

LETHAL AUTONOMOUS WEAPONS SYSTEMS: LEGAL AND ETHICAL DILEMMAS IN AN ERA OF ALGORITHMIC WARFARE

ARMAS AUTÔNOMAS LETAIS: DESAFIOS JURÍDICOS E ÉTICOS NA ERA DA GUERRA ALGORÍTMICA

Louise Amorim Beja¹

Universidade de Coimbra

"Only the dead have seen the end of war."

George Santayana, Soliloquies in England and Later Soliloquies (1922).

97

ABSTRACT

War has long been regulated by International Humanitarian Law (IHL) and International Criminal Law (ICL), aiming to limit violence and ensure accountability for crimes. However, the rise of artificial intelligence and autonomous systems introduces new challenges. Lethal Autonomous Weapons Systems (LAWs), capable of selecting and attacking targets without human intervention, threaten the normative foundations of the legal order. Unlike traditional or remotely operated weapons, LAWs rely on algorithms, raising ethical and legal dilemmas. Their use undermines principles of dignity, proportionality, and distinction between combatants and civilians, while complicating the attribution of responsibility for violations of IHL. From a Kantian perspective, delegating lethal decisions to machines disregards the intrinsic value of human life, collapsing the moral basis of justice. The lack of consensus on regulating LAWs has stalled treaty efforts, with technologically advanced states resisting limits. This paper examines their definition, risks, human rights impact, and proposes strengthening international regulation, including expanding the ICC's mandate.

Key-words: International Humanitarian Law; International Criminal Responsibility; Lethal Autonomous Weapons (LAWs); Ethical Dilemmas of Algorithmic Warfare.

RESUMO

A guerra sempre foi regulada pelo Direito Internacional Humanitário (DIH) e pelo Direito Penal Internacional (DPI), visando limitar a violência e responsabilizar autores de crimes. Contudo, o avanço da inteligência artificial e dos sistemas autônomos traz novos desafios. Os Sistemas de Armas Autônomas Letais (LAWs), capazes de selecionar e atacar alvos sem intervenção humana, ameaçam os fundamentos normativos da ordem jurídica internacional. Diferentes das armas tradicionais ou remotamente operadas, as LAWs dependem de algoritmos, levantando dilemas éticos e jurídicos. Seu uso compromete princípios de dignidade, proporcionalidade e distinção entre combatentes e civis, além de dificultar a atribuição de responsabilidade por violações do DIH. Sob uma perspectiva kantiana, delegar decisões letais a

¹PhD candidate in Legal and Criminal Sciences at the Faculty of Law of the University of Coimbra; Master's degree in Legal and Political Sciences with a specialisation in Public and European International Law and Specialist in Collaborative R&D&I Project Management from the same institution; Postgraduate in Armed Conflicts and Human Rights from the IGC (Human Rights Centre of the University of Coimbra) and the Military University Institute; and in Teaching and Management in Higher Education. Researcher at the University of Coimbra Institute for Legal Research (UCILeR); Scholarship holder from the Foundation for Science and Technology (2025.01733.BD); Member of the Eduardo Correia Institute; Human Rights Activist for Amnesty International Portugal, Coimbra Group. Orcid: <<https://orcid.org/0000-0003-4120-6868>>. E-mail: louisebeja@ij.uc.pt

máquinas ignora o valor intrínseco da vida humana, corroendo a base moral da justiça. A ausência de consenso sobre sua regulação trava tratados específicos, enquanto Estados tecnologicamente avançados resistem a limites. Este trabalho analisa riscos, impactos e propõe fortalecer normas internacionais, incluindo ampliar o mandato do Tribunal Penal Internacional.

Palavras-Chave: Direito Internacional Humanitário; Responsabilidade Penal Internacional; Armas Autônomas Letais (LAWs); Dilemas Éticos da Guerra Algorítmica.

INTRODUCTION

War is a social and political phenomenon that has always been subject to normative principles aimed at mitigating its destructive consequences and protecting those who are not directly involved in it. International Humanitarian Law (IHL) and International Criminal Law (ICL) have become central tools for limiting the use of force and holding perpetrators accountable for war crimes and serious human rights violations. However, the evolution of military technology, specifically in the field of artificial intelligence and autonomous systems, creates new and unique problems for the governance of war and the attribution of criminal responsibility.

One such challenge is the emergence of Lethal Autonomous Weapons Systems (LAWs), whose deployment in conflict theatres threatens the normative basis of the international legal order. LAWs are defined as weapons systems capable of selecting and engaging targets without human intervention in the decision-making process at the moment of firing. Unlike traditional weapons and even weapons operated remotely by human operators, Lethal Autonomous Weapons rely on algorithms and probabilistic models that process data in real time to optimise the execution of military objectives.

But this optimisation of lethality raises questions that go beyond the operationally imperative sphere. The autonomisation of lethal force undermines the classical conception of war as a phenomenon that can be regulated according to human rationality, which violates the principles of dignity, proportionality, and distinction between combatants and non-combatants. This, in turn, jeopardises the very concept of being able to blame and hold accountable those deemed guilty of violations of IHL, insofar as there may be systems operating without moral judgement.

From a Kantian perspective, war, although an undesirable reality, must be limited by rational principles that ensure the observance of humanity in its participants. Human dignity is an absolute value in Immanuel Kant's thinking, meaning that no

person can be treated as merely a means to external ends.² This ethical conception resonates with the legal framework that underpins both IHL and IPL, assuming that conduct during war is governed by objective and universally applicable rules. However, if lethal decisions are entrusted to machines incapable of understanding the intrinsic value of human life, then the moral basis on which the distinction between legitimate and illegitimate use of force rests collapses. Thus, the question arises: can international criminal responsibility accommodate the complexity of LAWs without distorting the fundamental principles of justice?

Faced with these uncertainties, the international community has been debating the need for a specific treaty to regulate LAWs. However, the lack of consensus on the legal status of LAWs hinders the creation of binding regulations. States with greater technological development tend to resist the imposition of limits, arguing that the use of these weapons can reduce risks to their own military personnel. However, regulatory inertia could result in an algorithmic arms race, in which war would become progressively dehumanised and less subject to legal control.

In this context, this paper analyses the legal and ethical challenges associated with LAWs from the perspective of International Criminal Law and human rights. The first chapter addresses the definition and classification of LAWs, as well as the risks inherent in their use in armed conflicts. The second chapter explores the impact of these weapons on the protection of human rights, discussing the loss of meaningful human control, the dehumanisation of conflicts, and the challenges related to criminal accountability. Finally, the third chapter presents proposals for strengthening international regulations applicable to LAWs, including the possibility of a new treaty and the expansion of the International Criminal Court's mandate to cover crimes committed by autonomous systems.

1. LETHAL AUTONOMOUS WEAPONS AND THE CHALLENGES FOR INTERNATIONAL CRIMINAL LAW

2.

²Kant, I. (1785/1997). *Fundamentação da metafísica dos costumes* [Groundwork of the Metaphysics of Morals] (P. Quintela, Trans.). Lisboa: Edições 70. Available at: <<https://id.bnportugal.gov.pt/bib/bibnacional/1874091>>, accessed on 1 March 2025 (Original work published 1785).

1.1. Definition and Classification of LAWs

Lethal Autonomous Weapons are weapon systems that operate independently of direct human intervention in the selection and neutralisation of targets. They differ from conventional drone-guided systems in that they do not require a human operator to make specific decisions in real time. This operational autonomy can be classified into different degrees, depending on the extent of human control over the system.³

The specialised literature distinguishes three main categories of LAWs: (i) human-supervised systems, where the machine performs actions upon the express authorisation of an operator; (ii) semi-autonomous systems, which have partial decision-making capacity, but with the possibility of human intervention for final approval; and (iii) fully autonomous systems, which carry out military operations without any human interference at the moment of lethal action. The latter category raises the most concerns, as it questions the very notion of morally justifiable decision-making in a conflict context, and is therefore the one that this paper seeks to examine in detail.⁴

The classification of LAWs, however, is not limited to a technical or operational problem. Kant argued that true autonomy is based on practical reason and moral self-determination.⁵ In any event, a weapon, however sophisticated, can never possess such a capacity, because it does not act in accordance with universal principles, but rather according to probabilistic calculations and pre-programmed commands. Thus, any attempt to attribute to LAWs a status analogous to that of a moral agent is a legal fiction that compromises the very foundations of criminal imputability.

1.2. Benefits and Risks Associated with the Use of LAWS

³Boulanin, V., & Verbruggen, M. (2017). *Mapping the development of autonomy in weapon systems*. Stockholm International Peace Research Institute (SIPRI). Available at: <<https://www.sipri.org/publications/2017/policy-reports/mapping-development-autonomy-weapon-systems>>, accessed on 1 March 2025.

