UNIVERSIDADE CATÓLICA PORTUGUESA

The Role of AI in Warfare

Autonomous Weapons Systems and International Community

Beatriz Pereira de Freitas Lima

Direito

Faculdade de Direito | Escola do Porto
2024

UNIVERSIDADE CATÓLICA PORTUGUESA

The Role of AI in Warfare:

Autonomous Weapons Systems and International Community

Beatriz Pereira de Freitas Lima

Orientadora: Prof. Maria Isabel Tavares

Direito

Faculdade de Direito | Escola do Porto
2024

‘Nascendo as guerras no espírito dos Homens é no espírito dos Homens que deve erguer-se a defesa da paz’

- Wall Unesco Paris, 2024

AGRADECIMENTOS

Aos meus pais que estiverem sempre lá para me apoiar, mesmo quando a distância nos separava, e acreditaram sempre em mim.
Senti sempre amor.

À minha pessoa favorita, a minha irmã, Catarina pelas novidades à distância, pelas piadas contadas, pela união e por nada ter mudado. A minha alma gémea.

Aos meus avós e padrinhos pela cumplicidade, pela confiança constante e pelo abraço apertado.

À minha orientadora, Maria Isabel Tavares, pela orientação, ajuda e preocupação.

Ao Rui pelo carinho e apoio constante, companheirismo e tolerância.

Às minhas amigas pela partilha de momentos, pelos fins de semana em Paris e por nunca me deixarem sentir sozinha neste percurso.

E, por último, ao Consulado Geral de Portugal em Paris e a Paris por terem sido a minha segunda casa nestes seis meses. Pelas pessoas fantásticas com quem tive o privilégio de me cruzar, pela experiência profissional e por me ter mudado tanto enquanto pessoa. Em especial, ao Samuel que será para sempre o meu querido amigo parisiense.

Obrigada.

Resumo

A inteligência artificial invadiu em todos os aspetos da sociedade. Os conflitos recentes, como o de Israel-Hamas, têm colocado uma atenção redobrada no papel que hoje a inteligência artificial desempenha nestes conflitos. Hoje, discutem-se várias aplicações deste mecanismo, mas o foco aqui será o da utilização e o desenvolvimento de Sistemas Autónomos de Guerra.

Estas armas têm suscitado opiniões diversas entre os Estados que as procuram enfrentar. O objetivo será, assim, compreender a posição da Comunidade Internacional sob este assunto, tanto na definição como na sua regulação jurídica sob o quadro do Direito Internacional Humanitário.

Palavras-chave: Inteligência Artificial; Sistemas Autónomos de Guerra; Direito Internacional Humanitário; Regulação Internacional.

Abstract

Artificial intelligence has invaded every aspect of society. Recent conflicts, such as Israel-Hamas, have placed increased attention on the role that artificial intelligence now plays in these conflicts. Today, various applications of this mechanism are being discussed, but the focus here will be on the use and development of Autonomous Weapons Systems.

These weapons have aroused diverse opinions among the states that seek to confront them. The aim will, therefore, be to understand the position of the International Community on this issue, both in terms of definition and legal regulation under the framework of International Humanitarian Law.

Keywords: Artificial Intelligence; Autonomous Warfare Systems; International Humanitarian Law; International Regulat

TABLE OF CONTENTS

ABBREVIATIONS

AC – Armed conflicts

AI – Artificial Intelligence

AWS – Autonomous weapons systems

CCW – Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to Be Excessively Injurious or to Have Indiscriminate Effects

Dod – U.S. Departament of Defense

GGE – Group of Governmental Experts of the High Contracting Partis

HRC – Human Rights Council

IHL – International Humanitarian Law

IHRL – International Human Rights Law

IL – International Law

LAWS – Lethal autonomous weapons

ML – Machine Learning

UN – United Nations

UNGA – United Nations General Assembly

1. INTRODUCTION

In every moment of history, armed conflicts have shaped the destines of nations, influencing their development and historical trajectory. Despite the atrocities inherents of war, it cannot be denied that it also served as a catalyst for technological advances that profoundly impacted military tactics and strategies[1]

Technological advancement has a close interconnection with military power. The long race to military prosperity, due to the international context of tensions and the rise of artificial intelligence, has made states look to it as a transformative force in the military[2].

The positive potential of AI in various areas is undeniable. Vladmir Putin, president of Russia, expressed recently that "whoever leads artificial intelligence will dominate the world". This idea is shared with other main powers that are now investing, hardly, in policies aimed at integrating AI into their military operations.

However, the role of AI for military purposes and the tendency of reliance in mechanisms with AI creates new challenges.

This need of States to overcome and prosper the enemy led to the constant development of new weapons. Inevitably when facing AI, one of the most important discussions in international community, has been Autonomous weapons systems, consider their rise, by the report of the Meeting of the American Association for the Scientific Advancement of Science (2019), as the third revolution in warfare, after gunpowder and nuclear weapons.[3]

The aim of this thesis is to determine the complex role of AI and AWS. The development and use of systems, during war, that are capable of operating without human supervision, raises significant ethical, legal and political concerns, making imperative identified points of intersection between these systems and norms of International Humanitarian Law (IHL).

For being a technically complex subject in itself, it will be analyzed the essential characteristics of artificial intelligence, focusing the detailed analysis on autonomous

[1] Roland, Alex (2016): War and Technology, Oxford Press
[2] *Anand, Vinod (2008). Impact of Technology on conduct of Warfare, Strategic Analysis.*
[3] Russel, Stuart J. (2019), Human Compatible: Artificial Intelligence and the Problem of Control, Viking

systems. The different perspectives within international panorama about AWS begin with the difficulty of reaching an agreement on peremptory issues, such as the definition and characterization of an autonomous system.

The purpose is to make a reasonable and comprehensive analysis that incorporates diverse perspectives and disciplines of States to address the technical, ethical and legal complexity of technology in military, balancing the crucial role of IHL and the legitimate defense interests of States in relation with these systems.

These perspectives are stamped in the formal and informal discussions by GGE in CCW, where States have space to reflect about the negative and positive aspects of AWS, but also stamped in the first resolution approved by 152 states on Lethal Autonomous Weapons (2023) by the General Assembly.

The reality in which we live makes it even more prime to evaluate the position and solutions of States, International Organizations, Academics and Civil Society, since these emerging technologies could be used in conflicts in the not-too-distant future.

2. TECHNOLOGY AS A MILITARY ALLIE

Alex Roland defends that using technology for military purposes is seen as the black side of that. Even though, technology can define and structures the way warfare proceed, are the armed conflicts that define and structure this technology for satisfy States' necessities.

This development of weapons, from artillery to anti-areas weapons, is in constant refinement, but also in the way warfare is conducted, since war is inevitable[4]. The new era is known as 'knowledge age', since is it presided with a technological evolution, without precedents[5]. The emerge of drones (Unmanned Aerial Vehicles – UAV) began in the conflict of Lebanon, when Israel used them in 1982, gaining prominence in the following conflicts.

The conducts carried out during the wars in the Middle East have brought with them the important change of strategy of warfare. The Gulf war, for Vinod Anand, was the turning point from a warfare based in elements of mass destruction to the new elements of war, abandoning the idea of "an armed conflict of e exhaustion or attrition to an idea of a conflict of individual maneuvers and strategies".

---

[4] Sabriya Alam, Olivia Jimenez, Julia Taylor, Dwaine Jengelley (2020): Unmanned and Autonomous Weapons Systems: Practices and Related Policy, https://core.ac.uk/download/pdf/359175783.pdf

[5] *Anand, Vinod (2008). Impact of Technology on conduct of Warfare, Strategic Analysis.*

The use of MQ-1 Predator in the attack against Mullah Omar, a Taliban leader, by the United States in 2001 stood out as an enabler to a 'game changer' system, as consider by Peter Warren Singer, since created the precedent for what is called *target killing*, without the necessity of the human operator to be in the active combat zone.

The changes are also evident in the increase of other actors in armed conflicts. Private companies play, nowadays, a recurrent and relevant role behind the curtain, influencing the direction taken.

2.1. ARTIFICIAL INTELLIGENCE

Despite of new emergent technologies cannot be determined of the result of war, as the more recent and modern armed conflict becomes, it is uncontrollable that technology defines the course and strategies of war[6].

Melvin Kranzberg, an American historian, views technology as a process of manipulating mechanisms to achieve human goals, not depending on itself, but on what humans want to attain. However, what happens if human intelligence and experience is set aside, and the conduct and decision of systems based on AI gained an important role on armed conflicts?

Artificial Intelligence has increased the role in every aspect of life at the present time. For that reason and along with the need to keep improving and advancing, armed forces are focusing in adapting their capabilities for military purposes using AI on battlefield, changing the perspective of States, Organizations and Non-State armed groups.

AI is about possessing characteristics of human intelligence, such as perception, think, learn, invention and so on. As said by Stuart Russel, machines are considered intelligent and beneficial to the extent that their actions can be expected to achieve human's objectives, since they don't have objectives of their own[7].

