

TENUOUS HARMONY: CRIMINAL LIABILITY IN THE FACE OF THE AUTONOMY OF ARMS ON THE THRESHOLD OF HUMANITARIAN LAW

Louise Amorim Beja¹

University of Coimbra

DOI: <https://doi.org/10.62140/LAB912024>

Summary: 1.Introduction;2.Artificial Intelligence in the context of lethal autonomous weapons; 3. Reflections on international criminal responsibility in the face of crimes committed by Autonomous Weapons; 4. Final Considerations.5.References.

Abstract:The present investigation examines the phenomenon of Autonomous Weapons Systems (AWS) and its implications for international criminal responsibility. Defining AWS as systems designed for military missions, capable of identifying, selecting, and engaging targets without direct human intervention, the study focuses on criminal responsibility for offenses committed by these weapons equipped with Artificial Intelligence (AI). The adopted methodology is bibliographic and documentary, employing deductive, empirical, and critically reflective approaches. The research explores the favorable and unfavorable aspects of applying this technology in armed conflicts, highlighting responsibility gaps that may arise from autonomous decisions. The central issue addressed is how International Criminal Law can hold criminals accountable for errors committed by complex AI systems in military operations.

Keywords: Artificial Intelligence; Criminal Liability; Deadly Autonomous Weapons; Humanitarian Law; International Criminal Law.

1. Introduction

Weapons systems can be defined as weapons systems designed and configured for a mission, with the adaptability and aptitude to identify, select and engage military targets without the need for human intervention. This definition encompasses three essential

¹University Professor; Researcher at the Legal Institute of the Faculty of Law of the University of Coimbra; Developer of Online Educational Materials; Evaluator of the Scientific Journal of the OAB sub-section of Montes Claros - MG and the Criminalis Journal; Associate Editor of the Erga Omnes Journal; Researcher at the Critical International Law Study and Research Group of the Federal University of Uberlândia; Researcher at the National Agency for the Study of the Right to Development (ANEDD); Human Rights Activist at the Coimbra Group of Amnesty International Portugal; Master's Degree in Legal and Political Sciences with a specialisation in Public and European International Law and PhD student in Legal and Criminal Sciences, both at the Faculty of Law of the University of Coimbra; Specialist in the Management of Collaborative I&D&I Projects; Postgraduate in Teaching and Management in Higher Education. E-mail: <louisebeja@gmail.com; louisebeja@ij.uc.pt>; Orcid: <<https://orcid.org/0000-0003-4120-6868>>; CienciaVitae: <<https://www.cienciavitae.pt/pt/F51D-E8E4-4D06>>; Lattes: <<http://lattes.cnpq.br/3433182208171067>>.

elements that give an Autonomous Weapon of War its distinctive characterisation. A priori, deadly autonomous weapons must be recognised as weapons systems, rather than being qualified as new actors in the theatre of military operations. Therefore, in order to dispel any fanciful science fiction conjectures, it is essential to realise that Deadly Autonomous Weapons are always created and programmed by human beings.

The concept of weapons systems operating without human intervention must be treated with caution. From one perspective, Autonomous Weapons of War open up a wide range of new opportunities, since they dispense with the presence or direct participation of human operators in the target selection process. In this sense, this research is fundamentally focused on international criminal liability for crimes committed using lethal autonomous weapons equipped with Artificial Intelligence (machine learning and independent decision-making power, humans out of the loop - HOO TL). To do this, it is necessary to understand Artificial Intelligence, looking at it in the context of "black box problems", bearing in mind that the use of this technology can result in failures and decisions that lead to violations of people's human rights.

In order to understand this issue, it is also necessary to analyse the favourable and unfavourable aspects of applying this type of technology in a context of armed conflict, clarifying which crimes could be committed by a lethal autonomous weapon. In this way, the focus of the research is on the following problem: How can International Criminal Law hold criminals responsible for errors committed by complex Artificial Intelligence computer systems - known in the literature as "responsability gaps" - in military operations?

The aim is to reflect on the important role of international criminal law in the face of these new challenges in our society. Finally, the aim is to present a reflective proposal on criminal responsibility in contemporary times, demonstrating the possible problems that may arise in the context of international criminal law and deadly autonomous weapons. With regard to methodology, this research will be bibliographical and documentary, descriptive, using the following methods: deductive; empirical; and critical reflexive based on theoretical research.

2. Artificial Intelligence in the context of lethal autonomous weapons

Autonomous

In the course of a substantial period, national armed forces have incorporated restricted autonomy weaponry, displaying limited capabilities in the identification of predefined military targets.² However, in the contemporary scenario, such weaponry outlines a significant shift in the traditional approach to warfare.³ From this perspective, autonomous systems emerge as transformative elements, relegating human actors to the role of mere "supervisors" of the actions of Autonomous Weapon Systems on the battlefield.

The development of Artificial Intelligence is regarded as the 4th industrial revolution; however, lethal autonomous weapons are considered the third major revolution in military history, following gunpowder and nuclear weapons.⁴ Since 2012, the discourse within intergovernmental bodies, public spheres, academic circles, and scientific communities has intensified on this subject. The Human Rights Watch, one of the most esteemed human rights organizations globally, actively engaged in the field of the laws of war, in collaboration with the International and Human Rights Clinic at Harvard Law School, has issued a report titled "Losing Humanity: The Case Against Killer Robots," advocating for the cessation of lethal autonomous weapons.⁵

This debate has resurfaced in recent times, gaining momentum with the Human Rights Watch and the Campaign to Stop Killer Robots (a global coalition of civil society coordinated by the Human Rights Watch), presenting a more tangible vision for establishing a treaty to ban lethal autonomous weapons. This envisioned treaty would be supported by a significant number of states.⁶

²Ronald C. Arkin, "Governing Lethal Behavior in Autonomous Robots," 7—26 (2009). "Several automatic and semi-autonomous systems are already in operation and used by States, such as Samsung Sentris, Iron Dome, AEGIS, and the Brimstone Missile, among others." Also, refer to the European Parliament's Draft Resolution on Autonomous Weapon Systems, available at: https://www.europarl.europa.eu/doceo/document/B-8-2018-0355_PT.html, accessed on January 5, 2024.