⁴Valente, M. M. G. (2024). *A utilização de robots na guerra: A responsabilidade penal* [The use of robots in war: Criminal liability]. In M. M. G. Valente (Ed.), *Ciências Criminais e Inteligência Artificial* [Criminal Sciences and Artificial Intelligence] (p. 14). Almedina. See also: Lemos, M., & Costa, M. J. (2022). *Inteligência Artificial e Direito da Guerra: Reflexões sobre as Armas Autônomas Mortíferas* [Artificial Intelligence and the Law of War: Reflections on Lethal Autonomous Weapons]. In A. M. Rodrigues (Ed.), *A Inteligência Artificial no Direito Penal. Vol. II* [Artificial Intelligence in Criminal Law. Vol. II] (pp. 91–124). Almedina.

⁵Heyns, C. (2013). *Report of the Special Rapporteur on extrajudicial, summary or arbitrary executions*. United Nations. Available at: <<https://digitallibrary.un.org/record/755741?v=pdf>>, accessed on 01 February 2026.

Technological advances applied to the military field have historically been accompanied by arguments extolling their ability to reduce collateral damage and make war “cleaner.” In the context of Lethal Autonomous Weapons (LAWs), this rhetoric manifests itself in the belief that such systems, being free from emotions and irrational tendencies, can act with greater precision, minimizing errors and, consequently, reducing civilian and military casualties.⁶ In addition, its proponents argue that by eliminating the need for human presence in combat scenarios, LAWs could preserve the lives of soldiers and operators, reducing human exposure to unnecessary risks. However, this technocratic view of war ignores a number of structural problems that compromise the ethical and legal viability of using LAWs.

First, there is a fundamental problem related to the lack of meaningful human control. Boulanin & Verbruggen warn that removing or reducing human oversight in lethal decision-making can lead to situations where attacks are conducted based on opaque criteria, making it impossible to assign responsibility in the event of an error.⁷ Although advanced machine learning systems and predictive algorithms can optimize target selection, the impossibility of immediate human intervention compromises the ethics of lethal decision-making. By transferring the power to decide on life and death to a machine, a logic of mechanized warfare is established that undermines the moral autonomy of the human agent.

The second concern is the violation of the principle of human dignity. International humanitarian law is based on the premise that acts of war must always be conducted by agents capable of distinguishing between right and wrong.⁸ However, LAWs operate according to probabilistic rules and predictive models that do not consider the moral complexity of combat situations. A system that determines lethality based on data, without any understanding of the dignity of the other, contradicts Kant's categorical imperative, which requires that every human being be treated as an end in itself. What is at stake, therefore, is not only the technical accuracy of LAWs, but their ontological incompatibility with a legal system based on moral imputability.

⁶ Arkin, R. C. (2010). *Governing lethal behavior in autonomous robots*. CRC Press.

⁷ Boulanin, V., y Verbruggen, M. (2017). *Mapping the development of autonomy in weapon systems*. Instituto Internacional de Investigación para la Paz de Estocolmo (SIPRI). Disponible en: <<https://www.sipri.org/publications/2017/policy-reports/mapping-development-autonomy-weapon-systems>>, consultado el 1 de marzo de 2025.

⁸ Heyns, C. (2013). *Report of the Special Rapporteur on extrajudicial, summary or arbitrary executions*. United Nations. Disponible em: <https://digitallibrary.un.org/record/755741?v=pdf>, acesso em 24 de fevereiro de 2026.

Another significant risk is the possibility of errors in target identification. LAWs rely on databases and sensors to operate, but these systems are not immune to failure. The misidentification of combatants, the inability to interpret complex scenarios and susceptibility to external interference can lead to attacks on civilians or protected infrastructure. The problem is exacerbated by vulnerability to cyberattacks, as AI-based military systems can be targeted by *hacking*, being reprogrammed to act against their operators or cause unforeseeable collateral damage.

The lack of *accountability* is another key issue⁹. If an LAWs commits a serious violation of International Humanitarian Law, who should be held responsible? The programmer who wrote the algorithm? The military commander who authorised the use of the system? The state that funded its development? This lack of clarity creates a grey area where responsibility is diluted, making it difficult, if not almost impossible, to attribute blame to a specific individual.

Finally, there is the risk of conflict escalation. Sharkey observes that the proliferation of LAWs could alter the dynamics of war, making it more accessible and less costly for technologically advanced states. By reducing the number of military casualties, policy-makers may feel less pressure to avoid armed conflicts, thereby encouraging a scenario of global instability and technological asymmetry.¹⁰ This could result in an algorithmic arms race, in which military power ceases to depend on strategic capability and comes to be determined by the sophistication of AI systems.¹¹

⁹In the legal sphere, *accountability* refers to the responsibility and obligation to give an account of one's actions; in the context of Lethal Autonomous Weapons (LAWs), the absence of a direct human agent at the moment of the lethal decision calls into question the possibility of attributing blame, since the actions of these systems derive from algorithmic processes that are often opaque and unpredictable. Cf. Kesavan, Sanjay, *Autonomous Weapon System: Accountability Under International Criminal Law* (15 April 2024). Available at SSRN: <<https://ssrn.com/abstract=4819857>>, accessed 15 February 2025; See also: Jain, N. (2023). *The accountability gap in international law: Who is responsible for AI-driven war crimes?* *International Law Studies*, 99(4), 567–589.

¹⁰Sharkey, N. (2018). The impact of autonomous military systems on international peace and security. In Bhuta, N., Beck, S., Geiß, R., Liu, & H., C. Kreß (Eds.), *Autonomous weapons systems: Law, ethics, policy* (pp. 333-352). Cambridge University Press.

¹¹Despite fears of an uncontrolled escalation of conflicts due to the autonomy of weapons, the current political climate reveals a contrary trend at a global level, where political discourse has been characterised by calls to strengthen conventional armaments. In the case of Europe, this phenomenon has become particularly evident with the recent statement by Ursula von der Leyen, President of the European Commission, in which she argued for the urgent need to 'rearm Europe' and to significantly increase investment in defence due to the complex contemporary geopolitical landscape (Observador, 2025). This stance reflects the growing perception that states must prepare to face new security challenges, which may mean bolstering traditional arsenals rather than placing an unrestricted focus on the autonomisation of weaponry. (RTP, 2025). Thus, one must consider the possibility that, in a future scenario, LAWs may be viewed not as substitutes for conventional armies, but as complementary components of a broader military strategy.

Therefore, the purported benefits of LAWs reveal that their risks are not merely secondary or collateral, but structural. The replacement of human moral judgement with algorithmic processes renders the attribution of criminal responsibility unfeasible and undermines the fundamental principles of international humanitarian law. Kant maintained that morality is an inalienable condition of freedom, and that any system which disregards this prerogative is destined to dehumanise existence itself. By transferring lethal decision-making to entities devoid of conscience and a sense of justice, LAWs represent, rather than an evolution, a regression in the way conflicts are conducted.¹²

If war, however undesirable, must be regulated by norms aimed at preserving what is human even on the battlefield, then it is imperative to question to what extent the autonomisation of lethality is compatible with legal civilisation. It is not, therefore, merely a matter of discussing the effectiveness of LAWs, but of questioning whether international criminal law is prepared to deal with a form of warfare in which responsibility becomes, paradoxically, increasingly impersonal and intangible. These issues will be explored in greater detail below.

1.3. Legal Framework and Limitations of the Rome Statute

The Rome Statute of the International Criminal Court (ICC)¹³ constitutes the cornerstone of international criminal law, establishing the parameters for accountability for war crimes, crimes against humanity and genocide. However, the document reflects a legal paradigm that presupposes the existence of a human agent as the perpetrator of such crimes. The autonomisation of warfare systems, embodied by Lethal Autonomous Weapons (LAWs), challenges this traditional conception, as it shifts the decision-making regarding lethal force from humans to machines, creating a regulatory vacuum that undermines the attribution of criminal responsibility. Without specific provisions for LAWs, the Rome Statute proves inadequate to address the complexities of algorithmic warfare.

The decision-making autonomy of LAWs raises far-reaching ethical and legal issues. Firstly, the problem of algorithmic transparency known as the “black box”¹⁴

¹²Kant, I. (1785/2003). *Groundwork of the Metaphysics of Morals* (trans. P. Quintela). Ed. 70.

¹³International Criminal Court. (1998). *Rome Statute of the International Criminal Court*. United Nations. Available at: <https://www.icc-cpi.int/resource-library>, accessed on 1 March 2026.

¹⁴Or, as it is also known in the literature as the “*black box problem*” – the opacity of algorithmic systems prevents a clear reconstruction of the causal chain that led to the attack. See Sousa, S. A. (2022). “It

makes it difficult to understand how lethal decisions are processed.¹⁵ The opacity of artificial intelligence systems prevents the reconstruction of causal chains, making it virtually impossible to determine with precision whether an error stemmed from a programmer's mistake, the way the data was processed, or a structural flaw in the algorithm. This problem is not merely technical, but rather a challenge to *accountability* in International Criminal Law¹⁶.

