International humanitarian law and autonomy in weapons

A normative analysis of the notion of human control

July 2023

2676843O
233478
6639818

Presented in partial fulfilment of the requirements for the Degree of

International Master in Security, Intelligence and Strategic Studies

Word Count: 21737

Supervisor: Marco Pertile, University of Trento

Date of Submission: 26/07/2023

To all the kind and patient people
whom I have encountered on my way

Abstract

Lethal autonomous weapons systems, as defined by the International Committee of the Red Cross, can select and apply force to targets without human intervention. The relationship between lethal autonomous weapons systems and international humanitarian law is herein studied from the perspective of human control. While autonomy in weapons is progressively increasing, it leads to the decrease of human control. This thesis finds that human control is required under the existing regulatory framework. Such a conclusion is reached through a normative analysis, carried out in two phases. First, the notion of human control is developed by examining the positions of states expressed in 2023 under the framework of the Convention on Certain Conventional Weapons. The convergences identified among states' statements allow to conceptualize human control. Second, the analysis of fundamental principles of International Humanitarian Law proves that human control is an intrinsic legal requirement. Normative analysis shows that the notion of human control includes human judgment, monitoring and capacity to intervene during the entire cycle of the performance of a weapon. It is demonstrated that a degree of human control must be maintained for lethal autonomous weapons systems to comply with International Humanitarian Law.

Keywords: Lethal Autonomous Weapons Systems, Human Control, Convention on Certain Conventional Weapons, International Humanitarian Law, Distinction, Proportionality, Precautions in Attack.

Index

List of abbreviations

AP	Additional Protocol
CCW	Convention on Certain Conventional Weapons
CIHL	Customary International Humanitarian Law
GC	Geneva Convention
GGE	Group of Governmental Experts
HC	Human Control
HITL	Human-in-the-loop
HOOTL	Human-out-of-the-loop
HOTL	Human-on-the-loop
IAC	International Armed Conflict
ICC	International Criminal Court
ICJ	International Court of Justice
ICL	International Criminal Law
ICRC	International Committee of the Red Cross
ICTY	International Criminal Tribunal for the former Yugoslavia
IHL	International Humanitarian Law
IHRL	International Human Rights Law
IT	Information Technology
LAWS	Lethal Autonomous Weapons System
MHC	Meaningful Human Control
NHI	Nominal Human Input
NIAC	Non-international armed conflict
UN	United Nations
UNA	United Nations Association
UNIDIR	United Nations Institute for Disarmament Research
UNODA	United Nations Office for Disarmament Affairs

Introduction

On the 27th of March 2020, as the forces affiliated with the Libyan commander Khalifa Haftar were retracting, they were targeted and engaged by several Kargu-2 quadcopter drones (Panel of Experts on Libya, 2021). Fabricated by the Turkish STM, Kargu-2 is a loiter drone, capable of performing its functions in two modes: either with the intervention of a human operator or autonomously. The UN Panel of Experts on Libya documented that incident in their report, describing the Kargu-2 used in the offensive as lethal autonomous weapons that "were programmed to attack targets without requiring data connectivity between the operator and the munition" (Panel of Experts on Libya, 2021: §63). These allegations of autonomy were publicly rebutted by the CEO of STM, who claimed that a human operator was always in the loop while engaging with targets (Tavsan, 2021). The United Nations did not provide any further comment to the report (Cramer, 2021). Yet, what happened in Libya in spring 2020 is frequently considered as the first ever documented use of lethal autonomous weapons in combat. Whether it was or not depends, among other considerations, on what is understood as lethal autonomous weapons. To date there is no commonly adopted definition, and there is no clear consensus on what autonomy in weapons means.

Since the turn of the century, the development of autonomous weapons has been conquering international attention. Lethal Autonomous Weapons Systems (LAWS) belong to the intersection of several fields of expertise, including the military, ethics, technology, and international law. Debates, that tend to be interconnected, arise from all these spheres. From the legal point of view, it is unclear whether fully autonomous weapons can comply with the principles of International Humanitarian Law (IHL). Ethically, it is contested whether the machine's capacity to decide on life and death is compatible with human dignity. Military-oriented discussions review strategic and tactical advantages that autonomous weapons can offer, if compared to non-autonomous

weapons and to combatants (Etzioni & Etzioni, 2018). In regard to technology, certain levels of autonomy in weapons are only possible with the employment of artificial intelligence (AI), thereby extending the discussions on LAWS to a broader topic, related to the limits of the use of AI for military purposes.

These and other questions are being addressed by the academia and by states. On the interstate level, a formal focus on LAWS was established in 2013, seven years prior to the Libyan attack. That year, the High Contracting Parties to the Convention on Certain Conventional Weapons (CCW) agreed to incorporate LAWS as one of the points on their agenda. In May 2014, the parties to the CCW held their first meeting on LAWS. That meeting led to the establishment of the Group of Governmental Experts (GGE) on emerging technologies in the area of lethal autonomous weapons systems. Under this setting, representatives of states have been convening annually since 2015. Because the CCW is strongly anchored on International Humanitarian Law, discussions of the GGE are closely IHL-related. In 2019, the GGE developed 11 non-binding guiding principles that set the ground for further negotiations on LAWS. The principles convey limitations, obligations, prohibitions, and risk mitigation measures. For example, they set the prohibition of the development of anthropomorphic LAWS, and the requirements for ensuring accountability and human responsibility, compliance with the international law and the need to establish mechanisms of risk assessment and mitigation.

While the guiding principles constitute an important step in building international consensus on LAWS, many questions remain open. For instance, the very way of referring to lethal autonomous weapons systems is not homogeneous across the states and non-state actors. Lethal Autonomous Weapons Systems or LAWS is the term first used in the framework of the CCW, which was later adopted commonly in other fora. The category "weapons system", subject to a separate definitional discussion of its own, encompasses one or several weapons, together with the munitions and other means that these weapons require to operate (Weapons Law Encyclopaedia, 2023). In other cases,

references to autonomous weapons exclude the word systems, and this is reflected in the acronym LAWs. Due to the fact that lethality is not the only possible outcome of an attack, the ICRC is one of the proponents to use the acronym AWS, standing for Autonomous Weapons Systems. Some alternative denominations, closer related to civil society, carry a negative connotation. The campaign Stop Killer Robots is one of the primary advocates of referring to autonomous weapons as killer robots. Another similarly negative term is slaughterbots. Because of its long trajectory at the CCW and the wide use by states and non-state actors, this thesis will employ the denomination LAWS, indistinctively with the shortened version 'autonomous weapons'.

In addition to this, when addressing autonomous weapons, states and non-state actors operate a wide range of definitions. This leads to different understanding of what LAWS are. The novelty of LAWS clearly lies in their autonomy, but depending on the definition adopted, autonomy itself can have different meanings. A minimal definition could explain autonomy as the absence of human intervention in engaging with a target. Landmines and booby traps are a common reference: once activated, they do not require a human operator to engage with a target, nor does the human operator have a role in selecting the concrete target of the attack. Therefore, a booby trap or a landmine could be considered a lethal autonomous weapon. A definition supporting this stance is the one proposed by the International Committee of the Red Cross: it defines autonomous weapons as "any weapons that select and apply force to targets without human intervention" (ICRC, 2022). For the purpose of this thesis, this definition will be employed.