According to the International Committee of Red Cross, Artificial Intelligence can be defined as "involves the use of computer systems that carry out actions that would originally require human cognition, planning and reason" and raises significant questions about the "integration in weapon systems, use in cyber and information operations and underpinning military 'decision support systems'. [8]

---

[6] Roland, Alex (2016): War and Technology, Oxford Press

[7] Russel, Stuart J. (2019), Human Compatible: Artificial Intelligence and the Problem of Control, Viking P. 24 e 25

[8] ICRC

The intricacies of the underlying technologies shaping their decision-making lies on algorithms. These algorithms can represent instructions given to systems to guide them to executing a specific task or solving problems, although these algorithms, that are given by humans to machines, can influence possible responses, being capable of creating bias and discriminative solutions. Nevertheless, the existing difficulties, in the specific area of armed conflicts, is a subject for later analysis.

AI-based technologies for military purposes can also represent an adaptation of a system created to civilian uses, including speech recognition, autonomous flight, facial and object recognition[9]

To better understand the capabilities of AI is crucial a technical complementary view about machine learning. Machine learning, as explained by Sarah Brown[10], is a subfield of AI.

This mechanism is capable to be a defensive or an offensive mechanism since it has competencies of understanding the origin and scope of an attack, improving the targeting domain in battlefield by identify the existence, the location – by processing images and video- and vulnerability of targets[11].

ML represents not only a descriptive system, as mentioned, but also have the predictive capability, while it can - according to the data sets given – realize possible reasons, intentions and predicts what is possible to happen, learning from precedents of trial and error or *sample data,* without any programed solution, as said by the pioneer Arthur Samuel – what is now called 'deep learning'.

In a final moment, this system will use the data, it has been given, to make recommendations on the best way to act according to the circumstances[12]having the ability to learn patterns. Boris Katz calls this whole process "intelligent behaviors".

This data, as explained by R. Verschae, that train these systems can be previously collected, but also 'real-time input from sensors and cameras to be used for images processing and object detection and recognition'[13].

Facing this process, it is correct to say that ML can revolutionize the opportunities that AI create to be use by armed forces, in context of armed conflicts, creating a vast

[9] Pascale Fung, Director of the Centre of AI Research, Hong Kong University of Science and Technology

[10] Brown, S. (2021) Machine Learning, Explained. MIT Sloan School of Management, Cambridge. https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained

[11] SIRPI Russia AI's Strategy

[12] Brown, Sarah

[13] (2018) *Verschae, R. & Ruiz-Del-Solar, J. (2015). Object Detention: Current and Future Direction s*

number of complexities that imply accuracy and innovation, but also difficulties, for instance, in the compliance with norms of International Humanitarian Law, such as distinction.

ML can also be divided in three distinction types: supervised, unsupervised and reinforcement. The first one portrays a situation in which ML systems are trained with a set of data, provided by operators, to make systems capable to learn and become more precise. In this situation, is given input features and corresponding output labels.

The unsupervised is about a system that function with openness in creating pardons and analyzing information, that was not asked explicitly. The data, contrary to ML supervised, is not labeled by the programmer and it main purpose is to find patterns and structures.

The latter seems to be the most controversial. ML, in the case of reinforcement, will go through a trial-and-error process via a reward system that will give the idea to the system that it has made the right decision.

So, even though, AI and the emerging technologies based in it, for military purposes, are primarily for non-lethal tasks - such as telemedicine or robotic mine-clearing – it can also be weaponized, in order to originate autonomous weapons.

A mention can be made that the relevance of these technical aspects can be more interesting when facing the specific case of autonomous weapons to understand the difficulties created in an ethical, moral and legal view.

3. AUTONOMOUS WEAPONS SYSTEMS

Since AI have the capability to enhance effectiveness, efficacy and decision-making, this is viewed as an opportunity to military purposes in different areas, such as what is so-called Autonomous Weapons Systems.

Autonomous weapons systems, abbreviated hereafter as AWS, are a central military AI-based technologies that represents a radical transformation for warfare. The relation to the activities of machines, algorithms and ML can create loopholes to the application of IHL and IHRL norms.

These systems have capabilities of reconnaissance, surveillance and target acquisition and, what is seen nowadays is not a new type of weapons, but just an adaptation of existing weapons. As mentioned by U.S. Deputy Secretary of Defense, William J. Lynn III, “few weapons in the history of warfare, once created, have gone

unused". This is the case of MQ-Reaper 9, an American drone, that have now the capacity of using AI in their actions[14]

These systems have the unique characteristic, highlighting autonomy that gives the independent competency of select and attack a target without a human operator to supervise, control or intervene[15]and without the conscience of the implications of their actions, due to the fact that "machines will never be able to bring a genuine humanity to their interactions, no matter how good they get at faking it"[16]

In this specific case, ML, as explained by Peter Margulies, have four key domains for adaptability and efficacy of these systems. In a concise approach these four domains have to do with probability valuation that helps with the selection of targets, helping with the recognition of patterns and visual for face identification, but also with the distinction between civilians and combatants. The last two domains are the movement or instability of dynamics '(…) which entails an awareness of the dynamic state caused by an AWS's movement and actions taken by the AWS in response to encounters in the field (..)' and lastly, the capacity of interpretation that explains the reason that identify the target[17]

To introduce the topic is relevant to have already understood some specifics, such as the difference between automation and autonomy. Autonomy is not about automatization, considering that already exist automatic systems in course. Remotely piloted vehicles or guided missiles can have a certain degree of automation in their behavior, but it cannot be equated with AWS because the use of these 'drones' continue to be centered on human scrutiny, even if they are launched remotely.

Prof. Mary Cummings, professor in Duke Univeristy[18], makes a comparation between drones and AWS, by saying that in the latter, the core functions in weapons – decision, execution and validation – can be all in the computer and not, as in drones, divided between a human decider and validator and a robotic executor.

So, although it seems to prevail an idea about these weapons, it becomes difficult to arrange a unanimous definition. In the opinion of International Committee for Robot Arms Control, a definition is not required, but rather a focus on concepts of human-

[14] https://www.forbes.com/sites/davidhambling/2020/12/11/new-project-will-give-us-mq-9-reaper-drones-artificial-intelligence/?sh=5c7d46857a8e
[15] Righetti et al.: (2018): Lethal Autonomous Weapon Systems (Ethical, Legal and Societal Issues), p. 200
[16] Google, Perspectives on Issues in AI Governance, January 2019, pp. 22 https://ai.google/static/documents/perspectives-on-issues-in-ai-governance.pdf
[17] Margulies, Peter. Making Autonomous Weapons Accountable: Command Responsibility for Computer-Guided Lethal Force in Armed Conflicts, p.8
[18] CCW/GGE: Chair Summary Report 2018

machine interaction, as mentioned, autonomy and human control or secondary characteristics, citing predictability and reliability[19]

The main concerns of these weapons, as UNSG's Advisory Board on Disarmament Matters mentioned in 2013, are three: "(…) the ability to conform to existing law, including IHL, IHRL or general international law, potential problems associated with the design of future fully AW that could require disarmament action or the ethical limits to robot autonomy in deciding on the life and death of a human"[20]

Autonomous weapons systems can vary between lethal autonomous weapons systems and non-lethal autonomous systems, but also from fully autonomous to partially autonomous. All of these sub-categories change the way this system is approached, since it changes the underlying complications and dynamics.

3.1 TYPES OF LETHAL AUTONOMOUS WEAPONS AND PARTIALLY AUTONOMOUS WEAPONS

The concept of lethality brings different purposes and capabilities to these systems and, for that reason, makes an important distinction into two categories of AWS. Even though, is not a concept to analyze autonomous weapons as a whole, is a main concept to characterize what is Lethal Autonomous Weapons Systems.

The concept of fully and partially have more compliance with the concept and spectrum of autonomy and independence from the human intervention. LAWS are the most problematic, and States do differentiate them for the partially autonomous weapons mainly because of the effects and consequences on the target – even though this approach is not unanimous, as we will see later from the Statement of Ireland in 2019.

Non-lethal AWS or partially AWS do not mean that lethality is not a part of their system, instead is only that 'lethality' is not the primary purpose. That being said, these systems are designed to incapacitate, deter or disrupt targets without causing significant physical harm or permanent damage, function mainly as a defensive system.

In the case of partially autonomous weapons, or human-in-the-loop systems, the approach is altered because it can still select and attack a target, but in this case the decision is made by a framework that is defined by a human operator. That means that

[19] (Chair Summary of Agenda 6(a) April 2018 GGW).

[20] United Nations, General Assembly, 'Work of the Advisory Board on Disarmament Matters', Report of the SG, A/68/206, 26 July 2013, para. 42)

the systems do not have the capacity of individual adaptation, self-learning and 'self-research' beyond what was already defined.

In words of France diplomacy "(12) (…) a PALWS is a LETHAL weapon system whose decision-making functions are defined according to the robotics meaning of decision-making autonomy, i.e. within a specific framework of action. A PALWS cannot take lethal initiatives that would result in it altering its functional scope."

For instance, the Irom Dome, an Israel's missel defense system used to protect Israel from the recent attack of Iran is gained new capabilities based on AI that can improve the efficiency of the system.[21]

The LAWS are fully AWS designed with destructive mechanisms, used in offensive contexts that can cause harm or inflict damage to identified and selected target, such as humans or infrastructures, without direct human intervention.