³See *BRIMSTONE ADVANCE ANTI-ARMOUR MISSILE*, Army Technology. Available at: <https://www.army-technology.com/projects/brimstone/>. Accessed on January 5, 2024. This source highlights the Brimstone Missile as a perfect example, indirectly targeting armored vehicles autonomously.

⁴See "Autonomous Weapons: An Open Letter from AI & Robotics Researchers" (2015), signed by 4,985 AI and Robotics researchers and endorsed by an additional 27,800 signatories, including Stephen Hawking, Elon Musk, Steve Wozniak, and Noam Chomsky. Available at: <https://futureoflife.org/open-letter/open-letter-autonomous-weapons-ai-robotics/>, accessed on January 27, 2024.

⁵Lemos, Miguel; Costa, Miguel João. *Inteligência Artificial e Direito da Guerra: Reflexões sobre as Armas Autônomas Mortíferas*. In: *A Inteligência Artificial no Direito Penal Vol. II*. Anabela Miranda Rodrigues (coord.), Coimbra: Almedina, 2022.

⁶Key Elements of a Treaty on Fully Autonomous Weapons (2019). Available at: <https://www.stopkillerrobots.org/wp-content/uploads/2020/03/Key-Elements-of-a-Treaty-on-Fully-Autonomous-Weapons.pdf>, accessed on January 29, 2024; and *New Weapons, Proven Precedent, Elements of and Models for a Treaty on Killer Robots* (2020). Available at: <https://www.hrw.org/report/2020/10/20/new-weapons-proven-precedent/elements-and-models-treaty-killer-robots>, accessed on January 29, 2024.

While there is yet to be a universally accepted definition of lethal autonomous weapons in international law, one may consider, in line with Lewis, that they are "Weapons that, once activated, can designate, select, and employ force (lethal) against targets without human intervention"⁷. However, Schmitt, in his discussion, contends that:

"The crucial point when discussing full autonomy is the capability to identify, fixate on a target, and engage a person or object without human interface. Although a human operator may retain the ability to take control of the system, it can operate without any human oversight. Of course, a fully autonomous system is never entirely free from human intervention. The system's design or the operator must, at the very least, program it to operate according to specified parameters."⁸

The definition proposed by Afonso Seixas-Nunes characterizes Autonomous Weapon Systems as armaments conceived for specific missions, endowed with adaptive capability and aptitude to identify, select, and target military objectives without human intervention. This definition encompasses three crucial elements, conferring a distinctive identity upon Autonomous War Weapons.⁹ It is imperative to recognize that, fundamentally, lethal autonomous weapons are weapon systems, thereby avoiding the ascription of a fantastical autonomy typical of science fiction.¹⁰ It is noteworthy to emphasize that Lethal Autonomous Weapons are always conceived and programmed by humans, dispelling any speculations about autonomy devoid of human control.¹¹

The conception of weaponry systems operating without human intervention warrants cautious consideration. From one perspective, Autonomous Weapons usher in a broad range of new opportunities, dispensing with the direct involvement of human

⁷Lewis, Dustin. «An Enduring Impasse on Autonomous Weapons», *Just Security*, 28 de setembro de 2020. Available at: <<https://www.justsecurity.org/72610/an-enduring-impasse-on-autonomous-weapons/>>, accessed on 27 de January de 2024.

⁸Schmitt, Michael N., «Autonomous Weapons Systems and International Humanitarian Law: A Reply to the Critics», *Harvard National Security Journal Features* (2013), pp. 1-37

⁹Seixas-Nunes, Afonso. *Autonomous Weapons Systems and the Procedural Accountability Gap*, 46 *Brook. J. Int'l L.* 421 (2021). Available at: <<https://brooklynworks.brooklaw.edu/bjil/vol46/iss2/3>>. Accessed on January 5, 2024.

¹⁰Brent J. Steele & Eric A. Heinze, *From Smart To Autonomous Systems: Confounding Territoriality And Moral Agency*, In *The Future of Just War: New Critical Essays* 98, 99 (Caron E. Gentry & Amy E. Eckert eds., 2014); e Tim McFarland & Tim McCormack, *MIND THE GAP: CAN Developers Of Autonomous Weapons Systems Be Liable For War Crimes?*, 90 *Int. L. Stud.* 361, 362 (2014).

¹¹Simon Chesterman, *Artificial Intelligence and the problem of autonomy*, 1 *Notre Dame J. Emerging Tech.* 210, 232 (2020).

operators in the target selection process. Numerous arguments have been articulated by researchers and states in support of the development of these lethal weapons, including the reduction of the human burden resulting from armed conflicts, especially concerning the physical and psychological health of combatants; the accessibility of these AIs to inaccessible or inhospitable areas for undertaking missions deemed too perilous for human combatants, thereby enhancing the strategic position of ground forces and facilitating a more agile response to enemy forces; the decrease in the likelihood of casualties in high-risk operations; a lower cost involved in deploying armed conflicts for states; a high capacity for quick and precise decision-making as opposed to human capability; greater adherence to International Humanitarian Law and enhanced ease in investigating and documenting incidents, contributing to the ability to adopt corrective measures and automatically generate information related to unexploded ordnance.¹²

However, from another perspective, advancements in Artificial Intelligence, when incorporated into the realm of lethal autonomous weapons, bring forth the possibility of multiple unintended and, in some cases, illegal consequences on the theater of operations, requiring special attention. As Louise Beja discusses, these include technical issues (encompassing faults, malfunctions, and errors in programming); susceptibility to hacking attacks or loss of control over such machines; the threat of dissemination and appropriation by non-state forces; dehumanization and trivialization of conflict, thereby promoting an increasing incentive for the use of force; and the potential infringement of human rights.¹³

To better comprehend this type of technology, it is crucial for this investigation to examine some of the armaments currently available on the market (for military purposes). In March 2020, for instance, the Libyan government employed a Kargu-2 quadcopter in the civil war, according to the UN report "Letter dated 8 March 2021 from the Panel of Experts on Libya established pursuant to resolution 1973 (2011) addressed to the President of the Security Council." This drone "hunted" a human target without being instructed to do so. It

¹²Beja, Louise Amorim. *Aplicação militar da inteligência artificial nos sistemas de armas autónomas letais em tempos de guerra*. Santos, N., & Barbosa, L. (coord.). I DoctoralStudents Network: (DocS-Net): "Desafios Sociais, Incerteza e Direito" - Livro de Resumos, Instituto Jurídico da Faculdade de Direito da Universidade de Coimbra, 2023, pág. 19, available at: <<https://hdl.handle.net/10316/107469>>, accessed on 05 January 2024.