On the opposite side of the definitional spectrum, autonomy could instead require a rather sophisticated exercise, in which targets are selected following complex algorithms, and the autonomous weapons might have the capacity to learn and evolve in this process. For example, the definition proposed by Brazil in 2020 describes LAWS as "An intelligent weapon system with autonomous operation mode (i.e., without human input after activation)

capable of recognizing patterns in combat environments, and of learning to operate and make decisions […] based on uploaded databases, acquired experiences and its own calculations and conclusions" (GGE on LAWS, 2023[a]: 6). Such a position understands LAWS as intrinsically tied to AI.

Decade-long discussions on the legal, ethical, military, technological and other characteristics of LAWS still largely revolve around the definition and the denomination of autonomous weapons, where consensus is yet to be reached. In the meantime, beyond the international fora, the very object of study has been evolving. Discussions on LAWS started before complex levels of autonomy in weaponry existed. In 2013 and prior, LAWS were almost a futuristic subject, possible but undefined and, as such, belonging to an undefined future. This was also reflected in the academia: before 2020, references and mentions to the Terminator — a fictional character represented by Arnold Schwarzenegger in 1984 — appeared in over five thousand academic articles related to LAWS[1]. Yet, the new lethal autonomous technologies that have actually emerged are far from anthropomorphic killer robots, and they are no longer a matter of future. As weapons with rising levels of autonomy are being progressively deployed in combat, their compliance with international law becomes an increasingly pressing concern.

This perception is shared within the international community. The most recent meetings of the GGE took place in Geneva, in March and May 2023. The resulting final report advances in the understanding of the key legal principles challenged by autonomy in weapons and looks into potential ways forward. Earlier that year, the Latin-American and Caribbean Conference addressed the issue of LAWS in the Costa Rican city of Belen. That was the first regional

---

[1] Calculated by the author using the search engine of Google Scholar. More interestingly, before 2020, publications mentioning the Terminator represented 18,78% or one-fifth of all the publications on autonomous weapons featured in Google Scholar. Up to the date, this figure has varied significantly. Only 5,28% of the publications dated from 2021 on mention the word *Terminator*. This calculation does not aim to present an exact number, but rather a visualisation of a tendency.

interstate forum that specifically focused on autonomous weapons. As a result, the Belén Communiqué urges the rest of the international community to develop a legally binding instrument that would establish prohibitions and regulations (Ministry of Foreign Affairs and Religion of Costa Rica, 2023). The same year, in April 2023 Luxemburg hosted an international conference covering four thematic areas, providing expertise on military, legal, technological, and ethical considerations related to LAWS. In addition to that, some states are progressively involving their population in discussions on lethal autonomy in weapons. For example, also in 2023, the Parliament of the United Kingdom launched a call for answers on seven questions related to LAWS, ranging from what the lethal autonomous weapons are to what international regulations, if any, must be developed to address them (UK Parliament, 2023).

Enveloping the 2023 reports, discussions, and panels on the lethal autonomous weapons systems, is the word *momentum*. From the states and experts' interventions, progress made and projected from the CCW and the successful harmonisation of regional positions in Belén, the year that marks a decade of international discussions on LAWS appears surrounded by the expectation of a new stride towards achieving a global understanding on this matter. *Momentum* was perceived at the 2023 meeting of the GGE on LAWS by the ICRC (Spoljaric, 2023), *momentum* was celebrated in Belén (UNA-UK, 2023), and announced by the panellists in Luxemburg.

This thesis belongs to this *momentum*. After more than a decade of discussions on national and international levels, many debates related to LAWS remain open. The anticipation of progress still requires a clear understanding of what autonomy in weapons means and how the international law applies to the lethal autonomous weapons systems. This thesis will address this issue by focusing on a notion that emerged as a counterbalance to autonomy in weapons: human control (HC). Increasing autonomy in weapons reduces the degree of human control. At the same time, a specific requirement related to human control can establish limits to the autonomous features. HC has been widely

discussed by states and scholars. Some, such as Sharkey (2019), believe that human control or meaningful human control should constitute a guiding principle that would determine the legality of a weapon. Others, including Jensen (2020) believe that the focus on human control is erroneous. In order to assess these positions, it is first necessary to understand what human control entails, and why it could be seen as a requirement. By shifting the focus from autonomous weapons to human control, this thesis seeks to provide a technology-neutral view on autonomy in weapons and IHL. Overall, this thesis will examine how the notion of human control contributes to the understanding of the relationship between LAWS and the IHL. In order to do so, it will focus on states' positions and the international legal framework. The most recent meeting of the GGE on LAWS will nurture the analysis of states' positions on human control. For the assessment of the legal framework, fundamental IHL principles will be examined in their application to LAWS.

This analysis will be structured as follows. First, the literature review will outline the existing research, delving into the interdisciplinarity of this topic, the potential categorisations of LAWS, different understandings of autonomy, the comparisons of the present case with previous cases of arms regulation, risks and the opportunities associated with LAWS. After this, a methodological section will describe the design of this research. Subsequently, analysis will take place in three chapters. First the most recent GGE on LAWS will be studied through content analysis, focusing on human control. Second, legal interpretation will expand on the IHL requirements that apply to LAWS, in order to see whether a human element is necessary under the current regulatory framework. The findings will be combined in third chapter. In this way, this thesis will provide a legal analysis oriented towards the CCW, focused on the notion of human control, and developed over the most recent data. It is written with the hope of enhancing the academic knowledge and of participating in the construction of necessary common understandings.

Literature review

Early literature on lethal autonomous weapons systems dates from the beginning of the 2000s. The first publications were written in future tense. For example, in addressing killer robots in 2007, Sparrow (2007) mentioned several types of weapons that were under development at that time, but prognosed that the new generation of truly autonomous weapons was yet to come. When Krishnan, also among the pioneers, wrote about autonomous weapons two years later, he could already constate that over forty countries were developing "robotic weapons" with increasing levels of autonomy (2009:11). This numerical increase shows the swiftness of technological developments, and also the practical implications of different criteria for defining autonomy in weapons.

Academia is a crucial source of common understandings in discussing key concepts, including the definitions of autonomy and LAWS. By thematically reviewing the existing literature on LAWS, this chapter will deepen on the main, closely intertwined, discussions surrounding the topic and the research that has been done up to the moment. In addition to this, literature review plays a contextualizing role, because academic research on LAWS is being developed at the same time as autonomy in weapons is evolving. Research illustrates technical advancements and the current state of development of LAWS, but also the context around them, including the most pressing debates and the interstate deliberations on their regulation.

The first sections of this chapter will allow to explain the relevance of LAWS for the research and deepen on their definitions and categorisations. After this, the literature review will move towards the most pressing debates surrounding LAWS, with the aim of understanding from which angles the topic has been embraced by scholars. Finally, the normative side of these debates will be examined, in order to provide clarity on the main legal aspects subject to academic research on LAWS. This includes considerations on human control. Structured in this way, the literature review should constitute a sufficient basis

for understanding the current state of the art and the sliding balance between autonomy in weapons and human control.