France and Germany, consider that these weapons, by their nature, are incapable to comply with principles of IHL, such as distinction and proportionality, but also, susceptible of creating an issue with the principle of human dignity. Yet, as explained in CCW/GGE report (22d)[22] lethality does not 'prejudice the application of and respect for all rules relevant to the conduct of hostilities.'

However, some believe that the boundary and difference between lethal and non-lethal can vanish, depending on the circumstances and context of use, because it is possible to create equally catastrophic consequences and outcomes.

In a statement by Ireland's delegation in 2019 on a meeting of the Convention on Certain Convention Weapons (CCW)[23], a subject that will be analyzed later, it was defended that the use of the term 'weapon systems' encompasses the notion of a weapon that produces the intended effect in relation to a given target, as such, they include both the elements associated with holding and targeting processes, as well as 'the application of the effect to the target'.

In that case, the delegation considered that the concept of lethality, even agreed by States as categorizing this type of system, could not be the gateway to autonomous systems for the CCW panel. In their perspective, this notion raises difficulties about the

[21] Gautam Ramachandra: How Artificial Intelligence is improving Iron Dome https://medium.com/@gautamrbharadwaj/how-ai-is-improving-iron-dome-3894cd3668f9
[22] CCW/GGE report (22d)
[23] CCW/GGE, Statement of Ireland, 25-29 march 2018 - agenda item 5(b)

focus placed on ‘less lethal’ systems, since it prompted a reevaluation of whether the discussed weapon system should be exclusively intended to cause lethal effects.

In that case, the delegation encouraged the consideration also of “non-lethal weapons systems that could have lethal impacts in certain circumstances but where the lethal effect is not primary purpose of the system (…) lethality does not prejudice the application of and respect for all rules relevant to the conduct of hostilities”

In 2022, Sierra Leone, the State of Palestine and Uruguay also made a statement defending that lethality is not inherent to these weapons, but only an ‘effect or manner of use’, concluding that ‘any weapon system can be contrary to international law regardless of whether it is lethal or not.’ However, China, in 2022, considers that a characteristic of unacceptable AWS, by saying that lethality means ‘sufficient lethal payload (charge) and means’.

The first time that was mentioned the (possible) use of a LAWS in an armed conflict was Libya conflict the use of Kargu-2, a Turkish system. In 2021, a letter made by the Panel of Experts on Libya established pursuant to resolution 1973/2011 addressed to the President of UNSC indicates on paragraph 63 that STM Kargu-2 was used against members of Libyan National Army, referring to this system as a LAWS. In this case the report says that *“The lethal autonomous weapons systems were programmed to attack targets without requiring data connectivity between the operator and the munition: in effect, a true “fire, forget and find” capability.”*

This capability is based on machine-learning, able to classify objects and (apparently to) autonomously assumes that, once launched, the drone needs no further control or guidance, guiding itself individually to the target[24]

Any example is also the Harpy system used by Israel forces that is programmed before launch and autonomously navigate to a specific area, detecting and neutralizing these enemy radar signals with a high accuracy for extent periods.[25]

3.2.SPECIFIC CHARACTERISTICS OF AWS

3.2.1 AUTONOMY

According to OTAN, autonomy can be considered “*the ability of a system to respond to uncertain situations by independently composing and selecting among different*

[24] (Collins dictionary, 01/2024).
[25] Autonomous Weapons | Harpy https://www.iai.co.il/p/harpy

courses of action in order to accomplish goals based on knowledge and a contextual understanding of the world, itself, and the situation" [26].

'Greater autonomy means more independence' was said by Frank Groomer[27] from Ireland delegation, emphasizing that the level of autonomy and independency must be the cornerstone of discussions between States, since determinations of the type and level of autonomy must be in line with IHL.

In that case, the level and type of autonomy depends on the type of function given to the systems, as defended by Tsvetelina van Benthem[28]. Birgitta Dresp-Langley[29] also defends this idea by saying that in the case of autonomous weapons, this definition must include the mention that their behavior and decision will be based and controlled by algorithms. So, the easiest way to understand what the concept of autonomy means is by understanding, at a functional level, the type of actions that are carried out by the systems[30].

In the case of LAWS, autonomy is face-to-face with different aspects of robotic and engine, but also with the 'human element', so it is necessary to approach three diverse categories of human-machine relation.

Human-on-the-loop system is an autonomous system designed to provide the human operator with the ability to intervene and determine possible events before the danger occurs. In this case, the human supervises all operations and have the power to cancel them, if necessary.

In contrast to this, is the situation of semi or partially AWS also known as human-in-the-loop systems. These systems, when activated, have the sole purpose of interacting with an individual target or a specific group of targets that have been previously selected by a human operator.

The third level is the most relevant when it comes to LAWS, Human-out-of-the-loop. This is a system that, once activated, is able to select and engage a target without human intervention. There are examples of these systems already in operation, such as

[26] Definition given by NATO Science and Technology, 2020
[27] Statement of Ireland in CCW/GGE 2019
[28] Tsvetelina van Benthem (2020): Exploring Changing Battlefields, Unintended Engagements and the Law of Armed Conflicts
[29] Birgitta Dresp-Langley (2023), The Weaponization of Artificial Intelligence: What the public needs to be aware of
[30] Paul Scharre 2016, Autonomous Weapons and Human Control

the use of autonomous systems that independently use 'electronic jamming' to interrupt enemy communications[31]

Despite a more in-depth analysis of the discussions within the CCW/GGE, is crucial to mention that, in a meeting of 2019, the international community consider autonomy as a basilar concept of LAWS, consider that autonomy is based in a spectrum that differentiate, as seen before, semi and fully AWS.

In respond to that, Russian Federation consider that autonomy is not an exclusive characteristic of LAWS, defending that the discussion should be focus on "(...) *general understanding of what the future LAWS could be with the 'advanced' level of AI, rather than discussing existing operation systems with a high autonomy/automation degree*". 11.[12]

3.2.2. HUMAN ELEMENT

All of the concepts are related in some way. The spectrum of autonomy is also a reflection of the different levels of human intervention in these systems, that according to UK House of Lords, depends on the type of system, objectives of the mission and operational context. (01/12/2023).

Ms. Mary Wareham[20] defines autonomous weapon systems as 'systems that can select target and use force with a meaningful human control', but what is in fact consider 'meaningful human control'?

So, to address this question, Sierra Leone, the State of Palestine and Uruguay propose a series of requirements in 2022. These three States consider that it involves the application of human judgement and intervention to uphold human agency, but also responsibility and accountability in making decisions concerning the use of weapons. This dynamic includes, also enabling human operators to actively supervise weapons, engage in directive or preventive necessary interactions and have the capability to abort weapon operations when needed.

Even though, it is considered the main concept, it is relevant to mention some struggles, starting with the disagreements between States about the denomination. United States consider that the use of 'control' can cover certain situations, which, in the end, can be seen as contradictor[32].

[31] Birgitta Dresp-Langley (2023), The Weaponization of Artificial Intelligence: What the public needs to be aware of
[32] 2019 CCW GGE, https://documents.unoda.org/wp-

The American delegation estates that the human element is related only to controlling the effects of the systems, and not to direct manual intervention in these systems - even if that is what most states maintain.

The concept adopted by United States relies on human judgment. The Dod 3000.09 refers that "A*utonomous and semi-autonomous weapon systems shall be designed to allow commanders and operators to exercise appropriate levels of human judgment over the use of force".*

Russian Federation goes beyond this and, in November of 2019 on the meetings of CCW, that concepts of *'human control'* or *'human envolvement'* requires subjective characteristics, thus forces the 'risk of politicizing the discussion' (CCW/GGE, 2017).

In the Russian view, this is the major problem with the development of the concept, even if at first glance it seems to be the one with the most practical consensus among states.

Despite the efforts of other countries to reach a consensus on the terminology of the human element - such as human responsibility or human intervention - the term that was used in 2019 was simply 'human-machine interaction', which can take various forms and is implemented at different times in the system's used cycle.

Even though, the definition is not unanimous, the key elements are. These key elements can be considered the: context-control, understanding the weapon system, understanding the environment, predictability and reliability, human supervision and the ability to intervene, accountability and ethical considerations.

This element has a major relevance to 'defining where to draw the line between lawful and unlawful use of autonomous weapon systems.' This valuable concept is the bridge to articulate the standards that systems must respect in their development and use to be compliant with international law. Nevertheless, as explained by Amanda Musco Eklund, reaching general agreement posed legislative challenges in defining this concept.[33]

According to iPRAW, human control is the idea of comprehension of the situation or of the intervention operations during the attack. This capacity is achieved with two different ways: *control by design,* that is about the way the system is construction and

content/uploads/2020/09/CCW_GGE.1_2019_3_E.pdf

[33] Musco Eklund, Amanda (2020), Meaningful Human Control of Autonomous Weapon Systems: Definitions and Key Elements in the Light of International Humanitarian Law and International Human Rights Law.

control in use, that is relatable with the implementation of the system during their development.[23]

This is complementing by ICRC that consider the three stages of human control. The first one is the development stage, where the system is being tested, followed by a stage of activation of the weapon system and, finally, the last one is the operation stage where the conduct of weapon is purely autonomous and independent in relation to the targets.