Among the main challenges of the current legal framework, the following stand out: Algorithmic Discrimination: AI systems, far from being neutral, are conditioned by the data that feeds them. This means that if algorithms are trained with biased information, LAWs may reproduce discriminatory patterns when selecting targets. Studies show that such biases can disproportionately affect vulnerable populations¹⁷, raising concerns about systematic violations of International Humanitarian Law.

Lack of Direct Criminal Liability: The Rome Statute is structured around individual accountability for international crimes, as set out in Article 25. This provision stipulates that only individuals can be held criminally liable, which creates an impasse in cases where unlawful acts are committed by autonomous systems. As international criminal law does not provide for the possibility of holding a machine directly liable, there are also no provisions for attributing blame to developers and manufacturers of LAWs. This gap may lead to impunity, insofar as crimes committed through such systems may not have a directly liable criminal subject¹⁸. Consequently, the absence of a clearly identifiable human perpetrator poses a significant normative challenge to the administration of international justice.

wasn't me, it was the machine": Theory of crime, liability and artificial intelligence. In A. M. Rodrigues (Ed.), *Artificial Intelligence in Criminal Law* (pp. 66–68). Almedina. This raises the following question: if it is not possible to determine precisely whether an error resulted from a programmer's mistake, algorithmic programming or an unforeseeable automated decision, how can the culpability of a human agent be established? Criminal liability in the context of AALs therefore requires a broader review of the concept of imputability in international criminal law, to include technical parameters that allow us to identify at what point human supervision failed and, thus, to define legal liability.

¹⁵Marchant, G. E., Allenby, B. y Herkert, J. (2020). *La creciente complejidad de los sistemas autónomos y el problema de la rendición de cuentas*. *Ethics and Information Technology*, 22(3), 217-228.

¹⁶Jain, N. (2023). *La brecha de responsabilidad en el derecho internacional: ¿Quién es responsable de los crímenes de guerra cometidos por la IA?* *International Law Studies*, 99(4), 567-589.

¹⁷Ekelhof, M. A. (2019). Sistemas de armas autónomas letales, control humano y el estatus especial de la toma de decisiones militares. *Revista de Ética y Asuntos Internacionales*, 33(1), 37–52. <https://doi.org/10.1017/S089267941800097X>

¹⁸Acquaviva, G. (2023). Crimes without Humanity? Artificial Intelligence, Meaningful Human Control, and International Criminal Law. *Journal of International Criminal Justice*, 21(5), 981–1004. <https://doi.org/10.1093/jicj/mqad024>.

Circumvention of State Responsibility: States that invest in and deploy LAWs within their military arsenals often delegate operational decisions to private companies and contractors. This outsourcing of warfare creates an additional problem: the dilution of state responsibility. As a result, human rights violations committed by LAWs may be attributed to non-state actors, thereby evading conventional international obligations.

The limitations of the Rome Statute in addressing algorithmic warfare reflect a fundamental dilemma for international criminal law. Kant argued that justice must be applied according to universal and impartial principles, ensuring that all morally reprehensible acts are punishable.¹⁹ However, if war crimes committed by humans can be prosecuted, there is no justification for those committed by autonomous systems to fall outside the reach of justice. The Rome Statute's omission in this regard creates a paradox: whilst International Humanitarian Law reinforces the need for control over lethal means, it allows weapons operated by artificial intelligence to act without a clear regime of accountability.

Against this backdrop, the issue is not merely to adapt International Criminal Law to new technologies, but rather to rethink the very concept of guilt and accountability in times of automated warfare. If the current system is maintained, the world may witness a new era of impunity, in which crimes are committed without the possibility of attributing responsibility to a specific individual. This leads to a deeper reflection on the limits of the legal system and on the need to establish new paradigms to hold accountable those involved in the creation, development and use of LAWs.²⁰

In short, the Rome Statute, as conceived, faces difficulties in holding the true perpetrators accountable. The identified regulatory gaps highlight the need for a broader debate on the adequacy of international criminal law to contemporary technological realities. The problem is not merely the absence of a criminally liable

¹⁹Kant, I. (1795/2008). *Perpetual Peace and Other Essays* (trans. A. Morão). Edições 70.

²⁰The use of artificial intelligence for military purposes is no longer a mere theoretical abstraction, but an operational reality in some conflict zones. In the context of the war between Israel and Palestine, investigations reveal that an AI system known as *Lavender* has been used to select targets for air strikes in Gaza. These algorithmic decisions, in addition to lacking transparency, are made based on prior calculations that estimate the acceptable number of civilian casualties, which raises serious concerns about the compliance of this model with International Humanitarian Law (IHL, 2024; O Globo, 2024). The growing use of these technologies in war zones indicates that the absence of clear regulation may lead to the normalisation of practices that undermine the distinction between combatants and non-combatants, as well as the principle of proportionality in armed conflicts (A Pública, 2024).

party, but rather the lack of a regulatory framework that prevents such systems from operating without control and without legal consequences.

1.4. The Legal Framework for LAWs in International Criminal Law

Currently, there is no specific regulation covering LAWs in international law. The Rome Statute, a fundamental document of international criminal law (ICL), presents challenges in its application to crimes committed by LAWs, particularly with regard to the accountability of individuals and entities involved in their development and use.

Article 8 of the Rome Statute defines war crimes, but does not expressly provide for liability for the use of LAWs. However, certain acts that may be committed using these weapons fall within the scope of existing war crimes, such as intentional attacks against civilian populations, the indiscriminate destruction of protected property, and the use of means of combat that cause unnecessary suffering. The lack of a specific regulatory framework for LAWs prevents their use from being properly analysed within the current IHL regime.

Commander responsibility, as provided for in Article 28^o of the Rome Statute, establishes that superiors may be held accountable when they permit or fail to prevent crimes committed by subordinates. In the context of LAWs, this article could be interpreted broadly to cover commanders who authorise the use of such weapons without ensuring effective oversight and control mechanisms. However, the difficulty in proving intent (*mens rea*) in errors committed by LAWs limits its application²¹. If an autonomous system acts unexpectedly, becoming unpredictable even to its operators, it may be difficult to attribute direct blame to a commander, as the latter would have no way of predicting or controlling all the variables of the weapon's algorithmic behaviour.

In the context of international criminal law and accountability for war crimes, the concept of *mens rea* plays a central role in the criminal imputation of agents involved in unlawful acts²². Traditionally, *mens rea* refers to the mental state of the perpetrator at the time the crime is committed, encompassing different degrees of culpability, ranging from direct intent to gross negligence²³. However, the autonomisation of military systems, such as Lethal Autonomous Weapons (LAWs), challenges this

²¹Schmitt, M. (2013). *Autonomous weapon systems and international humanitarian law: A reply to the critics*. Harvard National Security Journal, 4(1), 1-20.

²²Cassese, A. (2008). *International Criminal Law (2nd ed.)*. Oxford University Press.

²³Ambos, K. (2016). *Treatise on International Criminal Law: Volume I – Foundations and General Part*. Oxford University Press.

dogmatic framework, as the lethal decision is delegated to an algorithmic system, which is often opaque and unpredictable²⁴. Thus, the difficulty lies in determining whether and how a military commander, a remote operator, or even an AI developer can be considered to possess *mens rea* when the lethal outcome stems from algorithmic processes²⁵.

Authors such as Schmitt argue that, in the case of LAWs, the attribution of *mens rea* to the superior officer may be compromised due to the difficulty in demonstrating that the latter had actual knowledge of how the AI would process the data to make lethal decisions. This issue is particularly sensitive in light of Article 28 of the Rome Statute, which requires that the commander “had or ought to have had knowledge” of the crime committed by subordinates.²⁶ In the context of LAWs, the legal status of AI as a “subordinate” is not clearly defined, creating a legal vacuum.

On the other hand, Jain emphasises that criminal responsibility in the context of algorithmic warfare may require a reconfiguration of the traditional concept of *mens rea*, so as to include not only the subjective intent of the human agent, but also the degree of foreseeability and control that the agent had over the machine.²⁷ This would entail that intent could be interpreted more broadly, treating failures of supervision and of algorithmic design as relevant elements for criminal attribution.

Another significant legal challenge concerns the liability of developers. At present, programmers and companies responsible for the development of AI are not held criminally responsible for war crimes committed through their technologies, creating an “accountability vacuum”.²⁸ Article 25 of the Rome Statute provides for individual criminal responsibility, but applies only to natural persons, excluding companies and legal entities. A normative model designed to hold political and military leaders accountable proves inadequate to address the realities of LAWs, since these systems result from the interaction of diverse actors developers, manufacturers,

²⁴Dubber, M. D. (2020). *Theories of Crime and Artificial Intelligence*. Cambridge University Press.

