Winter 1-31-2023

Legality of Autonomous Weapons: Where to Draw the Line?

Nayra Abdeltawab Ibrahim Abdeltawab
The American University in Cairo AUC, nayragadallah@aucegypt.edu

The American University in Cairo

School of Global Affairs and Public Policy

LEGALITY OF AUTONOMOUS WEAPONS: WHERE TO DRAW THE LINE?

A Thesis Submitted by

Nayra Abdeltawab Ibrahim Gadallah

To the Department of Law

Fall 2022

in partial fulfillment of the requirements for
LL.M. Degree in International and Comparative Law

The American University in Cairo

School of Global Affairs and Public Policy

LEGALITY OF AUTONOMOUS WEAPONS: WHERE TO DRAW THE LINE?

A Thesis Submitted by

Nayra Abdeltawab Ibrahim Gadallah

to the Department of Law

Fall 2022

in partial fulfillment of the requirements for the LL.M. Degree in International & Comparative Law has been approved by

Professor Thomas Skouteris (Supervisor) _____
Associate Professor
Law Department
The American University in Cairo

Professor Hany Sayed (First Reader) _____
Associate Professor
Law Department
The American University in Cairo

Professor Jason Beckett (Second Reader) _____
Assistant Professor
Law Department
The American University in Cairo

Graduate Program Director	Date	GAPP Dean	Date
_____	_____	_____	_____
Professor Thomas Skouteris		Dr. Noha El-Mikawy	

DEDICATION

To those who strive to be a better version of themselves . . .Keep going!

ACKNOWLEDGMENTS

In acknowledgment of the endless efforts of my mother, Mervat, and my father, I want to start. Thanks for the love, support, and empowerment you provide me with. I shall not also forget to thank you for the valuable gift; my siblings, my companions, Noha, Nehal, and Ibrahim. You are all truly my source of strength. I also owe a deep sense of gratitude to my lifetime best friend, Maram Mohsen, and her family.

I want also to deeply thank all the people who directly and indirectly supported me throughout my LLM program at the AUC. I am grateful to the Yousef Jameel GAPP Public Leadership Program for funding my studies at the AUC. I want to thank specifically Dr. Laila El-Baradei and Ms. Amira El-Biltagy for all the support they extend to the Jameel family.
I wish also to express my extreme appreciation to all the professors who taught me in the Law Department. I am grateful to Professor Thomas Skouteris for all the knowledge and guidance he extended to me as my thesis supervisor. My deep thanks to Professor Hany Sayed and Professor Jason Beckett for the feedback they offered as my thesis readers. A special thanks to Ms. Diana Van Bogaert for the generous help and support throughout the program and for the special role she plays at the Law Department; you truly positively impacted my experience at AUC.
I would like also to express my warmest thanks to my study peers, Menna Moataz, Nourhane Awad, Moataz Aidaros, Waleed El-Farrs, Ali Moussa, Mohamed Maged, Mariam Salem, and Heba Habib.

I also want to acknowledge the support of my ex-supervisors at UNHCR, Egypt who encouraged me to pursue this degree, a special thanks to Ms. Rania El-Gindy, Ms. Seza Kirishdjian, and Ms. Florentina Covaliu.

I also owe a deep appreciation to my professors at Ain Shams University. Thanks to Dr. Ahmed Khalifa and Dr. Yassin El-Shazly, Dr. Doaa Abo Taleb for introducing me to the Moot Court experience at Oxford University, which I will continue to cherish for the rest of my life. Thanks for expanding my horizons and for believing in your message.

Last but not least, I cannot find enough words to thank my husband, Hesham, for all the misery he endured with me to finish this paper. We made it together to prove that when we have each other, we have everything!

The American University in Cairo
School of Global Affairs and Public Policy
Department of Law

LEGALITY OF AUTONOMOUS WEAPONS: WHERE TO DRAW THE LINE?

Nayra Abdeltawab Ibrahim Gadallah

Supervised by Professor Thomas Skouteris

ABSTRACT

Inspired by Koskenniemi's work, From Apology to Utopia, this paper attempts to engage in the discussion on the legality of autonomous weapons by showing the conflicting arguments presented by advocates of each side of the debate. The paper does not aim at finding the answer to whether autonomous weapons can be lawfully deployed or not, but rather its main interest is to highlight the indeterminacy within international law that allows both advocates and opponents of banning autonomous weapons to hold to their arguments and legally defend them on basis of the same legal rules used by their adversaries to refute their arguments and to build conflicting arguments. The paper will also be investigating the efforts made to define autonomous weapons and how definitions play an important role in giving international law this indeterminate character.

KEYWORDS: Killer Robots – Lethal Autonomous Weapons – Fully Autonomous Weapons – Martens Clause – Law Indeterminacy – Artificial Intelligence – Meaningful Human Control – Weapon Review.

TABLE OF CONTENTS

INTRODUCTION **1**

I- DEFINING AUTONOMOUS WEAPONS: **3**

A. Efforts for definitions: **3**

B. Autonomous weapons and human control: **5**

C. Autonomous weapons and level of intelligence **10**

D. Comprehensive approaches to define autonomous weapons **12**

II- LEGALITY OF AUTONOMOUS WEAPONS – AUTONOMOUS WEAPONS DEVELOPMENT AND WEAPON REVIEW UNDER ARTICLE 36: **14**

A. An introduction to Article 36; and its scope **15**

B. Applying Article 36 on autonomous weapons **20**
a) Principle of Distinction 22
b) Proportionality Principle 24
c) Principle of distinction and principle of proportionality and autonomous weapons 26
d) Autonomous weapons and Martens Clause 29
i. The Historical Background of the Martens Clause 30
ii. The legal scope and value of Martens Clause 31
iii. Principle of Humanity as a prong of the Martens Clause 37
iv. Impact of humanity considerations on disarmament efforts (Humanitarian Disarmament) 39
v. Morality of autonomous weapons and humanity principle 40
vi. Dictates of public conscience as the second prong of the Martens Clause 42

III- QUASI-LEGAL ARGUMENTS ON AUTONOMOUS WEAPONS: **44**

A. Inevitability of autonomous weapons and arms race **44**

B. Reaching consensus on the lawfulness of autonomous weapons and lessons learned from banning weapons. **47**

CONCLUSION **50**

Introduction

The inherent function of international humanitarian law as an area of law that governs the conduct of warfare while striking the balance between military necessity and humanity considerations has influenced debates over the legality of autonomous weapons. On one side, the advocates of the ban on the development and deployment of autonomous weapons submit that the necessity of the ban emanates from the urgency of protecting humanitarian considerations that are threatened by autonomous weapons. Such a position elevates humanity considerations over military necessity, the latter constitutes the main interest of states. On the other side, opponents of the ban view that such weapons and their development provide for aspiring promises of military advancements that are translated into a military necessity that shall be respected as the ultimate aim of states on the battlefield. At a first glance, one might think that embracing these two competing notions, humanity and military necessity, would render international humanitarian law ineffective. However, it is this quality of international law that allows for the survival of international humanitarian law as a body of law to which both supporters of each side subscribe. If the law favors the side of humanitarian considerations, it will propose a utopian legal framework that will collapse by states unwilling to commit to rules that limit their interests in preserving military survival. Conversely, if the law only regards military necessity, it will strip international humanitarian of its function of reducing human suffering during the war and it will shift to be a tool of apology for state practices regardless of their compliance with humanitarian considerations.[1] Owing to such characteristics of international law, both opponents and advocates of a ban on autonomous weapons can weave their contesting arguments on the legality of autonomous weapons on basis of legal foundations that are equally respected and protected. However, the ultimate decision of regulating a hard case within international law is a decision that finds its legal foundations in law while at the same time striking a balance between other interdisciplinary considerations, political, economic, moral, etc.

The question of assessing the legality of new weapons is primarily regularized by article 36 of Additional Protocol I, which reads as follows:

[1]Elliot Winter, *Pillars not Principles: The Status of Humanity and Military Necessity in the Law of Armed Conflict*, 25 Journal of conflict & security law 1–31 (2020).

> In the study, development, acquisition, or adoption of a new weapon, means, or method of warfare, a High Contracting Party is under an obligation to determine whether its employment would, in some or all circumstances, be prohibited by this Protocol or by any other rule of international law applicable to the High Contracting Party.[2]

To assess the legality of autonomous weapons, the paper will be investigating it through the lens of article 36. The decision to adopt article 36 as the primary legal foundation of the autonomous weapons review test is that article 36 within it encompasses the legality checks of all international law rules governing the legality of deploying new weapons including the obligation to respect proportionality and distinction principles as well as Martens Clause with its two prongs, the principle of humanity and dictates of public conscience. However, it is important before delving into the legal debate on the legality of autonomous weapons to engage in the debate of defining what constitutes autonomous weapons; the debate which contributes to the further complexity of reaching a consensus on regulating the deployment of autonomous weapons.

Accordingly, the first part of this paper will be presenting the current debate of finding a definition for fully autonomous weapons and within this part, a dichotomy of defining autonomous weapons based on their autonomy level and level of human control over such weapons will be explored. The second part of the paper will be concerned with the legality test within article 36 as the starting point for the legality test and within this part, the arguments advanced by both advocates of the ban of autonomous weapons and opponents of banning such weapons will be presented. In the third part, the aim will be to present the quasi-legal arguments presented by both sides, which are influenced by other interdisciplinary considerations, such as international relations, international security, and political and economic considerations. Ultimately, the fourth part will be interested in examining precedents from the history of banning weapons which gained the consensus of the international community to stand on the reasons and the factors that influenced the decision of banning such weapons as the concluding remarks of this paper.

[2] Article 36 of the Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts (Protocol I), 8 June 1977. *Available at*: https://ihl-databases.icrc.org/ihl/WebART/470-750045.

I- Defining autonomous weapons:

A. Efforts for definitions:

Agreeing on a definition for any weapon has been the primary task during the efforts of regulating the use of any weapons; however, this task seems to be a complex one when the discussion shifted to autonomous weapons. To date, there is no unified agreed definition of what constitutes autonomous weapons.[3] This is a result of the fast-paced technological developments attributed to this field and the complexity of defining key terms that formulate the characteristics of autonomous weapons. It is however noted that the lack of consensus on the definition did not prevent stakeholders from using the term autonomous weapons. In March 2021, the Panel of Experts on Libya reported the usage of autonomous weapons, particularly the STM Kargu-2, was used in the civil war in Libya in March 2020,[4] and around 30 states have expressly declared their support to ban on autonomous weapons.[5] The absence of an agreed definition has led to the prevalence of multiple definitions used and adopted by various stakeholders. In their paper on analyzing the definitions of autonomous weapons systems, Taddeo and Blanchard have listed twelve existing definitions of such systems[6] that were driven by political and strategic motivations. Some of the identified definitions were described as premature and others were criticized for being unrealistic. For example, the UK adopted in 2016 a definition of autonomous weapons "to be one which is capable of understanding, interpreting and applying higher level intent and direction based on a precise understanding and appreciation of what a commander intends to do and perhaps more importantly why."[7] A definition which was criticized for focusing on "intent" which describes futuristic sci-fi weapon systems that are not realistic to exist in the foreseen future and excludes autonomous weapons functions that are presently being produced and developed. It is however important to note that the UK has altered its position in its June 2022 publication in which it shifted

[3] Congress Research Service, Defense Primer: U.S. Policy on Lethal Autonomous Weapon Systems, Updated September 12, 2022, *available at*: https://sgp.fas.org/crs/natsec/IF11150.pdf.
[4] United Nations, Security Council, S/2021/229, 8 March 2021, *available at*: https://documents-dds-ny.un.org/doc/UNDOC/GEN/N21/037/72/PDF/N2103772.pdf?OpenElement .
[5] Human Rights Watch, Stopping Killer Robots: Country Positions on Banning Fully Autonomous Weapons and Retaining Human Control, August 2020, *available at*: https://www.stopkillerrobots.org/wp-content/uploads/2020/08/arms0820_web_final.pdf .
[6] Taddeo, M., Blanchard, A. A Comparative Analysis of the Definitions of Autonomous Weapons Systems. Sci Eng Ethics 28, 37, published on 23 August 2022, *available at*: https://doi-org.libproxy.aucegypt.edu/10.1007/s11948-022-00392-3 .
[7] I*d.*

its language to focus on human control on autonomy.[8] Another approach, primarily adopted by Russia, deviates from the discussion of defining autonomous weapons arguing that such discussion is early to be addressed noting their future development and that a working definition cannot be limited to the current understanding of such weapons.[9] The United States, on the other hand, adopted a definition for autonomous weapons in 2012 as:

> A weapon system that, once activated, can select and engage targets without further intervention by a human operator. This includes human-supervised autonomous weapon systems that are designed to allow human operators to override operation of the weapon system, but can select and engage targets without further human input after activation.[10]

The definition adopted by the U.S. is primarily focusing on the relationship between the human operator and the machine, with the focus on the level of human supervision and control over the system as the defining factor. In line with the U.S. approach, the ICRC also adopted a similar approach in defining autonomous weapons with a focus on human intervention in the process. The ICRC defined such weapons as

> [A]ny weapon system with autonomy in its critical functions. That is, a weapon system that can select (i.e. search for or detect, identify, track, select) and attack (i.e. use force against, neutralize, damage or destroy) targets without human intervention.[11]

Throughout the past eight years, efforts have been made within the meetings of the Convention on Certain Conventional Weapons (CCW) and its group of governmental experts on lethal autonomous weapons systems (GGE) to reach a consensus on the definition of autonomous weapons. Faced with the lack of consensus on the matter, in 2018, the group shifted the discussion to technical characteristics of the systems such as level of autonomy, artificial intelligence, and human control over such weapons with hopes to reach an agreement on these matters resulting in a clear definition. However, discussions over technical characteristics have not yet yielded to consensus on the definition of autonomous weapons and the contentions were consequently transferred to defining what constitutes human control and autonomous systems. It is, however,

---

[8] UK Ministry of Defence, Ambitious, Safe, Responsible: Our approach to the delivery of AI-enabled capability in Defence, June 2022, *available at:* https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1082991/20220614-Ambitious_Safe_and_Responsible.pdf

[9] CCW/GGE.1/2021/WP.1, dated 27 September 2021, *available at:* https://documents-dds-ny.un.org/doc/UNDOC/GEN/G21/263/18/PDF/G2126318.pdf?OpenElement

[10] United States of America, Department of Defense, DoDD 3000.09, 21 November 2012, *available at:* https://www.esd.whs.mil/portals/54/documents/dd/issuances/dodd/300009p.pdf

[11] ICRC, Autonomy, artificial intelligence and robotics: Technical aspects of human control, August 2019, *available at:* https://www.icrc.org/en/document/autonomy-artificial-intelligence-and-robotics-technical-aspects-human-control

not to be understood that the value of such discussions is undermined. On the contrary, they contribute to enriching the debate over the topic and provide a platform for policy sharing, and in case of adopting a governing treaty on the matter, they would act as travaux preparatoires for the treaty. Yet, the lack of consensus within the meetings of the CCW/GGE due to indeterminacy hoofing around the terminologies relevant to the discussion has led to modest-paced progress that is certainly unable to catch up with the hasty development of autonomous weapons. To show the indeterminacy of reaching a consensus definition of autonomous weapons-related terminologies, the following section will engage in the discussions of defining (i) human control over autonomous weapons and (ii) autonomous characteristic of such weapons. It is important to reiterate at this level that reaching a working definition that would gain consensus by the international society is not the interest of this part, but the purpose of engaging in discussion on terminology is to demonstrate that indeterminacy related to definitions contributes to the indeterminacy on the legality of autonomous weapons and how the interpretations of definitions are articulated within the same terminologies used by each side of the debates to advance their legal arguments.