Nevertheless, ICRC adds that human control is essential in accordance with the principles and norms of IHL. In that case, it has to have human-on-the loop supervision, knowing the consequences of their use and also the possibility of failure, and lastly, the operational constraints, that can increase human control, such as knowing the task the weapon is used for or the types of targets it attacks[34]

Peter Asaro defends the basis is that the moral and ethically responsible agent must retain control over the functions of any system that has an impact and possesses violent force, so reaching this consensus on a definition could be relevant in practical situations - where algorithms are applied to decisions with the potential to restrict or deny fundamental rights to human beings. It cannot be ignored that this is a situation of analyzing the potential to delegate authority to a machine that directly impacts a person's moral and legal rights.

Brazil, Austria and Mexico are very reticent about the type and level of control that would be necessary at the various stages of the systems' action, stating that there must be not only military, legal and moral concerns, but also ethical ones. These ethical considerations translate into a necessary human element in decisions on the use of force[35]

3.2.3. PREDICTABILITY

In the report made by the House of Lords entitled "AI in weapon system committee", one of the conclusions was that unpredictability is one of the characteristics inherent of systems AI-based, mainly when comforted by new or complex contexts.

In CCW/GGE 2019 report, the point 5(b) of the agenda is also mentioned as seven 'different potential characteristics' of LAWS, including predictability, reliability and self-initiate.

[34] Musco Eklund, Amanda (2020), Meaningful Human Control of Autonomous Weapon Systems: Definitions and Key Elements in the Light of International Humanitarian Law and International Human Rights Law

[35] 2019 Non proliferation Arms control CCW/GGE

The concept of predictability is thus adapted to LAWS as the knowledge of how the system will behave during use, under any circumstances, and what effects may result, according to ICRC experts[36]. So, complementing this, is also idea based on trust that assumes that the behavior is consistent and will work without any failure or unintended effects[37]

ICRAC considers that the unpredictability of these systems can be revealed in the constant need, due to the environment in which it is use - to adapt and learn - can lead to these unexpected decisions. But not only the environment of warfare causes these unexpected solutions, since to the nature of machine learning is intrinsic a notion of diversity of responses, creating this openness to unpredictability.

So, the conclusion of ICRC, is that this characteristic is the cause for LAWS cannot be developed, since "*A weapon system that is unpredictable by design would be unlawful by its nature",* because of the risks that are possible to be cause to protected people or objects[38]

This can be considered, alongside with reliability, the cores components of human control. The human operator needs to have a high level of confidence upon systems, and mainly in systems that are used on military contexts. In the situations of LAWS this confidence is not certain because of freedom given to achieve the tasks.

During the CCW's 2019 meeting, iPRAW reached the conclusion that the autonomous functions, exhibited in time and space, would be best characterized as '*boxed autonomy*', where the predictability of the system would be a technical feature that inly allows the system operator to mitigate the challenging effects of autonomous operation 'within the box', and therefore, predictability *per se* is not increased[39].

Merely, this unpredictability can be also a starting point to an abrupt escalation of conflicts (Birgita Dresp-Langley). There a risk of losing control of these weapons, being hacked by other armed forces or, even by NSAG.

3.3. THE POSITIVE IMPACTS IN ARMED CONFLICTS

[36] ICRC Expert Meeting, Autonomous Weapons Systems: Implications of Increasing Autonomy in the Critical Functions of Weapons, Versoix, Switzerland, 15-16 March 2016, page 9
[37] Ibid. page 13
[38] Statement of IPRAW, CCW/GGE, 2019
[39] Statement of IPRAW, CCW/GGE, 2019
https://reachingcriticalwill.org/images/documents/Disarmament-fora/ccw/2019/gge/statements/26March_iPRAW.pdf

The interest of these weapons is no difficult to understand. In the modern era of warfare renowned for target killings, drones, long distance and strategic points, the idea of mass indetermined destruction is vanished.

In a paradoxical sense, it is interesting to notice that the intersection of society, technology and security within advanced technology can contribute to 'make war more humane'[40].

The impact of using these weapons lies in diverse reasons that can be consider positive or negative, depends on the perspective. In military operations, precision can be the key for success, calculating the best way to cause the least collateral damage to achieve a give war aim.

The precision achieve by these systems can be revealed even in instable moments, during armed conflicts, and can be a factor to minimize the risk and harm for civilians during these times[41]

As explained by Prof. Pshikhopov, one of the experts that participate in CCW discussions[42], in war is always situations of inhumane conducts, and for that reason, humans can commit errors during AC. The question that he asked is 'what would be better a choice from the point of IHL – To let the human operator take the wrong decision and let people die or to use smart LAWS with much less casualties?'

So, in the opinion of United States and Russia Federation, these systems are capable of increase the security for civilians, because of a better accuracy in target selection and engagement that can comply with IHL.

Reaction time and speed also seem to be a factor in favor for this type of weapon, since it can process a great amount of information, in a much faster time, than human capacity allows.

A test was carried out with U.S. Air Forces to ascertain the competitive capacity between human pilots and pilots commanded by AI. Despite the experience, knowledge and strategy of the human pilots, it was clear that AI outperformed them, achieving faster and more effective responses. The test concluded that this is due to a simple fact: human action cannot be carried out without fear of the consequences that may arise; AI action is carried out without fear.

---

[40] Sabriya Alam, Olivia Jimenez, Julia Taylor, Dwaine Jengelley (2020): Unmanned and Autonomous Weapons Systems: Practices and Related Policy, https://core.ac.uk/download/pdf/359175783.pdf

[41] (Kirkpatrick, 2003).

[42] 2018 Chair's Summary of items 6

U.S.A is one of the States is one of the countries that invests the most in the development of artificial intelligence systems for military purposes, such as AWS[43]. The development of these weapons now seems to be the real cost, however, as times goes by - and similar to drones - development and producing them later can be less expensive, turning them attractive for States - obviously there is the other side of the coin.

3.4. THE NEGATIVE IMPACTS IN ARMED CONFLICTS

3.4.1. To the conflict

For Stuart Russel, these weapons can be considered scalable weapons of mass destruction. Russel explains this idea by saying that scalable is a technical term referring to the ability to process significantly larger data by increasing hardware resources. In LAWS, this means that it is possible of doing "do a million times more killing by buying a million times more weapons, precisely because the weapons are autonomous." They don't need human oversight, unlikely, as mentioned, AK-47.

This represent advantages if it is compared with other known mass destruction weapons, such as nuclear weapons and carpet bombing. The main difference to these weapons is the level of destruction they cause: in the case of AWS is not a widespread destruction but, instead, a specific damage to certain individuals or property that constitutes a threat. That does not mean, as explained by Russel, that they are not capable of "wipe out an entire ethic group or all the adherents of a particular religion (if adherents have visible indicia)".

In this context, Birgitta Dresp-Langley raises the idea of the risk of the emergence of new forms of existing weapons of mass destruction - such as drone swarms and autonomous nuclear, chemical, biological and radiological weapons - based on artificial intelligence, which could become more sophisticated and radically different.

The use of these weapons can represent a rapid escalation of future armed conflicts. India, in a statement made to General Assembly alert for the risk of proliferation and arms race, being a plausible risk that these weapons can fall on the hands of terrorist groups, that do not comply with international law.

[43] Haner. J, 2019: The Artificial Intelligence Arms Race: Trends and World Leaders in Autonomous Weapons Development.

In addition to NSAG, Peter Asaro have raised the issue of authoritarian regimes acquiring these weapons. Dictators and authoritarian rulers can use these types of weapons not only against their own people, but also to genocidal ends.

Adding to acquisition and proliferation, these systems can also be hacked by enemies: other State or NSAG. The so-called 'Adversary hacking' represents the loss of control of the system, which is now in the hands of the enemy, serving their needs and objectives, constituting a threat[44].

In the context of armed conflicts, these weapons can also represents situation of unjust war. In the letter made in 2021 about the conflict of Libya and the use of some weapons, one of the ideas was that the use of AW can caused an inequality and disproportional in the way the conflict was developing. The Libyan army was not 'trained or motivated' to defend itself against the use of this new technology and was retreating from its offensive for this reason.

In a situation of armed conflicts, with instability circumstances and 'when a vast amount of external data is being processed in real time', it is common, depending on the technology, that could exist false positives or false negatives, as explained by Sabriya Alam[45], in the implementation of AI algorithms in AWS.

3.4.2.Risks for humans

The main specificity of these weapons is the life-and-death decision-making, without a human intervention, constituting a new level of AI application[46]

One of the biggest concerns of the international community is the possible risks to human rights, not representing the IHL principles. The HRC Resolution 51/22 recognize that technologies that rely in "*data sets, algorithm-based programing and machine-learning processes, may, in certain circumstances, cause, contribute to or facilitate the commission of human rights violations and abuses, as violations of IHL and recognizing further the risk that these technologies could be acquired by NS actors"*

The new and emergent technologies can create difficulties by relying in mechanisms that reproduce and exacerbate existing patterns of discrimination, marginalization, stereotypes and bias [47].

