, which may be difficult to establish when dealing with sophisticated AI systems whose risks are not fully understood even by their creators.[23]

The rapid advancement of AI technology has outpaced the development of legal standards and regulations, particularly in the realm of criminal liability. There are currently no comprehensive legal frameworks that specifically address the criminal liability of programmers and manufacturers of AWS. This regulatory gap leaves courts without clear guidance on how to approach cases involving harm caused by AI systems.[24] Without explicit laws or regulations, it becomes difficult to establish a legal basis for holding individuals or companies criminally liable for the actions of autonomous systems. Furthermore, the novelty of AI technology means that there are few, if any, legal precedents that courts can rely on when dealing with cases involving AWS. The lack of precedent creates uncertainty in how legal principles should be applied to AI-related cases. Judges may be reluctant to extend criminal liability to programmers and manufacturers in the absence of clear legal guidance, fearing that such rulings could have far-reaching and unintended consequences.

Finally, imposing criminal liability on programmers and manufacturers could also have broader policy implications for innovation in AI technology.[25] Fear of criminal prosecution could stifle innovation, particularly in fields like defence and security, where the development of advanced AI systems is seen as a priority. Policymakers must balance the need to hold individuals accountable for harm caused by AI weapons with the need to encourage innovation and technological progress. Overly broad criminal liability could deter skilled professionals from

---

[22] M. Bo, "Are Programmers In or 'Out of' Control? The Individual Criminal Responsibility of Programmers of Autonomous Weapons and Self-Driving Cars", in S. Gless & H. Whalen-Bridge, (eds), *Human-Robot Interaction in Law and its Narratives: Legal Blame, Criminal Law, and Procedure,* Cambridge University Press, 2022, pp. 15–17.

[23] Ibid.

[24] T. Chengeta, "Autonomous Weapon Systems and the Inadequacies of Existing Law: The Case for a New Treaty", *Journal of Law & Cyber Warfare*, Vol. 8, 2020, p. 104.

[25] D. J. Baker & P. H. Robinson, "Emerging Technologies and the Criminal Law", *Artificial Intelligence and the Law*, 2020, pp. 1–30.

working in AI development, leading to slower progress in a field that is critical for national security.

From a corporate perspective, manufacturers could face legal consequences if their AWS fail owing to negligent programming, insufficient testing, or inadequate safety mechanisms. In such cases, the company itself could be held criminally liable, resulting in financial penalties, restrictions on its ability to produce or sell military technology, or even forced dissolution in extreme circumstances. However, corporate liability alone is often insufficient to ensure meaningful accountability. Holding individual executives, engineers, and compliance officers responsible would help prevent companies from treating financial penalties as a mere cost of doing business. Executives and decision-makers, particularly CEOs, CTOs, and board members, play a crucial role in setting corporate priorities. If a company knowingly rushes an AWS product to market without proper safety checks, these individuals could be liable for their failure to exercise due diligence. Similarly, lead engineers and product managers involved in programming and testing AWS could face responsibility if they ignored foreseeable risks or designed systems without adequate fail-safes. Compliance officers and internal review committees may also bear liability if they failed to flag legal or ethical concerns about the product's capabilities. However, Establishing a direct causal link between a manufacturer's negligence and harm caused by an AWS can be difficult, particularly if the system's AI evolves beyond its original programming. Courts will need clear legal tests to determine foreseeability and due diligence in AWS deployment.

II. Problem Solution: New Criminal Offence

Enforcing liability requires a robust legal framework that can effectively attribute responsibility within a corporate structure. A corporate criminal liability model similar to those found in financial and environmental regulations could be applied. The legal framework proposed in this article is primarily directed towards international criminal law, with a particular focus on potential amendments to the Rome Statute. This approach is rooted in the idea that international legal instruments, rather than fragmented national regulations, offer the most effective means

of addressing liability for the negligent deployment of AWS. By integrating negligence-based liability into the Rome Statute, the framework would create a unified standard applicable across different legal traditions, thereby resolving inconsistencies between common law and civil law approaches to culpability. While some national jurisdictions recognise criminal negligence, others adhere strictly to intent-based liability, as currently reflected in Article 30 of the Rome Statute. A harmonized international solution would ensure that liability for AWS-related harm is not subject to jurisdictional gaps or conflicting domestic interpretations. Moreover, embedding this framework within international criminal law would facilitate enforcement through institutions like the International Criminal Court (ICC), ensuring accountability even in cases where domestic legal systems are unwilling or unable to prosecute. By advocating for a legal amendment at the international level, the proposed framework aims to establish clear, universally applicable principles for holding AI manufacturers and military commanders accountable in the context of autonomous warfare.