²⁵Schmitt, M. (2013). *Autonomous Weapon Systems and International Humanitarian Law: A Reply to the Critics*. Harvard National Security Journal, 4(1), 1–20; See also: Sousa, S. A. (2022). “It wasn’t me, it was the machine”: Theory of crime, liability and artificial intelligence. In A. M. Rodrigues (Ed.), *Artificial Intelligence in Criminal Law* (pp. 59–95). Almedina.

²⁶Schmitt, M. (2013). *Autonomous weapon systems and international humanitarian law: A reply to the critics*. Harvard National Security Journal, 4(1), 1–20.

²⁷Jain, N. (2023). The accountability gap in international law: Who is responsible for AI-driven war crimes? *International Law Studies*, 99(4), 567-589.

²⁸Jain, N. (2023). The accountability gap in international law: Who is responsible for AI-driven war crimes? *International Law Studies*, 99(4), 567-589.

suppliers, and military personnel thus complicating the attribution of blame to a single agent. Against this backdrop, two possible approaches can be discussed within international criminal law in response to new technological realities:

Expansion of the doctrine of command responsibility: Article 28 could be interpreted so as to encompass responsibility of commanders who authorise the use of LAWs without adequate control. This would require revisiting ICL doctrine to include a duty of continuous supervision over autonomous systems, as well as requirements for auditing AI-generated decisions. This approach, however, faces practical obstacles, as commanders may claim they lack sufficient technical knowledge to understand and anticipate failures in systems operating through complex algorithms.

Adoption of specific norms for crimes committed through LAWs: a more effective solution could be the development of an international norm that attributes responsibility directly to States and companies involved in the development, production, and use of LAWs. Such a framework would allow developers and manufacturers to be held accountable in cases of systemic failures or design defects resulting in serious violations of IHL. This approach would address the gap in Article 25, which currently does not permit the criminal responsibility of legal persons.

Wherever the chain of command still includes a human being who orders or controls the actions of a LAWs, that person may be held responsible for its unlawful use. The ICC, operating under the Rome Statute, could establish guidelines for accountability not only of military commanders, but also of operators and engineers involved in the development of LAWs. Nevertheless, proving effective control over the machine's lethal decision remains a central difficulty.

Accordingly, algorithmic warfare is not only a military and ethical challenge, but also one that requires rethinking the current model of international justice. Developing legal mechanisms to hold States, commanders, and AI developers accountable is essential to prevent impunity and to ensure that technological progress does not erode the foundational principles of IHL.

More serious than the legal uncertainty resulting from the absence of specific regulation is the impact that LAWs may have on the very manner in which conflicts are conducted. Throughout history, war has been associated with a moral calculus: violence, however devastating, was accompanied by restraints intended to preserve the humanity of those involved. The autonomisation of lethality, however, places that

balance at risk, turning destruction into a purely technical operation, devoid of ethical judgement and moral responsibility.

If human dignity must always be treated as an end in itself and never merely as a means to external ends, delegating lethal decisions to machines constitutes an affront to this fundamental maxim. The mechanistic logic governing autonomous systems may dehumanise the very act of war, increasingly distancing those who initiate and control it from the devastating effects it produces. Beyond legal questions, a deeper ethical reflection is therefore required on the impact of LAWs on the protection of fundamental rights in times of conflict.

Thus, the debate cannot remain confined to the need for normative revision of international criminal law. The real problem lies in the compatibility between algorithmic warfare and the protection of human rights. Accordingly, this issue will be examined in greater depth below, analysing the risks inherent in the use of LAWs for the foundational principles of IHL and for human dignity itself.

2. THE IMPACT OF LETHAL AUTONOMOUS WEAPONS ON THE PROTECTION OF HUMAN RIGHTS

The defining feature of LAWs is their ability to select and attack targets without direct human intervention. This fact creates a legal and ethical paradox: on the one hand, the use of these weapons may be justified as a means of reducing civilian casualties by eliminating human error and increasing the precision of attacks;²⁹ on the other, the total absence of direct human intervention compromises the ability to assess complex contexts, resulting in opaque and unpredictable lethal decisions.

The Universal Declaration of Human Rights (UDHR) and international treaties establish that human life must be protected against arbitrary deprivation.³⁰ Yet, when an autonomous system makes a decision to kill, can one guarantee that such a decision is not arbitrary? The absence of a subjective moral criterion makes this question even more urgent, as algorithms possess no ethical conscience, but only statistical patterns derived from training data.³¹ There is therefore a risk that LAWs

²⁹Arkin, R. (2013). Lethal autonomous systems and the plight of the non-combatant. *AISB Quarterly*, (137), 1-9.

³⁰Assembleia Geral das Nações Unidas. (1948). Declaração Universal dos Direitos Humanos.

³¹Marchant, G. et al. (2011). International governance of autonomous military robots. *Columbia Science and Technology Law Review*, 12, 273-315.

reproduce algorithmic bias, leading to discrimination in target selection and violating the principle of human dignity.³²

The growing autonomy of LAWs drastically reduces meaningful human control over life-and-death decisions. The International Committee of the Red Cross (ICRC) emphasises that human oversight of the use of force is essential to ensure compliance with IHL.³³ The Martens Clause, enshrined in the Additional Protocols to the Geneva Conventions, provides that even where a situation of war is not explicitly regulated by positive law, the principles of humanity and the dictates of public conscience must prevail.³⁴

In international tribunals, the Martens Clause has been invoked to argue that any use of weapons that escapes human control may be incompatible with IHL.³⁵ Accordingly, the introduction of LAWs without guarantees of meaningful human control may be interpreted as a breach of that principle, opening the way for claims that their use is unlawful.

2.1. Breach of the Principles of International Humanitarian Law

International humanitarian law is grounded in a set of norms intended to limit the effects of armed conflict, ensuring that war is conducted with the minimum possible human suffering. For the purposes of this analysis, the following principles are particularly relevant:

- Humanity: limits war, ensuring that military violence does not exceed the demands of human dignity, as reflected in the Geneva Conventions and the Rome Statute of the International Criminal Court;³⁶
- Distinction: requires that combatants be distinguished from civilians, ensuring that only the former are legitimate targets;³⁷

³²Ekelhof, M. A. (2019). Lethal autonomous weapons and targeting processes: The moral implications of human-machine interaction. *International Review of the Red Cross*, 100(1), 119-144.

³³Comité Internacional da Cruz Vermelha (CICV). (2006). *A Guide to the Legal Review of Weapons, Means and Methods of Warfare*. Geneva: ICRC.

³⁴Protocolos Adicionais às Convenções de Genebra de 1949, Protocolo I, Artigo 1.º, parágrafo 2.º (1977).

³⁵Weizmann, N. (2014). *Autonomous Weapons Systems under International Law*. Academy Briefing no. 8, Geneva Academy.

³⁶Valente, M. M. G. (2024). A utilização de robots na guerra: A responsabilidade penal. In M. M. G. Valente (Coord.), *Ciências Criminais e Inteligência Artificial* (p. 18). Almedina.

³⁷CICV. (2016). *Autonomous Weapon Systems: Implication of Increasing Autonomy in the Critical Functions of Weapons*. Geneva: International Committee of the Red Cross.

- Proportionality: requires that incidental civilian harm not be excessive in relation to the concrete and direct military advantage anticipated from an attack;³⁸
- Military Necessity: requires that the force employed be strictly necessary to achieve a legitimate objective, avoiding disproportionate actions.³⁹

However, the operationalisation of these principles becomes highly problematic when applied to LAWS. Their capacity to make decisions independently raises serious doubts as to whether these foundational directives can be complied with. Unlike human beings, who can assess subjective and contextual factors on the battlefield, AI systems rely on predictive algorithms and statistical patterns that lack the capacity to interpret legal and ethical nuances.⁴⁰

The distinction between combatants and non-combatants is among the most critical challenges. IHL requires that attacks be directed exclusively against military objectives, but LAWS, operating on the basis of data and pre-programmed classifications, may err in identifying a lawful target. Issues such as combatants wearing civilian clothing, the presence of human shields, or chaotic movement in urban areas create highly complex scenarios that an AI may not be able to interpret correctly.⁴¹

Concerns about distinction become even more acute when one considers concrete cases of AI use in military operations. The Lavender system, reportedly used by Israel in the Gaza conflict, is a paradigmatic example of the challenges posed by algorithmic decision-making in contemporary warfare. Reports indicate that this AI was employed to identify and flag targets, relying on large volumes of data collected through aerial surveillance and intercepted communications. There are, however, indications that the system's criteria led to the overestimation of threats and the misclassification of civilians as combatants, potentially resulting in attacks on unlawful targets.⁴²

³⁸Schmitt, M. (2013). Autonomous weapon systems and international humanitarian law: A reply to the critics. *Harvard National Security Journal*, 4(1), 1-20.

³⁹Convenção de Genebra de 1949 e Protocolos Adicionais (1977).

⁴⁰Boularin, V., & Verbruggen, M. (2017). Mapping the development of autonomy in weapon systems. Stockholm International Peace Research Institute (SIPRI). Available at: <<https://www.sipri.org/publications/2017/policy-reports/mapping-development-autonomy-weapon-systems>>, accessed on 4 March 2025.