---

[44] (Birgitta Dresp-Lansley p.33)
[45] Sabriya Alam, Olivia Jimenez, Julia Taylor, Dwaine Jengelley (2020): Unmanned and Autonomous Weapons Systems: Practices and Related Policy, https://core.ac.uk/download/pdf/359175783.pdf
[46] Ibid.
[47] Stop Killer Robots, https://www.stopkillerrobots.org/stop-killer-robots/facts-about-autonomous-weapons/

So, even though, one of the good aspects of these weapons could be precision, some believe that this precision can deflagrate accordance with IHL principles because it is possible that they might inadvertently increase civilian harm and suffering by enabling attacks that were prohibited.

Even though, these weapons can have mechanisms, such as facial recognition, remains an uncertainty if they are capable of respect the principle of distinction, discriminating, in an efficiently way, between civilians and military targets. In any hypothesis can civilians become targets of war and, the problem of these weapons, are exactly the fact that is not so certain that they are capable of respect that on any occasion[48]

In June 2023, Lorraine Finlay, an Australia's Human Right commissioner, portraited a hypothetical example of sending a LAWS to a specific location to seek a target known as driving a white sedan. The systems proceeded to identify a white sedan with a letters written in the car but without the knowledge to understand what is painted on there, determining simply that there is a white sedan driving by the target and executes the attack on basis of AI and facial recognition, without human oversight.

Later, Finlay returns to the scene, but this time with a human monitoring the situation. In this case, it was possible for the operator to understand a crucial aspect: the paintings in the white sedan recognizes the phrase 'Just Married' in the native language of the civilians. This situation resumes the idea that human operator can understand the additional context and significance that can be crucial in an armed conflict situation.[49]

In addition to the fact that is difficult to respect the principle of distinction, it is also relevant to mention that there may be a disproportionate nature of the attacks in relation to the military objective, as said by Peter Asaro[50]. These weapons might become more aggressive or being unable to have a compassionate view of certain situations due to their focus on programmed objectives, obviously lacking a human psychological aspect that is inherent to us.

The role of a 12 years old child during an armed conflict can exemplify this case. The human knowledge and experience can tell that children can also represent a threat in some cases of war, through the dynamics, the way it behaves and the whole involvement of the situation. A machine is either prepared for a specific threat or, perhaps, it wouldn't

[48] Guersenzvaig, A. 2018: Autonomous Weapon Systems: Failing the Principle of Discrimination
[49] Link of the video: https://www.youtube.com/watch?v=WDujP4XPgQY

[50] Peter Asaro, p.4

even be able to neutralize the threat because it's programmed in a certain way, allowing the attack to happen with catastrophic consequences.

The new AI capabilities, many systems have acquired bias, mostly concerning gender and race (ICRC). In the case of AWS, they are 'trained' by humans, prejudices will be embedded in their roots that will be relevant, in some way it is possible to create solutions that perpetuate forms of discrimination[51]

This, then, is a sensitive topic when we talk about algorithm-based programmers, but - because these systems have the capacity of life-and-death decision - their conducts have much greater impacts due to the consequences they can represent when 'triggered'.

Thus, the inherent bias is considered by all the States, even USA. In their Dod[52] mentioned the necessity of a responsible use of these weapons in the way that it possible to bridging the ethical risks and ambiguities that arise, such as prejudices.

It is also defended that the conduct of American forces has to obey to five principles: accountability, equity, traceability, reliability and governance. The principle of equity is the most important one in this case, because it prevents and combats possible discrimination that these systems are based on.

3.5. DEFINITIONS OF THE STATES

Considering now all the characteristics, aspects and dynamics is time to consider what definition are States and Organisations given to Autonomous Weapons, more specifically to Lethal Autonomous Weapons.

The starting point is that creating a definition to these systems can be a hard task due to its character too current. As the Russian Federation defends the lack of practical examples to reach the individual characteristics can turning this unanimous definition something that will still take your time to achieve.

Yet, based on the concepts mentioned before, it also possible for States to achieve national definitions. In the case of United States, the definition is based on the human role, characterized as "*A weapon system that, once activated, can select and engage targets without further intervention by an operator. This includes, but is not limited to, operator-supervised autonomous weapon systems that are designed to allow operators to override operation of the weapon system, but can select and engage targets without*

[51] Sarah Brown, 2021
[52] Dod 3000.09, jan.2023 https://www.esd.whs.mil/portals/54/documents/dd/issuances/dodd/300009p.pdf

further operator input after activation"[53]

The definition of China can be considered very different from other States, since 2011 that China have been adapting their view on LAWS. In 2022, China decided to change their perspective from the definition given in 2011 and includes five basics, but not limited, characteristics.

China estates that LAWS need to have lethality, autonomy, the impossibility of termination, meaning that once started it is impossible to terminate the device, indiscriminative effect, executing a task regardless of conditions, scenarios and targets. The last is evolution that is relatable with the capacity of self-learning capacity, exceeding human expectations.

The ICRC is more simple and centralize the definition in the concept of autonomy by consider that is '*any weapon with autonomy in its critical functions – that is, a weapon system that can select (search for, detect, identify, track or select) and attack (use force against, neutralize, damage and destroy) targets without human intervention.*[36] In the perspective of Peter Asaro this definition goes beyond any other because of the reference of 'critical functions' .

Mention lastly the definition given by Russian Federation that tries to not be limitative to the present reality. For that reason: "a lethal autonomous weapon system is a fully autonomous unmanned technical means other than ordnance that is intended for carrying out combat and support missions without any involvement of the operator"

4. INTERNATIONAL HUMANITARIAN LAW: LEGAL VIEW OF LAWS

The concordance theorical and practical between new emergent technologies and international humanitarian law is the base of this chapter. The attentive and cautions nature of IHL, as mentioned by Philip Spoerri[54], has the power to adapt to a reality that does yet exist.

When the 'Hague Conventions' were signed between 1899 and 1907, the military practice of aerial bombing did not yet exist. However, at the time, Russia suggested that a mention need to be made of "*the prohibition of the discharge of any kind of projectile*

[53] (p. 21 Dod),
[54] Director of International Law and Cooperation: ICRC

or explosive from balloons or by similar means.", having already the speculation that aircraft could be used in the near future armed conflict.

So, IHL is a compilation of legal norms capable to limit the effects of armed conflicts for humanitarian reason, based on multiple principles, such as distinction, proportionality, military necessity, precaution on the attacks and humanity. These principles are operational and imperative keys that States must respect in their conducts for the use of these weapons[55]

IHL thus plays a fundamental role in uniting the discussion on autonomous systems. As Neil Davison [56] points out, autonomous systems must be capable of being used and, if they are, their use by states must follow IHL precedents.

The legal debate involving LAWS is around three fundamental points, as said by R. Shashank Reddy, international humanitarian law's rules of legal review, distinction and proportionality.

4.1.PRINCIPLES OF IHL

The principle of distinction is considered a cornerstone of IHL (ICRC). In the Advisory opinion of ICJ about the legality of Nuclear Weapons, the principle of distinction was consider one of the two cardinal principles that aimed to distinct between combatants and not-combatants during an armed conflict. In their perspective, this principle enables civilians to be an object of attack and 'must, consequently, never use weapons that are incapable of distinguish between civilian and military targets'[57]

To this advisory opinion, is adding the legal base: article 48.º and article 52.º of Additional Protocol I to the Geneva Conventions, that represents an important general obligation in relation to the use of human-controlled force in terms of humanitarian law.

The great doubt of these weapons is if they are capable of this distinction or not, even though they are programmed with AI and facial recognition, for instance. The difficulty of understanding if they are capable or not turns even more frightful if it is considered that these type of weapons are using in swarm mechanisms, possessing a great abundance in armed conflicts[58].

---

[55] CCW 2019 report 17 (a) e (d))

[56] Davison, Neil (2017): A legal perspective: Autonomous weapon systems under international humanitarian law, Arms Unit, Legal Division, ICRC

[57] .*Legality of the Threat or Use of Nuclear Weapons*, Advisory Opinion, 1996, paras 78–79.(Guersenzvaig,2018).

[58] (Kallenborn, 2021).

The fact that LAWS are in an early phase, makes the ambiguity of these systems soaring. It is possible that, as said earlier, based on algorithms and AI, it becomes possible to understand the conducts that are legal under IHL, however, as estate by Egypt in the 78th session of UNGA, maybe 'it can never determine what is ethical'.

The level of autonomy of LAWS can contribute to the failure of satisfying what is stipulated by military ethical[59], but also the possible existence of bias in data sets. To understand who are who, it is necessary that the weapons system can be guided for some factors, such as gender, race, gender and ability. Biased data sets and poorly programmed algorithms, as a recent UNIDIR report shows, can exposed specific groups to relevant risks[60]

Other important principle in the protection of civilians is the principle of proportionality. This principle is also mentioned in the ICJ Advisory opinion as a cardinal principle.