B. Autonomous weapons and human control:

One way of defining autonomous weapons is to look into the human-machine relationship and assess the level of control practiced by the operator on the machine and its functions. The obligation to maintain weapons under human control in the abstract is not a contested concept; whereas the disagreement is revolving around the degree of control to be exercised over weapons. The Campaign to Stop Killer Robots along with activists calling for the ban of autonomous weapons plea for humans to exercise meaningful human control over the use of force.[12] Whereas the United States within the Department of Defense Directive 3000.09 (DoDD) adopted the notion of exercising appropriate levels of human judgment over the use of force.[13] The conflicting terms used to describe control as well as the ambiguity surrounding the terms "meaningful" and "appropriate levels of human control" reflect the indeterminacy of the discussion on what constitutes autonomous weapons. In an attempt to clarify concepts, the loop notion was introduced into the discussion to mimic the "observe, orient, decide, act" (OODA) loop used by air forces.[14]

[12] Campaign to Stop Killer Robots, Response to GGE Chair's Guiding Questions, September 2021, *available at:* https://www.stopkillerrobots.org/wp-content/uploads/2021/09/CSKR_Response-to-GGE-Chairs-Guiding-Questions.pdf
[13] *Supra note* 3.
[14] Scharre P. Army of none. Autonomous weapons and the future of war. New York/London: W.W. Norton; 2018.

The position of the human operator within the "sense, decide, act" loop classifies the levels of control exercised by humans on the weapons system to humans in the loop, humans on the loop, and humans out of the loop. In its report on losing humanity,[15] Human Rights Watch defined three categories:

> • **Human-in-the-Loop Weapons**: Robots that can select targets and deliver force only with a human command;
>
> • **Human-on-the-Loop Weapons:** Robots that can select targets and deliver force under the oversight of a human operator who can override the robots' actions; and
>
> • **Human-out-of-the-Loop Weapons:** Robots that are capable of selecting targets and delivering force without any human input or interaction.

The report identified autonomous weapons, which they plea to ban, as weapons where humans are out of the loop and weapons where humans are on the loop but with limited human supervision. By applying the loop classifications on terminologies used by the United States in the DoDD, the weapons with humans on the loop correspond to human-supervised autonomous weapon systems which are defined as "An autonomous weapon system that is designed to provide human operators with the ability to intervene and terminate engagements, including in the event of a weapon system failure, before unacceptable levels of damage occur."[16] The DoDD considered human supervised autonomous weapons to be one type of autonomous weapons, which are defined as "a weapon system that, once activated, can select and engage targets without further intervention by a human operator", human out of the loop weapon within the language of HRW. In addition to that, DoDD also adopted a definition for semi-autonomous weapons:

> A weapon system that, once activated, is intended to only engage individual targets or specific target groups that have been selected by a human operator. This includes: Semi-autonomous weapon systems that employ autonomy for engagement-related functions including, but not limited to, acquiring, tracking, and identifying potential targets; cueing potential targets to human operators; prioritizing selected targets; timing of when to fire; or providing terminal guidance to home in on selected targets, provided that human control is retained over the decision to select individual targets and specific target groups for engagement. "Fire and forget" or lock-on-after-launch homing munitions that rely on TTPs to maximize the probability that the only targets within the seeker's acquisition basket when the seeker activates are those individual targets or specific target groups that have been selected by a human operator.[17]

[15] Human Rights Watch & The International Human Rights Clinic, Losing Humanity: The Case against Killer Robots, 2012.
[16] *Supra note* 10.
[17] *Supra note* 10.

Owing to the different languages spoken by the HRW report and DoDD situating semi-autonomous weapons within the spectrum of the control loop is a challenging task. The first sentence of the definition, a weapon system that, once activated, is intended to only engage individual targets or specific target groups that have been selected by a human operator, implies that such weapons would fall within the human in the loop category. However, the rest of the definition shows that functions assigned to such weapons would situate them within the blurry area between humans in the loop and humans on the loop. The DoDD's main objective is to regulate the development of semi-autonomous and autonomous weapons to ensure that control over such weapons would not be lost. In doing so, the Directive identified types of semi-autonomous weapons and autonomous weapons that require a senior review process, *see* Appendix I. Going back to the discussion on levels of control, advocates of the ban on autonomous weapons call for an absolute ban on the development and deployment of such weapons due to the absence of meaningful control over their critical functions.[18] To clarify what would constitute critical functions of a weapon, HRW cited Article 36 NGO, "This means when, where and how weapons are used; what or whom they are used against; and the effects of their use", concluding that "[h]umans should exercise control over individual attacks, not simply overall operations".[19] However, this approach is criticized on basis of two main grounds. First, adopting the meaningful control criteria to define autonomous weapons leads to the inclusion of types of weapons that neither fall within the understood concept of autonomous weapons, nor are targeted by the ban of autonomous weapons; such as anti-tank mines, which humans have no control over what or whom they are used against. The second major criticism is that the requirement to have control over individual attacks, not the overall operations, overlooked the changes in the nature of the targeting process in modern warfare where the targeting process has become more dynamic when compared to the prevalence of deliberate targeting in the past. In this regard, Ekelhof[20] presented the case study of executing an F-16 air mission which was preplanned following six steps of find, fix, track, target, engage and assess, known as the F2T2EA cycle,[21] where the F-16 operator, tasked to attack

[18] Human Rights Watch, Killer Robots and the Concept of Meaningful Human Control, April 2016, Memorandum to Convention on Conventional Weapons (CCW) Delegates.
[19] *Id.*
[20] Merel Ekelhof, *'Moving Beyond Semantics on Autonomous Weapons: Meaningful Human Control in Operation'*, Global Policy, vol. 10/no. 3, (2019), at 343-348.
[21] F2T2EA cycle is within the mission planning and execution phase of the six phases of NATO targeting operations policies, for more on this topic see: Ekelhof, M. A. C. 'LIFTING THE FOG OF TARGETING: "Autonomous Weapons" and Human Control through the Lens of Military Targeting', Naval War College Review, vol. 71/no. 3, (2018), at 61-95, and Mark Roorda, NATOs Targeting Process: Ensuring Human Control over (and Lawful Use of) 'Autonomous' Weapons, 2015, *available at:*

a military compound using GPS-guided weapons, was not involved in the pre-planned process of finding and validating the target and the target and engage phases were conducted by GPS-guided weapons without the need to visually confirm the targeting and engaging by the operator. In this scenario, the operator himself did not control the whole process and relied on information passed to him from his superiors. Accordingly, such conduct lacks meaningful human control; however, a holistic overview of the whole process supports the argument that control was exercised throughout the process. On basis of this argument, opponents of the ban on autonomous weapons build their case that weapons are never truly autonomous as a human is always in the loop during the planning stage to set goals, gather intelligence, select and validate targets, and decide on the suitable weapons to be used and the circumstances of their deployment.[22] This approach stretches the scope of the loop so that only weapons that could function without human control throughout all the six phases of the targeting[23] would be functioning out of the loop. A type of weapon that is not technically realistic to exist and states have no interest to develop.

On the premise of stretching the loop of control, the United States, one of the major opponents of banning autonomous weapons, ascertains that the design of autonomous weapons should always allow commanders and operators to exercise appropriate levels of human judgment over the use of force while keeping in mind that appropriate level of judgment is purely a context-based term that cannot be defined to fit all scenarios.[24] Owing to this obligation, the United States issued the DoDD to define autonomous and semi-autonomous weapons that should go through a senior-level review process in 2012, and most recently in March 2022, it co-submitted a proposal to the CCW/GGE meeting on the principles and good practices on emerging technologies in the area of lethal autonomous weapons systems.[25] The proposal identified some safeguards for the lawful use

https://pure.uva.nl/ws/files/2566151/167993_Roorda_NATO_s_Targeting_Process_Ensuring_Human_Control_Over_and_Lawful_Use_of_Autonomous_Weapons_October_2015_.pdf

[22] Mark Roorda, NATOs Targeting Process: Ensuring Human Control over (and Lawful Use of) 'Autonomous' Weapons, 2015, *available at:* https://pure.uva.nl/ws/files/2566151/167993_Roorda_NATO_s_Targeting_Process_Ensuring_Human_Control_Over_and_Lawful_Use_of_Autonomous_Weapons_October_2015_.pdf

[23] The six phases of targeting are: (1) Commander's Intent, Objectives, and Guidance, (2) Target Development, Validation, Nomination, and Prioritization, (3) Capabilities Analysis, (4) Commander's Decision, Force Planning, and Assignment, (5) Mission Planning and Execution, (6) Combat Assessment, see Ekelhof, M. A. C. 'LIFTING THE FOG OF TARGETING: "Autonomous Weapons" and Human Control through the Lens of Military Targeting', Naval War College Review, vol. 71/no. 3, (2018), at 61-95

[24] *Supra note* 3.

[25] Convention on Certain Conventional Weapons – Group of Governmental Experts on Lethal Autonomous Weapons Systems, Principles and Good Practices on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems - Proposed by Australia, Canada, Japan, the Republic of Korea, the United Kingdom and the United States, March 2022, *available at:* https://meetings.unoda.org/section/ccw-gge-2022_documents_18542_proposals_19869/.

of autonomous weapons including, but not limited to, retaining human responsibility for the use of autonomous weapons throughout the entire life-cycle of the weapon system, weapons cannot be designed to be used against the civilian population, maintain control over the duration, types of targets, geographical scope and scale of operation of autonomous weapons and equip them with self-destruct, self-deactivation, or self-neutralization mechanisms, incorporate readily understandable human-machine interfaces and controls, conduct operations under responsible command, and train personnel on autonomous systems in realistic operational conditions.[26]

On the other hand, supporters of the ban on autonomous weapons, in their early efforts to advocate for their cause, focused on maintaining meaningful human control over weapons to show the disastrous humanitarian implications of developing and deploying autonomous weapons with no meaningful human control without clearly putting meaningful human control into context.[27] However, as the discussion over the topic evolved and the ambiguity of the term meaningful human control was framed, their position was altered to a less utopian one by acknowledging that maintaining control is a context-oriented process.[28] A thorough reading of publications by HRW and the Campaign to Stop Killer Robots shows that advocates of the ban on autonomous weapons started to discuss practical and operational frameworks to ensure control over weapons systems.[29] For example, the ICRC in June 2020 issued a publication[30] focusing on addressing the practical limitations of the use of autonomous weapons by introducing certain measures to insure human control over weapons. One of the measures suggested within this publication is to maintain a certain level of control for all human users using the weapon and its effects. The publication also suggested maintaining a level of autonomy that would allow human operators to exercise reasonable certainty of the effects of the weapon, and to maintain control over the parameters of use of the weapon including control over target type and profile, and spatial and temporal scope of operation. While it is clear that the report involved an in-depth discussion of what constitutes

[26] Id see also: Convention on Certain Conventional Weapons – Group of Governmental Experts on Lethal Autonomous Weapons Systems, Humanitarian benefits of emerging technologies in the area of lethal autonomous weapon systems, submitted by the United States, April 2018.
[27] See for example: *Supra note 15*, see also, *infra note* 87.
[28] See for example Limits on Autonomy in Weapon Systems: Identifying Practical Elements of Human Control, June 2020 - Executive Summary', International Review of the Red Cross (2005), vol. 102/no. 913, (2021); see also Recommendations on the Normative and Operational Framework for Autonomous Weapon Systems, Campaign to Stop Killer Robots, June 2021; and *supra note* 12.
[29] *Id.*
[30]Limits on Autonomy in Weapon Systems: Identifying Practical Elements of Human Control, June 2020 - Executive Summary', International Review of the Red Cross (2005), vol. 102/no. 913, (2021),

human control over autonomous weapons by reference for example to numerical equations to determine the number of human operators needed for a certain weapon, such measures did not solve the debate. A synopsis of the development of arguments by the pro-ban side and anti-ban side shows that the debate trickled down to the details of such measures suggested by both sides. For example, both sides agree that there should be temporal scope for the deployment of autonomous weapons; but what constitutes an acceptable time limitation for the operation is still contested. What would be considered a "reasonable" level of certainty? From the anti-ban perspective, all weapons including medieval weapons such as the arrow and stone would have a fire-and-forget moment after which they cannot be recalled after release[31] and their final target could be altered by other external factors such as the movement of the target, wind, etc. In conclusion, the determinative factor concerning the level of control that should be maintained is the position that advocates take about determining their priorities between military necessity and humanitarian considerations.