The proposal for a new criminal offence (NDAWS), raises an important legal challenge when analysed in the light of Article 30 of the Rome Statute, which establishes intent as the primary standard of criminal responsibility in international law. Article 30 states that a person is criminally responsible only if they engage in conduct with intent and knowledge – meaning they must either intend to cause a particular consequence or be aware that it will occur in the ordinary course of events. This provision sets a high threshold for liability and largely excludes negligence as a sufficient *mens rea* for international crimes.[26] The NDAWS proposal, however, seeks to introduce negligence-based liability for programmers, manufacturers, and military personnel who fail to exercise due diligence in the deployment of AWS. This represents a significant departure from the traditional standard in international criminal law, as it would impose criminal responsibility on actors who did not necessarily intend to cause harm, but failed to prevent foreseeable risks. While such an approach aligns with domestic legal systems that recognise negligence-based criminal liability, its compatibility with the

[26] W. A. Schabas, *The International Criminal Court: A Commentary on the Rome Statute*, Oxford University Press, 2010, pp. 472–480.

Rome Statute remains uncertain. One possible way to reconcile this conflict would be to push for an amendment to the Statute or to advocate for a parallel legal framework, such as a treaty-based mechanism, that explicitly recognises negligence as a basis for liability in the context of autonomous weapons. However, such efforts would likely face resistance from states that prefer the existing intent-based standard, particularly in military and defence contexts where proving intent is already a contentious issue.

The liability of AI manufacturers and command liability under the Rome Statute differ fundamentally in their legal foundations, scope, and standards of culpability. While both relate to responsibility for harmful outcomes, they operate under distinct legal doctrines with different evidentiary and conceptual requirements.

Liability of AI Manufacturers is primarily a matter of product liability, negligence, and corporate criminal responsibility. It concerns those who design, develop, and distribute autonomous weapon systems (AWS). Under this framework, responsibility arises from a failure to foresee and mitigate foreseeable risks associated with the deployment of AI-powered weapons. This form of liability does not require direct involvement in the act causing harm, but instead focuses on whether the manufacturer took reasonable steps to ensure that AWS function within acceptable legal and ethical boundaries. The NDAWS proposal in the text suggests extending criminal liability to AI developers and manufacturers based on negligence – meaning that even in the absence of intent, individuals or corporations could be held accountable if they failed to take necessary precautions to prevent unlawful harm.

In contrast, command liability under the Rome Statute (Article 28) applies to military commanders and civilian superiors who fail to prevent or punish crimes committed by their subordinates. This form of liability is distinct from direct perpetration, as it is based on a superior's duty to exercise effective control over forces under their command.[27] Command responsibility does not require proof that the superior personally ordered the crime, but instead hinges on whether they "knew or should have known" that their subordinates were committing or about to commit international crimes and failed to take necessary and rea-

[27] *Ibid.*, p. 459.

sonable measures to prevent or repress them.[28] Importantly, unlike the proposed manufacturer liability, which includes negligence as a basis for culpability, the Rome Statute explicitly limits criminal responsibility under Article 30 to intent, except in the case of command liability, which allows for responsibility based on omission and constructive knowledge ("should have known" standard). A strict legal distinction, therefore, exists between these two forms of responsibility. AI manufacturers, as non-combatants, do not have a duty analogous to that of military commanders, nor do they exercise control over the deployment of AWS in battlefield conditions. Their liability would derive from failing to adhere to legal and ethical obligations in the design and production of AWS, rather than from a failure to oversee and control subordinates' actions. While command liability under international law has long-standing precedent, extending criminal liability to AWS manufacturers would require new legal mechanisms, either through amendments to the Rome Statute or the creation of an independent legal framework tailored to the challenges posed by autonomous warfare.

The foundation of criminal liability for individuals involved in the creation and implementation of algorithms, particularly in the context of AI weapons, lies in the existence of a duty of care – often referred to as a "guarantor's duty" – to prevent harm that such technology might cause to third parties.[29] This duty is primarily based on the concept of risk management, which is within the control of the responsible parties. In criminal proceedings, the first step is to ascertain on whom the burden of this duty falls and at what point this burden ceases or is transferred to another party.[30]

In terms of criminal liability, this duty of care applies equally to the programmers and manufacturers of AWS. These parties are legally obliged to prevent harmful outcomes, and criminal liability becomes relevant if they fail to conduct all necessary checks and tests, or if they make errors during the processes of production, programming,

---

[28] *Ibid*., p. 462.