¹³Beja, Louise Amorim. *Aplicação militar da inteligência artificial nos sistemas de armas autónomas letais em tempos de guerra*. Santos, N., & Barbosa, L. (coord.). I DoctoralStudents Network: (DocS-Net): "Desafios Sociais, Incerteza e Direito" - Livro de Resumos, Instituto Jurídico da Faculdade de Direito da Universidade de Coimbra, 2023, p. 19, available at: <<https://hdl.handle.net/10316/107469>>, accessed on 05 January 2024.

marked the first time in history that a lethal autonomous weapon – also referred to as a "killer" robot – was used.¹⁴

The SWARM DRONE, which operates as a swarm of drones, is another noteworthy example. Let's examine the description that one of the responsible companies provides on its commercial website about the product:

"In just a few minutes, a swarm of up to 20 drones can approach your battlefield equipped with special sensors and payloads. All equipment and communication components used in our system are of military quality. Our swarm of drones equipped with radiation sensors can scan large areas to search for radioactive threats, help monitor situations in real-time, and provide a more comprehensive analysis. With swarms ranging from 4 to 20 drones and an automated scanning feature, the swarm can collect data across hundreds of acres."¹⁵

Regarding recent armed conflicts, for example, this type of AI swarm technology in Ukraine offered a revealing perspective on the transformative potential of drones, including those operating outside of swarms. These devices played crucial roles in intelligence gathering, surveillance, and the execution of precision attacks. The effectiveness of these functions, enhanced by the incorporation of artificial intelligence and space technology, constitutes concrete evidence that, in the context of modern warfare, technological superiority can confer a direct strategic advantage. Simultaneously, innovations promoted by China, notably the development of missiles aimed at high-value military targets, indicate a significant shift in the global military mindset. These advances underscore the intrinsic vulnerabilities associated with excessive reliance on conventional high-cost military hardware. This scenario clearly outlines a direction for the future, oriented towards agile, economically viable, and strategically versatile technologies, exemplified by the application of swarm drones.¹⁶

¹⁴(Retd), General Dushyant Singh. Swarm Drones - New Frontier of Warfare. *Aero India* 2021. LAND FORCES. Available at: <https://www.spslandforces.com/story/?id=747&h=Swarm-Drones---New-Frontier-of-Warfare>. Accessed on January 30, 2024.

¹⁵AUTONOMOUS DRONE SWARMS: Made for Military, Police, and Emergency Response Forces. ICARUSS WARMS. Available at: <https://www.icaruswarms.ai>. Accessed on January 30, 2024.

¹⁶Cheek, Robert. *Autonomous Swarm Drones New Face of Warfare*, in National Defense.

The SPUR – Special Purpose Unmanned, is an unmanned rifle designed for special use. This Artificial Intelligence was specifically crafted to:

“Provide precision firepower from an unmanned platform. Featuring secure chamber and clear fire capabilities, it allows the safe and reliable deployment of the weapon system, granting the operator the ability to handle and protect the weapon remotely. These features also provide the operator with the capability to clear malfunctions and safely unload the platform before recovery. Chambered in 6.5 Creedmoor, enabling precision firing up to 1200m, the SPUR™ can also utilize the NATO 7.62×51 cartridge for ammunition flexibility. Due to its highly capable sensors, the SPUR™ can operate in a myriad of conditions, both day and night”.¹⁷

Certainly, in this lethal weapon, death will come, and it will have four "legs." The SPUR made its debut on the show floor at a U.S. Association, notably the annual Army convention (AUSA) in Washington, D.C. Although Ghost Robotics (the company behind it, in partnership with SWORD International) has collaborated with various other companies to explore defense and security applications for its Q-UGVs, this one is considered one of the pioneers in the realm of unmanned systems with a real weapon mounted on it. Unarmed versions of the Q-UGV are already in limited use with the U.S. The 325th Security Forces Squadron at Tyndall Air Force Base in Florida was, until now, in the testing phase by other units within this service.¹⁸

The Iron Dome, a missile defense system, gained widespread recognition for its consolidated use by the State of Israel in its national security framework. In operation for just over a decade, this system was developed with substantial financial and technical support from the United States and became operational around 2011, originating in 2007. According

¹⁷ SWORD DEFENSE SYSTEMS. *Spur™ Mod 0 Pat. Pend. Special Purpose Unmanned Rifle*. Available at: <<https://sworddefense.com/spur/>>. Accessed on January 30, 2024.

¹⁸Trevithick, Joseph. *Robot Dogs Now Have Assault Rifles Mounted On Their Backs (Updated)*, in *The War Zone*, 2021. Available at: <<https://www.twz.com/42717/robot-dogs-can-now-have-6-5mm-assault-rifles-mounted-on-their-backs>>. Accessed on January 30, 2024.

to Israeli authorities, the device is approximately 90% effective in intercepting short-range rockets commonly used by Hamas and other groups in the region.¹⁹

Functioning as an air defense system designed to intercept rockets targeting Israeli territory, the Iron Dome follows the typical structure of air defense systems, comprising three fundamental elements. Initially, a radar is responsible for detecting incoming rockets. Subsequently, a command and control system processes the gathered information, triggering the activation of the third component: the interceptor. This interceptor is a missile designed to neutralize the adversary's projectile by destroying it. Essentially, the interceptor detonates the rocket in the air. It's important to note that this technology cannot intercept everything, such as ballistic missiles from other states. Israel is developing other systems, like the Sling or Arrow of David, to intercept such projectiles.²⁰

The previously listed models of lethal weapons represent just a fraction of the myriad developments and commercializations within the military sphere. In this investigation, it is imperative to comprehend how Artificial Intelligence is embedded in lethal weapons and what impact this entails.