C. Autonomous weapons and level of intelligence

Another approach to defining autonomous weapons is by referring to the level of intelligence; where automation is a spectrum that has automatic machines at the bottom level of intelligence and automated at the middle and autonomous at the top of the spectrum. Automatic machines are the ones that exhibit a simple process of decision-making by sensing the environment in a simple way predictable to humans.[32] A classic example of an automatic machine is the automatic washing machine, which predictably reacts to its environment through the inputs of the user by pressing its buttons within the threshold of functions assigned to it. Whereas automated machines are more complex machines that can take more sophisticated actions by weighing certain variables within a process that can be traced by the user. Autonomous machines are however machines that are goal-oriented, and their directions are determined by the machine through cognitive processes done by the machine and are less traceable to users.[33] The complexity of situating a machine within the autonomy spectrum is augmented due to the blurring lines between different types of autonomy on the spectrum and when this is applied to weapons the distinction is not a straightforward task and even some argue that an autonomous system would eventually turn into an automated system

[31] *Supra note* 14.
[32] *Id.*
[33] *Id.*

after its functions become more predictable to users.[34] Autonomous systems are mostly machines that function relying on algorithms and artificial intelligence. From the perspective of pro-ban on autonomous weapons advocates, the unpredictability of functions of autonomous weapons could lead to unforeseen results[35] and provide examples of the failure of autonomous systems to support their case. Whereas on the other side, anti-ban advocates argue that autonomous weapons do not act by their own will, but they act in a goal-oriented scheme to achieve the goals set by the user; yet the machine has the flexibility to decide how to achieve this goal. Scharre even argued that autonomous systems are necessary for the uncontrolled environment; where variables and surroundings are not predictable and accordingly the autonomous machine can react to them and achieve its goal. He gave an example of a self-driving car that operates on a closed track with no obstacles where its actions are all pre-determined on basis of the anticipated environment, the machine would then be deemed automatic or automated, but such a machine cannot function and is not effective in complex and unpredictable environments.[36] Yet, the pro-ban group argues that such autonomous functions can be fooled by their surroundings as they lack the cognitive abilities of humans to process and analyze data. An embarrassing example of the failure of autonomous systems is an incident when a computer vision system was tricked by a fruit with an iPod label on it to recognize it as an iPod.[37] Whereas on the other side, anti-ban advocates promote the promising capabilities of autonomous systems to outperform humans in image recognition functions, Scharre presented the findings by the Microsoft team in 2015 after they developed a deep neural network that was able to perform image recognition with 4.94% error rate and then raised to 3.57% error rate as compared to 5.1% error rate by humans. Despite such promising advances, Scharre also recognized the vulnerability of deep neural networks to adversarial images that are fed to the system to spoof the networks; networks falsely identified objects in a way that is not comprehensible to humans and even counterintuitive and unpredictable.[38] Whether machines would evolve to perform cognitive functions in a way similar to humans or surpass them is still a matter of contention that might be proven right or wrong by time; however, what is certain in the meantime is that classifying machines as autonomous or not based on their level of intelligence is

[34] *Id.*
[35] Eugenio V. Garcia, Killed by algorithms: Do autonomous weapons reduce risks?, Beyond the Horizon, *available at:* https://behorizon.org/killed-by-algorithms-do-autonomous-weapons-reduce-risks/
[36] *Supra note* 14.
[37] Kelsey Atherton, Loitering munitions preview the autonomous future of warfare, August 2021, *available at:* https://www.brookings.edu/techstream/loitering-munitions-preview-the-autonomous-future-of-warfare/
[38] *Supra note* 14.

a grey zone that would not yield to concrete findings noting the blurry zones within the automation spectrum. Accordingly, some scholars resorted to defining autonomous weapons by taking a comprehensive approach that defines the weapon by merging more than one automation feature; which will be discussed in the following section.

D. Comprehensive approaches to define autonomous weapons

For a purpose of providing a more conclusive definition of autonomous weapons, some have resorted to defining such systems by reference to multiple features of the system. Scharre[39] for example referred to the three dimensions of autonomy which are: the task performed by the machine, the relationship between the human operator and the machine (human control), and the level of sophistication by which the machine performs the task (level of intelligence). To these three spectra, Blanchard and Taddeo[40] added adaptive capabilities of the machine to the equation, although adaptive capabilities intersect with the level of sophistication as identified by Scharre, Blanchard, and Taddeo wanted to stress the role of artificial intelligence in the process. Scharre concluded that an increase in autonomy can be achieved by an increase in any of the three dimensions he identified. He introduced the task being performed by the machine as one of the spectrums that have to be explored when deciding on the level of autonomy of the machine acknowledging that we cannot deal with a machine a simple task such as the thermostat the same way we deal with machines that have control over nuclear weapons. The concept of investigating the task of the machine was touched upon by pro-ban activists and this can be seen for example in their choice of the name of the Stop Killer Robots Campaign which focuses on the lethality function of the machine. Blanchard and Taddeo suggested a definition for autonomous that reflects the four features they identified, control, level of autonomy, adaptive capabilities of the machine, and its purpose of use, the definition reads as follows:

> [A]n artificial agent which, at the very minimum, is able to change its own internal states to achieve a given goal, or set of goals, within its dynamic operating environment and without the direct intervention of another agent and may also be endowed with some abilities for changing its own transition rules without the intervention of another agent, and which is deployed with the purpose of exerting kinetic force against a physical entity (whether an object or a human being) and to this end is able to identify, select or attack the target without the intervention of another agent is an AWS. Once deployed, AWS can be operated with or without some forms of human control (in, on or out the loop). A lethal

[39] *Id.*
[40] *Supra note* 6.

> AWS is specific subset of an AWS with the goal of exerting kinetic force against human beings.[41]

The attempt to define and identify autonomous weapons by merging the different features of the machine including the level of control, the sophistication of automation, and the task of the system has advanced the discussion on finding a conclusive definition for autonomous weapons. However, such an attempt did not solve the dilemma of finding a conclusive definition for each of the defining features, most importantly level of control; which reverts the discussion to what constitutes meaningful human control over machines. The conclusion that can be drawn from the debates on the definition of autonomous weapons is that owing to the relativity of the concepts related to autonomy, the indeterminacy of reaching a definition is contemplated; which opens the door widely for each of the contesting sides to shape the definition of autonomous weapons in a manner that serves their driving motives being humanitarian considerations or military necessity.

[41] *Id.*

II- Legality of autonomous weapons – Autonomous weapons development and weapon review under Article 36:

As explained in the introductory part of this paper, the legality of autonomous weapons is examined through the lens of article 36 of Additional Protocol I. The choice of examining the legality of autonomous weapons through the lens of Article 36 or the Martens Clause is a continuation of positivists-naturalists debate which I will be further elaborating upon throughout this paper. It is the extension of the indeterminate character of international law by which hierarchy of its sources is often a topic for a heated debate. Although the Martens Clause has evolved over decades to retain acceptance as a widely recognized customary international rule; for the purpose of the discussion on the legality of autonomous weapons, I choose Article 36 to be the starting point of my engagement with the legality of autonomous weapons. This decision is primarily based on the fact that Article 36 constitutes the *lex specialis* rule on weapon review when compared to the Martens Clause. It further provides, through its wordings, a more inclusive tool to examine the legality of autonomous weapons against rules of international law applicable to the states, including the Martens Clause. This, however, does not negate that the interpretation of Article 36 and the application of weapon review, in reality, involves legal indeterminacy which will be unveiled later in this paper.

The bindingness of article 36 as a customary value is not well-established; however, the article provides a guide for the legal review of weapons which is conclusive. For this reason, the author chooses it to be the starting point of the legal review process which encompasses the relevant legal instrument and concepts that need to be investigated to decide on the legality of autonomous weapons.

Weapon review obligation appeared in article 36 of the Additional Protocol I to Geneva Conventions which was adopted in 1977. The body of the article stipulates that:

> In the study, development, acquisition or adoption of a new weapon, means or method of warfare, a High Contracting Party is under an obligation to determine whether its employment would, in some or all circumstances, be prohibited by this Protocol or by any other rule of international law applicable to the High Contracting Party.[42]

[42] *Supra note* 2.

The article constitutes solid legal grounds for advocates of the ban of autonomous weapons to argue that they fail to respect this weapon review obligation due to the unpredictable nature of autonomous systems which would exclude them from the effective weapon review process. At the same time, the article provides objectors against the ban with the guiding manual to legally establish their support for the continuous development of autonomous weapons shall they succeed to prove that such systems fully respect or are expected to respect the weapon review tests illustrated by this article. For these reasons and to set the parameters for this legal debate, it is important to stand first on the scope of application of article 36, its binding value, and states practices in this regard.

A. An introduction to Article 36; and its scope

The way in which article 36 was articulated provided broad guidelines for the weapon review process that should ideally be performed by state parties; however reference should also be made in this context to the 1987 ICRC commentary on the 1977 Additional Protocols[43] and the 2006 ICRC guide to the Legal Review of New Weapons, Means and Methods of Warfare[44], which provide an elaborate interpretation of the article. States' practices concerning weapon review are somehow scarce for reasons related to the costs of establishing weapon review mechanisms and the classified nature of the process which discourage states from sharing information on the national weapon review process. The United States of America, despite not ratifying API, is one of the leading states in the field of weapon review and offers a high number of published reports on weapon review owing to the existence of the Freedom of Information Act.[45]

Perhaps the most straightforward question that can be answered by only reading through the wording of the article is the timing in which states should start the implementation of the weapon review process. The article clearly states that the review process should start as early as the study phase for the weapon development and should continue through the development process, its design, and at the time of acquisition or adoption of new weapons, means, and methods of warfare;

[43] International Committee of the Red Cross (ICRC) commentary on the 1977 Additional Protocols, Article 36, Geneva 1987, *available at*: https://ihl-databases.icrc.org/applic/ihl/ihl.nsf/Comment.xsp?action=openDocument&documentId=F095453E41336B76C12563CD00432AA1.
[44] International Committee of the Red Cross (ICRC), A Guide to the Legal Review of Weapons, Means and Methods of Warfare, Volume 88 Number 864 December 2006, (ICRC: Geneva, 2006), *available at:* https://www.icrc.org/en/doc/assets/files/other/irrc_864_icrc_geneva.pdf.
[45] *Id* at 955.

asserting that the review process should be undertaken in all the stages of the weapon life cycle.[46] The obligation to conduct a weapon review at the time of acquisition or adoption is a confirmation that the weapon review process is not only an obligation on the manufacturing states and purchasing states; an obligation that fills the gap of purchasing weapons from the private sector which does not necessarily abide by weapon review obligations.[47]

The bindingness of Article 36 as a customary rule is subject to debate. 2006 ICRC guide on Article 36 stated that it "arguably" applies to all States, regardless of whether or not they are party to Additional Protocol I; favoring the argument that the rule is a customary one but cautiously used the word arguably as a recognition of the existence of the debate. Additional Protocol I was ratified by 174 states,[48] yet only less than two dozen states are known to have weapon review mechanisms within their national systems.[49] The fact on which Jevglevskaja built her argument that law review obligation has not yet crystallized into customary rule due to the absence of extensive and uniform state practice and the absence of *Opinio Juris*.[50] Additionally, the weapon review obligation was not listed within the rules of the customary IHL database by ICRC.[51] However, one cannot overlook the significance of the weapon review obligation since it overlaps with other IHL rules that gained the customary status. Article 36 was adopted to set a practical restriction on states' rights to use any means or methods of war;[52] a basic principle that was enshrined in Article 35 of the API[53] and the obligation of precautions in the choice of means and methods of warfare was recognized within rule 17 of ICRC database of IHL customary rules.[54] Article 36 also overlaps with Article 82 of the API; the latter was also crystallized into a customary rule.[55] The obligation to make legal advisors available to military commanders provides a minimum safeguard to respect

[46] *Id* at 951&952.
[47] *Id* at 951&952.
[48] Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts (Protocol I), 8 June 1977, https://ihl-databases.icrc.org/applic/ihl/ihl.nsf/Treaty.xsp?action=openDocument&documentId=D9E6B6264D7723C3C12563CD002D6CE4 .
[49] Natalia Jevglevskaja, 'Weapons Review Obligation under Customary International Law' (2018) 94 International Law Studies 186, *available at*: https://digital-commons.usnwc.edu/cgi/viewcontent.cgi?article=1724&context=ils
[50] *Id.*
[51] ICRC, IHL database, customary IHL, accessed on 24 Sep. 2022, *available at*: https://ihl-databases.icrc.org/customary-ihl/eng/docs/v1_rul
[52] *Supra note* 43.
[53] *Supra note* 48.
[54] Chapter 5, Rule 17 of the customary IHL database of the ICRC, *available at:* https://ihl-databases.icrc.org/customary-ihl/eng/docs/v1_rul_rule17
[55] *Supra note* 48; Chapter 40, Rule 141 of the customary IHL database of the ICRC, *available at*: https://ihl-databases.icrc.org/customary-ihl/eng/docs/v1_rul_rule141 .

the obligation of precautions in the choice of means and methods of warfare; yet it should be noted that legal advisors lacking the technical expertise to assess the legality of certain weapons must be taken into consideration and accordingly cannot always substitute a thorough weapon review process. In conclusion, while not everyone would agree that article 36 is a customary IHL rule, one cannot deny its necessity to ensure compliance with other customary IHL obligations as illustrated.

Article 36 is primarily addressing state parties to the API, whereas non-state actors, such as armed groups, were not covered by the weapon review obligation. However, an absence of such obligation on non-state actors cannot be used as a ground to elude responsibility for deploying weapons that violate IHL customary rules.[56]

In determining types of weapons that should be subject to weapon review procedures, Article 36 included weapons, means, and methods of warfare. The terms means and methods of warfare were added to widen the scope of application of the article to include all types of weapons be they lethal or non-lethal, defensive or offensive. However, Germany for example made a reservation that it would only apply weapon review procedures to conventional weapons.[57] In this regard, the term means of warfare is an umbrella term that was added to include other systems that complement the usage of weapons and are not necessarily of destructive nature; such as navigation systems and surveillance systems that collect data contributing to the targeting process;[58] while the term methods of warfare serves the purpose of ensuring that the methods in which weapons and means of warfare are deployed respect IHL. Accordingly ensures that off-the-shelf warfare means and weapons are assessed with methods of their deployment in mind. In this regard, the commentary on API noted that:

> [T]he article is intended to require States to analyse whether the employment of a weapon for its normal or expected use would be prohibited under some or all circumstances. A State is not required to foresee or analyse all possible misuses of a weapon, for almost any weapon can be misused in ways that would be prohibited.[59]

[56] An analogy was made in this regard by reference to ICRC commentary on rule 141 concerning legal advisors for armed forces, see Chapter 40, Rule 141 of the customary IHL database of the ICRC, *available at:* https://ihl-databases.icrc.org/customary-ihl/eng/docs/v1_rul_rule141 .
[57] Vincent Boulanin, implementing Article 36 Weapon Reviews in the Light of Increasing Autonomy in Weapon Systems', States News Service,2015.
[58] *Id.*
[59] *Supra note* 43.