The novelty of LAWS

It is generally established that the distinctive feature of LAWS, if compared to other kinds of weapons, is their autonomy. Autonomy can be understood as the ability of the weapon to engage with a target without the intervention of a human operator (Leveringhaus, 2016: 3-4). Should the former sentence be read in broader terms, it would mean that lethal autonomous weapons have long existed in past. Once a trap is set up, it does not require any further action from an operator to perform its function. Once a mine is placed, it does not require any further action by the operator in order to explode. Thus, the proponents of a broader understanding of LAWS open this category to a variety of currently existing weapons. For example, following the official position of the ICRC, "mines can be considered rudimentary autonomous weapons" (2022).

Contrarily to this view, LAWS can be expected to incorporate sensors that allow them to assess the surrounding context, together with an "algorithmic process of target acquisition and decision to kill" (Righetti *et al.,* 2018: 124). Algorithms are sometimes compared with the exercise of intelligence, in the sense that they enable machines to determine what targets to select (Krishnan, 2009: 3). Noone and Noone remind, in this regard, that machine intelligence is distinct from human intelligence: an algorithm cannot reflect emotional, social, or moral considerations (2015: 27). However, machine intelligence can involve, as suggested by McFarland and Assaad, "replicating some aspect of biologically observable intelligence such as cognition, decision-making or, most importantly here, learning" (2023: 2). Seixas-Nunes adds that, for a weapon to be considered autonomous, it must be capable to assess both the context and its position within it (2022: 13). This view is consistent with the idea that a truly autonomous weapon is necessarily AI-enabled.

Differing degrees of autonomy and their aftermaths will be discussed in the following sections. For the nonce, it can be considered that the novelty of LAWS lies in the open horizon of possibilities in which autonomy can contribute to military technologies. Autonomous capacities are the reason why LAWS can and should be studied as a distinct, emerging type of weapons.

Understanding LAWS through autonomy in weapons

Research by Boulanin and Verbruggen provides a comprehensive compendium on autonomy in weapons, intended to facilitate the GGE discussions related to LAWS. Among their key contributions is the proposal of the definition of autonomy, which is understood as "the ability of a machine to perform an intended task without human intervention using interaction of its sensors and computer programming with the environment" (Boulanin & Verbruggen, 2017: vii). A machine can be more or less autonomous, because autonomy is a scale: "the smaller the need for human supervision and intervention, the greater the autonomy of the machine" (Krishnan, 2009: 4). In other words, autonomy in weapons can be seen as a counterpoise to human control.

Krishnan suggests assessing different degrees of autonomy by providing answers to six questions: "Is the weapon triggered by the operator or by itself? Are the targets chosen by an operator or by the weapon itself? Is the weapon stationary or mobile? If it is mobile, does the weapon need external input for finding its target, or can it find its target by itself?" and, in addition to this, whether the weapon can self-repair and self-replicate (Krishnan, 2009: 45). In this scale, a landmine would be seen as an autonomous weapon with a very limited autonomous capacity, because it only fully corresponds with one of the six factors proposed by Krishnan.

Categorisations of LAWS following the degree of autonomy
Landmines constitute an important angle for understanding autonomy and LAWS, because they are a common example in the discussions enriched by the

di- or trichotomy between autonomous, automatic and, sometimes, the middle realm of automated weapons. Strictly following this distinction narrows the category of LAWS: landmines and similar weapons are no longer seen as autonomous weapons. The difference between autonomous and automated weapons lays in the fact that "autonomous includes the notion of an algorithmic decision-making process difficult to predict, while automated is related to a response deterministically triggered by well-defined events" (Righetti *et al.,* 2018: 124). Within the latter, Homayounnejad (2017: 10) suggests distinguishing between automatic and automated weapons: automated weapons are more complex, but still predictable.

In a similar classification, Boulanin and Verbruggen (2017) distinguish between reactive and deliberative systems. Reactive systems can be simple or model-based, implying a higher level of complexity. A simple reactive system responds to a trigger in an action-reaction way. A model-based reactive system incorporates databases that allow the machine to assess the situation beyond its sensors. Differently, a deliberative system represents the next level of complexity. Its decision-making process requires the machine to "weigh the consequences of possible actions and measure whether and to what extent they will serve the achievement of the goal" (Boulanin & Verbruggen, 2017: 9-10). In other words, deliberative systems present a sophisticated level of autonomy, enabled by the AI.

Following this classification, a landmine would be categorised as an automatic or a simple reactive system, on the lowest level of autonomy, or even below the threshold of autonomy required for a weapon to be categorised as LAWS. Within the rest of the spectrum the level of autonomy can be assessed depending on the interaction between the machine and the human operator. A common theoretical approach to this scale is the use of three models: human-in-the-loop (HITL), human-on-the-loop (HOTL), and human-out-of-the-loop (HOOTL). Weapons following the model human-in-the loop require a human operator in the selection and engagement with a target, where the human

provides either input or authorisation. Human-on-the-loop implies that a weapon can carry out these functions on its own, but under the supervision of a human operator, who has the capacity to detain the weapon if deemed necessary (Massacci & Vidor, 2022: 103). In the human-out-of-the-loop model, a human operator is not present and cannot intervene (Noone & Noone, 2015: 28). Boulanin and Verbruggen refer to these categories as semi-autonomous, human-supervised autonomous and fully autonomous weapons (2017: 113).

Application of categories of LAWS to the debate

Categories outlined above contribute to the definition of LAWS and open the discussion to legal and ethical implications of different levels of autonomy. Autonomy simultaneously counteracts both the human control and the predictability of the weapon. As Roff and Danks ascertain, LAWS with a high level of autonomy are less predictable, and thus they generate more uncertainty for their operators than other types of weapons (2018: 8). Because of that, weapons with higher levels of autonomy might be deemed unreliable, dissuading the military use of systems where the human is not in the loop. For Nahavandi (2017: 12), the way forward fostering the use of human-supervised weapons necessarily implies building trust. Trust is a requisite that weapons developers have to meet, by enhancing the autonomous capacities of the weapons, but also ensuring that the LAWS comply with the applicable regulations.

In regard to the human-out-of-the-loop model, literature tends to point out that such degree of autonomy in weapons does not exist up to the date. The autonomous features that already exist in current weapons are only the precursors of the potential future autonomy (Solovyeva & Hynek, 2018: 170). As Sharkey wrote, "The autonomous robots being discussed for military applications are closer in operation to your washing machine than to a science fiction Terminator" (2010: 376). Thirteen years after Sharkey's publication, this is still the case: fully autonomous weapons with lethal capacities do not exist

yet (Christie *et al.,* 2023: 2). If developed, out-of-the-loop systems will come along with a number of legal and ethical challenges. Such a weapon would potentially have the technical capacity to select and engage with a target without human intervention, in a process that cannot be aborted by an operator, and where machine learning leads the weapon to develop new processes on its own. A machine "might learn something it was not intended to learn or do something that humans do not want it to do" (Boulanin & Verbruggen, 2017: 17). Despite these challenges, academia and states agree that fully autonomous weapons might be developed in future. This assumption, frequently mentioned in allusion to forthcoming arms races, motivates ethical and legal debates related to the prospective regulations of weapons with sophisticated degrees of autonomy (Congress of the United States, 2023: 1). In these debates, the sliding level of autonomy is key for establishing considerations on the development and use of LAWS. As mentioned earlier, human control is viewed by some as the key element that should establish what is permissible. Moreover, the categorisations of autonomy are related to the definitions and the characterisations of LAWS.