[29] S. Fahim & G. S. Bajpai, "AI and Criminal Liability," *Indian Journal of Artificial Intelligence & Law*, Vol. 1, 2020, p. 64.

[30] J. K. Kingston, *Artificial Intelligence and Legal Liability, in Research and Development in Intelligent Systems XXXIII: Incorporating Applications and Innovations in Intelligent Systems XXIV*, Springer International Publishing, 2016, pp. 269–279.

or in informing distributors and the public.[31] The programmer is typically a natural person, whereas the manufacturer often involves joint responsibility (complicity) between a responsible natural person and a legal entity. Thus, the criminal liability of legal entities might be invoked under the applicable law governing the criminal liability of legal persons. The basis for liability is the criminal act committed by the responsible person. To hold a legal entity accountable, it must be proved that the responsible person violated some duty of the legal entity or that the legal entity unlawfully obtained a pecuniary benefit for itself or another party.[32]

The liability of a legal entity is predicated on the fault of the responsible person, and a unified procedure is conducted against both the legal entity and the responsible person, resulting in a single judgment. *Argumentum a contrario*, a legal entity will not be held liable if the responsible person was *non compos mentis*, if they acted under an insurmountable error, or under the influence of exculpatory reasons.[33] It is presumed that such situations will be rare in these cases. However, there may be situations where the responsible person cannot be tried for certain reasons, such as actual barriers (e.g., death or incapacity to stand trial) or legal barriers (e.g., immunity). In such cases, the proceedings may be conducted solely against the legal entity.

Criminal liability in these contexts will hinge on the presence of either conscious or unconscious negligence. In determining negligence, both objective and subjective criteria must be employed. The objective criterion of negligence involves the breach of the objective standard of care, which consists of the duty to foresee danger with the care of a diligent and reasonable person (often referred to as "internal care") and to adapt one's subsequent behaviour accordingly ("external care"). The subjective criterion involves the breach of the standard of care, which exists if the perpetrator, considering their personal characteristics (such

---

[31] S. A. S. Nunes, "Scapegoats!: Assessing the Liability of Programmers and Designers for Autonomous Weapons Systems," in: *Responsible Use of AI in Military Systems*, CRC Press and the Dutch Ministry of Defence, 2024, pp. 192–210.

[32] M. Engelhart, "Corporate Criminal Liability from a Comparative Perspective", in D. Brodowski, M. Espinoza de los Monteros de la Parra, K. Tiedemann, & J. Vogel (eds), *Regulate Corporate Criminal Liability*, Springer, 2014, pp. 53–76.

[33] Ibid.

as intelligence, education, experience, etc.), failed to foresee the danger and adjust their behaviour accordingly.[34] Negligence is established only if both the objective and subjective criteria are cumulatively satisfied.

In the context of AWS, it can be assumed that any such system must have embedded software or an operational system through which it assesses newly emerging situations not originally accounted for by the program or system. These programs or systems must include safeguards that prevent erroneous judgements and ensure the safe operation of the AWS. Should AWS cause harm owing to an inability to accurately assess a newly arising situation, criminal liability would need to be sought in the actions of the programmer who failed to create a system capable of correctly assessing the situation and preventing undesirable consequences.

Developing a *mens rea* test for courts to distinguish when harm from AWS was foreseeable involves, in my opinion, creating criteria that balance the unique characteristics of AI with established legal principles. The challenge lies in adapting traditional negligence concepts to systems that operate autonomously. Below is a proposed foreseeability-based *mens rea* test tailored for AWS deployment, focusing on negligence rather than intent.

I believe that the court should apply a three-step test to determine whether the harm caused by AWS was foreseeable, and consequently, whether the responsible parties exhibited the requisite *mens rea* (knowledge or recklessness) in deploying or failing to prevent the harm. This test consists of several components. The first of these relates to knowledge of system capabilities and limitations (objective component). Here, it is necessary to determine whether the responsible party (e.g., developer, commander, operator) possesses, or whether they have reasonably possessed knowledge of the AWS's capabilities, limitations, and potential risks? This prong uses an objective "reasonable person" standard[35] within the industry or military context. The court assesses whether a reason-

---

[34] H.-H. Jescheck & T. Weigend, *Lehrbuch des Strafrechts: Allgemeiner Teil* (5th ed., fully revised and expanded), Duncker & Humblot, 1996, p. 577. See also F. Blomsma, "Fault Elements in EU Criminal Law: The Case for Recklessness", in A. Klip (ed.), *Substantive Criminal Law of the European Union*, Maklu, 2011, pp. 139–159.