In defining the rules to be investigated to ensure respect of IHL during the weapon review process, Article 36 referred to rules protected by the Protocol itself and any other rule of international law applicable to the High Contracting Party. The wording of the article allowed for the inclusiveness of all envisioned obligations that result from treaty obligations or customary law obligations. Accordingly for a weapon to respect the weapon review process, the reviewing body should tick all the boxes of the checklist of prohibited weapons and means of warfare as prescribed by treaties which the state has joined as well as weapons prohibited by customary rules and general customary rules governing the deployment of weapons and methods of warfare. Throughout the history of IHL, international efforts succeeded to ban certain weapons through specific treaties, such as banning blinding laser weapons.[60] Ban on certain weapons has even gained the status of customary international law and accordingly, the ban on such weapons gained its authority from customary rules, such as the use of biological weapons,[61] and their usage is prohibited for all states whether they signed the prohibiting treaty or not. In addition to banning specific weapons by treaties or rules of custom, warring parties are obliged to respect certain customary rules and rules prescribed within the API itself when deploying any weapons and ensure that such weapons comply with these rules. These rules include:[62]

1. Prohibition of weapons and methods of warfare of a nature that would cause superfluous injury or unnecessary suffering (Art. 35(2) of API), Rule 70 of IHL customary rules database,
2. Prohibition of weapons and means of warfare which are intended, or may be expected to cause widespread, long-term, and severe damage to the natural environment (Articles 35(3) and 55 of API), Rule 44 of IHL customary rules database,
3. Prohibition of weapons, methods, or means of warfare that do not respect the principle of distinction between military and civilian objects (Art. 51(4)(b) & (c) and Art. 51(5)(a) of API), Rules 1, 7, 11, 71 of IHL customary rules database,
4. Prohibition of weapons, methods, or means of warfare that do not respect the proportionality principle and result in excessive incidental civilian loss or damages when

[60] *Supra note* 44 at 940.
[61] *Supra note* 44 at 941.
[62] Id at 943-945, see also *supra note* 51.

compared to the military advantage anticipated by the attack (Art. 51(5)(b) of API), Rules 14, 84 of IHL customary rules database

5. Prohibition of weapons, means, and methods of warfare that do not respect the Martens Clause with its two prongs principles of humanity and dictates of public conscience (Art.1 (2) of API).

The weapon review process is accordingly a comprehensive, multi-disciplinary one that should collect and consider data from different fields and has to be guided by expertise such as technical, legal, medical, environmental, and military experts. The weapon review process under article 36 accordingly requires a weapon to respect all treaty obligations on the state as well as all customary rules as explained above. A requirement that seems hard to achieve; but at the same time the usage of the word "in some or all circumstances" implies that if only one possible lawful usage or deployment of the weapon appears possible, such weapon should be deemed lawful under this article.[63]

Article 36 did not clearly specify the mechanism by which the weapon review process should be governed and left this task to each state to regulate it within its domestic procedures. Delegates of some states expressed during the deliberations of the API the need to establish by the Protocol itself a special body, a committee responsible for drawing up a list of weapons or methods of use that would fall under the prohibition. However, this proposal did not survive because of political and military considerations.[64] Furthermore, under article 36, it is not clear if the weapon review obligation is an obligation of conduct or an obligation of result. The conventional interpretation tends to consider the weapon review obligation an obligation of conduct.[65] The result of this interpretation is that states are under obligation to conduct the weapon review but if the results of deployment do not match the results of the review, the state is not considered violating its weapon review obligation. However, this gap of responsibility is solved by state responsibility obligation shall such deployment result in international law violations.[66] Another unsettled issue for conducting the weapon review process is to determine the phase in which the weapon is considered

[63] Dienelt, Anne. "The Shadowy Existence of the Weapons Review and Its Impact on Disarmament." Sicherheit Und Frieden (S+F) / Security and Peace 36, no. 3 (2018), *available at*: https://www.jstor.org/stable/26630036.
[64] *Supra note* 43.
[65] *Supra note* 63.
[66] Responsibility of States for Internationally Wrongful Acts, Yearbook of the International Law Commission, 2001, vol. II (Part Two), *available at*: https://legal.un.org/ilc/texts/instruments/english/draft_articles/9_6_2001.pdf

"new". There is no doubt that such a weapon review process applies to future weapons, but for already existing weapons, the decision to determine whether the weapon is new or not is left to the state conducting the review.[67] Although the majority of autonomous weapons are relatively novel for the majority of states' arsenals, the application might be quite problematic for certain types of weapons that have some autonomous features and existed before the adoption of weapon review policies within the domestic laws.

The unregulated weapon review mechanism opened the door for unanswered questions on how the process should be domestically undertaken and the consequent reality that a decision made by one country on the legality of a specific weapon is not subsequently binding internationally.[68] It is up to the state to determine how the process should be regulated whether through issuing regulating documents or by allocating a reviewing authority for this task.[69] Additionally, states are free to decide on how the process should be initiated; whether by request of the concerned authority probably the military or to consider the process a condition that should be fulfilled before developing or purchasing any weapons. The bindingness of the reviewing authority's decision on the legality of a weapon is also another element left to domestic procedures whether such decisions are issued in a form of recommendations or legally binding resolutions[70] and whether such decisions should be publicly published or classified.

B. Applying Article 36 on autonomous weapons

A focus on the legal review phase of autonomous weapons would require to assess this type of weapons against the legal rules as prescribed by Art.36. In the same order adopted by article 36, the legal review of autonomous weapons should start by looking into the rules prescribed by the additional protocol itself as illustrated earlier. The article then referred to "any other rule of international law applicable to the High Contracting Party"; which involves treaties and customary rules in the legal review process. To reiterate, the rules embedded within API are:

[67] *Supra note* 63.
[68] *Supra note* 43.
[69] See Natalia Jevglevskaja & Rain Liivoja, Weapons Review, last updated on 27 July 2022 and accessed on 25 September 2022, *available at:* https://www.premt.net/weapons-review/
[70] However, a recommendation concluding the illegality of certain weapons is arguably politically binding and development and acquisition of such weapons should be halted by the state, see: *Supra note* 44 at 954.

1. Prohibition of weapons and methods of warfare of a nature that would cause superfluous injury or unnecessary suffering (Art. 35(2) of API), Rule 70 of IHL customary rules database,
2. Prohibition of weapons and means of warfare which are intended, or may be expected to cause widespread, long-term, and severe damage to the natural environment (Articles 35(3) and 55 of API), Rule 44 of IHL customary rules database,
3. Prohibition of weapons, methods, or means of warfare that do not respect the principle of distinction between military and civilian objects (Art. 51(4)(b) & (c) and Art. 51(5)(a) of API), Rules 1, 7, 11, 71 of IHL customary rules database,
4. Prohibition of weapons, methods, or means of warfare that do not respect the proportionality principle and result in excessive incidental civilian loss or damages when compared to the military advantage anticipated by the attack (Art. 51(5)(b) of API), Rules 14, 84 of IHL customary rules database
5. Prohibition of weapons, means, and methods of warfare that do not respect the Martens Clause with its two prongs principles of humanity and dictates of public conscience

Whereas customary rules that apply to autonomous weapons are principles of distinction and proportionality and the Martens Clause with its two prongs humanity and dictates of public conscience.

For treaty obligations, it is important to note that there is no specific treaty governing the use of autonomous weapons and that the discussions to adopt such a treaty are still ongoing under the auspices of the CCW as illustrated above. Accordingly, for this legal review of autonomous weapons, only the rules embedded within the API and customary rules will be examined.

A careful consideration of the five rules embedded in the API and listed above shows that the first two rules, the prohibition of weapons that cause superfluous injury or unnecessary suffering as well as the prohibition of weapons that cause widespread, long-term and severe damage to the natural environment are of less relevance to the examination of the legality of autonomous weapons since autonomous weapons are seen as a platform for deploying means of warfare in the battlefield. Accordingly, if the kind of weapon deployed by autonomous weapons causes unnecessary suffering or superfluous injury or is expected to cause widespread, long-term, and severe damage to the environment, the weapon is accordingly illegal not because of the level of automation but because of the kind of weapon in usage. For example, if an autonomous weapon is

used to deploy blinding laser, the weapon is illegal because of the ban on the blinding laser which causes unnecessary suffering not because of the level of autonomy. The description of autonomous weapons as a platform for deploying weapons should focus the discussion on the capability of such autonomous systems to respect principles of distinction, proportionality principle, humanity principle, and dictates of public conscience (the two prongs of Martens Clause). For this reason, the legal review of autonomous weapons shall focus on the principle of distinction, proportionality principle, and Martens Clause being the relevant principles that would govern the legality of autonomous weapons.

a) Principle of Distinction

The principle of distinction is at the core of international humanitarian law and is one of the customary principles of the law that stipulates that: "The parties to the conflict must at all times distinguish between civilians and combatants. Attacks may only be directed against combatants. Attacks must not be directed against civilians."[71] The 4th Geneva Convention was adopted on basis of this principle to protect the rights of civilians in times of war. The authority of the principle of distinction as a positive law obligation is not contested in light of its customary law status and its repeated mention in various international treaties and conventions and the violation of this principle is prosecuted as a war crime.[72] While civilians are the preliminary beneficiaries of the principle of distinction, other categories also enjoy protection from being attacked such as hors de combat, individuals no longer participating in hostilities by choice (surrendering) or by circumstances.[73] The principle of distinction played a role in the history of weapon prohibition; it was on the basis of distinction that antipersonnel mines were prohibited for their violation of the principle of distinction.[74]

Theoretically, the principle seems straightforward aiming at protecting civilians from war atrocities and that they should not be treated as military targets. However, the application of the principle has posed practical challenges for a long time, in relation to the challenge of "*francs-tireur*s" and such challenges continue to amplify in our modern times due to the changing nature

[71] Chapter 1, Rule 1 of customary IHL database of the ICRC, *available at:* https://ihl-databases.icrc.org/customary-ihl/eng/docs/v1_cha_chapter1_rule1
[72] 1. Anicée Van Engeland, Civilian or Combatant?: A Challenge for the 21st Century (2011), https://go.exlibris.link/P6gzd0XW.
[73] Volume II, Chapter 15, Section B, Rule 47 of customary IHL database of the ICRC, *available at:* https://ihl-databases.icrc.org/customary-ihl/eng/docs/v1_rul_rule47
[74] *Supra note* 72 at 27.

of conflicts where the distinction between civilians and military targeted becomes blurrier. The aim of discussing the principle of distinction in reference to autonomous weapons is not to find answers to the contentious questions of applying the principle of distinction but rather to shed light on the challenges of practical applications of this principle by humans themselves; which stands as one of the main arguments made by proponents of an absolute ban on autonomous weapons and to inquire whether advocates of autonomous weapons can provide technical solutions to overcome such challenges.

The challenges attributed to the application of the principle of distinction mainly result from the negative definition of civilians as being non-combatants and from the modern warfare reality that battlefields witness a muddle of military and civilian presence due to the commonplace of fighting in densely populated urban zones. The fact that civilians are defined as being non-combatants serves the goal to ensure that no categories are excluded from any definition that would limit the protection to those who are ought to be considered civilians.[75] Although the definition of combatants offers more clarity by listing the categories which are considered combatants, such definition does not exclude the existence of grey zones where the decision of whether the person enjoys civilian protection or not is very difficult and sometimes not possible to be taken. Real-life scenarios where the distinction is diluted existed and continue to evolve. The case of civilians directly participating in hostilities is one of the clearest examples of how the distinction is blurred. Civilians carrying arms and directly participating in battlefields lose their immunity against being military targets, however, the complexity rests in defining the forms of direct participation in hostilities. Debates regarding the civilians who work at military factories, and civilians who work at electricity companies that provide electricity to both civilians and military institutes, are all examples of how the distinction becomes a critical decision to make. The ICRC in this regard elaborated on criteria for direct participation in hostilities as follows:

> 1. The act must be likely to adversely affect the military operations or military capacity of a party to an armed conflict or, alternatively, to inflict death, injury, or destruction on persons or objects protected against direct attack (threshold of harm), and
>
> 2. there must be a direct causal link between the act and the harm likely to result either from that act, or from a coordinated military operation of which that act constitutes an integral part (direct causation), and

[75] *Supra note* 72 at 30.