Understanding LAWS through their definitions

Despite the many advancements in establishing common understandings on LAWS, there is yet no consensus on their definition. Mentions to this shortcoming are common in the research on autonomous weapons, and sometimes become the very object of study. There are at least four reasons why the definitional approach is relevant for the academia. First, a definition provides a clearer understanding of the reference object. Secondly, the lack of a common definition leads to the uncertainty on whether the referent object of the research currently exists, or if it might only be developed in future. Thirdly, a definition is frequently understood as a pre-requisite for the development of international regulations, such as a treaty (Kayser, 2023: 1). Finally, the pursuit of a common definition can be intentional, and not necessarily neutral. For example, it could contribute to the distinction between permissible and unlawful

levels of autonomy (Chernyavsky & Sibileva, 2020: 237). But it could also attempt to promote, as UNIDIR alerts in its report, particular viewpoints and interests related to LAWS, "driven by political and strategic motivations" (2017: 22). For example, the need for a binding regulation can be disqualified if the suggested definition of LAWS leads to the understanding that such weapons cannot feasibly be developed in the near future.

This is why a wide segment of literature focuses on the definitions of LAWS. Boulanin and Verbruggem (2017) distinguish between three types of definitions that have been developed by the stakeholders. The first type focuses on the interaction between the operator and the machine. The second type rather centres on the capabilities of the weapon. For example, it could be the weapons' capacity to evaluate the context and design its own course of action. The third type of definitions revolves around the tasks and functions of the weapon, and tends to involve a normative angle, distinguishing between permissible and not permissible functions (Boulanin & Verbruggen, 2017). The three suggested types of definitions are common and are operated by states and non-state actors concerned with the prospective regulations of LAWS.

Taddeo and Blanchard (2022) analyse how the different definitions of LAWS inform the approaches to ethical and political viewpoints on autonomy in weapons. In order to do so, the authors identify twelve definitions employed by states and non-state actors, along with their embedded political, ethical, and strategic components. This allows them to develop their own, value-neutral definition of LAWS, as follows:

> "An 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 AWS is specific subset of an AWS with the goal of exerting kinetic force against human beings" (Taddeo and Blanchard, 2022: 15).

Likewise focused on the intentional use of specific definitions, Ekelhof (2017) studied them within the discourses of the stakeholders. Discourse analysis allowed her to examine how different narratives are constructed around the commonly used vocabulary, such as the definitions of autonomy and LAWS. It also proved that a common vocabulary does not necessarily set common grounds in the discussions, because stakeholders might use the same terms with different implications. In this sense, the lack of a common definition can be seen as one more obstacle in the discussions on LAWS.

Horowitz (2016), instead, focuses on the practical consequences conveyed by the different definitions. Having underlined the undesirable implications of too narrow and too broad approaches to LAWS, Horowitz suggests a functional solution, pointing out that autonomy in weapons differs depending on the specific function of every element that forms part of a given weapons system. Different functions bring along different challenges related to autonomy. For example, within a lethal autonomous weapons system, the munitions and the platforms present completely different characteristics, sets of functions and challenges (Horowitz, 2016: 94-95). Each of them contributes to the autonomous functions of the weapons system, but in different ways. Therefore, Horowitz suggests shifting the focus from trying to understand the meaning of autonomy towards looking into the functional capacities that autonomy brings to different elements.

Understanding LAWS through comparison

A rather common approach to understanding LAWS is their comparison with other types of weapons and with previous cases of arms regulation. For example, LAWS have been frequently compared with drones. Both the lethal autonomous

weapons systems and drones are a modern kind of weapons, and they partially present similar challenges. Unlike in LAWS, human control remains present in drones (Egeland, 2016: 97). However, some drones already present autonomous features, such as autonomous navigation and landing (Petman, 2017: 6). In regard to their ethical and normative implications, Espada and Hortal (2013) notice that LAWS can be used in future similarly to the ways in which drones are already used today. For example, their deployment could become normalised in counterterrorist activities. Moreover, the lethal use of force both via drones or LAWS can constitute inhuman or degrading treatment (Espada & Hortal, 2013: 15, 23).

Comparison is a useful resource for research that centres on the prospective regulation of LAWS. Previous cases of weapons regulations, such as the landmines and cluster munitions, set the ground for understanding the ways in which LAWS might challenge the international legal system. For example, the principle of distinction led to the regulation of landmines in the past, and today it is key in the legal discussions on LAWS. A comparative analysis can show why and how LAWS could be regulated. In addressing the 'how', a useful comparative approach can focus on the regulative processes themselves. Understanding the way in which past cases of arms regulations were concluded contributes to the envisioning of the normative future of LAWS. Akkuş (2023) focuses on LAWS and the 'legal transplants', the processes in which regulations related to one case are customised, adapted and adopted for other cases. This approach allows him to compare the case of LAWS with the regulations related to landmines, incendiary weapons, and cluster munitions. He then suggests a regulatory framework based on those past experiences. While Akkuş concludes by recommending the development of an IHL manual that would cover the main issues related to LAWS (2023: 139), a similar comparative approach leads Kayser to determine that the future of LAWS is a legally binding treaty, which could be initiated by a smaller group of states and later embraced by the international community (2023: 4).

An interesting and perhaps less common angle of comparison is the one between LAWS, on the one hand, and blinding lasers and exploding bullets, on the other hand. In the two latter cases, regulations were developed pre-emptively, "before its use on the battlefield and before a stream of victims gave visible proof of its tragic effects" (ICRC, 1995). With limitations, a similar statement can be made in relation to LAWS: despite the current technological advancements, consensus remains that fully autonomous weapons do not exist yet. In this way, LAWS can be subject to a comparative analysis in regard to future-oriented and pre-emptive regulations. This approach also allows to distinguish between preventive and reactive regulations.

Rosert and Sauer (2018) show how, similarly to LAWS, blinding lasers were also understood as a futuristic matter at the time of their regulation. They outline the mixed perspective from which blinding lasers were examined, that went beyond a strictly legal approach. For example, psychological and socioeconomic consequences of their use were part of the deliberations within the CCW (Rosert & Sauer, 2018: 12-14). Research linking autonomous weapons and exploding bullets is rather limited, but relevant considerations are explored by Cass (2015). He notes how a pre-emptive regulation of a newly developed weapon allowed to set the line between lawful and unlawful uses of emerging technologies, thereby "allowing technology to progress and the technology's benefits to be recognised" (Cass, 2015: 1045).