Within the military operational landscape, there are three distinctions concerning the application of artificial intelligence in weaponry. The first scenario is Humans in the Loop (HITL), wherein humans are directly involved in real-time decision-making and control of autonomous weapon operations. They play an active and continuous role in monitoring, assessing, and, if necessary, intervening in the actions of the autonomous system, maintaining direct and immediate control over AI operations to ensure appropriate supervision and intervention.²¹

The second scenario is Humans on the Loop (HOTL), in contrast to the previous scenario, where humans are more removed from the day-to-day operations of autonomous weapons. While still supervisory, their intervention is more sporadic and less direct than in the "in the loop" model. Artificial intelligence is responsible for autonomous decision-

¹⁹Kirby, Jen. *Israel's Iron Dome, explained by an expert*, in Vox, 2021. Available at: <<https://www.vox.com/22435973/israel-iron-dome-explained>>. Accessed on January 30, 2024.

²⁰Kirby, Jen. *Israel's Iron Dome, explained by an expert*, in Vox, 2021. Available at: <<https://www.vox.com/22435973/israel-iron-dome-explained>>. Accessed on January 30, 2024.

²¹Beja, Louise Amorim. *Aplicação militar da inteligência artificial nos sistemas de armas autónomas letais em tempos de guerra*. Santos, N., & Barbosa, L. (coord.). I Doctoral Students Network: (DocS-Net): "Desafios Sociais, Incerteza e Direito" - Livro de Resumos, Instituto Jurídico da Faculdade de Direito da Universidade de Coimbra, 2023, p. 19, available at: <<https://hdl.handle.net/10316/107469>>, accessed on 05 January 2024.

making in most situations, but humans retain the ability to intervene and adjust parameters, as well as establish general policies.²²

In both cases, human presence is acknowledged as essential to ensuring ethical, legal, and operational compliance. The choice between HITL and HOTL depends on ethical, security, and efficiency considerations in specific contexts. The central issue is to strike a balance between AI autonomy and human control to ensure responsible and ethical decisions in the use of autonomous weapons.

The third and final scenario is Humans out of the Loop (HOOTL), which allows machines to employ lethal force against both machines and humans based on their own assessment, without direct human intervention or supervision.²³

Regarding the structure and composition of the Armed Forces in a post-AI arrival context within military operations, four scenarios are outlined. In the first scenario, existing military structures are supplemented, with AI incorporating equipment and technology into pre-existing brigades and divisions without fundamentally altering them. In the second scenario, autonomous systems partially replace existing structures, with human presence completely eliminated in specific areas, such as submarine warfare. In the third scenario, autonomous systems constitute parallel forces, coexisting with human divisions and divisions composed entirely of AI. In the fourth and final scenario, autonomous systems are provided by external entities, with all AI infrastructure being granted by private initiative, transforming the operationalization of national security into a contracted service.²⁴

Therefore, the increasingly pervasive implementation of lethal autonomous weapons will inevitably impact the rights of combatants and civilians, as well as the obligations of the deploying State. In this context, it is necessary to have an understanding of the technical aspects inherent in such systems, enabling the law to keep pace, albeit more slowly, with the ever more frequent technological advancements. The most significant advancement in

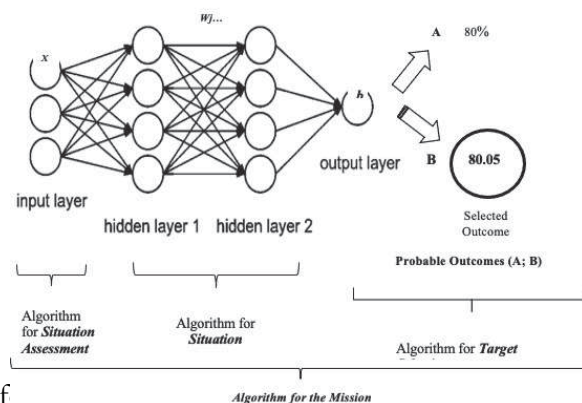
²²Beja, Louise Amorim. *Aplicação militar da inteligência artificial nos sistemas de armas autónomas letais em tempos de guerra*. Santos, N., & Barbosa, L. (coord.). I DoctoralStudents Network: (DocS-Net): “Desafios Sociais, Incerteza e Direito” - Livro de Resumos, Instituto Jurídico da Faculdade de Direito da Universidade de Coimbra, 2023, p. 19, availableat: <<https://hdl.handle.net/10316/107469>>, accessedon 05 January 2024.

²³Beja, Louise Amorim. *Aplicação militar da inteligência artificial nos sistemas de armas autónomas letais em tempos de guerra*. Santos, N., & Barbosa, L. (coord.). I DoctoralStudents Network: (DocS-Net): “Desafios Sociais, Incerteza e Direito” - Livro de Resumos, Instituto Jurídico da Faculdade de Direito da Universidade de Coimbra, 2023, p. 18 e 19, availableat: <<https://hdl.handle.net/10316/107469>>, accessedon 05 January 2024.

²⁴Beja, Louise Amorim. *Aplicação militar da inteligência artificial nos sistemas de armas autónomas letais em tempos de guerra*. Santos, N., & Barbosa, L. (coord.). I DoctoralStudents Network: (DocS-Net): “Desafios Sociais, Incerteza e Direito” - Livro de Resumos, Instituto Jurídico da Faculdade de Direito da Universidade de Coimbra, 2023, p. 19, availableat: <<https://hdl.handle.net/10316/107469>>, accessedon 05 January 2024.

artificial intelligence over the past two decades has come from a subfield of machine learning known as "Deep learning," also referred to as profound learning.

Figure 1 - AI Data Processing according to Afonso Seixas-Nunes (2021)



SOURCE: Af

Systems and the

Procedural Accountability Gap, 46 *Brook. J. Int'l L.* 421 (2021), p. 14. Available at: <https://brooklynworks.brooklaw.edu/bjil/vol46/iss2/3>. Accessed on January 5, 2024.