3. the act must be specifically designed to directly cause the required threshold of harm in support of a party to the conflict and to the detriment of another (belligerent nexus).[76]

Perhaps a more complex scenario is the case of human shields, where the military takes civilians as shields to protect military targets from being attacked, benefiting from the immunity provided to civilians against attacks. In such scenarios, the intent of the civilians whether they are willingly acting as human shields or if they are coerced to do so, in addition to the difficulty of standing on such intent, along with the protection offered to innocent civilians from being attacked, are all factors that contribute to the complexity of the situation.[77]

In this context, the distinction principle is not absolute, and the balance between military necessity considerations and the proportionality principle is introduced to weigh between civilian protection from one side and military necessity on the other side. The following part will be dedicated to explaining proportionality as a legal principle and its practical applications.

b) Proportionality Principle

Embedded within rules of customary rules of international humanitarian law by ICRC, the proportionality principle stipulates that: "Launching an attack 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, is prohibited."[78] The ICRC confirmed and provided evidence for the applicability of the principle in both international and non-international armed conflicts. The principle also appeared in Additional Protocol I to the Geneva Convention in article 51(5)(b) which describes an indiscriminate attack as "an attack 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."[79] This principle primarily aims at protecting civilians against targeting, yet at the same time it opened the doors for an exception on the immunity of civilians from military attacks. Put differently, the principle allows attacks on civilian populations and properties whenever such attacks do not result in excessive

[76]Nils Melzer, Interpretive Guidance On The Notion Of Direct Participation In Hostilities Under International Humanitarian Law, ICRC, May 2009.
[77] *Supra note* 72 at 110.
[78] Chapter 4, Rule 14 of the customary IHL database of the ICRC, *available at*: https://ihl-databases.icrc.org/customary-ihl/eng/docs/v1_cha_chapter4_rule14
[79] *Supra note* 48.

damages to civilian populations compared to military advantages sought by such attacks. Assessing the proportionality of attacks is a forward-looking process in which an average reasonable commander is expected to act in good faith by doing everything feasible to assess the information available at the time of the attack to ensure that no excessive collateral damage is inflicted on civilians while pursuing military advantage. In battlefields, black-and-white scenarios when it is possible to easily determine in favor or against the proportionality of the attack exist, however, grey zone scenarios are more common. The ambiguity of the principle and difficulties in the application are products of the elastic terms used in defining the principle, such as excessive, military advantage, attack, and expected.[80] The way in which the principle is phrased allows for discretionary power when assuming good faith applications, since a more restrictive rule would jeopardize military compliance by the proportionality principle. Such elastic terms however come with the cost of indeterminate interpretations of what constitutes excessive damage, the threshold of excessiveness, the importance of military advantage sought by the attack and how it can be weighed against the value of civilian damages, the nature of the attack and whether proportionality applies to all types of attacks including non-military attacks such as cyber-attacks. In an article on proportionality, Clarke proposed a need for a guidance document to clarify ambiguities related to the application of the proportionality principle. In the article he enumerated number of issues that should be clarified such as: assessing proportionality in cases when the protection of the attacking forces is at stake, how proportionality should be assessed in cases when the impacts of the attacks are of prolonged reverberating nature, assessing the value of military targets and military advantage during the attack, the legality of repeated attacks failing within the grey zone of proportionality assessment.[81] It is further noted that the implications of the forward-looking assessment for proportionality at the time of deciding on the attack add to the complexity of respecting the principle of proportionality. Commanders are obliged to fully assess the situation on basis of the information available at the time of making the decision. In cases when the damages of the attacks are excessive and could not be predicted by an average reasonable commander, no liability for violating the principle of proportionality can be pursued. In battlefields, when time is crucial to the decision-making process and when information available to commanders is sometimes misguided, the defense is easily made by commanders that they respected the

[80] Ben Clarke, *Proportionality in Armed Conflicts: A Principle in Need of Clarification?*, 3 Journal of international humanitarian legal studies 73–123 (2012).
[81] *Id.*

proportionality assessment but the outcomes of the attack could not be reasonably predicted.[82] Assessing proportionality is not a clear-cut process and its loosely defined terminologies allow for contesting arguments that are reflective of the arguments between those who view military necessity as the goal worthy of utmost respect on battlefields and those who favor humanitarian considerations over any other considerations.

c) Principle of distinction and principle of proportionality and autonomous weapons

Advocates of the ban on autonomous weapons build their arguments against autonomous weapons principally on the challenges related to the application of principles of distinction and proportionality in real-life scenarios. They acknowledge that the obligation to respect principles of distinction and proportionality is an obligation that must be respected by the human users, not the machine itself. Decisions on distinction and proportionality require qualitative and evaluative judgments by humans that cannot be translated to numerical values pre-programmed into the software of the machine.[83] Sharkey, one of the advocates of the ban on autonomous weapons, has expressed the same doubts that machines would be able to determine on critical matters such as distinction during military attacks. He emphasized the fact that the definition of civilians is not clear enough for humans to be programmed into machines. Civilians being defined as non-combatants would be a complex concept to be determined by machines.[84] The process is further complicated when deciding on other protected categories such as hors de combat and civilians actively participating in hostilities. In his paper on the compatibility of autonomous weapons with the principle of distinction, Winter divided the distinction process into three phases: observation, recognition, and judgment. Winter argued that while machines could outperform humans in the observation and recognition phases, the judgment phase remains tricky for machines to perform owing to its nature as context-based. He illustrated his point of view by giving the example of soldiers raising their hands to surrender.[85] The same challenges, Scharre recognized from his experience on the battlefield in Afghanistan and Iraq when the distinction between villagers and Taliban fighters was a challenging task that was determined by the behavior of the individual. He

[82] Lippman, Matthew (2002) *"Aerial Attacks on Civilians and the Humanitarian Law of War: Technology and Terror from World War I to Afghanistan*," California Western International Law Journal: Vol. 33 : No. 1 , Article 2.
Available at: https://scholarlycommons.law.cwsl.edu/cwilj/vol33/iss1/2
[83] *Supra note* 30.
[84] Noel Sharkey, Why robots should not be delegated with the decision to kill, 29 Connection Science 177–186 (2017).
[85] Winter, E. (2020). the compatibility of autonomous weapons with the principle of distinction in the law of armed conflict. The International and Comparative Law Quarterly, 69(4), at 845-876. https://doi.org/10.1017/S0020589320000378.

also narrated a scenario from his experience in Iraq of a chaotic battlefield where the distinction between friendly troops and enemy insurgents was very hard.[86]

Additionally, for advocates of a ban on autonomous weapons, one of their main arguments related to distinction and proportionality is that actions of highly sophisticated autonomous weapon systems cannot be predicted. Accordingly, leading to an accountability gap in incidents when distinction and proportionality principles are violated. HRW concluded that war crimes cannot be attributed to machines due to the lack of mental state and intent for machines that is necessary to establish *mens rea* element of the crime. Similarly, indirect criminal responsibility of commanders or human operators launching the machines cannot be established when human operators are unable or unexpected to know of the commission of the crime.[87] The unpredictability of machine behavior and lack of effective control over them by human commanders results in an accountability gap since commanders cannot prevent the commission of the crime.

On the other hand, opponents of the absolute ban on autonomous weapons, respond to challenges related to respecting distinction and proportionality. They argue that autonomous weapons are not indiscriminate weapons by nature[88] to be deemed illegal in all circumstances, an example of an indiscriminate weapon by nature is the Balloon Bombs used by Japan in WWII.[89]

The decision on the legality of autonomous weapons from the perspective of ban opponents depends on the circumstances in which such weapons are used. For example, Air-to-air combat over an open ocean and with pressures of an arms race to increase the speed of aircraft resulting in torque and inertial pressures that cannot be endured by humans might be one day operated by autonomous systems and equipped by autonomous weapons to cope with the rest of the autonomous system can be envisioned as a battlefield scenario where autonomous systems can operate lawfully without challenges of distinction and proportionality due to presence of civilians in the field.

Furthermore, some of the opponents of the ban on fully autonomous weapons described that the decision to conclude that autonomous weapons cannot evolve to mimic human-like cognitive

---

[86] *Supra note* 14.
[87] Human Rights Watch & The International Human Rights Clinic, Mind the Gap The Lack of Accountability for Killer Robots, 2015.
[88] Kenneth Anderson, Daniel Reisner and Matthew Waxman, Adapting the Law of Armed Conflict to Autonomous Weapon Systems, 90 INT'L L. STUD. 386 (2014), *available at*: https://digital-commons.usnwc.edu/cgi/viewcontent.cgi?article=1015&context=ils
[89] For more on Balloon Bombs, see: Coen, R. A. (2014). Fu-go: The curious history of Japan's balloon bomb attack on America. University of Nebraska Press. https://doi.org/10.2307/j.ctt1d9nmg0

abilities to make better judgments on distinction and proportionality is a pessimistic view of the development of autonomous weapons since it overlooks the equally valid possibility that machines will be able to perform such tasks in the future.[90] Krishnan[91] for example argued that autonomous weapons if highly developed could be more successful at respecting distinction during targeting in warfare. He gave an example of future sophisticated technologies where military targets could be defined and attacked by machines on an individual basis instead of targeting the whole area in which the target is believed to be. He further argued that proportionality is based on calculations and that computerized weapons would be able to calculate faster and precisely blast and reduce weapons effects that result in collateral damage; calculations that would be complex for humans to make in real-time.[92] To overcome challenges of distinction and proportionality, opponents of the ban even proposed the philosophy of developing autonomous weapons that only aim at disarming soldiers without harming them, for example, weapons that would be programmed to scan for weapon profiles and neutralize them.[93]

Apart from the optimistic claims on the future capabilities of autonomous weapons, opponents of the ban view proportionality assessment as a holistic process in which the assessment is an ongoing process throughout the targeting process; which usually takes place in the pre-planning phase and depends on information gathered beforehand. From their perspective, those directly participating in attacks might have very low situational awareness whether they are humans or machines as they did not take part in the pre-planning phase.[94]

To fill the gaps and to respond to challenges posed by advocates of an absolute ban, the United States in one of the proposals presented during the meetings of CCW on autonomous weapons suggested some safeguards as good practices when dealing with autonomous weapons.[95] Examples of these safeguards are incorporating understandable human-machine interfaces and controls to enhance the predictability of the systems, retaining human control over the system and that autonomous weapons should not be designed to perform attacks that would exclude the responsibility of the human operator of the system to fill the responsibility gap. Additionally, the

[90] Marcus Hickleton, Shaky Foundations: Killer Robots and the Martens Clause, 4 PERTHILJ 31 (2019).
[91] Armin Krishnan, Killer robots: legality and ethicality of autonomous weapons (2016).
[92] *Id.*
[93] *Supra note* 14.
[94] *Supra note* 88.
[95] *Supra note* 25.

proposal stressed the importance of incorporating measures to mitigate risks of unintended attacks, such as measures to reduce automation bias, measures of self-destruct, self-deactivation, and self-neutralizations mechanisms to combat system spoofing. Additionally, the DoDD required anti-personnel autonomous weapons to undergo a senior-level review process.[96] The initiatives to propose such safeguards and to set such principles limiting the use of autonomous weapons are all efforts to prove the fallacy of the call for an absolute ban on autonomous weapons.

A synthesis of the conversations between advocates of the absolute ban of autonomous weapons and opponents of such a ban shows that both sides of the conversation are taking place on different premises. The driving motive of advocates of the absolute ban is the humanitarian considerations that are worthy of being protected through aborting the development of autonomous weapons even if such developments hold the promises of advancing autonomous weapon systems that are capable of outperforming humans in distinction and proportionality calculations which would arguably reduce human suffering. On the other hand, opponents of the absolute ban are not willing to compromise military advancement aspirations of developing autonomous weapons that would eventually result in reducing suffering on the battlefield, at least for their troops. The conflicting perceptions of competing values, humanity considerations, and military necessity, on both sides of the conversations have impacted their glossary of defining terms such as human control over the machine, and their acceptance of the scope of the attack which starts when the autonomous weapon is activated;[97] whereas opponents of the ban embrace the position that an attack is a holistic process of targeting.

d) Autonomous weapons and Martens Clause

In this section, the debate on the legality of autonomous weapons under the Martens Clause is presented. With the morality complications of deploying autonomous weapons being the undertone of the voices calling for the ban of autonomous weapons,[98] the two prongs of the Martens Clause, humanity and dictates of public conscience should be investigated. Considering the absence of a clear and specific body of law governing autonomous weapons, advocates of the ban invoke the applicability of the Martens Clause on the case of autonomous weapons. The

[96] See Annex 1.
[97] *Supra note* 30.
[98] The Campaign to Stop Killer Robots identified delegating life and death decisions to be crossing moral threshold on the top of the problems of using autonomous weapons. see: https://www.stopkillerrobots.org/learn/

Martens Clause in its version adopted in the preamble to the 1899 Hague Convention II reads as follows:

> Until a more complete code of the laws of war is issued, the High Contracting Parties think it right to declare that in cases not included in the Regulations adopted by them, populations and belligerents remain under the protection and empire of the principles of international law, as they result from the usages established between civilized nations, from the laws of humanity, and the requirements of the public conscience.[99]

The Clause since its adoption has resulted in a wide debate over its interpretation and scope of applications. Accordingly, before delving into the debates that are founded on Martens Clause to ban autonomous weapons, it is vital first to understand the historical context of the Martens Clause, and its development since it was introduced in 1899 which in turn impacted the debate over its interpretation and the scope of application.

i. The Historical Background of the Martens Clause

Deeply rooted in the norms of natural law,[100] Martens Clause was the first codified form of the clause adopted in a widely recognized international instrument as the Hague Convention. The trigger of embracing the clause at that time by its godfather, Fyodor Fyodorovich Martens, was the conflict between the Russian delegation, voicing concerns by powerful states, and the Belgium delegation, representing the smaller states. The conflict emerged at that time around the status of "*francs-tireur*s", a similar concept to civilians directly participating in hostilities in our modern days, and the protection that they should enjoy under laws of war if any. Russia submitted that they should be deprived of protection and when caught they deserve severe punishments amounting to death as they do not commit to duties of belligerents as they do not have responsible command hierarchy, they do not have fixed emblems and they do not carry weapons openly. On the other hand, Belgium wanted to preserve the rights of those fighting against occupation to pursue patriotic endeavors of carrying weapons against their occupiers.[101] Fearing threats of failure

[99] Convention (II) with Respect to the Laws and Customs of War on Land and its annex: Regulations concerning the Laws and Customs of War on Land. The Hague, 29 July 1899, *available at:* https://ihl-databases.icrc.org/applic/ihl/ihl.nsf/Article.xsp?action=openDocument&documentId=9FE084CDAC63D10FC12563CD00515C4D

[100] Theodor Meron, *The Martens Clause, Principles of Humanity, and Dictates of Public Conscience*, 94 The American journal of international law 78–89 (2000), establishing that the roots of Martens Clause existed as early as 1643 in natural law and chivalry norms.

[101] Jeffrey Kahn, *"Protection and empire": the Martens clause, state sovereignty, and individual rights*, 56 Virginia journal of international law 1 (2016).

of the Hague Conference owing to this contention, Martens suggested the Clause in its form that we know today and as appeared in the preamble of the Convention as stated above.

Since that time, the clause, and sometimes shortened versions of it, continued to appear in various international legal instruments such as the 1907 Hague Convention, the 1949 four Geneva Conventions, and the two additional Protocols of 1977. However, the slight variations of the wordings of the Clause and its positions in these instruments have further fed into the debate on the legal value of the Clause and its interpretations. It is however not contested that the Martens Clause gained the status of customary international law. This has been asserted by the ICJ in its Advisory Opinion on the Legality of The Threat or Use of Nuclear Weapons that the applicability of the Martens Clause is not to be doubted and the Court also listed the Clause as one of the examples of Customary Law.[102] Although the applicability of the Clause and its relevance are not contested in cases when there is no specific treaty or customary law governing the matter, its interpretation and its legal scope and value constitute the foundations of the debates between natural law adherents and positivists. Accordingly, it is important to expose these debates in greater detail before presenting the foundations of the arguments for banning autonomous weapons based on the Martens Clause.

ii. The legal scope and value of Martens Clause

The aim of determining the legal scope and the value of the Martens Clause is to stand on its bindingness vis-à-vis other legal principles. The contentious status of the Martens Clause in international law emanates from its conflict with the Lotus Principle. This debate is an extension of the classical naturalists-positivists debate within international law on the sources governing international law; the ongoing inquiry on whether international law should be primarily based on the sovereignty of the states being the actors of the international community or on the supremacy of absolute norms that are worthy of being protected by the international community. In other words, should international law be shaped to protect the interests of the states and principles for which the states have given consent to protect, or should it work to protect morals and absolute norms that are of concern for the international community irrespective of states' interests and their will? The Lotus Principle underpinned the argument that states sovereignty should be the driving power of the creation of international law and that more value should be given to what states have

[102] Legality of the Threat or Use of Nuclear Weapons, Advisory Opinion, 1. C.J. Reports 1996, at 226.

agreed to bind themselves with, i.e. treaties and conventions. According to the Lotus Principle, whenever a particular act is not explicitly prohibited by international law, states are free to act in any way as long as their actions are not explicitly prohibited by rules.