[35] S. Rane, "The Reasonable Person Standard for AI", arXiv preprint arXiv:2406.04671, 2024, available at: https://arxiv.org/html/2406.04671v1 [last accessed 19.03.2025].

able person in the defendant's position, with similar knowledge and expertise, would have been aware of the risks inherent in the AI system.

The evidence that will be of significance in this context includes, for example, the following: system design and documentation, risk assessments and testing data, industry standards and best practices, known incidents of similar AI behaviour or failures etc. If the defendant knew or should have known about the AI system's limitations or potential risks, then foreseeability is in my opinion established at this stage.

The second component relates to adequacy of risk mitigation measures (preventive component). At this point it is necessary to establish whether the responsible party has taken reasonable steps to mitigate foreseeable risks and prevent harm, given the party's knowledge of the AWS's potential to cause harm. The court therefore examines whether the party implemented sufficient preventive measures to address foreseeable risks, such as system testing, human oversight, or fail-safe mechanisms. Failure to implement reasonable safeguards could indicate reckless disregard for the risks. Here, the court can consider evidence such as: implementation of safety protocols and oversight mechanisms, documentation of testing, simulation, and risk mitigation strategies, compliance with regulatory or ethical standards for AWS deployment, the presence of any contingency plans or shutdown protocols in case of malfunction, and similar. Of course, if the defendant failed to take reasonable preventive measures, this suggests a reckless disregard for the foreseeable harm.

The next component relates to post-deployment conduct and response to known issues (reactive component). Here, the court needs to test, after deployment, whether the responsible party appropriately monitored the AI system and responded to any emerging issues or malfunctions that could lead to foreseeable harm. In other words, this step examines the party's conduct post-deployment. Even if harm was not foreseeable initially, courts assess whether the party responded reasonably to signs of malfunction, unpredictable behaviour, or early warnings that could lead to harm. Ignoring warning signs may indicate a negligent or reckless state of mind. Evidence that should be considered here among other includes: monitoring and reporting mechanisms in place during and after deployment, response to early signs of malfunction or unexpected behaviour, whether there was a failure to intervene or deac-

tivate the system when risks became apparent etc. Failure to address or mitigate known issues post-deployment suggests a heightened degree of negligence or recklessness.

In the following, I will attempt to illustrate the aforementioned through a hypothetical practical example. A military commander deploys an AI-controlled drone with knowledge that the system struggles to distinguish between combatants and civilians in certain environments. First, the court would assess whether a reasonable commander with similar training should have been aware of this limitation, potentially establishing foreseeability. Second, despite this knowledge, the commander fails to implement additional safeguards, such as human oversight or operational restrictions in civilian-heavy areas. The court examines whether reasonable preventive measures were taken to mitigate known risks. Third, once deployed, the AI drone begins showing signs of erratic behaviour, but the commander does not take any action to recall or monitor the drone more closely. If harm occurs owing to this failure to act, the court would consider whether the commander exhibited a reckless disregard for foreseeable harm. This offers a structured approach for courts to determine whether harm caused by AWS was foreseeable. It shifts the focus from traditional intent to negligence, emphasizing a party's knowledge, actions (or inactions), and responses to emerging risks. By breaking foreseeability into three components – knowledge, prevention, and response – this test provides a comprehensive framework that can account for the unique challenges posed by AWS systems.

This linkage of criminal liability to the quality of programming and design of AWS, as well as the actions or omissions of programmers and manufacturers during the development and implementation stages, is crucial. For instance, if AWS cause harm because they were unable to appropriately assess danger in a given situation, this may indicate a lapse in the programming or testing phases. The programmer has a duty to foresee various scenarios in which AWS may interact with humans or property and ensure that the system responds in a manner that minimizes the risk of harm.[36] Should it be established that the programmer neglected these obligations, he/she could be held criminally liable.

---

[36] See e.g. J. Turner, *Robot Rules: Regulating Artificial Intelligence*, 2019, pp. 81–132 (Chapter "Responsibility for AI").

Similarly, the manufacturers of AWS may face criminal liability if it is determined that they failed to ensure adequate testing before bringing the system to market, or did not properly inform distributors and the public of the associated risks.[37] If such omissions lead to harm, the manufacturer could also be held liable.