As illustrated in Figure 1, deep learning involves the use of artificial neural networks inspired by the way it is believed that neurons in the human brain interact with each other. Such neural networks form the backbone of current and future military technology. These neural networks are goal-oriented algorithms developed through a trial-and-error process called "reinforcement learning" (RL), aiming to optimize the most promising outcome. They can establish non-causal but predictive patterns between the data provided to the system and, based on probabilities, accurately identify military targets. Any form of "deep learning" has the potential to enhance operations on the battlefield, allowing the system to "learn" and adapt during the assigned mission, adjusting to environmental uncertainties without the need for human intervention.²⁵

However, various experts in Artificial Intelligence argue that RL-based systems pose considerably higher risks of causing substantial real-world harm compared to other machine learning methods, such as classification.²⁶ Autonomous technology, with its inherent

²⁵Afonso Seixas-Nunes, *Autonomous Weapons Systems and the Procedural Accountability Gap*, 46 *Brook. J. Int'l L.* 421 (2021), p. 14. Available at: <https://brooklynworks.brooklaw.edu/bjil/vol46/iss2/3>. Accessed on January 5, 2024.

²⁶Tardif, Antoine. Vahid Behzadan, Director Of Secured And Assured Intelligent Learning (Sail) Lab — Interview Series, UNITE.AI, 2020. Available at: <https://www.unite.ai/vahid-behzadan-director-of-secured-and-assured-intelligent-learning-sail-lab-interview-series/>. Accessed on January 5, 2024.

unpredictability, opens a "Pandora's box".²⁷ This is partly due to the inability to predict how new inputs will be processed in different layers of the system and, additionally, the uncertainty regarding the accessibility of the why and how of decisions resulting from the selection process for human understanding. Such reasons explain why autonomous systems are commonly referred to as "black-box" systems - systems whose operation is opaque to human operators and military commanders, where inputs and outputs are observable, but the intermediate processes, i.e., the structural interrelationships between the data, remain invisible, much like the neural connections in our brains.²⁸

With the introduction of machine learning algorithms, systems can "learn" directly from the environment, acquiring new data at distinct speeds and volumes that would be practically unattainable for a human operator. In this regard, as discussed by Afonso Seixas-Nunes, the situation management algorithm can discern which data resources to select and under what conditions to allow an Autonomous War Weapon to choose the most appropriate course of action to achieve the target (through the target selection algorithm), all without direct human intervention.²⁹

The situation assessment algorithm plays a crucial role, as the algorithms of Autonomous War Weapons need to adapt to new circumstances on the battlefield to ensure that the selection algorithm is accurate in identifying military targets. This explains why such weapons rely on machine learning algorithms that, by definition, are non-deterministic but unpredictable, for two essential reasons. First, algorithms need to be open to the environment, continuously collecting new relevant data and adapting to the mission. However, the challenge arises in controlling the acquired data after the system is activated and mitigating the risk of biases created by the algorithms.³⁰ Secondly, machine learning algorithms select and attack military targets based on probabilities. The chosen result is the most likely among several possibilities, seeking to achieve the best possible outcome.

²⁷Afonso Seixas-Nunes, *Autonomous Weapons Systems and the Procedural Accountability Gap*, 46 *Brook. J. Int'l L.* 421 (2021), p. 14. Available at: <https://brooklynworks.brooklaw.edu/bjil/vol46/iss2/3>. Accessed on January 5, 2024.

²⁸Sousa, Susana Aires. "Não fui eu, foi a máquina": Teoria do Crime, Responsabilidade e Inteligência Artificial. In Rodrigues, A. (coord.), *A Inteligência Artificial no Direito Penal*, Almedina, Coimbra, 2022, p. 89.

²⁹Afonso Seixas-Nunes, *Autonomous Weapons Systems and the Procedural Accountability Gap*, 46 *Brook. J. Int'l L.* 421 (2021), p. 14. Available at: <https://brooklynworks.brooklaw.edu/bjil/vol46/iss2/3>. Accessed on January 5, 2024.

³⁰Afonso Seixas-Nunes, *Autonomous Weapons Systems and the Procedural Accountability Gap*, 46 *Brook. J. Int'l L.* 421 (2021), p. 14. Available at: <https://brooklynworks.brooklaw.edu/bjil/vol46/iss2/3>. Accessed on January 5, 2024.

Figure 1 demonstrates that the result does not stem from discernible human logic but rather from the "reasoning" inherent to the system, involving a qualitative human judgment that culminates in a quantifiable choice. If an Autonomous War Weapon commits illegal actions on the battlefield, the primary concern, both for robotic engineers and legal experts, will be to explain and understand what occurs within the system's "black box."³¹ For this reason, it is necessary to reflect on the responsibility for these errors or potential human rights violations.

However, if automated AI makes a decision resulting in war crimes/human rights violations, who will be criminally responsible? The answer to this question, along with other issues inherent to this research, will be addressed in the following section.

3. Reflections on international criminal responsibility in the face of crimes committed by Autonomous Weapons

International Criminal Law follows the logic that criminal responsibility is addressed as established in the Rome Statute, which not only deals with war crimes but also encompasses human rights violations in armed conflicts as a whole. The statute also includes provisions regarding crimes such as genocide, aggression, and crimes against humanity.³²

The Rome Statute provides a comprehensive framework to deal with various forms of criminal conduct in the context of armed conflicts. It establishes the International Criminal Court (ICC) as a permanent international tribunal responsible for prosecuting individuals accused of committing serious violations of international law. These crimes encompass a wide range of acts that result in severe harm to individuals and communities, especially during armed conflicts. The statute sets forth specific elements and definitions for each crime, ensuring a comprehensive legal framework for prosecuting those responsible.³³

Crimes against humanity, as defined in Article 7 of the Rome Statute, include acts committed as part of a widespread or systematic attack directed against any civilian

³¹Afonso Seixas-Nunes, *Autonomous Weapons Systems and the Procedural Accountability Gap*, 46 *Brook. J. Int'l L.* 421 (2021), p. 14. Available at: <https://brooklynworks.brooklaw.edu/bjil/vol46/iss2/3>. Accessed on January 5, 2024.