The circumstances in which the Clause came to life, as a way to untangle the conflict between states during the drafting of the Hauge Convention, have impacted the interpretation of the Clause. On basis of the historical context of the adoption of the Clause, some argued that the value of the Clause is only to act as a reminder for states of their obligations under customary international law.[103] As Cassase described this argument as a *contrario* to reject the presupposition that what is not prohibited by law is not necessarily lawful.[104] Supporters of this view built their argument on legal certainty positivistic approach. Additionally, the fact that the Clause appeared first in the preamble of the Convention contested its quality as a legal rule.[105] This position was supported by Russia in its written submission to the ICJ in the case of the legality of the use of nuclear weapons. Russia argued that the Clause was inserted as a temporary measure until a more complete code of law of war is issued and by the adoption of the Geneva Conventions, the Clause is no longer applicable to the case of nuclear weapons.[106] However, this argument is criticized on basis of the role that the Clause played in practice after it was adopted. The Clause was mentioned in numerous international legal instruments, and it was later mentioned in the main bodies of the treaties such as in Article 1(2) of Additional Protocol I. Additionally, The Nuremberg Trials built its prosecution of war crimes on Natural Law rules to qualify certain acts as war crimes despite the absence of codified rules criminalizing them at the time of their commission.[107] The development of international humanitarian law since the first appearance of the Clause in 1899 until now suggests that even if the Clause was adopted to solve the Hague Conference from failure by solving the conflict at that time, the Clause has developed since then to serve a higher role than merely acting as a reminder for the states to abide by their obligations under customary law.[108]

---

[103] M. Salter, *Reinterpreting Competing Interpretations of the Scope and Potential of the Martens Clause*, 17 Journal of conflict & security law 403–437 (2012).
[104] A. Cassese, *The Martens Clause: half a loaf or simply pie in the sky?*, 11 European journal of international law 187–216 (2000).
[105] *Supra note* 103.
[106] Written submissions by Russia to the ICJ in the case of the Legality of Use of Nuclear weapons, mentioned in letter dated 19 June 1995 from the Ambassador of the Russian Federation to the Court.
[107] Rupert Ticehurst, The Martens Clause and the Laws of Armed Conflict, International Review of the ICRC no. 317, 1997, *available at*: https://www.icrc.org/en/doc/resources/documents/article/other/57jnhy.htm
[108] *Supra note* 103.

The second approach to determine the normative value of the Clause argues that it should serve as a tool of judicial aide when interpreting ambiguous rules that have conflicting applications. In such cases, applications that are more consistent with humanity considerations should elevate over others. One of the vocal supporters of this interpretation of the value of the Clause is Antonio Cassese and he also attributed to the Clause the role of lowering the threshold of state practice for a rule to become a customary rule, which serves to fill the gap between emerging legal issues and lagging state practices.[109] Cassase also in the decision of *Kupreskic*, in which he acted as the presiding judge, referred to human rights rules as a guideline to interpret humanity considerations. The decision also excluded that the components of the Martens Clause, principles of humanity and dictates of public conscience, have elevated to the rank of independent sources of international law. In light of interpreting rules consistently with principles of humanity, the Court concluded that in case of repeated military conducts falling within the grey zone of unlawfulness:

> [I]t might be warranted to conclude that the cumulative effect of such acts entails that they may not be in keeping with international law. Indeed, this pattern of military conduct may turn out to jeopardise excessively the lives and assets of civilians, contrary to the demands of humanity.[110]

The interpretative aide approach opened the door for limiting apologetic attitudes that consider state practices as the compass for international law-making by empowering judges to shield their judgments favoring humanity considerations on basis of the Martens Clause. Judges are capable of building their judgments, which would otherwise be deemed political, on basis of legal grounds.[111] However, the interpretative aide approach did not justify the fact that the Martens Clause itself has become a customary rule, yet there are no guiding principles as to when judges are obliged to resort to it or favor it over other international law principles. Additionally, the complexity of the interpretative aide approach in the context of armed conflicts is that the essence of international humanitarian law is that it attempts to balance humanity including the protection of civilians and reducing suffering during wars and military necessity. Accordingly, the aide approach does not necessarily offer a clear guiding principle for the application of the Martens Clause.[112] In other words, in times of war, international human rights principles, *lex generalis*

[109] *Supra note* 104.
[110] Prosecutor v Kupresˇkic´ et al (Judgment) ICTY 95-16-T paragraph 526.
[111] *Supra note* 103.
[112] *Id.*

rules, are derogated and the application of international humanitarian law, *lex specialis* rules, override humanity considerations applicable in peacetime.

The third and the most radical approach to interpret the Martens Clause invokes the components of the Clause, principles of humanity, and dictates of public consciences, as peremptory sources of international law. Support for this approach finds its basis in precedents of international criminal court jurisprudence. Salter is of the view that this approach should be adopted when assessing the value of the Clause as it supports the humanitarian disposition of international humanitarian law which has been affirmed in the wording of the Clause when it was created, regardless of the intention of its drafters at the time of its adoption.[113] Salter argued that the development of the Clause since it was created supports the view that the Clause in itself stands as a separate source of law. The basis of his argument is the repetitive court decisions that relied on the Clause to criminalize acts that were not criminalized at the time of their commission, contrary to the principle of *"nullum crimen sine lege"*. The Nuremberg trials were an evident example of embracing the Clause to ensure that perpetrators of atrocities against civilians would not escape prosecution. Salter further viewed that both the interpretative aide approach and the norm approach should co-exist in a manner that the interpretative aide approach should expand. In cases when the interpretative aide approach was invoked the ultimate outcome was the recognition of principles of humanity and the dictate of public conscience were adopted as the sources of obligation. [114] The Clause is an important tool to provide a positivist ground for adopting normative values as humanity and dictates of public conscience when state practices are lagging behind the development of methods of warfare in armed conflicts. Such function of the Clause acts as a proactive safeguard to prevent human sufferings which result from legal loopholes.[115] However, noting the ambiguity of the Clause and its two prongs, the principle of humanity and dictates of public conscience, this approach is criticized by positivists for not following the process of imposing obligations on states when state practices are scarce or not consistent.[116]

The indeterminacy of concluding on the value of the Martens Clause vis-à-vis other legal principles or sources has also impacted the judicial intercourse. The ICJ advisory opinion on the legality of the threat or use of nuclear weapons provides a clear example of the split of opinions

[113] *Id.*
[114] *Id.*
[115] Mitchell Stapleton-Coory, The Enduring Legacy of the Martens Clause: Resolving the Conflict of Morality in International Humanitarian Law, 40 ADEL. L. REV. 471 (2019).
[116] *Id.*

regarding the status of the Clause. The advisory opinion itself briefly asserted that the Martens Clause proves to be an "effective means of addressing the rapid evolution of military technology", the Court also listed the Clause as a pre-existing customary law.[117] The Court however did not offer a clear analysis of the value that the Clause owns when in conflict with other international law principles. It is yet implied from the decision that humanity principles and dictates of public conscience would not always supersede when competing with other principles preserving state sovereignty such as self-defense by the non-liquet decision of the Court to conclude on the legality of the use of nuclear weapons in extreme cases of self-defense when the survival of the state is at stake. However, judge Shahabuddeen in his dissenting opinion offered a detailed analysis of the Martens Clause. He articulated his opinion by rejecting both the restrictive and interpretative aide approaches to stand on the rank of the Clause. Citing his decision, he stated:

> I am not persuaded that the purpose of the Martens Clause was confined to supplying a humanitarian standard by which to interpret separately existing rules of conventional or customary international law on the subject of the conduct of hostilities… It is also difficult to accept that all that the Martens Clause did was to remind States of their obligations under separately existing rules of customary international law.[118]

He supported this opinion by the change in the position of the delegate of Belgium after the Clause was adopted which confirms that the Clause was introduced to put obligations on states to act according to principles of humanity and dictates of public conscience. Judge Shahabuddeen asserted that:

> [T]he Martens Clause provided its own self-sufficient and conclusive authority for the proposition that there were already in existence principles of international law under which considerations of humanity could themselves exert legal force to govern military conduct in cases in which no relevant rule was provided by conventional law. Accordingly, it was not necessary to locate elsewhere the independent existence of such principles of international law; the source of the principles lay in the Clause itself.[119]

The same position taken by judge Shahabuddeen was also supported by the ICJ in the *Corfu Channel Case* when declaring that the obligations on Albania were based "on certain general and well-recognized principles, namely: elementary considerations of humanity."[120]

---

[117] *Supra note* 102.
[118] Dissenting Opinion Of Judge Shahabuddeen, Legality of the Threat or Use of Nuclear Weapons, Advisory Opinion, 1. C.J. Reports 1996, English version, at 408, *available at*: https://www.icj-cij.org/public/files/case-related/95/095-19960708-ADV-01-11-EN.pdf
[119] *Id.*
[120] "Corfu Channel case, Judgment of April 9th, 1949 : I.C. J. Reports 1949, at 4.

Within the international criminal judiciary jurisprudence, the contention between the interpretative aide approach and the peremptory source approach continues. As illustrated above, the Nuremberg trials relied on rules of natural law to prosecute figures of the Nazi regime; the Court stated that if "'faced with two interpretations—one in keeping with the principles of humanity and moral standards, and one which is against these principles—then we should, of course, give priority to the former interpretation."[121] The Court, in this case, found the rules of humanity and moral standards as separate legal principles that would prevail over other rules that are not in conformity with these principles. However, on the other hand, the ICTY in the *Kupreskic* Case the Court stated that:

> …this Clause enjoins, as a minimum, reference to those principles and dictates any time a rule of international humanitarian law is not sufficiently rigorous or precise: in those instances the scope and purport of the rule must be defined with reference to those principles and dictates.[122]

The development of the Martens Clause since its first appearance could perhaps tell that the restrictive approach is no longer a popular interpretation of the value of the Clause. However, disagreement still exists on whether the Clause is of an interpretative aide value or its components stand as separate sources of obligations in international law. This shift is a result of the evolving status of individuals from being objects of international law to subjects of international law. In our modern world, where the focus on human rights law elevated humanity considerations to a higher status compared to the time when the Clause was adopted, state sovereignty is no longer absolute.[123] However, answers to questions such as when the Clause should be invoked, the anticipated outcomes of invoking the Clause, and the status that the Clause endures when conflicting with competing other international law principles remain indeterminate. While keeping this in mind, it is perhaps relevant to shift the discussion to the interpretation of the components of the two prongs of the Clause, the principle of humanity, and dictates of public conscience in the hope to stand on higher clarity of the Clause.

[121] *Supra note* 103.
[122] *Supra note* 110 at para. 525.
[123] *Supra note* 101.

iii. Principle of Humanity as a prong of the Martens Clause

It is perhaps not a coincidence that one of the alternative names of laws of armed conflict is international humanitarian law; which aims to attribute humanitarianism to the conduct of warfare in the hope of limiting its brutality.

Humanity principles were listed by the Martens Clause as one of the limiting principles of conduct of belligerents in warfare. Noting the legal value of the Martens Clause as demonstrated above, as at least acting as an interpretative aide for rules of armed conflict, it is important to seek clarity of the definition of humanity principles.

An investigation of the literature shows that there are several attempts to define this principle. However, such definitions do not negate the ambiguity character attributed to the term itself.[124] The attempts to define humanity are often made by defining conducts that are consistent or inconsistent with principles of humanity or by defining other principles which would characterize the conduct as humane when respected, such as proportionality, reduction of unnecessary human suffering, distinction, protection of civilians and lawful collateral damage to civilians. In all cases, the term humanity remains relative and changes depending on different factors, including the surrounding circumstances, time, and those targeted by an attack. In 1863, the Lieber Code which was driven by humanitarian considerations at its time, some of its provisions are lacking the humanity character in our modern time.[125] The inseparable ambiguity of the term humanity has an important impact on constructing an argument of legality or illegality on basis of the Martens Clause. Since the term remains vague and hard to define, it is should be seen as an inspiration for lawmakers rather than a principle that stands on its own and results in legal impacts.[126] For this reason, Meron argued that when contesting the legality of weapons, arguments should be made on more concrete principles such as proportionality, distinction, and unnecessary suffering.

> Nevertheless, the Martens clause does not allow one to build castles of sand. Except in extreme cases, its references to principles of humanity and dictates of public conscience cannot, alone, delegitimize weapons and methods of war, especially in contested cases… Additional prohibitions of particularly objectionable weapons and methods of war can better be attained by applying such generally accepted principles of humanitarian law as

[124] Robin Coupland, Humanity: What is it and how does it influence international law?, ICRC, December, Vol. 83 No 844, *available at*: https://www.icrc.org/en/doc/assets/files/other/irrc-844-coupland.pdf; see also, *infra note* 129.
[125] *Infra note* 129, Article 17 of the Code stipulated that: "It is lawful to starve the hostile belligerent, armed or unarmed, so that it leads to the speedier subjection of the enemy." Instructions for the Government of Armies of the United States in the Field (Lieber Code). 24 April 1863.
[126] *Supra note* 1.

> the requirements of distinction and proportionality and the prohibition of unnecessary suffering than by pushing the Martens clause beyond reasonable limits.[127]

Such conclusions however do not attempt to claim that humanity considerations do not have an impact on the process of weapon review in our modern time. The rising interest in embedding humanity considerations in international humanitarian law resulted in a new system of humanitarian disarmament. The following paragraphs will aim to present the attempts of defining humanity and to demonstrate the impacts of humanity on the disarmament efforts through the scheme of humanitarian disarmament.