This principle appears in codified in Article 51(5)(b) of the 1977 Additional Protocol I, and represents customary international law. The idea of this principle is translate in the prohibition of attacks 'which may be expected to cause incidental loss of civilian life, injury to civilians, damage to civilian objects, or a combination thereof, which would be excessive in relation to the concrete and direct military advantage anticipated'[61]

In LAWS the concern is the disproportionate risk that these weapons can cause because of the lack of human supervision. Some academics consider that this risk can also be analyzed by a reveal of human dignity as a fundamental characteristic of humanitarian law. Not only will it be questionable how the person was killed or assaulted, but the entire process to the result.

Even though, the defensive capacity that these weapons represent are mentioned by many States, such as India or China, this principle is clear by consider that States do not have an unlimited freedom of choice the type of weapons they can use during an AC. Many weapons are banned and inaccessible to States, such as chemical weapons.

4.2.LEGAL SOLUTIONS

States have achieved small conquests in terms of LAWS. It is true that, during this time, it was possible to discuss and understand the position of States and organisations,

[59] Birgitta Dresp-Langley, 2023; Walzer, 1977
[60] **CCW/GGE.1/2024/WP.5**
[61] https://www.icrc.org/sites/default/files/wysiwyg/war-and-law/04_proportionality-0.pdf

but it was not possible to understand what is a plausible solution to this threat - that is not a very far future.

International community are divided. United States and Russia accept the view that existing IHL is sufficient to regulate LAWS, due to its attentive and sophisticated character. Other delegations consider that the specifies of these weapons, unlike any others already developed, requires to reach a new legal binding instrument and/or a preemptive ban or, as India, a political declaration that stress the importance of respect IHL and establish important measures, but do not have a binding characteristic.[62]

Article 36 of AP I is at the centre of discussions, and estates that: "*in the study, development, acquisition or adoption of a new weapon, means or method of warfare . . . to determine whether its employment would, in some or all circumstances, be prohibited by [Additional Protocol I] or by any other [applicable] rule of international law*." Because of the adapting characteristic of the 1977 Protocol I, many States believe that can be an effective method of ensuring that autonomous systems fulfil the requirements of international law[63]

Along with article 36, article 48 also represents, as mentioned earlier, an important general obligation regarding the use of force with human control in terms of humanitarian law: the principle of distinction.

In addition, other articles, such as 51(4) and (5), 52(2) or 57, prove that humanitarian principles are interrelated, stipulating the prohibition of disproportionate and indiscriminate attacks and reaffirming that parties to a conflict must take precautions in the way they conduct attacks.[64]

Other treaties, such as Arms Trade Treaty, is a way of given a precedent of how emergent weapons must be developed and use to be in accordance with law. Even so, the application of humanitarian law norms does not mean a plausible solution in every case. The dynamics of an armed conflict are unstable and, for this reason, there are not always certain solutions. So, reviews of article 36 could be a valuable method to ensure that LAWS are in accordance with IHL.

---

[62] 2018 Report of the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems

[63] ICRC New Weapons, New Precedent New Weapons, Proven Precedent: Elements of and Models for a Treaty on Killer Robots | HRW

[64] Ibid.

Existing international law in the view of certain states - such as Israel, the US and Russia (2018) - appears to be capable and sufficient to overcome these difficulties imposed by LAWS. The legal reviews conducted by international bodies, such as CCW, are also part of this 'sufficiency', because it is a way to understand the legal and ethical issues that these systems purposes.

However, Austria consider that legal reviews alone might not be sufficient to beware of the dangers posed by these weapons. The European Union, Argentina and the Netherlands all emphasize the idea that these reviews must be transparent, giving states the chance to improve them. This results in the fulfilment of IHL standards through the influence and cooperation of states.

Another view is that advocated by non-governmental organisations and some states, such as Costa Rica or Pakistan, of the need to negotiate a legal instrument. This could take the form of an additional protocol to the CCW or an independent treaty capable of prohibiting, regulating or simply imposing certain obligations on states when using these systems.

In July, the UN Secretary-General recommended that states should come together and finalize, by 2026, a binding legal instrument to "(...) prohibit lethal autonomous weapon systems that function without human control or oversight."

The nature of these systems has a significant impact on and delays the process of new regulation. There is therefore a need to continue looking at the existing principles and instruments of IED in order to arrive at a response model - even realising that there is no perfect one.

The International Panel on the Regulation of Autonomous weapons (IPRAW) believes that the innovative and peculiar nature of this type of system - such as the functions made possible by machine learning - means that there needs to be a regulatory process of its own in order to be able to respond to all the important challenges.

The lack of accountability for the use, inability to comply with the principles and potential operations risks are one of the arguments for States want a preemptive ban of these weapons. Brazil consider that the use of these weapons could put at risk a dilution of the very concept of accountability. However, this decision is not inhibiting research into technologies that may provide civilians with benefits?

A report created by HRW deals with the elements and the model that should be followed for the creation of a treaty[65]. The starting point is the three obligations that states must fulfil: the general obligation is that meaningful human control must be maintained over the conduct of the use of force - not the system itself.

The next obligation has to do with the prohibitions that must be created by states in relation to the development, production and use of autonomous systems which, due to their nature, impose moral and legal problems that are fundamental. based on date to select targets and, positive obligations that guarantee that significant human control is maintained in these systems.

And finally, the report mentions the positive obligations that must be established by States in order to guarantee that meaningful human control is thus maintained in all the systems created.

This report thus brings together the essential standards of IHL, interlinking them with the technological principles that guide the safe use of Artificial Intelligence, taking into account both ethical and legal aspects. A total of 36 principles have been created.

The lack of human control in other weapon system motivated several prior disarmament treaties, such as Biological and Chemical Weapons. In 1969, the report from SG mention that 'controllability' was the most relevant consideration in the use of these weapons.

Different from the legal reviews and new legal instrument is the idea presented by Portugal of a new political document, such as Montreux Document establish in 2008. In the opinion of Portugal this could be a better instrument to the adaptation of challenges, since it is a non-legal binding instrument, do not affect the existing obligations under IHL, serving as a technical outcome and can be easier to have a cooperative implementation between States.

India, France and Germany are also in favor of this perspective that only imposes political obligations on states and not legal ones, functioning (almost) as a code of conduct based on guiding principles.

4.3.CCW/GGE – A WAY FOR A REAL SOLUTION?

As is widely recognized, the norms of IHL lies in an inherent adaptability to diverse circumstances, as previously underscored. The UN Convention on Certain Weapons

[65] IHRC, HRW (2020), New Weapons, Proven Precedent: Elements of and Models for Treaty on Killer Robots

(CCW), established in 1981, stands as a practical and respect demonstration of collective effort towards arms control and a key instrument of IHL, with its primarily aim of "ban or restrict the use of specific types of weapons that are considered to cause unnecessary or unjustifiable suffering to combatants or to affect civilians indiscriminately", thus relating to important humanitarian principles and the *Martens Clause*.

In order to pursue its purpose and as an instrument that facilitates discussions on appropriate regulations of new emergent weapons (UNODA – Informal note), since 2014, LAWS have been theme of legal, ethnical, technological and military (sharing) analysis and, more recently, was established an open-ended formal group of Government Experts (GGE).

The idea of open-ended group is interesting since any international organization, non-governmental organizations or a State as an observer can become member of the discussion. This 'openness' is not only happening to achieve a consensus between the States, but also because of the ambiguity of the subject, that pursue an explicit clarification by anyone that can have important knowledge of the technological advancements.

The CCW/GGE on LAWS is tasked to be a 'fast-pace' to States, giving them space to estate their position about these new technologies. To address them it was establish five agenda items. The initial trio is concerned with the obstacles to IHL (5(a)), the delineation to a common understanding (5(b)), and the human-element (5 c) on the use of lethal force by these weapons. The latter pair concerns the potential military use of this emergent technology (5(d)) and avenues to confront humanitarian and global security challenges implicated (5(e))[66]

These discussions are the central forum to surfacing various perceptions regarding the extent and characteristics of the concerns posed by LAWS. It was already achieved some unanimity in aspects, such as the impossibility of an unlimited autonomy of these systems, as mentioned in the chapter 2 of this thesis.

Given the fact that CCW/GGE decisions requires unanimity and consensus, because all of the High Contracting Parties have the right to veto, makes somehow difficult to achieve some progress in the subject. Some organisations, such as NGO Stop Killer Robots consider that is now time to an alternative forum to these discussions.

[66] (paragraph 11 Report 2019 - https://documents.un.org/doc/undoc/gen/g19/285/69/pdf/g1928569.pdf?token=lnAF1s8QcAV3ix5UZg&fe=true)

But there are other obstacles to these meetings, such as the relevance of the power military opinions. Ray Acheson[67] considers that there is some bias in relation to military powers, defending that their interests and convictions are always taking more space that other countries. For instance, the fact that prevails the desire of United States to change the concept of 'human control'

In the Chair's summary of the characterization of LAWS, four approaches were mentioned to highlight and served as a reference point. The first one was a 'separative approach' where States delineate, in a flexible and individual form, concepts as positive or negative vias, depending on the relevance to the objectives and purposes of CCW.

Contrastingly, the second approach is called the 'cumulative'. This approach entails the creation of a comprehensive list comprising categories of concepts, such as human-machine relationship or reliability, that are later evaluated against certain technical, legal-humanitarian or political-security criteria, including targeting performance, to assess their relevance to the objectives of CCW.