³²Ambos, Kai. *A parte geral do Direito Penal Internacional – Bases para uma elaboração dogmática*. São Paulo: Revista dos Tribunais, 2008.

³³Ambos, Kai; Carvalho, Salo de. *O Direito Penal no Estatuto de Roma: Leituras sobre os fundamentos e a aplicabilidade do Tribunal Penal Internacional*. Rio de Janeiro: LúmenJúris, 2005.

population, with knowledge of the attack. Genocide, outlined in Article 6 of the Statute, involves specific acts committed with the intent to destroy, in whole or in part, a national, ethnic, racial, or religious group. The crime of aggression, as defined in the amendments adopted in 2010, addresses the use of armed force by a state against the sovereignty, integrity, or independence of another state, constituting a manifest violation of the United Nations Charter.

Individuals considered responsible for these crimes under the Rome Statute may face proceedings before the ICC. The statute underscores the importance of accountability and justice in the face of serious violations of human rights and international law, establishing a legal framework that seeks to deter such conduct and provide justice to the victims. But in the end, when such a crime is committed through the use of a lethal autonomous weapon, who could be held responsible? The machine operator? The operation commander? The AI manufacturer? The machine itself?

In the realm of International Criminal Law, assigning responsibility for war crimes committed through the use of lethal autonomous weapons involves specific considerations. As established by the Rome Statute, which serves as a fundamental reference for the jurisdiction of the International Criminal Court (ICC), criminal responsibility falls upon individuals who have committed or ordered such crimes.

The theory of crime, within the framework of International Criminal Law, emphasizes the individualization of responsibility, requiring the identification of specific individuals who played a direct role in the criminal acts. Therefore, responsibility cannot be attributed to the machine or artificial intelligence itself but rather to the human agents involved in its programming, operation, or command.

The machine operator, the operation commander, and even the AI manufacturer can be subjects to criminal responsibility, depending on their respective contributions to the criminal acts. The Rome Statute addresses individual responsibility and establishes that individuals who commit war crimes, crimes against humanity, genocide, or aggression can be held accountable before the ICC.

Thus, determining the responsible party will depend on the specific analysis of each case, identifying individuals who actively participated in the design, operation, or control of lethal autonomous weapons in violation of international law. The application of criminal responsibility aims to ensure that those who have committed crimes are properly judged and

held accountable before international justice, in accordance with the standards established by the Rome Statute.

The analysis of responsibility under the Rome Statute, in the light of International Criminal Law, is grounded in several articles delineating categories of those accountable for serious crimes. It is noteworthy that the interpretation and application of these articles may vary depending on the specific context of each situation.

Article 28 of the Rome Statute addresses the responsibility of commanders and superior officers. It establishes that military commanders and superiors may be held accountable for crimes committed by subordinates under their authority and effective control. Responsibility is attributed when the commander or superior officer knew or should have known that forces under their command were committing such crimes and failed to take adequate measures to prevent or punish them.

For the machine operator, responsibility can be analyzed based on general principles of international criminal law, as the Rome Statute does not explicitly specify the category of machine operators (who could be considered soldiers within the military hierarchy). This leaves a gap as to whether Article 28 of the Statute would effectively apply in these cases, solely to higher-ranking commanders in military commands. However, accountability may be considered for individuals who play a direct role in the operation of instruments used to commit crimes.

Regarding the responsibility of the artificial intelligence (AI) manufacturer, it can be considered in light of Article 25 of the Rome Statute, addressing complicity in crimes. If an individual provides assistance, instigation, or facilitation for the commission of a crime by another, they can be held accountable. In this context, an AI manufacturer could be held responsible if their contribution is deemed complicit in the commission of crimes, such as war crimes.

The effective application of these standards will depend on the specific circumstances of each case, the extent of each agent's involvement in criminal acts, and the ability to demonstrate individual responsibility before the International Criminal Court (ICC). This terrain becomes a fertile ground for human rights violations and the trivialization of war/armed conflicts.

In the context of Human Rights, this trivialization poses challenges to the protection of fundamental principles, namely the principles of distinction between combatants and non-combatants, proportionality, and precaution. The implementation of lethal autonomous weapons powered by AI raises questions about responsibility for violations of these principles, as decision-making may escape direct human control.

In the ethical sphere, the principles of robotics, outlined by Isaac Asimov, emphasize the importance of preserving human life and the autonomy of machines. However, the application of these principles is challenged by the moral complexities associated with the development and use of autonomous weapons.

Yuval Harari and Stephen Hawking have expressed significant concerns about lethal autonomous weapons. Harari underscores the need for strict regulation to prevent indiscriminate use of these technologies. Hawking, on the other hand, has warned of the risks of an arms race in artificial intelligence, emphasizing the urgency of establishing robust safeguards to prevent catastrophic consequences.³⁴

In light of these considerations, it becomes imperative, from a legal and formal perspective, to establish stringent regulations and oversight mechanisms to ensure that the integration of AI into military operations is guided by unquestionable ethical principles and unwavering respect for Human Rights.

However, this discussion and regulation pertaining to the matter of autonomous weapons have been conspicuously omitted from the European Regulation Proposal. The explicit exclusion of AI systems for military purposes in the Regulation Proposal raises legitimate questions about the necessity and effectiveness of such exclusion. The justification provided in the latest version of the proposal emphasizes the autonomy of defense and national security areas, arguing that these areas are subject to international public law.