⁴¹Melzer, N. (2016). Human Rights Implications of the Usage of Drones and Unmanned Robots in Warfare. United Nations Office of the High Commissioner for Human Rights (OHCHR).

⁴²IHU. (2024, abril 4). Lavender: A inteligência artificial de Israel que decide quem é bombardeado em Gaza. Instituto Humanitas Unisinos. Available at: <<https://www.ihu.unisinos.br/categorias/638566-lavender-a-inteligencia-artificial-de-israel-que-decide-quem-e-bombardeado-em-gaza>>, accessed on 17 February 2025.

The lack of transparency in the way Lavender's algorithms operate raises fundamental questions about responsibility and compliance with IHL. In the absence of clarity concerning the precise parameters used to determine a target's identity, it becomes extremely difficult to verify whether the decisions taken respect the principle of distinction. Moreover, reports suggest that the system was designed to operate with a considerable error threshold, permitting decision-making that tolerated a pre-calculated number of civilian casualties. Such an approach directly conflicts with humanitarian principles, as it runs counter to the duty of precaution in the conduct of attacks.⁴³

Proportionality, in turn, is even more difficult to guarantee through autonomous systems. The principle requires a moral and legal judgement to weigh whether a particular attack will cause greater harm than the concrete military advantage anticipated. A LAWs cannot subjectively assess the tactical significance of a target, nor can it weigh humanitarian consequences in an ethically meaningful way. Studies show that AI systems may be manipulated through external interference or programmed with bias, leading to erroneous interpretations that produce disproportionate attacks.⁴⁴

The principle of military necessity is likewise compromised by the functioning of LAWs. If an autonomous system is programmed to eliminate any potential threat within an operational perimeter, it may make attack decisions without considering less lethal alternatives, such as the surrender of combatants or the use of non-lethal force. In addition, the opacity of the algorithms used in these systems often described as the "*black box*" problem prevents transparent auditing of the decisions made, making it difficult to attribute responsibility for failures or violations of IHL.⁴⁵

The Martens Clause, set out in the Preamble to the 1899 Hague Convention and reaffirmed in the Additional Protocols to the Geneva Conventions, provides that in situations where there is no specific regulation, "civilians and combatants remain under the protection and authority of the principles of the law of nations, as they result from

⁴³O Globo. (2024, abril 4). Israel usou IA para definir 37 mil alvos, com cálculo de permissão prévia de morte de civis, diz investigação. Available at: <<https://oglobo.globo.com/mundo/noticia/2024/04/04/israel-usou-ia-para-definir-37-mil-alvos-com-calculo-de-permissao-previa-de-morte-de-civis-diz-investigacao.ghtml>>, accessed on 17 February 2025. ⁴⁴Russell, S. (2019). Human Compatible: Artificial Intelligence and the Problem of Control. Viking Press.

⁴⁴Russell, S. (2019). Human Compatible: Artificial Intelligence and the Problem of Control. Viking Press.

⁴⁵Docherty, B. (2012). Losing Humanity: The Case Against Killer Robots. Human Rights Watch.

established custom, from the principles of humanity and from the dictates of public conscience”.⁴⁶ This clause has been invoked before international tribunals to argue that, even in the absence of specific norms for LAWs, they must be governed by fundamental principles of humanity and ethics in war.⁴⁷ Accordingly, any use of weapons that escapes meaningful human control may be interpreted as a serious breach of the Martens Clause, calling into question the legality of their use in contemporary conflicts.

Thus, the advance of LAWs without adequate human control represents a critical challenge to the effective application of IHL. The risk of systematic breaches of distinction, disproportionate attacks, and military decisions without justified necessity reinforces the need for an in-depth debate on the legitimacy of employing such technologies on the battlefield. Without guarantees that autonomous systems can operate within the legal and ethical parameters demanded by IHL, their deployment may lead to a significant erosion of fundamental rights in armed conflict.

Replacing human discernment with algorithmic processes in the use of lethal force constitutes a degradation of the ethics of war, as it shifts responsibility for the decision to a mechanised system that operates without any intrinsic regard for the dignity of those targeted.⁴⁸ If war is conducted without moral judgement, in accordance with the mechanistic logic of LAWs, there is a risk of transforming armed conflict into a process of indiscriminate annihilation, in which legal norms cease to have practical applicability. Human dignity, a foundational principle of international human rights law, thus becomes increasingly vulnerable in the face of the growing autonomisation of weaponry and the lack of control over its operation in theatres of military operations.

2.2. The Logic of “Legal by Design” in Algorithmic Warfare: Between Programming and Human Morality

The concept of “legal by design”, applied to LAWs, proceeds from the premise that a robotic system can be designed with a normative and ethical framework embedded in its algorithms, enabling it to decide autonomously whether to fire or not,

⁴⁶Convenção de Haia, Preâmbulo, 1899.

⁴⁷Weizmann, N. (2014). Autonomous Weapons Systems under International Law. Academy Briefing no. 8, Geneva Academy.

⁴⁸Bhuta, N., Beck, S., Geiß, R., Liu, & H., C. Kreß (Eds.), Autonomous weapons systems: Law, ethics, policy (pp. 333-352). Cambridge University Press.

between life and death.⁴⁹ The underlying ideal is the possibility of modelling a machine that, through the parametrisation of legal rules and ethical principles, could conduct military operations while respecting the limits imposed by IHL and human rights. This model, however, confronts fundamental contradictions, particularly when analysed in light of the UDHR and international treaties that protect dignity and human life.

At the technical level, designing lethal systems based on legal by design would mean that military decisions are taken by an algorithm structured around programmed norms to distinguish combatants from civilians, assess proportionality, and evaluate military necessity. This model seeks to minimise human error and ensure that the use of lethal force fits within a pre-defined normative framework. Yet such an approach ignores the moral and contextual complexity of war, reducing the application of law and ethics to binary calculations.⁵⁰ War, as a human phenomenon, is marked by unpredictable factors that cannot be translated solely into zeros and ones.

Article 3 of the UDHR provides that “everyone has the right to life, liberty and security of person”.⁵¹ On this basis, one may argue that delegating lethal decisions to an algorithmic system undermines the requirement of universal protection of human life, since there is no moral subject responsible for the choice to attack. AI, however advanced, does not comprehend the moral value of human life; it only processes data on the basis of statistics and patterns. Accordingly, the introduction of LAWs that autonomously decide upon targets creates a scenario of depersonalised responsibility, in which deaths result from probabilistic calculations rather than conscious ethical judgement.

The logic of legal by design also raises serious concerns regarding the protection of civilian objects and essential infrastructure. Article 52 of Additional Protocol I to the Geneva Conventions provides that civilian objects shall not be the object of attack unless and for such time as they are used for concrete military purposes.⁵² Determining whether a building is being used for military purposes requires contextual interpretation and an analysis of intent something an algorithmic system

⁴⁹Valente, M. M. G. (2024). A utilização de robots na guerra: A responsabilidade penal. In M. M. G. Valente (Coord.), *Ciências Criminais e Inteligência Artificial* (p. 18 e 19). Almedina.

⁵⁰Valente, M. M. G. (2024). A utilização de robots na guerra: A responsabilidade penal. In M. M. G. Valente (Coord.), *Ciências Criminais e Inteligência Artificial* (p. 18 e 19). Almedina.

⁵¹Assembleia Geral das Nações Unidas. (1948). Declaração Universal dos Direitos Humanos. Nações Unidas.

⁵²Comité Internacional da Cruz Vermelha (CICV). (1977). Protocolo Adicional I às Convenções de Genebra, Artigo 52.º.

cannot reliably perform. As a result, programming a robot on the basis of fixed rules may lead to the unjustified destruction of hospitals, schools, and humanitarian shelters, violating the international protection afforded to such infrastructure.

Another critical issue in applying legal by design is the possibility of algorithmic bias in identifying threats. Studies show that AI systems may reflect the prejudices embedded in training data, disproportionately classifying certain populations as potential threats.⁵³ This creates heightened risks of discrimination in target selection, especially in conflict zones characterised by significant ethnic or religious segregation. If an autonomous lethal system operates on biased data, it may contribute to systematic extrajudicial killings, violating Article 6 of the International Covenant on Civil and Political Rights, which similarly guarantees that no one shall be arbitrarily deprived of life.⁵⁴

From the standpoint of accountability, applying legal by design to LAWs creates a grey zone within international criminal law, as discussed in Chapter 1. Article 25 of the Rome Statute provides that only individuals may be prosecuted before the ICC.⁵⁵ A machine cannot be tried, and its programmers cannot readily be held responsible for a decision taken through opaque statistical processes.