The Accountability approach have other focus, shifting towards a set of characteristics related to the functions and decision-making capabilities delegated to autonomous machines. This approach avoids using levels of autonomy and other technical characteristics related to the loss of human control, instead it emphasizes contextual considerations and scenario evaluations that would involve a combination of technical and human-interface evaluations center on accountability.

The latter is focused on desirable and undesirable consequences of possible LAWS based on emerging autonomous intelligent systems. The last approach is consequential and translates into the boarder ramifications of potential LAWS development, considering both desirable and undesirable outcomes arising from the emergence of this weapons, according to societal, ethical and humanitarian implications from the proliferation of LAWS.

In this view, 2019 was an important year to the achievement of LAWS. States agreed to a set of important 'guiding principles' for LAWS, such as that States must be observant to the risks of LAWS, including the proliferation to NASG. Despite the need to mention states perspectives earlier in some of these principles, other principles need a mention. The (a) principle is an important confirmation about the value of IHL: this continues to be applicable fully to all weapons systems, not only in their use, but also in

[67] Ray Acheson, Report CCW/GGE August 2019

their development. The cases that are not covered by CCW and other protocols, should always remain under the protection and authority of the principles of customary law, as said in 2019 Report, paragraph 17(G) .

These principles have different preoccupations, such as responsibility, human-machine interaction, weapons reviews and risk assessments and mitigation measures. The idea is understanding which actions with autonomy do not representing an ethical and legal risk. So, Principle (b) was about human responsibility that have to be retained, since accountability cannot be transferred to machines and principle (c), the last principle to be agreed by States, complement the idea about human-machine interaction by saying that could take various forms and be implemented in various stages of the life cycle.

The strategies to diminish the risk adopt determine practical measures that can reconcile responsible use and innovation, such as develop rigorous testing and a continued system assessment[68]or adding theoretical measures that may complement this mitigation, including the establishment of procedures, doctrine and national legal reviews (principle E and G)

The last agreed principle shows that, even though CCW and States are concerning about the risks of these weapons, it is important to reach a balance between military necessity and humanitarian concerns.

Since 2019, other reports have been done by CCW/GGE. Focusing, however, on the most relevant particularities it must be mentioned the draft articles submitted by Australia, Canada, Japan, Republic of Korea, United Kingdom and United Sates that present some prohibitions ant regulatory measures according to IHL.

In these articles is taking into considerations not only the norms of existing IHL and existing protocols within the Convention, but also recommendations, guiding principles and conclusions made by GGE.

This framework, different from what has been done, mention not LAWS but only autonomous weapons, considering in a first moment the positive and negative aspects of these mechanisms.

According to article 7 of the drafts, these articles are only to clarify some possible doubts of imposing law, not creating any obligation. This can work as a precaution measure, since it gives some possible conducts to States to adopt relatable to the use of LAWS, taking into account the context presiding at the time.

[68] point 23(b) CCW/GGE 2019

The first article is not a definition of what is an AWS, instead is the measures that need to be respected in their design, developing and use for AWS being in accordance with IL.

The article estates which characteristics contributes for the prohibitions of their use: '(..) cause superfluous injury and unnecessary suffering, if it is inherently indiscriminate or if it is otherwise incapable of being use in accordance with IHL.'

The article is also concerning with the way AWS are designed and developed, contributing in this sense to play a signaling role serving as a guidepost to States in the way they develop these weapons, according to Eric Talbot[69], if there is a careful deliberation and consideration of the advancements in the future. The 3 paragraphs touch upon the respect for civilians, proportionality and responsibility, mentioning the rigorous test to give reliably for their use and precautions measures in attacks, such as (iii) reduce automation bias in systems operators. The other articles are relatable with possible measures that can be implemented in accordance with principles of IHL, such as accountability (article 6).

In the opinion of ICRAC, CCW is now capable of reaching and advance to the discussion of a legally binding instrument, since these discussions have achieved some important conclusions about context and human-machine interaction about these systems.

4.4.UNGA RESOLUTION 78/241: THE BEGINNING OF THE UNITED NATIONS' SOLUTION?

In the 78th session of the General Convention, the first committee was pledged to discussing issues related to disarmament and collective security of States. Under this debate, which brought together topics such as nuclear disarmament or Arms Trade Treaty, the role of new technologies AI-based and autonomous weapons in the military future also featured prominently, and it was possible for States to reach and vote on the first resolution on this topic.

The first resolution (78/241), approved in the United Nations General Assembly, starts with a common idea shared by States that IHL and IHRL continues to be applicable to Autonomous Weapons. This resolution, therefore, mentions in an undetailed manner,

[69] MAJ Ronald T.P Alcala, Talbot Jensen, Eric (2019). The Impact of Emerging technologies on the Law of Armed Conflict, Oxford University Express

the positive impacts, but also the risks and concerns that LAWS may represent to the global security and stability, sufficient fit, according to some authors, to consider that indirect obligations are already being created here for States regarding their use and development.

Without ignoring the efforts made so far, at national, regional and international level, the relevance and timeliness of LAWS makes the need for continued attention by United Nations. For this reason, this resolution gives to the Secretary-General a more active role to understand the discussion and the points of view of the States, International and Regional Organizations, ICRC, civil society, the scientific community and industry, bringing them together in a substantive report to present to UNGA in the next session.

Even though, this resolution was approved by a majority– 152 approved votes, 4 against (Russia Federation, Mali, Belarus and India) and 11 abstentions - it is important to be aware that States, within the debate, exposed some differences about this subject.

The impossibility of a detailed analysis of the resolution's paragraphs and arguments given by the states means that a more comprehensive assessment of the essential points is required.

The first accomplishment of this resolution is on preambular paragraph 1 that affirms the application, without doubt, of norms of international law, such as the Charter of UN, IHL and IHRL. This provision was only had one vote against by India with 12 abstentions. The problem is not the application of IHL, since India and other States agree with this affirmation in the conclusions of CCW, the problem may relay in mentioning human rights law. The delegation of Iran stated that the issues of human rights and the reference to HRC Resolution 51/22, analyzed before, is not relevant in this case, although supporting "human rights in principle"

During the statements, Iran, Turkey, India, China and others begun pointed out the non-existent definition and scope of the terminology. The element of 'lethality', as seen before, is an important characterization factor of the topic to a legal solution, according to some delegations, and the resolution, in the point of view of China, cannot confused LAWS and autonomous weapons. Israel considers that removing this factor, could lead to "broadening the debate in a manner that could over-complicate our discussions and significantly slow down any progress we wish to make"

In this sense, even so the delegations of United States and Australia voted favor to the resolution, they defended that it needed certain modifications to become more inclusive. In the present situation, the ambiguity of the issue does not give space for the

resolution to be a reliable response for the effective concerns and risks, however, States look at it more as a whole "because of its comprehensive and meaningful posture" (South Africa).

The risks of an emerging arms race, lowering the threshold for conflict and proliferation, including NSAG, are at the center of delegations' concerns, making special mention to the statement of India. Despite that, India and Russian Federation also recognize the need of an approach that give space for the positive impacts of the evolution of these weapons.

This idea for Nations is connected to the progress made by CCW/GGE, that was the center of the discussion in the approval of this resolution. This CCW 'is the appropriate forum to discuss issues relating to LAWS, with a view to striking a balance between military necessity and humanitarian imperatives' (Statement by India), as the inclusion of other mechanism, including UN, could lead to 'counterproductive'.(Russian Federation) and stressed the importance of the IHL to LAWS (Israel).

The inclusion of operative paragraph 2 that request UNSG a report based in the views of the issue creates some tension even in the delegations that vote in favor of the resolution, such as South Africa, United States and Poland.

The majority have some skepticism about transferring the discussion within the UN machinery, since the discussion of GGE have a progress done. Poland believes that other international forums don't have the same capabilities of GGE, "as they often lack of technical and diplomatic capacity" and with reasonable balances. The Republic of Korea, for instance, defend that these measures must be done gingerly for not jeopardize the ongoing efforts of GGE, since States parties of CCW 'already made their views known on the issue' (South Africa)

The progress of GGE appointed by France as tangible is not accompanied by the Egyptian delegation that enforce this opinion by mentioning the necessity of international cooperation, capacity on the fields of AI and to address the gaps between developed and in developing States. Concluding that the only mechanism that properly can face this is UN "as the most inclusive, equitable, and effective platform for the development of international rules in this domain"[70]

The report of UNSG can function as a complementary mechanism to face the challenges of LAWS, avoiding duplication as stated by Pakistan. The complexity of this

[70] Statement of the Arab Republic of Egypt United Nations Disarmament Commission General Exchange of Views 1-2 April 2024

subject presupposes a ‘holistic multilateral response within the UN machinery’(Pakistan) that can represent wider and fresher (Brazil) views of all member states and denote “an pragmatic approach to opportunities for progress in CCW that can achieve consensus” (Statement of Australia).

Discussing LAWS in the General Assembly contributes to a greater political impact and could represent meaningful progress, even without the support of Russia and India and with the abstention of China, the major powers.

This resolution is not the solution for LAWS, but, as said by Mary Wareham, director of HRW, the prerequisites have been established for states to advance in determining international regulations for these systems.