Nevertheless, the exclusion raises concerns about regulatory gaps and the need to ensure that AI technologies in the military context are addressed in an ethical and secure manner. The proposal to exclude these areas from the scope of the regulation can be interpreted as a waiver of specific regulation, leaving room for a potential lack of adequate oversight. The regulatory proposal seems to acknowledge the possibility of AI systems

³⁴Reference: "Autonomous Weapons: An Open Letter from AI & Robotics Researchers" (2015), signed by 4,985 AI and Robotics researchers and with an additional 27,800 signatories, including Stephen Hawking, Elon Musk, Steve Wozniak, and Noam Chomsky. Available at: <https://futureoflife.org/open-letter/open-letter-autonomous-weapons-ai-robotics/>, accessed on January 27, 2024.

initially developed for military purposes being subsequently used for civilian purposes. However, the effectiveness of this approach in ensuring compliance with ethical and safety standards remains a concern. Supervision and compliance with regulations can be challenging, especially when dealing with complex AI systems with broad applications in both military and civilian environments.³⁵

The lack of clarity in the definition of "military purposes" in the legislation and the reliance on Title V of the Treaty on European Union may be points of criticism. The inclusion of a more precise definition could strengthen the foundation for the proposed exclusion. Not having the issue of autonomous weapons in the EU AI ACT is a significant setback, as it is considered by many to be a global milestone in the regulation of Artificial Intelligence. The ethical and legal implications of autonomous weapons may have been considered a separate domain or outside the initial scope of the regulation under consideration.³⁶

Ultimately, in this matter, there is a need to balance the imperatives of national security with the need for ethical and responsible regulation in the development and use of AI systems, especially in military contexts. The effectiveness and coherence of the proposed regulation will be crucial in addressing these concerns satisfactorily.

The preservation of human rights in the face of technological advancements, particularly in the context of lethal autonomous weapons, emerges as a paramount concern. The assignment of responsibility for crimes perpetrated by these weapons is a complex challenge, requiring a meticulous approach within the framework of International Criminal Law. When considering the culpability of operators, commanders, and artificial intelligence manufacturers, the need to adapt existing legal standards to emerging technologies becomes evident. Simultaneously, ethical and human rights issues, influenced by the principles of robotics and warnings from experts such as Harari and Hawking, underscore the risks associated with the trivialization of armed conflicts. The omission of autonomous weapons

³⁵Costa, Miguel João. *Sistemas de Armas Autónomas e Respectiva Regulamentação. Direito em Mudança: A Proposta de Regulamento Europeu sobre Inteligência Artificial: Algumas Questões Jurídicas*, Coimbra, 2023. Available at: <https://estudogeral.uc.pt/bitstream/10316/108673/1/A%20proposta%20de%20Regulamento_ebook.pdf>, accessed on January 27, 2024.

³⁶Costa, Miguel João. *Sistemas de Armas Autónomas e Respectiva Regulamentação. Direito em Mudança: A Proposta de Regulamento Europeu sobre Inteligência Artificial: Algumas Questões Jurídicas*, Coimbra, 2023. Available at: <https://estudogeral.uc.pt/bitstream/10316/108673/1/A%20proposta%20de%20Regulamento_ebook.pdf>, accessed on January 27, 2024.

in the European Regulation Proposal raises concerns about regulatory gaps, reinforcing the urgency to balance national security requirements with ethical regulations. In this scenario, the implementation of stringent regulations and oversight mechanisms is advocated to guide the incorporation of artificial intelligence into military operations, emphasizing the importance of a comprehensive and coherent approach to satisfactorily address these ever-evolving challenges.

4. Final Considerations

In the face of the advancement of autonomous weapons, it is evident that the preservation of human rights confronts significant challenges. Reflection on international criminal responsibility reveals the need to adapt existing legal standards to encompass emerging technologies. The individualization of responsibility, as outlined in the Rome Statute, underscores the importance of identifying human agents involved in the programming, operation, or command of these autonomous weapons in violation of international law. The implementation of rigorous regulations and oversight mechanisms is crucial to guide the ethical integration of AI into military operations, balancing the demands of national security with ethical regulations.

The omission of autonomous weapons in regulatory proposals highlights the urgency to comprehensively and cohesively address these concerns in the international legal realm. It is essential to ensure that regulations keep pace with technological advancements, a critical element in preserving human rights in an increasingly technological military landscape.

Regarding autonomous weapons, the discussion about their role and impact on the landscape of armed conflicts raises considerable concerns. The autonomy of these weapons, often grounded in AI systems, raises questions of proportionality, distinction between combatants and non-combatants, and the principle of precaution. This directly impacts the Third Generation of Human Rights, which encompasses solidarity rights, such as the right to development, and collective rights, such as the right to peace.

The incorporation of generative AI and autonomous weapons into the scenario of armed conflicts highlights the need for a critical review and update of existing legal and ethical frameworks. The Fourth Generation of Human Rights, centered on digital rights, privacy, and protection against information manipulation, assumes a crucial role. Ensuring compliance of military operations with these emerging rights becomes imperative to maintain

respect for the fundamental principles of human dignity in an increasingly digitized and complex environment.

In an era of technological advances, how should we safeguard order and human rights? (Socrates). The indispensable harmony between machines and humanity lies in fair coexistence, where we find true virtue (Aristotle). Let us guide progress under the light of ethics, so that society achieves a just and harmonious coexistence (Plato). Wisdom lies in preserving human rights in technological advancement, thus ensuring security and order (Confucius). In the coexistence between machines and humanity, the preservation of human rights is a categorical imperative (Immanuel Kant).

It is necessary to investigate appropriate possibilities to promote a break with the paradigms perpetuated within the framework of international criminal responsibility, as the “backbone” of the punitive system, in order to broaden the discussion about the punishment of crimes that today's society places before us. As Foucault once argued, contemporary reformers must be pressurised into the task of re-founding justice. In the journey of technological progress, the preservation of human rights must be placed as the moral compass of this path, because it is understood that only in this way can we mould the future where machines and humanity coexist in harmony and justice.

REFERENCES:

(Retd), General Dushyant Singh. *Swarm Drones - New Frontier of Warfare*. Aero India 2021. LAND FORCES. Available at: <https://www.spslandforces.com/story/?id=747&h=Swarm-Drones---New-Frontier-of-Warfare>. Accessed on January 30, 2024.

Afonso Seixas-Nunes, *Autonomous Weapons Systems and the Procedural Accountability Gap*, 46 *Brook. J. Int'l L.* 421 (2021), p. 14. Available at: <https://brooklynworks.brooklaw.edu/bjil/vol46/iss2/3>. Accessed on January 5, 2024.