Understanding LAWS through for and against arguments

While some challenges intrinsic to LAWS appear undeniable, so are the opportunities that they present for the military field. Etzioni and Etzioni (2018), adopted this approach in summarizing the positive and the negative elements surrounding LAWS. Among the advantages, the force multiplier factor is undeniable from the military perspective. Deployed instead of combatants, LAWS can preserve combatants' lives. LAWS can also reach further and do so faster; they can be less expensive if compared to the deployment of human

soldiers in the long term, and there are difficult or monotonous tasks that robots could perform better than humans. The authors also find ethical and moral foundations in support of LAWS: their deployment would reduce emotional biases in battlefield, such as the ones related to stress, revenge, anger, or fear. Arguments against LAWS are summarised from a moral standpoint: Etzioni and Etzioni quote positions from the civil society, UN representatives and engineers denouncing the dehumanizing nature of LAWS, risks related to autonomy and the challenge of determining accountability and responsibility (2018: 253-259).

In a similar approach, Solovyeva and Hynek (2018) call attention to six dilemmas related to autonomous weapons systems, addressing the issues of predictability, dehumanisation in decision making, depersonalisation of the enemy, the coordination between humans and machines, strategic matters, and the compliance of autonomous weapons with international law. Within each dilemma, pro and against arguments are identified among multi-disciplinary academic sources. In this way, the authors establish a dialogue between the proponents and opponents of autonomous weapons systems.

Differently from the above-explored approach, academic literature frequently presents focuses only on the "pros" or only on the "cons". From the military and the technological areas of research, arguments can be found in support of the development of LAWS. Toscano (2015) sustains that the debate should depart from the assumption that states will develop LAWS either way. For this reason, LAWS should be embraced along with the military opportunities that they offer, and controlled by a set of comprehensive technical requirements that would ensure their compliance with the international law. If technical conditions are met, "autonomous weapons can perform warfighting functions in compliance with international humanitarian law more effectively than humans" (Toscano, 2015: 224).

While Toscano's argument departs from the military desirability of the development of LAWS, Smith (2019) embraces the topic from the angle of moral permissibility, asking whether it is acceptable for military engineers to

participate in the development of lethal autonomous technologies. Using the just war theory, Smith answers positively: technological research and advancements are desirable because LAWS have the potential to comply better with the rules of war than human soldiers. For example, "LAWS will not degrade, rape, pillage, or kill for pleasure" (Smith, 2019: 285). Overall, Smith argues that the use of LAWS in warfare will imply a pareto risk improvement, in the sense that "at least one morally relevant class of individuals will see their risk profile improve and that no other classes are made worse" (2019: 284). The combatants would be subject to less risk, and the civilian population would not be exposed to the potentially criminal behaviour from the combatants.

A rather nuanced approach to the permissibility of LAWS is the evaluation of their uses. Schmitt (2012) holds that the focus should not be placed on lawful or unlawful weapons, but on lawful or unlawful uses. Quoting Seneca on the idea that a sword cannot be a killer but merely a tool, Schmitt develops an argument in favour of LAWS when those are used in compliance with the international normative framework (2012: 1, 35). Sassòli's early research on LAWS adds on this that, in some circumstances, LAWS can not only be used in compliance with international law, but also perform better than human soldiers (2014: 320). Following this argumentative line, LAWS can be employed in tasks in which they equal or outstand humans in complying with the international humanitarian law.

Views opposing the development and use of LAWS also find a wide reflection in literature. In addition to the frequently employed moral arguments outlined by Etzioni and Etzioni (2018), researchers do not reach consensus on whether LAWS can actually be used lawfully. It is unclear whether weapons with sophisticated levels of autonomy and the capacity to kill comply with international requirements imposed by international humanitarian law and international human rights law. Dremliuga (2020) adds the issue of opacity. Opacity affects LAWS on three levels: first, the legal opacity refers to the intellectual rights protection over the algorithms that LAWS employ, and over

the military secrets. Such opacity could undermine the possibility of carrying out a comprehensive review of the machine's compliance with the law. Secondly, technical opacity makes LAWS hardly comprehensive for the end-users not trained in computer science. When the decision-making of a weapon is unclear for the operator and other actors, "even if the materials that explain the principles of the autonomous weapons system get published, no one outside the immediate circle of experts would be able to understand them" (Dremliuga, 2020: 119). Thirdly, machines capable to learn autonomously might develop such levels of complexity that their processes become uncomprehensive even for their very developers. The latter type of opacity will only grow if autonomy in weapons continues to increase with new technological advancements. For this reason, Dremliuga urges for the development of guidelines that would secure a degree of human participation in the deployment of LAWS.

Addressing the rather recurrent argument that 'LAWS will not rape', Sandvik and Lohne (2015) outline the problematic nature of this approach. References to rape are seen by the authors as a new instance of instrumentalization of sexual crimes with the purpose of developing policies and legislations aligned with the interests of governments and lobbies. Therefore, references to rape in arguments supporting LAWS create utopian visions of technologically enhanced wars and misleading narratives, where rape appears instrumentalised in the same way as "online child pornography has for a long time been used to legitimate widespread government censorship and surveillance in cyberspace" (Sandvik & Lohne, 2015). Too from a feminist perspective, Acheson (2022) studies LAWS within the patriarchal context and current political structures. This approach is informed by the consideration that both the military and the technological fields are male-dominated, and patriarchal structures can easily transpire to the developing autonomy in weapons. Patriarchy might affect the different levels of control and oversight over LAWS, and also the databases that the systems will use to select the target, reinforcing "existing norms of gender and power" (Acheson, 2022: 8).

Databases and biases also rise concerns in regard to the security of persons with disabilities. This risk has been foreworn by the UN Special Rapporteur on the Rights of Persons with Disabilities in 2021 (Quinn, 2021). Figueroa *et al.,* (2023) point out that persons with disabilities are not included in the international deliberations over LAWS. This only contributes to the biases that might affect the databases employed by the autonomous weapons. The authors prepare a set of examples: a weapon might fail to recognize non-verbal communication; a light signal can be misunderstood by a person suffering colour-blindness, and a special assistance device such as a walker might lead the weapon to wrongfully identify a civilian as a weapon bearer (Figueroa *et al.,* 2023: 287-288). These and other research cases show some of the risks associated with LAWS. Among further concerns are the risks of proliferation and nuclearization of LAWS, addressed, for example, by Szpak (2020: 126) and Boulanin (2019: 88-89). Altogether, arguments in favour and against LAWS are subject to the uncertainty, inevitable in addressing emerging technologies. Together with the concerns and opportunities explored above, stand the purely legal considerations, some of which have already been pointed out, and that will be explored in the section that follows.