The ICRC interprets the principle of humanity that it "forbids the infliction of all suffering, injury or destruction not necessary for achieving the legitimate purpose of a conflict."[128] Such a definition adopts the definition of humanity by referring to the principle of unnecessary suffering that does not achieve military necessity. Within this definition, the proportionality principle is invoked to balance military gains and damages inflicted during the attack. Following the same approach, the UK ministry of defense defines humanity as:

> [B]ased on the notion that once a military purpose has been achieved, the further infliction of suffering is unnecessary. Thus, if an enemy combatant has been put out of action by being wounded or captured, there is no military purpose to be achieved by continuing to attack him. For the same reason, the principle of humanity confirms the basic immunity of civilian populations and civilian objects from attack because civilians and civilian objects make no contribution to military action.[129]

This approach was also further confirmed by the UN Special Rapporteur reporting on the human rights situation in Kuwait under the Iraqi occupation. The report identified the principle of humanity embedded in the Martens Clause as encompassing three customary principles of human rights protection:

> (i) that the right of parties to choose the means and methods of warfare, i.e. the right of the parties to a conflict to adopt means of injuring the enemy, is not unlimited; (ii) that a distinction must be made between persons participating in military operations and those belonging to the civilian population to the effect that the latter be spared as much as

---

[127] *Supra note* 100.

[128] ICRC, What is IHL?, 18 September 2015, *available at*: https://www.icrc.org/en/document/what-ihl

[129]Kjetil Mujezinović Larsen, Camilla Guldahl Cooper & Gro Nystuen, Introduction by the editors:: Is there a 'principle of humanity' in international humanitarian law?, in Searching for a 'Principle of Humanity' in International Humanitarian Law 1–20 (Kjetil Mujezinović Larsen, Camilla Guldahl Cooper, & Gro Nystuen eds., 2012) citing the UK Ministry of Defence, the Manual of the Law of Armed Conflict (Oxford University Press, 2004) 23.

possible; and (iii) that it is prohibited to launch attacks against the civilian population as such.[130]

The other approach to defining the principles of humanity as stated above is by listing the conduct which would be consistent or inconsistent with the principles of humanity. Adopting this approach the ICRC defined what would be perceived as humane conduct on battlefields as follows:

> The principle of humanity enjoins that capture is to be preferred to wounding, and wounding to killing; that the wounding should be effectuated in the least serious manner – so that the wounded person may be treated and may recover – and in the least painful manner; that the captivity should be as bearable as possible.[131]

The same approach was also adopted by Common Article 3 in defining what would be a humane treatment in non-international armed conflicts. The Article provided a list of prohibited conducts against protected categories such as violence to life and persons such as murder, mutilation, torture, cruel treatment, taking of hostages, outrages upon personal dignity, in particular humiliating and degrading treatment, and extra-judicial killing.[132]

The opinion to define principles of humanity by reference to other principles of proportionality, distinction, prohibition of unnecessary suffering, and attacks against civilians, was the inspiration to codify these principles under the four Geneva Conventions as a translation of humanity principles. In earlier pages of this paper, details on obligations to respect principles of distinction and proportionality and how they impact the debate on the legality of autonomous weapons were illustrated. Accordingly, this section will focus on the impacts of humanity consideration on disarmament efforts as well as the moral arguments related to banning autonomous weapons on basis of the humanity principle will be presented.

iv. Impact of humanity considerations on disarmament efforts (Humanitarian Disarmament)

With a progressing humanitarian-centered approach toward regulating laws of war, new concepts have emerged that are inspired by humanitarianism; one of these concepts is humanitarian disarmament. Humanitarian disarmament is a departure from traditional disarmament where

[130] Report of the situation of human rights in Kuwait under Iraqi occupation / prepared by Walter Kälin, Special Rapporteur of the Commission on Human Rights, in accordance with Commission resolution 1991/67.
[131] ICRC, Conference of Government Experts on the Reaffirmation and Development of International Humanitarian Law Applicable in Armed Conflicts, Vol. IV, Rules Relative to Behaviour of Combatants (CE/4b) (1971).
[132] Common Article 3 to the Geneva Conventions, *available at:* https://ihl-databases.icrc.org/applic/ihl/ihl.nsf/WebART/365-570006?OpenDocument.

people's lives become the focus of the debate rather than state security. In her paper on humanitarian disarmament, Docherty demonstrated how humanitarian disarmament has changed the vocabulary of the debate on the legality of nuclear weapons from interests in nuclear weapons as tools of deterrence and means of maintaining strategic stability between states. From a state security perspective, nuclear war is avoided to protect state security interests according to the traditional approach; while the humanitarian approach is more interested in disarmament to protect human lives and reduce human suffering.[133] Humanitarian disarmament is characterized by putting human security at the center of the disarmament efforts, by being more inclusive and transparent by providing channels for international organizations and NGOs to advocate for a ban on certain weapons and by being more preemptive, efficient and quick to respond to the development of weapons that would be threatening to human security.[134] In the context of autonomous weapons, the campaign to ban killer robots is leading the efforts of disarmament of autonomous weapons from a humanitarian perspective. Humanitarian disarmament is a turning point in the weapons race as it allows other non-state actors, such as international organizations, to voice their arguments to ban certain weapons against state representatives. Humanitarian disarmament advocacy when done consistently and effectively could turn arguments for a ban on humanitarian considerations into a representation of public conscience.

v. Morality of autonomous weapons and humanity principle

Early efforts of advocacy to ban autonomous weapons focused on the moral implications of delegating decisions to kill people to machines. The main objection to machines taking life-and-death decisions is that they lack human compassion.[135] Furthermore, the notion that human lives are degraded to numerical data sensed and processed by machines undermines human dignity and is morally unacceptable.[136] Scharre elaborated on this concept of empathy in war and presented both views on consequentialism ethics and deontological ethic, the former presents the point of view that right and wrong depend on one's actions and that the outcome justifies the mean, whereas the latter focuses on determining what is right or wrong by only looking at the action regardless of

[133] Bonnie Docherty, A "light for all humanity": the treaty on the prohibition of nuclear weapons and the progress of humanitarian disarmament, 30 Global change, peace & security 163–186 (2018).
[134] *Id.*
[135] *Supra note* 87.
[136] *Supra note* 12.

the consequences.[137] These two different perspectives explain the different soldiers' reactions on battlefields when feelings of mercy and compassion are triggered. Deontological soldiers would hesitate to fire at enemy soldiers when they recognize the opponent's humanity. Walzer called these moments "naked soldier" when a snipper spots an enemy soldier bathing alone constituting no imminent threat, a moment which would recall to his mind the identity of the soldier as a fellow human, not an enemy, and owing to such moment the snipper hesitates to fire at him.[138] Removing humans from the targeting process and replacing them with machines means that the machine will in all cases choose to fire with no consideration for these feelings of compassion and mercy. On the other hand, a counterargument presented by consequentialists is that the decision not to fire owing to feelings of compassion and mercy in incidents similar to the one explained above means that the enemy soldier who was bathing the other day would be left to fight which will cost other fellow soldiers from your troops their lives.[139] Furthermore, the question of why the case of autonomous weapons would be different from remotely controlled weapons as they also disconnect human operators from feelings of mercy and compassion on battlefields is a puzzling one.

Opponents of the ban on autonomous weapons do not only reject the idea that deploying autonomous weapons is immoral, but they further argue that their deployment should be a moral preference. They build this argument on the fact that humans do not only experience compassion and mercy, but most commonly they sense stress, fear, anger, and pain which would affect their judgments on battlefields rendering them more likely to commit crimes. Owing to the responsibility of the state to protect its troops against war brutality, humanitarian and moral obligations vis-a-vis their soldiers should prevail.[140]

The debate surrounding humanitarian considerations shows how the shift to more elastic notions, such as ethics, wrong, and right which have relative interpretations that vary depending on one's perception has magnified indeterminacy in reaching a decision. A question to choose between a prolonged war in which mercy and empathy are regarded and a war that ends sooner but with less

[137] *Supra note* 14.
[138] *Id.*
[139] *Id.*
[140] Amitai Etzioni & Oren Etzioni, *Pros and Cons of Autonomous Weapons Systems*, Military Review Journal, May-June 2017, *available at:* https://www.armyupress.army.mil/Journals/Military-Review/English-Edition-Archives/May-June-2017/Pros-and-Cons-of-Autonomous-Weapons-Systems/

consideration of such principles has no right answer. Similarly, a question to decide on what is more worthy to be respected; to protect your soldiers against war brutality or ensuring that your enemies are killed with dignity by a human, not a machine, both sides will succeed in giving you convincing answers to align with the value they prioritize, military necessity or humanitarian considerations. The more the discussion shifts to blurry notions and concepts, the harder to reach a consensus and find common grounds for the conversation. The conclusion that can be drawn from this section is that humanity is primarily a moral term and to be seen as a legal term with legal implications depends on the position of the person as a positivist or not.[141] However, the magnificence of the Martens Clause is that it allowed naturalists to present their moral argument in legal vocabulary; and for this reason, whenever humanity is invoked as an argument it should be complemented and supported by other black letter law principles.[142]

vi. Dictates of public conscience as the second prong of the Martens Clause

The dictates of public conscience provide a source for moral standards in international law by which state practice is not exclusively seen as the component of public opinion. However, to stand on what constitutes public conscience or how to measure remains unclear.[143] Dictates of public conscience are seen as a tool to restrict the authoritative powers of states to shape international law; however, its application would raise questions of compliance with the principle of legality[144] for positivism adherents.

Advocates of a ban on autonomous weapons presented an open letter calling for a ban on military AI development and autonomous weapons signed by over 1000 leading experts in artificial intelligence including pioneers of the industry such as Hawking, Ellon Musk, and Google DeepMind chief executive.[145] The letter certainly gained international attention being supported by figures from artificial intelligence industry; however, would the letter be considered as reflective of public conscience? Furthermore, would a letter arguably reflecting public conscience

[141] *Supra note 1.*
[142] *Id.*
[143] Smith, T. (2019). Challenges in identifying binding martens clause rules from the 'dictates of the public conscience' to protect the environment in non-international armed conflict. Transnational Legal Theory, 10(2), 184-201. https://doi.org/10.1080/20414005.2019.1621737.
[144] *Id.*
[145] Samuel Gibbs, Musk, Wozniak and Hawking urge ban on warfare AI and autonomous weapons, the Guardian, July 2015, *available at*: https://www.theguardian.com/technology/2015/jul/27/musk-wozniak-hawking-ban-ai-autonomous-weapons

be enough to render a weapon unlawful? Scharre answered the latter question by confirming that Martens Clause, with its two prongs, has never been used to ban a weapon before.[146]

He further showed how measuring public conscience is context-based by presenting the findings of Horowitz when he asked people about their opinions on autonomous weapons in the abstract, the results showed that 48% opposed them, 38% supported them and the rest could not decide. When Horowitz revised the question asking about their opinions on using autonomous weapons knowing that they could be effective and protective to friendly troops, supporters rose to 60% and opponents fell to 27%.[147]

The conclusion that can be drawn from the value of public conscience is that it acts as an effective tool for advocacy and gaining support for the cause promoted; however, it is hard to survive as the sole legal foundation to determine the legality of the subject or to produce binding legal obligations due to the inconclusive and the dynamic character of what constitutes public conscience.

[146] *Supra note* 14.

[147] *Supra note* 14 & Michael C Horowitz, Public opinion and the politics of the killer robots debate, Research and Politics, January-March 2016, *available at:* https://journals.sagepub.com/doi/pdf/10.1177/2053168015627183

III- Quasi-Legal Arguments on Autonomous weapons:

As illustrated in the previous chapter of this paper that the debate around the lawfulness of autonomous weapons involves legal and ethical considerations that should be taken into account; in addition to these legal and ethical considerations, quasi-legal arguments exist. Quasi-legal arguments emanate from other political considerations that serve the agendas of both sides. Enriching the debate by provoking quasi-legal arguments serves the goal of reaching a consensus on the matter by linking legal obligations to realistic political considerations that need to be highlighted to avoid arguing in a vacuum.

One of the major concerns echoed by proponents of the ban on autonomous weapons is their fear of the proliferation of these weapons systems that would lead to an arms race that gets out of our control. They plea to learn from past lessons when the international community successfully joined forces to ban other weapon systems. On the other hand, opponents of the ban highlight the inevitability of autonomous weapons to advance their argument on the necessity of continuing the development of autonomous weapons. This part of the paper will be investigating the quasi-legal arguments related to the arms race and the inevitability of autonomous weapons and then will be looking into precedents of banning weapons by consensus to stand on whether the circumstances of reaching consensus are fulfilled in the case of autonomous weapons or not.

A. Inevitability of autonomous weapons and arms race

The scenario of deploying autonomous weapons on the battlefield and directing them against humans has for many years remained a scene from a sci-fi movie, far from our reality. This is no longer the case after reports of using autonomous weapons in Libya in 2020 were published and after the news was out that Israel used AI-assisted weapons to assassinate an Iranian nuclear scientist.[148] Such incidents, with ambiguity looming around the level of human control involved in these operations, triggered fears about the future of the arms race with autonomous capabilities.

Concerns related to the escalation of conflicts and threats to friendly relations between states due to the unpredictability of autonomous and their impacts on strategic global stability were raised.[149] Additionally, disadvantaged developing countries that cannot develop such weapons will suffer

[148] Will Knight, Autonomous Weapons Are Here, but the World Isn't Ready for Them, December 2021, *available at:* https://www.wired.com/story/autonomous-weapons-here-world-isnt-ready/
[149] *Supra note* 35.

from the militarization gap which will contribute to the imbalance of powers between states and people in poor countries will be the first to suffer the horrific consequences of such weapons.[150]

Ongoing concerns are also voiced that autonomous weapons will lead to lowering the threshold of the decision to resort to war. In this regard, Heyns argued that:

> Due to the low or lowered human costs of armed conflict to States with LARs in their arsenals, the national public may over time become increasingly disengaged and leave the decision to use force as a largely financial or diplomatic question for the State, leading to the "normalization" of armed conflict. LARs may thus lower the threshold for States for going to war or otherwise using lethal force, resulting in armed conflict no longer being a measure of last resort.[151]

However, Arkin[152] refuted this argument made by Asaro since it is not a particular issue limited to autonomous robots, but is typical for the advent of any significant technological advance in weapons and tactics.