In light of these difficulties, accepting that life-and-death decisions be left to algorithmic systems means relinquishing human control over the use of lethal force, thereby undermining the foundational pillars of the international system for the protection of human rights. War cannot be reduced to a mathematical calculation, as it involves moral and legal dimensions that cannot be translated into binary commands. Human dignity, a central principle of international law, requires that the decision to kill remain subject to human judgement, ensuring that morality and justice are not replaced by the cold logic of algorithms.

Accordingly, while legal by design may be useful as a decision-support tool, it cannot substitute the ethical and legal responsibility of human agents. Accepting that a machine may choose between preserving or extinguishing human lives represents an unacceptable risk to the foundations of IHL and to the modern conception of human rights. International regulation must therefore ensure that any lethal weapons system

⁵³Ekelhof, M. (2019). Lethal autonomous weapons and targeting processes: The moral implications of human-machine interaction. *International Review of the Red Cross*, 100(1), 119-144.

⁵⁴Assembleia Geral das Nações Unidas. (1966). Pacto Internacional sobre Direitos Civis e Políticos, Artigo 6.º.

⁵⁵Estatuto de Roma do Tribunal Penal Internacional, Artigo 25.º (1998).

remains under meaningful human supervision, so that morality and law prevail over algorithmic logic.

3. PROPOSALS FOR STRENGTHENING INTERNATIONAL NORMS APPLICABLE TO LETHAL AUTONOMOUS WEAPONS

The autonomisation of warfare systems and the rise of LAWs constitute one of the most complex challenges for contemporary international law. While such technologies may promise increased military effectiveness and reduced risk to combatants, they also raise ethical and legal questions that transcend the boundaries of traditional war.⁵⁶ The existing regulatory vacuum makes urgent the need for a specific international treaty governing the development, use, and limitations of LAWs.

Among the principal concerns stemming from the absence of regulation is the risk of uncontrolled proliferation, facilitating the integration of these weapons into non-conventional conflicts and their use by non-state actors.⁵⁷ Moreover, in the absence of clear rules establishing responsibility, accountability for operations conducted through LAWs remains nebulous, hindering the investigation of war crimes.

A specific international treaty should therefore provide for:

- Objective criteria for classifying LAWs, distinguishing systems operating under human supervision from those that are fully independent;
- Clear limits on their use, ensuring respect for the principles of distinction, proportionality, and military necessity;⁵⁸
- The creation of international oversight mechanisms, including a specialised body to monitor the use of these technologies;
- International sanctions for States and companies that violate the treaty's provisions.

Discussion of the need for such a treaty has been pursued in international forums, notably within the United Nations disarmament process, which since 2013 has

⁵⁶Bhuta, N., Beck, S., Geiß, R., Liu, H., & Kreß, C. (Eds.). (2016). *Autonomous Weapons Systems: Law, Ethics, Policy*. Cambridge University Press.

⁵⁷Boulain, V., & Verbruggen, M. (2017). *Mapping the development of autonomy in weapon systems*. Stockholm International Peace Research Institute (SIPRI). Available at: <<https://www.sipri.org/publications/2017/policy-reports/mapping-development-autonomy-weapon-systems>>, accessed on 1 March 2025.

⁵⁸Conforme o que estabelece o Estatuto de Roma do Tribunal Penal Internacional, Artigos 7 e 8 (1998). Available at: <<https://www.icc-cpi.int/resource-library/Documents/RS-Eng.pdf>>, accessed on 24 February 2025.

debated the feasibility of a partial or total ban on LAWs.⁵⁹ Nevertheless, resistance from military powers investing in the development of these weapons has hindered concrete progress.

3.1. *The Stop Killer Robots Initiative*

One of the principal international campaigns advocating the prohibition of LAWs is the Stop Killer Robots initiative, which calls for the adoption of an international treaty banning the development, production, and use of such weapons. Created by a coalition of human rights organisations, academics, and disarmament experts, the campaign is premised on the view that weapons operating without meaningful human control represent an inadmissible ethical, political, and humanitarian risk.⁶⁰

The initiative's central argument is that LAWs violate foundational principles of IHL and human rights, because they: (i) undermine distinction, as AI may fail to identify civilians and combatants, as well as the principle of humanity; (ii) undermine criminal accountability, since autonomous systems cannot be tried for war crimes; (iii) increase the risk of proliferation, enabling States and armed groups to acquire lethal technologies without international supervision.⁶¹

Although the initiative is supported by various States and organisations, some countries strongly resist its implementation, arguing that a general ban on LAWs may hinder military innovation and national security.⁶²

The Stop Killer Robots campaign is only one among several initiatives aimed at regulating LAWs. In 2021, the ICRC issued a formal recommendation calling for the adoption of rules ensuring meaningful human control over lethal decisions.⁶³ Similarly,

⁵⁹Comitê das Nações Unidas sobre Desarmamento. (2013). Report on Lethal Autonomous Weapons Systems. Available at: <<https://undocs.org/en/A/74/303>>, accessed on 3 March 2025.

⁶⁰Stop Killer Robots. (2024). About the Campaign. Available at: <<https://www.stopkillerrobots.org>>, accessed on 5 March 2025.

⁶¹Human Rights Watch. (2020). New Weapons, Proven Precedent - Elements of and Models for a Treaty on Killer Robots. Available at: <<https://www.hrw.org/report/2020/10/20/new-weapons-proven-precedent/elements-and-models-treaty-killer-robots>>, accessed on 6 March 2025; e Human Rights Watch. (2020). Stopping Killer Robots Country Positions on Banning Fully Autonomous Weapons and Retaining Human Control. Available at: <<https://www.hrw.org/report/2020/08/10/stopping-killer-robots/country-positions-banning-fully-autonomous-weapons-and>>, accessed on 6 March 2025.

⁶²Human Rights Watch. (2020). Stopping Killer Robots Country Positions on Banning Fully Autonomous Weapons and Retaining Human Control. Available at: <<https://www.hrw.org/report/2020/08/10/stopping-killer-robots/country-positions-banning-fully-autonomous-weapons-and>>, accessed on 6 March 2025.

⁶³Comitê Internacional da Cruz Vermelha (CICV). (2021). Recommendations on Lethal Autonomous Weapons. Available at: <<https://www.icrc.org/en/document/icrc-position-autonomous-weapon-systems>>, accessed on 2 March 2025. See also: Comitê Internacional da Cruz Vermelha (CICV). (2024).

the European Union has discussed guidelines to restrict the use of these AI technologies within its territory, although the development of weapons has remained outside the scope of the AI Act.⁶⁴

The growing international concern surrounding LAWs has prompted statements from various actors, including intergovernmental organisations and the Holy See. The Vatican, through its representation at the UN, has warned of the ethical risks of allowing autonomous systems to decide over life and death, reaffirming that such weapons affront human dignity and the principles of international law.⁶⁵ In addition, the UN Secretary-General, António Guterres, has argued that autonomous weapons should be prohibited or, at a minimum, strictly regulated.⁶⁶

In addition, Harvard Law School has developed a set of legal materials reflecting different State interpretations of legal questions associated with the use of algorithmic systems in armed conflict.⁶⁷ This analysis demonstrates that, notwithstanding the absence of global consensus, there is a growing recognition of the need to establish binding norms to ensure that the use of AI in war does not violate the fundamental principles of IHL.

International pressure for the regulation of LAWs has increased in recent years; however, absent global consensus, it is unlikely that a binding treaty will be adopted in the short term. In this context, international law must continue to adapt in order to respond to the new threats posed by algorithmic warfare.

3.2. *Debates Concerning the International Criminal Court*

Autonomous weapons: ICRC submits recommendations to UN Secretary-General. Available at: <<https://www.icrc.org/en/document/autonomous-weapons-icrc-submits-recommendations-un-secretary-general>>, accessed on 2 March 2025.

⁶⁴União Europeia. (2022). Artificial intelligence act. Available at: <[https://www.europarl.europa.eu/thinktank/en/document/EPRS_BRI\(2021\)698792](https://www.europarl.europa.eu/thinktank/en/document/EPRS_BRI(2021)698792)>, accessed on 1 March 2025.

⁶⁵Vatican News. (2021). ONU: Vaticano apela contra o uso de armas autónomas letais. Available at: <<https://www.vaticannews.va/pt/vaticano/news/2021-08/onu-vaticano-armas-automaticas.html>>, accessed on 6 March 2025.

⁶⁶United Nations. (2024). UN Secretary-General Report on Lethal Autonomous Weapons Systems. Available at: <<https://documents.un.org/doc/undoc/ltd/n24/305/45/pdf/n2430545.pdf>>, accessed on 5 March 2025.

⁶⁷Harvard Law School. (2020). A Compilation of Materials Reflecting States' Views on Algorithmic and Data-Reliant Socio-Technical Systems in Armed Conflict. Available at <<https://pilac.law.harvard.edu/a-compilation-of-materials-apparently-reflective-of-states-views-on-international-legal-issues-pertaining-to-the-use-of-algorithmic-and-data-reliant-socio-technical-systems-in-armed-conflict>>, accessed on 3 March 2025.