5. CONCLUSION

Artificial Intelligence has now the power to revolutionize every aspect of life, including warfare. The changing of paradigm, with the emerging on Autonomous weapons in battlefield have divided the international community.

The precedents of International Humanitarian Law are essential in the topic, if only to understand how it will be possible to create a new regulation for these systems, taking into account the military interests of the states, but also the implicit objectives of the law.

Analyzing the discussions and opinions of States serves as the main source for understanding what problem we are facing and what we can expect in the near future from something that still seems so ambiguous. We should not and cannot ignore the urgency of this issue, which has already been in the battlefield.

Articles in Electronic Support and Books

First Committee 78 th UN General Assembly, Statement of Mr. Benjamin Sharoni Director of Arms Control Department, MFA https://reachingcriticalwill.org/images/documents/Disarmament-fora/1com/1com23/eov/L56_Israel.pdf,

Statement of the Arab Republic of Egypt United Nations Disarmament Commission General Exchange of Views 1-2 April 2024 https://docs-library.unoda.org/United_Nations_Disarmament_Commission_-_(2024)/egypt_en.pdf

Draft Resolutions, Voting Results, and Explanations of Vote from First Committee 2023, 78th UN General Assembly https://reachingcriticalwill.org/disarmament-fora/unga/2023/resolutions

2018 Report of the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems ('GGE') https://documents.un.org/doc/undoc/gen/g18/323/29/pdf/g1832329.pdf?token=LostWQoA6dRkSAhN27&fe=true

Explanation of Vote of the Delegation of the Islamic Republic of Iran on the Draft Resolution L.56 Entitled: lethal Autonomous Weapons Systems 78th Session of the UNGA First Committee New York, 1 November 2023 https://reachingcriticalwill.org/images/documents/Disarmament-fora/1com/1com23/eov/L56_Iran.pdf

Resolution 78/241 of United Nations General Assembly https://www.un.org/en/ga/78/resolutions.shtml

Report 2019 CCW/GGE https://documents.un.org/doc/undoc/gen/g19/285/69/pdf/g1928569.pdf?token=lnAF1s8QcAV3ix5UZg&fe=true)

Commissioner Finlay presented this speech 'Stopping Killer Robots – prohibiting the use of lethal autonomous weapons systems (LAWS)' at RightsCon Costa Rica on 7 June 2023 (AEST).https://www.youtube.com/watch?v=WDujP4XPgQY

CCW/GGE (2023): Non-exhaustive compilation of definition and characterizations submitted by Chairperson (CCW/GGE.1/2023/CRP.1)

CCW/GGE (2019): Report of the 2019 session of the Group of Governamental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons (CCW/GGE.1/2019/3)

CCW/GGE (2024): Addressing Bias in Autonomous Weapons, Submitted by Austria, Belgium, Canada, Costa Rica, Germany, Ireland, Luxembourg, Mexico, Panama and Uruguay (CCW/GGE.1/2024/WP.5)

CCW/GGE (2023): Draft Articles on Autonomous Weapons Systems – prohibitions and other regulatory measures on the basis of IHL, submitted by Australia, Canada, Japan, Republic of Korea, the United Kingdom and the United States (CCW/GGE.1/2023/WP.4/Rev.1)

CCW/GGE (2017): Examination of various dimensions of emerging technologies in the area of the lethal autonomous weapons systems, in the context of the objectives and purposes of the Convention, submitted by Russian Federation (CCW/GGE.1/2017/WP.8)

Chair's Summary on the discussion on Agenda items 2018: https://unoda-documents-library.s3.amazonaws.com/Convention_on_Certain_Conventional_Weapons_-_Group_of_Governmental_Experts_(2018)/Summary%2Bof%2Bthe%2Bdiscussions%2Bduring%2BGGE%2Bon%2BLAWS%2BApril%2B2018.pdf

United Nations, General Assembly, 'Workn of the Advisory Board on Disarmament Matters', Report of the SG, A/68/206, 26 July 2013, para. 42

IHRC, HRW (2020), New Weapons, Proven Precedent: Elements of and Models for Treaty on Killer Robots

https://www.hrw.org/report/2020/10/20/new-weapons-proven-precedent/elements-and-models-treaty-killer-robots

Killer Robots: UN Vote Should Spur Action on Treaty, 01/2024 https://www.hrw.org/news/2024/01/03/killer-robots-un-vote-should-spur-action-treaty

ICRC, 2023 What you need to know about artificial intelligence in armed conflict

https://www.icrc.org/en/document/what-you-need-know-about-artificial-intelligence-armed-conflict

UN Human Rights Council, Resolution 51/22, October 2022, https://documents.un.org/doc/undoc/gen/g22/520/64/pdf/g2252064.pdf?token=G8BkS5wliWgr5mOfdI&fe=true

Acheson, Ray. Women's International League for Peace and Freedom , 2019 https://reachingcriticalwill.org/disarmament-fora/ccw/2019/laws/ccwreport/14022-ccw-report-vol-7-no-7

Asaro, Peter (2020). Autonomous Weapons and the Ethics of Artificial Intelligence, Oxford Press.

Sabriya Alam, Olivia Jimenez, Julia Taylor, Dwaine Jengelley (2020): Unmanned and Autonomous Weapons Systems: Practices and Related Policy, https://core.ac.uk/download/pdf/359175783.pdf

Brown, S. (2021) Machine Learning, Explained. MIT Sloan School of Management, Cambridge. https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained

Roland, Alex (2016): War and Technology, Oxford press https://www.wook.pt/livro/war-and-technology-a-very-short-introduction-alex-roland/17653824

Russel, Stuart J. (2019), Human Compatible: Artificial Intelligence and the Problem of Control, Viking

Boulanin, V., Carlsson, Moa Peldán (2019): Non-proliferation, arms control and Disarmament, chapter "The Group of Governamental experts on lethal autonomous weapons" https://www.sipri.org/sites/default/files/SIPRIYB20c13sII.pdf

Casley-Maslen, S., Weizmann, N., homayounnejad, M., Stauffer, H. (2018). Drones and Other Unmanned Weapons Systems under International Law, Leiden, Boston Brill Nijhoff

MAJ Ronald T.P Alcala, Talbot Jensen, Eric (2019). The Impact of Emerging technologies on the Law of Armed Conflict, Oxford University Express

Scharre, P. (2018). Army of None: Autonomous Weapons and the Future of War. W. W. Norton & Company.

Arkin, R. C. (2009). Governing Lethal Behavior in Autonomous Robots. Nova Iorque, EUA. Chapman and Hall Book.

Melzer, N. (2019). International Humanitarian Law: A Comprehensive Introduction. Genebra, Suíça: International Committee of the Red Cross.

Melzer, N. Interpretative Guidance on the notion of direct participation in hostilities under IHL. Genebra, Suíça: International Committee of the Red Cross.

Sharkey, N. (2014). Autonomous Weapons Systems and Human Supervisory Control.

(2018) *Verschae, R. & Ruiz-Del-Solar, J. (2015). Object Detention: Current and Future Direction*

Anand, Vinod (2008). Impact of Technology on conduct of Warfare, Strategic Analysis.

Taddeo, M. e Blanchard, A. (2022). A Comparative Analysis of the Definitions of Autonomous Weapons Systems. Science and Engineering Ethics. https://link.springer.com/content/pdf/10.1007/s11948-022-00392-3.pdf?pdf=button%20sticky.

Stefanocivh, Dimitry (2020), Artificial Intelligence Advances in Russian strategic weapons, SIPRI

Birgitta Dresp-Langley (2023), The Weaponization of Artificial Intelligence: What the public needs to be aware of, https://www.researchgate.net/publication/369084397_The_weaponization_of_artificial_intelligence_What_the_public_needs_to_be_aware_of

Elisabeth Hoffberger-Pippan, Vanessa Vohs, Paula Köhler (2022) Autonomous Weapons Systems: UN Expert Talks Facing Failure https://www.swp-berlin.org/10.18449/2022C43/

Ian Anthony and Chris Holland (2014), Armaments, Disarmament and International Security: The governance of Autonomous weapons

https://www.sipri.org/sites/default/files/SIPRIYB14c09sII.pdf

Davison, Neil (2017): A legal perspective: Autonomous weapon systems under international humanitarian law, Arms Unit, Legal Division, ICRC

ICRC, The Principle of Proportionality in the Rules Governing the Conduct of Hostilities under IHL, 2018 (ICRC Proportionality Report)

International Law and Jurisprudence

1868, St.Petersburg Declaration

1899 to 1907 Hague Conventions

2014, Arms Trade Treaty https://disarmament.unoda.org/convarms/arms-trade-treaty-2/

1977 Aditional Protocol I to Geneva Convention, 12 August of 1949 – protection of victims of international armed conflicts

Montreux Document, 2008 ICRC https://www.eda.admin.ch/eda/en/home/foreign-policy/international-law/international-humanitarian-law/private-military-security-companies/montreux-document.html

1996, International Court of Justice, Advisory Opinion on the legality of the threat or use of nuclear weapons: https://digitallibrary.un.org/record/230996?v=pdf