Such utopian arguments that value humanitarian considerations are countered by resistance from states that perceive the development of autonomous weapons as a matter of military necessity in light of their inevitable existence in our modern warfare. Scharre referred in his book to how autonomous technologies and robotics are currently being developed by students in high school labs. He pointed out how a project to build a robot, which would take two years to complete in the past, can now be completed by students at the school in nine weeks.[153] Scharre further spotlighted how autonomous weapons can be manufactured as DIY home projects. He cited a clip posted on Youtube in which a teenager revealed the capabilities of a drone that he armed by himself.[154] The prevalence of off-the-shelf autonomous systems and artificial intelligence technologies will further complicate the situation and seeing armed groups using autonomous weapons is a viable scenario. Owing to such fears, states stick to their right of self-defense against such threats and they see developing autonomous weapons as a deterring mechanism.

The inevitability of autonomous weapons in the near future should truly motivate the international community to speed up the drafting of a treaty regulating autonomous weapons. As demonstrated throughout the paper, the indeterminacy of IHL rules governing the topic will continue to lead the

---

[150] *Id.*
[151] C. Heyns, Report of the Special Rapporteur on extrajudicial, summary or arbitrary executions, Christof Heyns. UN General Assembly, A/HRC/23/47 (9 April 2013).
[152] Arkin, R. (2009). Governing Lethal Behavior in Autonomous Robots. New York: Chapman and Hall/CRC, at 37, available at: https://doi.org/10.1201/9781420085952.
[153] *Supra note* 14.
[154] *Id.*

debate into a vicious circle of a debate on what the law says rather than what the law is.[155] To refute Russia's position that it is too early to discuss a governing definition of autonomous weapons, a delay in a reaching a consensus on treaty governing autonomous weapons means that the final product of rules governing autonomous weapons more likely to be a descriptive of what states do, purely apologetic rules, rather than a reflection of the wishes of international community subjects.[156] For this reason, NGOs participating in the meetings of CCW/GGE should be mindful of the current Global Legal Order (GLO) where artificial intelligence technologies are embedded within our modern economy, trade and investment and impact our daily activities. In doing so, they should be mindful of the dichotomy between narrative and material international law rules.[157] A call for an absolute ban of autonomous weapons is a manifestation of an extreme utopian call that cannot be currently realized in material international law when considering the inevitability of autonomous systems. One cannot deny that they are founded on ethical considerations, yet such ethical considerations exist only in debates on narrative international law, deeming them as false contingency lacking authority in real life.[158] To put it clearly, a treaty governing autonomous weapons is not the magic wand that would find an answer to all the indeterminacies highlighted in this paper. It will, however, be one step closer to identifying what the law is when regulating autonomous weapons. It would be a manifestation of the law-making role of the international community, including states and NGOs, as a virtual sovereign.[159] Drafting a treaty regulating autonomous weapons is an important phase to untangle indeterminacy through law identification and creation[160] since existing IHL rules fall short to solve the identified indeterminacies and a definition of autonomous weapons remains contested. Yet, a certain level of indeterminacy will continue to exist even after the draft of the treaty when interpretation and application of the rules are invoked and at this stage "*jurispathic function*" of the court will contribute to the evolution of the rules by killing laws to create laws.[161] At this stage, consistency, and determinacy of the law will greatly depend on the consistency of the "*committed internal point of view"*[162] that reflects the politics of the system. In this sense, law will still be labeled as apologetic, but an apologetic

[155] Jason A. Beckett, *Countering Uncertainty and Ending Up/Down Arguments: Prolegomena to a Response to NAIL*, 16 European journal of international law 213 (2005).
[156] *Id.*
[157] Jason Beckett, *The Divisible College: A Day in the Lives of Public International Law*, 23 German law journal 1159 (2022).
[158] *Id.*
[159] *Supra note* 155.
[160] *Supra note* 157.
[161] *Id.*
[162] *Id.*

version of the law that reflects the wishes of the subjects of international community; in which NGOs, activists, public opinion are represented as part of the virtual sovereign in our modern world.

B. Reaching consensus on the lawfulness of autonomous weapons and lessons learned from banning weapons.

Advocates who heed the call to ban autonomous weapons often refer to successful incidents of banning weapons from the past. In one of the articles dedicated to banning fully autonomous weapons, Human Rights Watch focused on the analogy between autonomous weapons and blinding lasers to make a case for a preemptive ban on autonomous weapons by showing how the blinding lasers ban is a relevant precedent that should be pursued. The article argued that autonomous weapons like blinding lasers would violate Martens Clause with its two prongs, humanity and dictates of public conscience. The paper[163] recognized the different grounds of breaches of the Martens Clause when comparing the nature of autonomous weapons and blinding lasers, yet the ultimate conclusion of the paper is that autonomous weapons should, owing to violations of the Martens Clause, be banned in the same way blinding lasers were banned by consensus. The paper also highlighted the similarities between banning autonomous weapons and blinding lasers on grounds of violations of rules of armed conflict. It argued that autonomous weapons should be banned due to their violations of principles of distinction and proportionality, in the same way, blinding lasers were banned due to violations of the obligation not to cause superfluous injury or unnecessary suffering. The paper further argued that the ban on fully autonomous weapons should not affect the research and development of autonomous technologies for both civilian and military purposes, including remotely controlled drones, in the same way, the ban on blinding lasers did not restrain the development of laser technologies for both military and civilian purposes. The paper aimed to show, by presenting the case of blinding lasers, how the ban on autonomous weapons is a conceivable case by presenting how blinding lasers and autonomous weapons share the same grounds to ban and refute the arguments presented by anti-ban advocates by highlighting how such arguments did not restrain the consensus of banning blinding lasers. However, the case of banning blinding lasers as a precedent to ban autonomous weapons is

[163] Human Rights Watch & International Human Rights Clinic, Precedent for Preemption: The Ban on Blinding Lasers as a Model for a Killer Robots Prohibition Memorandum to Convention on Conventional Weapons Delegates, November 2015.

contested on the ground that banning blinding lasers was a case that gained consensus owing to factors that would render banning weapons by consensus more likely, whereas such factors do not exist in the case of autonomous weapons. Such factors include: that the weapon is ineffective, the existence of other weapons or means that would achieve similar military objectives, the scope of the ban is clear and narrowly tailored, states are sufficiently committed to enacting banning regulations, and the weapon has not been integrated into states' weapon arsenals.[164] Autonomous weapons offer promising military advances for which no substitutes exist; especially in terms of deploying arms in dangerous or inaccessible zones. Autonomous weapons, as shown in this paper, also pose challenges concerning defining them and defining terms related to their regulations such as control. Furthermore, states, specifically those that invest in researching autonomous technologies for military projects, resist attempts to obstruct such developments, especially in light of the ongoing tensions between world powers.[165] States that support calls for banning autonomous weapons for humanity considerations are primarily states that cannot catch the pace of developing autonomous weapons, and their poor human rights records deem their call to ban such weapons for humanity considerations cynical.[166] The history of regulating weapons shows that:

> [I]f a new weapon system greatly advantages a side, the tendency is for it gradually to be adopted by others that perceive they can benefit from it as well. In some cases, legal prohibitions on the weapon system as such erode, as happened with military submarines and aircraft; what survives is a set of legal rules for the use of the new weapon, with greater or lesser specificity.[167]

Blinding lasers might not be the most relevant case to present as a precedent for banning autonomous weapons for the reasons listed above. Perhaps advocates of banning the usage of autonomous weapons should investigate other cases of regulating weapons such as nuclear weapons that despite their abhorrent impacts, their absolute ban did not survive grounds justifying their use for self-defense and deterrence. However, this does not negate that states might still be open to discussing possible arenas of controlling the usage of autonomous weapons and choosing

[164] Rebecca Crootof, Why the Prohibition on Permanently Blinding Lasers is Poor Precedent for a Ban on Autonomous Weapon Systems, November 2015, *available at*: https://www.lawfareblog.com/why-prohibition-permanently-blinding-lasers-poor-precedent-ban-autonomous-weapon-systems .
[165] David Hambling, Efforts to regulate 'killer robots' are threatened by war in Ukraine, July 2022, *available at:* https://www.newscientist.com/article/2327965-efforts-to-regulate-killer-robots-are-threatened-by-war-in-ukraine/.
[166] *Supra note* 14.
[167] Kenneth Anderson & Matthew C. Waxman, Law and Ethics for Autonomous Weapon Systems: Why a Ban Won't Work and How the Laws of War Can, 2013, *available at:* https://scholarship.law.columbia.edu/cgi/viewcontent.cgi?article=2804&context=faculty_scholarship.

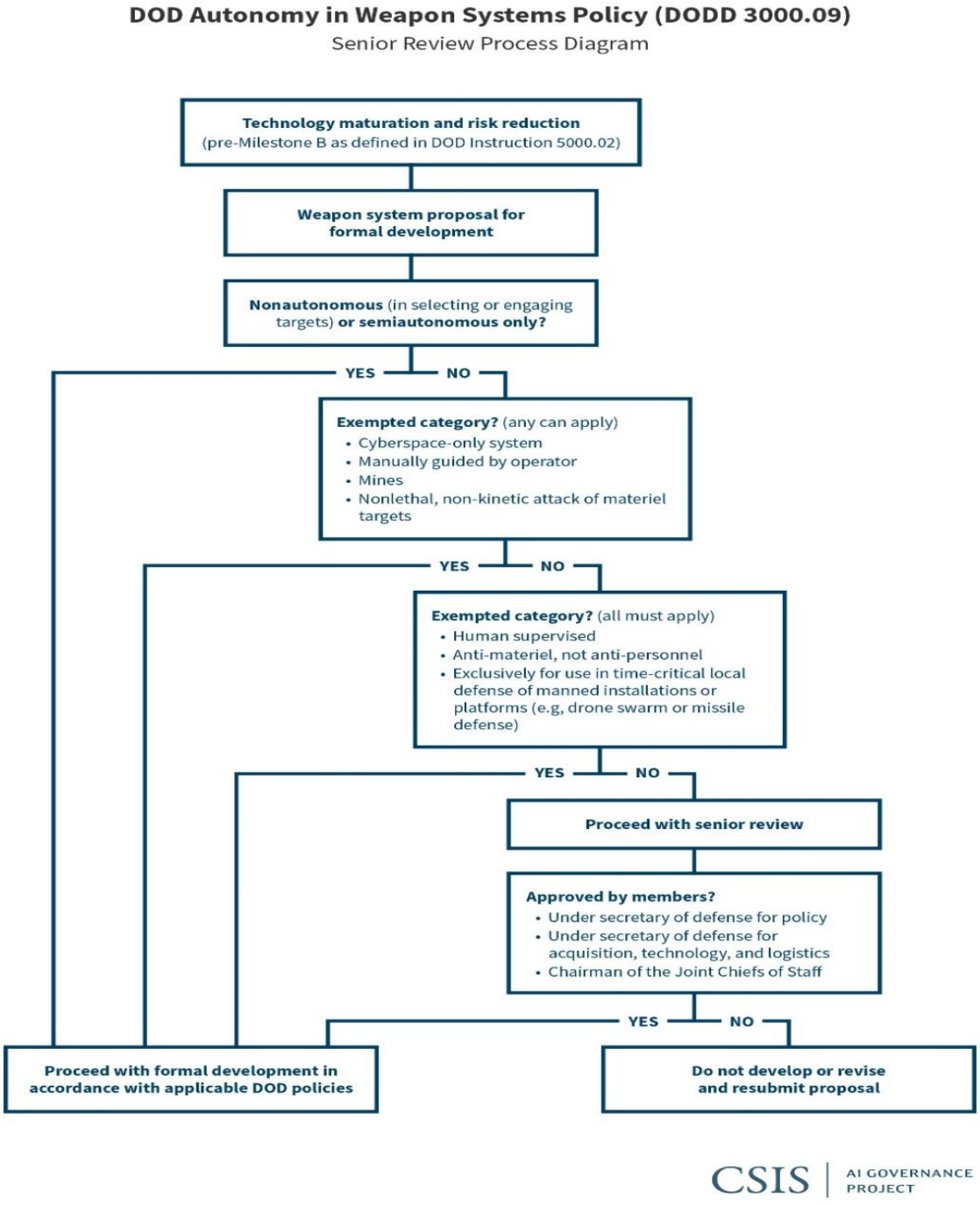
what could or could not be prohibited in this regard.[168] One of the feasible limitations on autonomous weapons is to limit their deployment against personnel until they evolve in a manner that would ensure respect of principles of distinction and proportionality, but a call for an absolute ban of this category of weapons is perhaps too utopian to achieve.

[168] *Supra note* 164.

Conclusion

We cannot predict if our generation would witness the era of robots and machines fighting against each other, excluding humans completely from the scene to achieve optimum promises of reducing human suffering and we cannot predict if this would ever be a viable scenario. What we witness nowadays could be a preliminary phase of either modern warfare coping with technological expansions of post-millennial generations where military superiority is determined by militaries abilities to incorporate artificial intelligence and robotics and where military advancement is not decided based on killing soldiers of your adversaries but by rendering their weapons ineffective; or a state of apocalyptic fighting where things get out of control and machines superiority would threaten our existence. The legal discourse surrounding these scenarios should overcome the challenges of competing interests which translate into tangled definitions of autonomous weapons and their components such as the concept of human control. In doing so, purely utopian calls for an absolute ban on autonomous weapons should be under-toned and become more grounded to the reality of technological developments of autonomous machines. On the other hand, states need to give up on apologetic claims that would justify the usage of any means for military necessity considerations and be mindful of the dark consequences of delegating machines with powers that would be unstoppable one day and turn to shoot on them. This paper attempted to engage in the discussion on the legality of autonomous weapons by showing how the debate is full of contentious issues that reflect the indeterminacy of international law when deciding on similar topics that fall within the grey zone, where law alone is incapable of giving a decisive answer. An area where stakeholders, with different motivations and interests, are capable of weaving their legal arguments to support their position. An area where perhaps the law is helpless to offer a solution, but for a solution to enjoy consensus it has to be supported by law and expressed in a legal language. The case of autonomous weapons is a perfect example of this situation; the debates on fundamental concepts governing autonomous weapons including their definition, and their legality have demonstrated indeterminacy as a character embedded within the concepts and the language of international law.

APPENDIX I



Source: Gregory C. Allen, DOD Is Updating Its Decade-Old Autonomous Weapons Policy, but Confusion Remains Widespread, June 2022, *available at*: https://www.csis.org/analysis/dod-updating-its-decade-old-autonomous-weapons-policy-confusion-remains-widespread