Understanding LAWS through laws

Legal analyses on LAWS are common, enriched by a variety of perspectives and entertained by the many combinations in which the words 'LAWS' and 'laws' can fit together in the title of an article. Legal discussions are naturally predominant in the international deliberations and interstate debates on LAWS. The CCW is intrinsically connected with IHL, and LAWS, as weapons, are clearly subject to IHL when employed in international or non-international armed conflicts. Beyond IHL and its principles, legal discussions also address the compliance of LAWS with, *inter alia*, the IHRL and other international law principles, such as the principle of respect for territorial integrity of the states. The normative approach also allows researchers to assess the adequacy of the

current international framework to the emerging technologies. A variety of publications propose solutions in regard to the interpretation of currently existing principles and sometimes the development of new ones.

Legal research typically departs from the following position: it is clear that IHL applies to LAWS, but it is unclear whether LAWS are capable to comply with IHL principles, nor whether the existing principles are sufficient to regulate emerging technologies (Righetti, 2018: 124). As Winter (2022: 10) reminds, the principles of IHL are not compilated nor grouped in a single comprehensive list, therefore the applicable principles must be identified for every specific case. In regard to LAWS, IHL principles that are most commonly explored are the principles of distinction, proportionality, precaution in the attack, humanity, and military necessity.

The principle of distinction receives particular attention from the academia. Compliance with this principle implies distinguishing "between military objectives and civilian objects, combatants and civilians, and active combatants and those hors de combat" (Davison, 2018: 7). On the technological side, it would require LAWS "to possess advanced skills in observation and recognition as well as sophisticated judgement-making ability" (Winter, 2022: 13). Autonomous weapons would need to match the human capacity to assess the context surrounding them and to interpret the verbal and non-verbal signs that allow to distinguish between the different statuses of persons involved in armed conflicts, in order to decide whether this person can become a lawful target. Winter (2022: 15) predicts that such technology, necessarily enabled by the AI, would only be developed towards the year 2040 or later. Similarly, and according to Sharkey (2019), LAWS cannot currently comply with the principles of military necessity and proportionality. For example, the AI cannot determine "whether the military advantage to be gained [in a particular situation within an armed conflict] would justify the use of any form of attack near a school" (Sharkey, 2019: 76). Therefore, moderation by a human operator, in or on-the-loop, is currently essential for a weapon to comply with IHL.

The principles outlined above, along with the fact that some human functions might be delegated to autonomous weapons, set requirements related to predictability. Predictability is key to understand whether LAWS can comply with IHL (Sassòli, 2019: 517). IHL does not require the weapon to be predictable in every possible way and in all its functions. However, a certain degree of predictability is indispensable inasmuch the operators must be able "to understand whether the system will comply with the law" (Schuller, 2017: 421). Predictability can also be required as part of taking the necessary precautions in the attack (Cherry & Johnson, 2021: 19). This sets legal limitations for the functions enabled by machine-learning, both in present and in view of future development of LAWS.

The principles of humanity also apply to LAWS. They derive from the Marten's Clause, which can be summarised as, "the fact that there is no law prohibiting a weapon does not mean that its use is permitted" (Righetti *et al.,* 2018: 125). Depending on its interpretation in relation to LAWS, the Marten's Clause can be used as an argument against their development and use. First, LAWS lack compassion, an integral component of the principles of humanity. Secondly, the use of a lethal autonomous weapon against a human target can be seen as incompatible with human dignity (Saket, 2020: 823). Interestingly, the opposite position can also derive from the Marten's Clause: the use of lethal autonomous weapons systems becomes "a moral obligation if they allow the reduction of harm" (Righetti *et al.,* 2018: 125). Through the references to human dignity, the principles of humanity open the discussion to ethical considerations and the IHRL. If LAWS are incompatible with human dignity, their use against humans can constitute inhuman and degrading treatment (Espada and Hortal, 2013: 23). For Rosert and Sauer, it is clear that LAWS violate human dignity: "Treating a human as an object is what happens when LAWS are allowed to kill. The victim, be she combatant or civilian, is reduced to a data point in an automated killing machinery that has no conception of what it means to take a human life" (2019: 372). From an ethical perspective, if a weapon

autonomously selects and engages with a human target, this can be harmful for both: the victim and the society that produced the weapon.

The variety of understandings and interpretations that the academia provides in regard to the application of IHL to LAWS is far from reaching common grounds. For example, Winter (2022) believes that the principles of humanity and military necessity should not be seen as an obstacle to the development of LAWS. Instead, he urges to examine LAWS' compliance with the principles of distinction, proportionality, and precaution. Contrarily, for Foy, distinction and proportionality are a wrong angle for understanding autonomy in weapons. Jensen (2020: 578) holds that LAWS present no novel challenge to the principle of precautions in the attack, because this principle does not necessarily "require human judgment in targeting decisions". Overall, the relationship between autonomy in weapons and the international humanitarian law is perceived by legal scholars from different and, sometimes, opposing positions.

Understanding LAWS by understanding human control

Having seen the new challenges that come along with LAWS, academic research on autonomy in weapons and its degrees, definitional discussions, arguments in favour and against autonomous weapons and the legal principles involved in their regulation, this review of the literature concludes by returning to the initial point: the novelty of LAWS. This novelty lays simultaneously in the growing autonomy and the decreasing levels of human control. Research on LAWS includes works that address this balance, focusing on human control in order to better understand autonomous weapons. This branch of research is closely related to the CCW, with the academia providing suggestions on why and how the notion of human control should be incorporated into the discussions on LAWS. Mirroring the discussions at the CCW, research tends to focus on the notion of meaningful human control (MHC).

The relationship between autonomy in weapons and human control is a balance, or a sliding scale. In 2014, Roff was asking "How much human control ought we delegate?" (2014: 222). Two years later, Roff and Moyes were presenting their findings to the CCW, in representation of the organisation Article 36. They defend the concept of MHC as "a threshold of human control that is considered necessary" (Roff & Moyes, 2016: 1). They also explain that the adjective 'meaningful' was only employed to highlight the need to establish an acceptable threshold (Roff & Moyes, 2016: 6). In their view, MHC should become a policy approach, that would play a regulatory role when the use of force is executed through autonomous weapons. However, not any degree of control is sufficient for a weapon to be acceptable. For example, MHC is not in place if the role of the operator is reduced to pressing a 'fire' button (Roff & Moyes, 2016: 1).

Amoroso and Tamburrini (2020) — a lawyer and an engineer — explored the notion of MHC departing from its functions. They argue that MHC should combine three roles: a guarantee against malfunction, an element of accountability and a moral oversight of the weapon. Contrarily, an illusion of human control, where "the role of human operators [is reduced] to a perfunctory supervision of decisions taken at superhuman speed" should not be considered an acceptable threshold (Amoroso & Tamburrini, 2020: 190). As a policy recommendation for the CCW, Amoroso and Tamburrini propose establishing a taxonomy of weapons depending on their degree of autonomy. This should lead to the development of differentiated policies for the varying levels of human control. For example, in highly autonomous weapons that select and engage a target without human involvement, there is no place for the MHC; therefore, they should not be developed (Amoroso & Tamburrini, 2020).