It is important to note that criminal liability for the handling of AWS may extend to distributors and users, depending on their role and level of awareness of the system. For example, a distributor who is aware of the potential risks, but nonetheless chooses to market AWS without the necessary warnings or additional safety measures might also share in the liability.[38]

One of the key challenges in the legal regulation of AWS is the fact that these systems are often highly complex and may make decisions based on algorithms that include elements of machine learning. This means that AWS can operate in ways that the programmers or manufacturers may not have directly anticipated, making it more difficult to establish liability.[39] Nonetheless, the legal system may impose liability where there is evidence that the creators and operators of AI weapons were aware of the risks, but failed to take the necessary measures to mitigate or eliminate them.

Moreover, the role of legal standards and regulatory frameworks concerning AI weapons is crucial. In some jurisdictions, there may be specific laws or regulations that govern the development, production, and use of AI weapons, which may also include provisions on criminal liability. For example, if the law requires that all AI weapons undergo certain certification procedures before being deployed, and the manufacturer neglects this process, this could lead to criminal liability in the event of harm. Conversely, there may also be liability at the international level, particularly if AI weapons cause harm beyond the borders of a single state. In such situations, responsibility under international law

---

[37] Ibid.

[38] Ibid. See also B. Zhang, "Accountability and Responsibility for AI-Enabled Conduct", in: H. Lahmann, R. Geiss (eds.), *Research Handbook on Warfare and Artificial Intelligence*, Edward Elgar Publishing, 2024, pp. 216–233.

[39] R. W. Bellaby, "Can AI Weapons Make Ethical Decisions?", *Criminal Justice Ethics*, Vol. 40, Issue 2, 2021, pp. 86–107.

may be considered, including the possibility of proceedings before international courts or tribunals.[40]

There is a growing concern about the "autonomy" of AWS and the extent to which human oversight is retained in their operation.[41] While AWS may be designed to operate with a degree of independence, it is crucial that legal standards ensure that there is always a level of human control that can intervene to prevent or mitigate harm. The concept of "meaningful human control" has been proposed as a standard that could be incorporated into legal frameworks to ensure that humans remain ultimately responsible for the actions of AWS.

Additionally, the development of AWS must be guided by principles of proportionality and necessity, which are central to international humanitarian law.[42] These principles require that the use of force be proportionate to the threat and necessary to achieve a legitimate military objective.[43] AWS must therefore be designed and programmed to comply with these principles, and any failure to do so could result in criminal liability. Moreover, the transparency and accountability of AWS are critical issues that must be addressed in legal frameworks. It is essential that the decision-making processes of AWS are transparent and that there are mechanisms in place to hold those responsible accountable for their actions. This includes ensuring that there is a clear chain of command and responsibility, as well as procedures for investigating and addressing any incidents involving AWS.

Finally, penalties under the proposed NDAWS offence would be stringent, reflecting the significant potential for harm arising from the negligent use of AWS. Criminal sanctions could encompass imprison-

---

[40] T. Weigend, "Convicting Autonomous Weapons? Criminal Responsibility of and for AWS under International Law", *Journal of International Criminal Justice*, Vol. 21, Issue 5, 2023, pp. 1137–1154.

[41] F. M. Hassan & N. D. Osman, "AI-based Autonomous Weapons and Individual Criminal Responsibility under the Rome Statute", *Journal of Digital Technologies and Law*, Issue 1(2), 2023.

[42] M. Brenneke, "Lethal Autonomous Weapon Systems and Their Compatibility with International Humanitarian Law: A Primer on the Debate", *Yearbook of International Humanitarian Law*, Vol. 21, 2018, pp. 59–98.

[43] *Ibid*. See also A. L. Schuller, "At the Crossroads of Control: The Intersection of Artificial Intelligence in Autonomous Weapon Systems with International Humanitarian Law," *Harvard National Security Journal*, Vol. 8, 2017, p. 379.

ment, fines, and prohibitions on engaging in the development or deployment of such technology. Additionally, the offence would allow for civil liability, enabling victims to seek redress for damages sustained. The offence would also extend to corporate entities, thereby holding companies accountable for negligence in the development, sale, or deployment of AI weapons. Corporate penalties could include substantial fines and operational restrictions.