Ambos, Kai; Carvalho, Salo de. *O Direito Penal no Estatuto de Roma: Leituras sobre os fundamentos e a aplicabilidade do Tribunal Penal Internacional*. Rio de Janeiro: LúmenJúris, 2005.

Arkin, Ronald C. "Governing Lethal Behavior in Autonomous Robots," 7—26 (2009). "Several automatic and semi-autonomous systems are already in operation and used by States, such as Samsung Sentries, Iron Dome, AEGIS, and the Brimstone Missile, among others." Also, refer to the European Parliament's Draft Resolution on Autonomous Weapon Systems, available at: https://www.europarl.europa.eu/doceo/document/B-8-2018-0355_PT.html, accessed on January 5, 2024.

AUTONOMOUS DRONE SWARMS: Made for Military, Police, and Emergency Response Forces. ICARUSS WARMS. Available at: <<https://www.icarusswarms.ai>>. Accessed on January 30, 2024.

Autonomous Weapons: An Open Letter from AI & Robotics Researchers (2015). Available at: <<https://futureoflife.org/open-letter/open-letter-autonomous-weapons-ai-robotics/>>, accessed on January 27, 2024.

Beja, Louise Amorim. *Aplicação militar da inteligência artificial nos sistemas de armas autónomas letais em tempos de guerra*. Santos, N., & Barbosa, L. (coord.). I DoctoralStudents Network: (DocS-Net): “Desafios Sociais, Incerteza e Direito” - Livro de Resumos, Instituto Jurídico da Faculdade de Direito da Universidade de Coimbra, 2023, p. 19, available at: <<https://hdl.handle.net/10316/107469>>, accessed on 05 January 2024.

Brent J. Steele & Eric A. Heinze, *From Smart To Autonomous Systems: Confounding Territoriality And Moral Agency*, In *The Future of Just War: New Critical Essays* 98, 99 (Caron E. Gentry & Amy E. Eckert eds., 2014); e Tim McFarland & Tim McCormack, *MIND THE GAP: CAN Developers Of Autonomous Weapons Systems Be Liable For War Crimes?*, 90 *Int. L. Stud.* 361, 362 (2014).

BRIMSTONE ADVANCE ANTI-ARMOUR MISSILE, Army Technology. Available at: <https://www.army-technology.com/projects/brimstone/>. Accessed on January 5, 2024. This source highlights the Brimstone Missile as a perfect example, indirectly targeting armored vehicles autonomously.

Cheek, Robert. *Autonomous Swarm Drones New Face of Warfare*, in *National Defense*, 2021.

Chesterman, Simon. *Artificial Intelligence and the problem of autonomy*, 1 *Notre Dame J. Emerging Tech.* 210, 232 (2020).

Costa, Miguel João. *Sistemas de Armas Autónomas e Respectiva Regulamentação. Direito em Mudança: A Proposta de Regulamento Europeu sobre Inteligência Artificial: Algumas Questões Jurídicas*, Coimbra, 2023. Available at: <https://estudogeral.uc.pt/bitstream/10316/108673/1/A%20proposta%20de%20Regulamento_ebook.pdf>, accessed on January 27, 2024.

Key Elements of a Treaty on Fully Autonomous Weapons (2019). Available at: <<https://www.stopkillerrobots.org/wp-content/uploads/2020/03/Key-Elements-of-a-Treaty-on-Fully-Autonomous-Weapons.pdf>>, accessed on January 29, 2024; and *New Weapons, Proven Precedent, Elements of and Models for a Treaty on Killer Robots* (2020). Available at: <<https://www.hrw.org/report/2020/10/20/new-weapons-proven-precedent/elements-and-models-treaty-killer-robots>>, accessed on January 29, 2024.

Kirby, Jen. *Israel's Iron Dome, explained by an expert*, in *Vox*, 2021. Available at: <<https://www.vox.com/22435973/israel-iron-dome-explained>>. Accessed on January 30, 2024.

Lemos, Miguel; Costa, Miguel João. *Inteligência Artificial e Direito da Guerra: Reflexões sobre as Armas Autónomas Mortíferas*. In: *A Inteligência Artificial no Direito Penal Vol. II*. Anabela Miranda Rodrigues (coord.), Coimbra: Almedina, 2022.

Lewis, Dustin. *An Enduring Impasse on Autonomous Weapons*, Just Security, 28 de setembro de 2020. Available at: <<https://www.justsecurity.org/72610/an-enduring-impasse-on-autonomous-weapons/>>, accessed on 27 de January de 2024.

Schmitt, Michael N., «Autonomous Weapons Systems and International Humanitarian Law: A Reply to the Critics», *Harvard National Security Journal Features* (2013), pp. 1-37

Seixas-Nunes, Afonso. *Autonomous Weapons Systems and the Procedural Accountability Gap*, 46 *Brook. J. Int'l L.* 421 (2021). Available at: <<https://brooklynworks.brooklaw.edu/bjil/vol46/iss2/3>>. Accessed on January 5, 2024.

Sousa, Susana Aires. “Não fui eu, foi a máquina”: Teoria do Crime, Responsabilidade e Inteligência Artificial. In Rodrigues, A. (coord.), *A Inteligência Artificial no Direito Penal*, Almedina, Coimbra, 2022, p. 89.

SWORD DEFENSE SYSTEMS. *Spur™ Mod 0 Pat. Pend. Special Purpose Unmanned Rifle*. Available at: <<https://sworddefense.com/spur/>>. Accessed on January 30, 2024.

Tardif, Antoine. *Vahid Behzadan, Director Of Secured And Assured Intelligent Learning (Sail) Lab — Interview Series*, UNITE.AI, 2020. Available at: <https://www.unite.ai/vahid-behzadan-director-of-secured-and-assured-intelligent-learning-sail-lab-interview-series/>. Accessed on January 5, 2024.

Trevithick, Joseph. *Robot Dogs Now Have Assault Rifles Mounted On Their Backs (Updated)*, in *The War Zone*, 2021. Available at: <<https://www.twz.com/42717/robot-dogs-can-now-have-6-5mm-assault-rifles-mounted-on-their-backs>>. Accessed on January 30, 2024.