Research on human control in LAWS ramifies into different spheres. Among others, human control requirements have an impact on the systems of values that guide the design of the new technologies. In order to understand the MHC requirements for developing value-sensitive design, Riebe *et al*., (2020)

performed a discourse analysis of the states' interventions at the CCW. They find that MHC depends on the reliability and the predictability of the weapon, and the time that the human operator has to intervene. This has implications on how the weapons should be designed. Moreover, research encountered conflicts of interest among the states, where the economy- and the military-oriented discourses were the most reluctant to integrate the notion of MHC in the interstate discussions on LAWS.

In May 2023, the initiative Automated Decision Research, launched under the framework of the coalition Stop Killer Robots, published a report on human control. This report constitutes a mapping of the convergences among states' positions on human control expressed during the March 2023 meeting of the GGE on LAWS. Researchers focused on three elements associated with human control: "Predictability, understandability, [and] temporal and geographic limitations on use" (Automated Decision Research, 2023: 5). Having assessed states' interventions at the GGE, they conclude that autonomous weapons must comply with the three elements. In addition to this, they observe alignment among the states and conclude that: "human control in the use of autonomous weapons systems - whatever terminology is used by each individual state - is of significant concern and importance" (Automated Decision Research, 2023: 18).

The report by Automated Decision Research sets the grounds for this thesis. First, it confirms the importance of human control as an appropriate angle for understanding LAWS, and of the GGE on LAWS as a central setting of interstate deliberations. Second, it sets an example of an analysis of states' positions, in a mapping exercise that simultaneously highlights the different understandings of human control and the states that hold these positions. Finally, it allows to set the scope of analysis beyond the already explored concepts of predictability, understandability, and temporal and geographic limits. In order to complement their findings, this research will focus on the normative aspect of human control and LAWS, in the way described in the chapter that follows.

Research design and methodology

States and scholars refer to the notion of human control in order to discuss the political, technological, and legal challenges related to LAWS. However, this notion remains undefined on the international level. There is no common understanding of what human control is, and what it is necessary for. This thesis attempts to address this shortcoming. In doing so, it will pursue the answer to the research question:

"In which way does the notion of human control contribute to the understanding of the relationship between lethal autonomous weapons systems and the international humanitarian law?".

The research question incorporates different components, that lead to the development of two sub-questions:

1. What is human control?
2. Does IHL require human control and, if so, in which way?

The structure of this thesis is designed to mirror these sub-questions and, thereby, provide an answer to the research question. The first chapter will focus on the notion of human control, as it is currently understood by the states. The second chapter will address the legal considerations related to the requirement of a human element in IHL, examined from the perspective of autonomy in weapons. The cumulative outcome of the first two chapters will provide an understanding of the expectations of human control under the international humanitarian law. In this way, this thesis will fit in and enhance the academic research on LAWS by focusing on the notion of human control and employing the most recent data, retrieved from the 2023 meeting of the Group of Governmental Experts on lethal autonomous weapon systems.

Ontology and epistemology

This thesis is grounded on the assumption that social realities are not given but developed: the object of this research is a notion that does not exist *per se*, but that is arising in debates and literature, with a potential to become a new barometer of the legality of autonomy in weapons. Such an ontological approach leads to a constructionist view on epistemology. Rather than seeking objective empirical understandings, this research focuses on constructing meanings. This implies that the knowledge produced will not necessarily be objective. However, rather than a limitation of the research design, the lack of objectivity is a consequence of the ontological understanding of the reality, described above. Within constructionist or contextual research, meanings are inter-subjective, and a single objective reality does not exist (Madill et al., 2000: 9).

Different from interpretivism, which constitutes a separate epistemology, interpretative social constructionism "focuses upon meanings" (Chen et al., 2011:132). Interpretative constructionist epistemology aligns with the object of this thesis: a juridical notion, such as human control, will only find a meaning through interpretation. Interpretation is necessarily a human, mainly intersubjective, exercise. Therefore, the answer to the research question should not be sought in an empirical reality but rather in the construct of human understandings. The methodology employed in this thesis has been developed based on these constructionist assumptions on reality and knowledge.

Methodology

By exploring the relationship between epistemology and methodologies, Darlaston-Jones (2007) establishes a natural link between constructionism and qualitative methodologies. Qualitative methodologies are better suited to inquire into the causes and the underlying conditions of a matter; they "provide the means to seek a deeper understanding and to explore the nuances of experiences not available through quantification" (Darlaston-Jones, 2007: 25). This thesis aligns with this idea, and hence employs qualitative methodology, with a complementary quantitative element.

In addition to this, the qualitative methodology here employed is intra-paradigmatic. Intra-paradigm research refers to a methodology where two or more qualitative methods are combined (O'Reilly *et al.*, 2021: 66). This methodological approach is comparable with mixed methods. However, mixed methods require a combination of quantitative and qualitative methods. Even though some samples of data will be analysed quantitatively, the methodology of this thesis is predominately qualitative. Therefore, it is intra-paradigmatic rather than inter-paradigmatic.

In order to understand what expectations of human control are expressed by states and could be required by IHL, two qualitative data analysis methods will be used. This is reflected in the structure of the thesis, where the first two chapters employs their own data set and a specific method of data analysis. Data will be analysed separately, but in a sequential, cumulative manner as depicted below (figure 1).
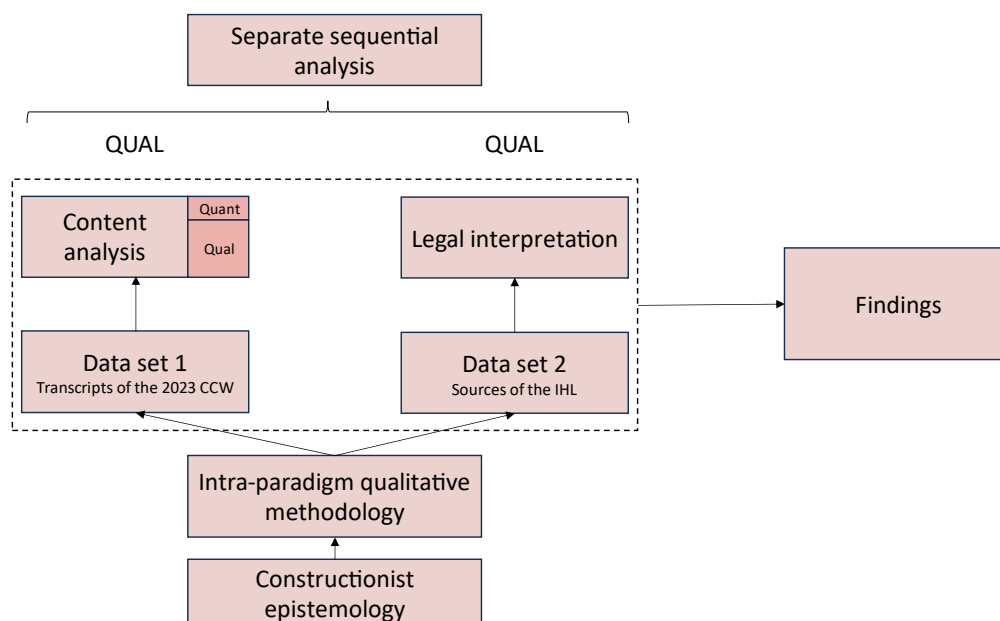

Figure 1. Research design, including the methodology and the methods used.