As discussed above, the Rome Statute provides that the ICC may prosecute only crimes committed by individuals, excluding the responsibility of collective entities such as corporations and States. This legal framework presents serious challenges in the current technological landscape. In order to ensure that crimes committed through LAWs may be prosecuted, the ICC's mandate could be expanded in three principal ways:

- Explicit criminalisation of the unlawful use of LAWs, by inserting into the Rome Statute provisions that characterise the use of these weapons in violations of IHL as international crimes;
- Inclusion of crimes committed through autonomous systems, establishing objective criteria for determining the responsibility of commanders, developers, and operators of LAWs;
- Creation of a regime of corporate responsibility, allowing companies involved in the production and commercialisation of LAWs to be investigated and sanctioned by the ICC.

Including such provisions in the Rome Statute would require normative revision a process that traditionally faces resistance from States Parties. Yet precedents demonstrate that adaptations are possible in order to reflect new legal and technological realities.⁶⁸ Accordingly, growing international pressure for the regulation of LAWs may create a political environment conducive to expanding the ICC's scope of action.

In light of these challenges, it becomes evident that the ICC's current mandate could be expanded to encompass crimes committed through LAWs. Adapting the Rome Statute to include such crimes, broadening the doctrine of command responsibility, imposing transparency obligations for autonomous systems, and holding developers and States accountable are fundamental steps to prevent algorithmic warfare from producing a condition of legal impunity.

Therefore, this study has shown that the regulation of LAWs is not merely a technical or military question, but an ethical and legal imperative for the protection of human rights and international humanitarian law. The conclusions that follow synthesise the principal reflections developed throughout the paper and present

⁶⁸CICV. (2021). Recommendations on Lethal Autonomous Weapons.

concrete recommendations to enable the international community to mitigate the risks posed by these weapons before their use becomes irreversible.

CONCLUSIONS

The analysis undertaken in this study has shown that LAWs represent one of the greatest contemporary challenges to international criminal law, international humanitarian law, and the modern conception of human rights itself. The autonomisation of warfare, once confined to the realm of science fiction, has become an operational reality whose ethical, legal, and geopolitical impacts remain insufficiently regulated. The debate concerning these technologies is not limited to their military effectiveness or their alleged potential to reduce casualties among armed forces; it forms part of a broader reflection on the limits of technological power and the role of humanity in the conduct of armed conflict. The issue cannot be approached with indifference, as decisions taken today will shape the future of war and the protection of fundamental rights in the twenty-first century.

The existing normative vacuum favours impunity and enables State and private actors to exploit legal gaps in order to employ LAWs without effective oversight. In its current configuration, the ICC lacks adequate means to address crimes committed through autonomous systems, as the Rome Statute is grounded in a traditional model of individual criminal responsibility that does not accommodate the complexity of algorithmic warfare.

The absence of a specific international treaty regulating LAWs is among the most evident legal deficits in this field. Unlike other military technologies that have been the subject of specific international regulation, LAWs remain in a normative limbo that permits accelerated development without a corresponding supervisory mechanism. As argued throughout this paper, creating a dedicated international treaty is an urgent necessity to prevent LAWs from becoming instruments of indiscriminate violence, dehumanisation of conflict, and the perpetuation of systematic human rights violations.

The principle of human dignity, a cornerstone of international human rights law, requires that the decision to take life not be delegated to machines, but remain under the responsibility of a human agent. However, the absence of a universally accepted definition of “meaningful human control” enables divergent interpretations and hinders effective regulation. While some States and international organisations advocate

continuous human supervision over the use of LAWs, others argue that ex post facto oversight would suffice. This study has shown that any regulatory approach allowing lethal decision-making without a human agent at the moment of action not only violates foundational principles of IHL, but also undermines any subsequent attempt at criminal accountability.

The analysis also highlighted the need to expand the ICC's mandate to encompass crimes committed through LAWs. Criminal responsibility in algorithmic warfare must be understood within a framework that includes military commanders, software developers, manufacturing companies, and States that promote unrestricted use of these weapons. The doctrine of command responsibility in the Rome Statute can and should be adapted to cover situations in which a commander authorises the use of LAWs without ensuring adequate supervision. Likewise, the responsibility of developers and manufacturers must be seriously considered, because in many cases these actors design algorithms without sufficiently clear criteria to prevent attacks against civilians or infrastructure protected by IHL.

The opacity of algorithmic systems used in LAWs often described as the “black box” problem was identified as one of the principal obstacles to criminal accountability. The lack of transparency in LAWs’ decision-making processes impedes the identification of failures and creates challenges for auditing the algorithms responsible for target selection. As argued above, any effective normative regime must include mechanisms ensuring that algorithms are auditable and that their use in military operations can be examined retrospectively. The principle of foreseeability, essential to criminal law, must be protected, preventing lethal decisions from being made by systems whose criteria cannot be understood by human operators or by the legal system itself.

Beyond legal considerations, the research showed that regulating LAWs is not only a normative necessity but also a moral and ethical requirement. The claim that LAWs could reduce casualties and make war “cleaner” ignores battlefield complexity and the risks associated with dehumanisation. Allowing machines to make life-and-death decisions without any human mediation not only undermines foundational IHL principles, but also redefines the very nature of war. The cold logic of statistics, applied blindly by algorithms, cannot replace human judgement: war cannot be reduced to a probabilistic calculation of military efficiency.

The Stop Killer Robots initiative and the ICRC's recommendations show that civil society and international organisations already recognise the urgency of regulating LAWs before their use becomes irreversible. However, resistance from major military powers, which regard such weapons as a strategic advantage, hinders the adoption of binding international treaties. The absence of global consensus on prohibiting or limiting LAWs may create a scenario of uncontrolled technological proliferation in which States and armed groups acquire autonomous systems without regulation. The risk of an algorithmic arms race in which AI sophistication replaces strategic and diplomatic capacity is an imminent reality requiring an immediate response from the international community.

In light of these reflections, the regulation of LAWs cannot be postponed. It is imperative that the international community advance the creation of specific regulations imposing clear restrictions on the development and use of these weapons, ensuring meaningful human control over their operations, and establishing a regime of criminal accountability for the actors involved. The ICC may refine its mandate to better prosecute crimes committed through LAWs, ensuring that impunity does not prevail in the face of new forms of warfare. Regulating algorithmic warfare is not only a matter of law, but also a matter of justice and of preserving the fundamental values of humanity.

REFERENCES

- A PÚBLICA. *Lavender: A máquina de inteligência artificial que comanda bombardeios de Israel em Gaza*. 4 abr. 2024. Disponível em: <<https://apublica.org/2024/04/lavender-a-maquina-de-inteligencia-artificial-que-comanda-bombardeios-de-israel-em-gaza>>. Acesso em: 17 fev. 2025.
- ACQUAVIVA, G. Crimes without Humanity? Artificial Intelligence, Meaningful Human Control, and International Criminal Law. *Journal of International Criminal Justice*, v. 21, n. 5, p. 981–1004, 2023. DOI: <https://doi.org/10.1093/jicj/mqad024>.
- AMBOS, K. *Treatise on International Criminal Law: Volume I - Foundations and General Part*. Oxford: Oxford University Press, 2016.
- ARKIN, R. Lethal autonomous systems and the plight of the non-combatant. *AISB Quarterly*, n. 137, p. 1-9, 2013.
- ARKIN, R. C. *Governing Lethal Behavior in Autonomous Robots*. Chapman & Hall, 2010.

ASSEMBLEIA GERAL DAS NAÇÕES UNIDAS. *Declaração Universal dos Direitos Humanos*. 1948.

ASSEMBLEIA GERAL DAS NAÇÕES UNIDAS. *Pacto Internacional sobre Direitos Cíveis e Políticos*. 1966.

BHUTA, N.; BECK, S.; GEIß, R.; LIU, H.; KREß, C. (Eds.). *Autonomous weapons systems: Law, ethics, policy*. Cambridge: Cambridge University Press, 2016.

BOULANIN, V.; VERBRUGGEN, M. *Mapping the Development of Autonomy in Weapon Systems*. SIPRI, 2017.

CASSESE, A. *International Criminal Law*. 2. ed. Oxford: Oxford University Press, 2008.

CICV. *Autonomous Weapon Systems: Implication of Increasing Autonomy in the Critical Functions of Weapons*. Geneva: International Committee of the Red Cross, 2016.

CICV. *Recommendations on Lethal Autonomous Weapons*. Geneva: International Committee of the Red Cross, 2021.

COMITÊ DAS NAÇÕES UNIDAS SOBRE DESARMAMENTO. *Report on Lethal Autonomous Weapons Systems*. 2013. Disponível em: <<https://undocs.org/en/A/74/303>>. Acesso em: 3 mar. 2025.

COMITÊ INTERNACIONAL DA CRUZ VERMELHA (CICV). *Protocolo Adicional I às Convenções de Genebra*. 1977.