To ensure fairness, the concept proposes defences such as a due diligence defence, where the accused party could demonstrate that all reasonable steps were taken to prevent the negligent outcome, including compliance with legal standards and the conduct of thorough testing.[44] Moreover, limited exceptions could be made for the deployment of AI weapons in unforeseen emergency situations, provided that such deployment was necessary and proportionate.

The concept also underscores the imperative for international cooperation and the harmonization of global standards to regulate the use of AWS. An international treaty or legal framework could be established to ensure the consistent application of the proposed offence across various jurisdictions. Furthermore, the concept advocates robust regulatory oversight, including regular audits of AWS, mandatory reporting of incidents, and stringent licensing requirements for entities involved in their development and deployment. Comprehensive training and education would also be crucial, ensuring that operators of AWS are fully cognizant of the legal and ethical responsibilities attendant to their use.

The proposed NDAWS offence represents a necessary and timely response to the evolving challenges posed by AI in warfare. By establishing clear legal standards for negligence in the deployment of AWS, the creation of this offence would address a critical gap in current legal frameworks, thereby ensuring accountability and averting unlawful harm in an increasingly autonomous battlefield.

---

[44] J. Kulesza, *Due Diligence in International Law*, Volume 26, Brill, 2016.

Conclusions

In conclusion, the criminal liability associated with AWS represents a complex and evolving area of law that requires careful consideration of both technical and ethical issues. Legal professionals and policymakers must work together to develop robust legal frameworks that ensure the safe and responsible use of AWS, while also addressing the broader moral questions they raise. As AWS become increasingly integrated into military and security operations, it is essential that the legal system adapts to meet the challenges they present, providing clear guidelines on liability and ensuring that those who develop, deploy, and operate these systems are held accountable for their actions.

The Negligent Deployment of Autonomous Weapon Systems (NDAWS) offence is anchored in traditional legal principles of negligence, adapted to accommodate the unique characteristics of AI technology. Central to the offence is the existence of a duty of care owed by military commanders, developers, and other relevant actors to ensure that AWS are utilized responsibly and within the confines of international law. A breach of this duty – whether through inadequate programming, insufficient testing, or poor oversight during deployment – would constitute the essence of the offence. Unlike offences requiring intent or recklessness, NDAWS would criminalize a failure to foresee and mitigate risks that a reasonable person in a similar position ought to have anticipated.

This offence would apply to a broad spectrum of parties, including military personnel who authorize the use of AWS, developers responsible for designing and programming these systems, and contractors involved in their deployment. The scope of the offence would be both domestic and international, ensuring accountability across different jurisdictions and extending to non-state actors engaged in the deployment of AWS.

Ultimately, criminal liability in relation to AWS necessitates a comprehensive legal analysis that takes into account various factors, including the technical aspects of the systems, the conduct of those involved in their development and production, and the legal obligations arising from relevant legal frameworks. Issues of liability will become increas-

ingly pertinent as AWS become more prevalent in military and security operations, requiring legal professionals to develop new legal doctrines and standards to effectively address the challenges posed by this advanced technology.

For these reasons, it is imperative that legal experts and legislators work intensively on developing appropriate legal frameworks capable of encompassing all aspects of liability associated with AWS. This includes not only standards for design and testing, but also clear guidelines on who bears responsibility in the event of harm, in order to ensure justice for victims and to provide incentives for the development of safer and more reliable systems.

Furthermore, it is necessary to develop international legal frameworks that will enable coordination and cooperation between states regarding the regulation and responsibility of AWS. Given the global nature of the challenges posed by AWS, it is essential that international agreements or treaties are established to set common standards and ensure accountability across borders.

To address the gap in criminal liability, governments and international organizations could develop specific regulations governing the development and deployment of AWS. These regulations could include mandatory safety standards, testing protocols, and oversight mechanisms to ensure that AI systems are designed and deployed responsibly. By establishing clear legal standards, regulators could help close the gap in criminal liability, and provide courts with the tools they need to hold individuals and companies accountable.

The gap in criminal liability for programmers and manufacturers of AWS stems from the difficulty of applying traditional legal principles to autonomous systems. The complexity of causation, the evolving nature of AI, the absence of clear regulatory standards, and ethical concerns about fairness all contribute to this challenge. Closing this gap will require a combination of legal innovation, regulatory frameworks, and international cooperation to ensure that those responsible for the development and deployment of AWS are held accountable for their actions. Without such measures, the legal system risks falling behind as technology continues to advance, leaving victims of AI-related harm without recourse and undermining public trust in the responsible development of autonomous systems.