COMITÊ INTERNACIONAL DA CRUZ VERMELHA (CICV). *A Guide to the Legal Review of Weapons, Means and Methods of Warfare*. Geneva: ICRC, 2006.

COMITÊ INTERNACIONAL DA CRUZ VERMELHA (CICV). *Recommendations on Lethal Autonomous Weapons*. 2021. Disponível em: <<https://www.icrc.org/en/document/icrc-position-autonomous-weapon-systems>>. Acesso em: 2 mar. 2025.

COMITÊ INTERNACIONAL DA CRUZ VERMELHA (CICV). *Autonomous weapons: ICRC submits recommendations to UN Secretary-General*. 2024. Disponível em: <<https://www.icrc.org/en/document/autonomous-weapons-icrc-submits-recommendations-un-secretary-general>>. Acesso em: 2 mar. 2025.

CONVENÇÃO DE GENEBRA. 1949.

CONVENÇÃO DE HAIA. 1899.

DOCHERTY, B. *Losing Humanity: The Case Against Killer Robots*. Human Rights Watch, 2012.

DUBBER, M. D. *Theories of Crime and Artificial Intelligence*. Cambridge: Cambridge University Press, 2020.

EKELHOF, M. A. Lethal Autonomous Weapon Systems, Human Control and the Special Status of Military Decision-Making. *Journal of Ethics & International Affairs*, v. 33, n. 1, p. 37–52, 2019. DOI: <https://doi.org/10.1017/S089267941800097X>.

EKELHOF, M. A. Lethal autonomous weapons and targeting processes: The moral implications of human-machine interaction. *International Review of the Red Cross*, v. 100, n. 1, p. 119-144, 2019.

ESTATUTO DE ROMA DO TRIBUNAL PENAL INTERNACIONAL. 1998. Disponível em: <<https://www.icc-cpi.int/resource-library/Documents/RS-Eng.pdf>>. Acesso em: 24 fev. 2025.

HARVARD LAW SCHOOL. *A Compilation of Materials Reflecting States' Views on Algorithmic and Data-Reliant Socio-Technical Systems in Armed Conflict*. 2020. Disponível em: <<https://pilac.law.harvard.edu/a-compilation-of-materials-apparently-reflective-of-states-views-on-international-legal-issues-pertaining-to-the-use-of-algorithmic-and-data-reliant-socio-technical-systems-in-armed-conflict>>. Acesso em: 3 mar. 2025.

HEYNS, C. *Report of the Special Rapporteur on Extrajudicial, Summary or Arbitrary Executions*. United Nations Human Rights Council, 2013.

HUMAN RIGHTS WATCH. *New Weapons, Proven Precedent - Elements of and Models for a Treaty on Killer Robots*. 2020. Disponível em: <<https://www.hrw.org/report/2020/10/20/new-weapons-proven-precedent/elements-and-models-treaty-killer-robots>>. Acesso em: 6 mar. 2025.

HUMAN RIGHTS WATCH. *Stopping Killer Robots: Country Positions on Banning Fully Autonomous Weapons and Retaining Human Control*. 2020. Disponível em: <<https://www.hrw.org/report/2020/08/10/stopping-killer-robots/country-positions-banning-fully-autonomous-weapons-and>>. Acesso em: 6 mar. 2025.

IHU. *Lavender: A inteligência artificial de Israel que decide quem é bombardeado em Gaza*. Instituto Humanitas Unisinos, 4 abr. 2024. Disponível em: <<https://www.ihu.unisinos.br/categorias/638566-lavender-a-inteligencia-artificial-de-israel-que-decide-quem-e-bombardeado-em-gaza>>. Acesso em: 17 fev. 2025.

JAIN, A. G. Autonomous Weapon Systems, Errors and Breaches of International Humanitarian Law. *Journal of International Criminal Justice*, v. 21, n. 5, p. 1005–1032, 2023. DOI: <https://doi.org/10.1093/jicj/mqad043>.

KANT, I. *Fundamentação da Metafísica dos Costumes*. Trad. P. Quintela. Lisboa: Edições 70, 2003 [1785].

KANT, I. *A Paz Perpétua e Outros Opúsculos*. Trad. A. Morão. Lisboa: Edições 70, 2008 [1795].

KESAVAN, S. Autonomous Weapon System: Accountability Under International Criminal Law. SSRN, 15 abr. 2024. Disponível em: <https://ssrn.com/abstract=4819857>. Acesso em: 15 fev. 2025.

LEMO, M.; COSTA, M. J. Inteligência Artificial e Direito da Guerra: Reflexões sobre as Armas Autônomas Mortíferas. In: RODRIGUES, A. M. (Coord.). *A Inteligência Artificial no Direito Penal*. Vol. II. Coimbra: Almedina, 2022. p. 91-124.

MARCHANT, G. E.; ALLENBY, B.; HERKERT, J. *The Growing Gap Between Emerging Technologies and Legal-Ethical Oversight*. Springer, 2020.

MARCHANT, G. et al. International governance of autonomous military robots. *Columbia Science and Technology Law Review*, v. 12, p. 273-315, 2011.

MELZER, N. Human Rights Implications of the Usage of Drones and Unmanned Robots in Warfare. United Nations Office of the High Commissioner for Human Rights (OHCHR), 2016.

OBSERVADOR. Von der Leyen defende aumento das despesas militares na Europa. 3 mar. 2025. Disponível em: <https://observador.pt/2025/03/03/temos-urgentemente-de-rearmar-a-europa-von-der-leyen-pede-aumento-acentuado-das-despesas-na-defesa/>. Acesso em: 3 mar. 2025.

O GLOBO. Israel usou IA para definir 37 mil alvos, com cálculo de permissão prévia de morte de civis, diz investigação. 4 abr. 2024. Disponível em: <https://oglobo.globo.com/mundo/noticia/2024/04/04/israel-usou-ia-para-definir-37-mil-alvos-com-calculo-de-permissao-previa-de-morte-de-civis-diz-investigacao.ghtml>. Acesso em: 17 fev. 2025.

RTP. Temos urgentemente de rearmar a Europa: Von der Leyen pede aumento acentuado das despesas na defesa. Rádio e Televisão de Portugal, 2025. Disponível em: https://www.rtp.pt/noticias/mundo/carta-de-von-der-leyen-traca-plano-para-rearmar-a-europa_a1638235. Acesso em: 3 mar. 2025.

RUSSELL, S. *Human Compatible: Artificial Intelligence and the Problem of Control*. Viking Press, 2019.

SCHMITT, M. N. Autonomous Weapons Systems and International Humanitarian Law: A Reply to the Critics. *Harvard National Security Journal Features*, p. 1–37, 2013.

SHARKEY, N. The Impact of AI on Warfare. *Journal of Peace Research*, v. 55, n. 2, p. 223–236, 2018. DOI: <https://doi.org/10.1177/0022343317750212>.

SOUSA, S. A. “Não fui eu, foi a máquina”: Teoria do crime, responsabilidade e inteligência artificial. In: RODRIGUES, A. M. (Coord.). *A Inteligência Artificial no Direito Penal*. Coimbra: Almedina, 2022. p. 66-68.

STOP KILLER ROBOTS. *About the Campaign*. 2024. Disponível em: <<https://www.stopkillerrobots.org>>. Acesso em: 5 mar. 2025.

TRIBUNAL PENAL INTERNACIONAL. *Estatuto de Roma do Tribunal Penal Internacional*. Nações Unidas, 1998. Disponível em: <<https://www.icc-cpi.int/resource-library>>. Acesso em: 1 mar. 2025.

UNIÃO EUROPEIA. *Artificial Intelligence Act*. 2022. Disponível em: <[https://www.europarl.europa.eu/thinktank/en/document/EPRS_BRI\(2021\)698792](https://www.europarl.europa.eu/thinktank/en/document/EPRS_BRI(2021)698792)>. Acesso em: 1 mar. 2025.

UNITED NATIONS. *UN Secretary-General Report on Lethal Autonomous Weapons Systems*. 2024. Disponível em: <<https://documents.un.org/doc/undoc/ltd/n24/305/45/pdf/n2430545.pdf>>. Acesso em: 5 mar. 2025.

VALENTE, M. M. G. A utilização de robots na guerra: A responsabilidade penal. In: VALENTE, M. M. G. (Coord.). *Ciências Criminais e Inteligência Artificial*. Coimbra: Almedina, 2024.

VATICAN NEWS. ONU: Vaticano apela contra o uso de armas autónomas letais. 2021. Disponível em: <<https://www.vaticannews.va/pt/vaticano/news/2021-08/onu-vaticano-armas-automaticas.html>>. Acesso em: 6 mar. 2025.

WEIGEND, T. Convicting Autonomous Weapons? Criminal Responsibility of and for AWS under International Law. *Journal of International Criminal Justice*, v. 21, n. 5, p. 1137–1154, 2023. DOI: <https://doi.org/10.1093/jicj/mqad037>.