Methods

Content analysis

Social and political understandings of what human control should mean can be retrieved from a variety of sources. Here, the focus is placed on states. This thesis will analyse the 2023 meeting of the Group of Governmental Experts on lethal autonomous weapon systems established within the CCW. This is a specialised and renown forum, with over a decade of experience in LAWS, where the notion of human control has been discussed substantively. Contributing to the data collection, the meetings of the GGE are transcribed. This allows to collect the data through textual analysis, with the textual transcripts gathered from the UN Digital Recordings Portal (UN Geneva, 2023). Transcripts are available for each of the twenty sessions that took place in 2023, and they will be divided into units of analysis. Data sampling and analysis will take place in two phases. The same qualitative textual data (the transcriptions of the GGE sessions) will be analysed twice: first, from a quantitative and then from a qualitative perspective. Such a sequence corresponds to explanatory sequential design, where qualitative analysis is developed over the findings of a quantitative analysis (Snelson, 2016: 9). Software-wise, the tool used to conduct this method is Atlas.ti, that allows manual coding of large text samples.

First, the development of units of analysis will take place linearly, following the timeline of the sessions. The 2023 meeting of the GGE on LAWS took place in March and May, over ten days, with ten morning and ten afternoon sessions. Each unit of analysis will correspond to one session. This approach will allow to see the evolution of the discussions on human control in LAWS, from a quantitative point of view. Inductive conceptual content analysis, where the presence of a concept is quantified, will allow to see how many times human control was mentioned during the sessions, and establish quantitative patterns. For example, inductive coding will allow to see what qualifiers or adjectives states tends to use when referring to human control. Quantitative content analysis will reflect the natural timeline of the sessions, setting the ground for qualitative analysis.

Secondly, the same sample of text will be analysed through inductive qualitative content analysis. As opposed to quantitative content analysis, qualitative analysis "goes beyond merely counting words or extracting objective content from texts to examine meanings, themes and patterns that may be manifest or latent in a particular text" (Zhang & Wildemuth, 2009: 2197). It also allows to underline both the message and the creator of the message (Baxter, 2020: 394). Here, the sample will be divided into units of analysis on a state-by-state basis, ensuring that each unit of analysis corresponds to the aggregation of the interventions of a state. By doing so, the research aims to understand how states address and understand human control in LAWS and focus simultaneously on the states and on their positions. The development of a new set of texts will allow to show how many times each state referred to the examined concepts, rather than the number of times that a concept has been pronounced during a particular session.

Inductive qualitative content analysis is an appropriate method to developing concepts (Kibiswa, 2019: 2061). Both the non-randomised selection of the text samples and the inductive approach to coding are consistent with the characteristics of qualitative content analysis (Zhang & Wildemuth, 2009: 2197). As this thesis aims to develop an understanding of the notion of human control, coding will be conducted inductively. This will allow to identify all the relevant positions related to the wider spectrum of human participation in relation to LAWS. Inductive coding means that codes will not be developed as a first step that precedes the analysis. Instead, coding will take place during the analysis, over the sample of text that will be analysed. Findings obtained through qualitative content analysis will be presented in a synthesised manner and organized thematically.

Legal interpretation

The second chapter will introduce a normative approach. Focus will be placed on interpretation, understood as a method of qualitative legal analysis. The use

of qualitative methods in legal research is consistent with the epistemological assumptions of this thesis: the findings produced through this research are not expected to be natural, but constructed in a certain context and within a certain system of values (Mitchell, 2023: 105).

This chapter will focus on the international legal framework. While, in the previous step, content analysis will have provided a minimal interstate understanding of the notion of human control, the normative analysis will focus on the relationship between this concept and IHL. As shown in the figure 1, the data set for legal analysis is typically the combination of sources of the international humanitarian law. Sassòli (2019) identifies these sources as follows: treaties, customary law, principles, unilateral acts, and hybrid soft law instruments, such as IHL manuals and the commentaries by the ICRC (Sassòli, 2019: 65). This thesis will focus on the treaty and customary law in order to understand whether the requirement of HC derives from the IHL and, what IHL provisions require a human element comparable with human control. Legal interpretation will depart from identifying the rules and principles applicable to LAWS. After this, from within the rules and principles identified, it will single out the rules that have a direct repercussion on human control.

This exercise will allow to see what the notion of human control is necessary for, and therefore what features it should integrate in order to cover these needs. A common practice in normative analysis is to employ footnotes in order to quote legal sources (Balkin, 1988: 276). For this reason, and consistently with this practice, footnotes will be used in order to refer to treaty law, customary law and jurisprudence.

Limitations

This research acknowledges a number of limitations that must be considered in assessing the findings. First, the scope of this thesis is not all-embracing. While analysing states' positions, this research will not delve into the historical, social, political, or other reasons behind these positions. It will not focus, for instance,

on power relations between the global north and the global south, despite the undeniable interest of such an approach, demonstrated, for instance by Bode (2019). It will examine states' positions verbally expressed at the CCW, but not states' practices. The scope of content analysis will be limited to the 130 states that are parties and signatories to the CCW, thereby excluding the positions of the rest of the states. Within the positions expressed, this thesis will focus on the meanings explicitly stated, but not on the discourse, on the implied meanings, nor the unsaid. Moreover, while centring on the states, this thesis does not extend the scope to the civil society or other relevant actors.

This thesis is written from the standpoint of awareness that the concept of human control is mainly derived from the positions of states that argue in favour of the development of this concept. However, this is not a consensual position within the international community. For this reason, the notion of human control is not designed to fit all the states' positions, but rather to propose a solution for the existing legal gaps related to LAWS. Moreover, while the notion of human control will be analysed in relation to IHL, this thesis will not extend to other, equally relevant considerations, such as ethics or International Human Rights Law.

Limitations stated above are primarily related to the extension of this thesis. Further research should address the topic of LAWS and human control in a deeper and a broader way.

Chapter 1: States' positions on human control at the 2023 meeting of the GGE on LAWS

Contextualisation: the CCW and the GGE

In 1983, the Convention on Certain Conventional Weapons (CCW), entered into force (United Nations Treaty Collection, 2023). The purpose of the CCW is "to ban or restrict the use of specific types of weapons that are considered to cause unnecessary or unjustifiable suffering to combatants or to affect civilians indiscriminately" (UNODA, 2023). Some examples of such weapons, regulated by the CCW Protocols, are mines, booby traps, incendiary weapons, and blinding laser weapons. The High Contracting Parties to the CCW hold regular meetings every year, and every five years. In addition to this, within the CCW framework, Groups of Governmental Experts can be created to discuss specific topics in representation of states. Participation in the GGE is not limited to states, but also to the "international organisations, civil society, academia and industry" (Jacobson, 2017: 2).

Starting in 2013, the High Contracting Parties have been discussing LAWS at the annual meetings. This led to the establishment, in 2016, of the GGE related to emerging technologies in the area of LAWS, who met for the first time in 2017 (Jacobson, 2017: 2). Up to the date, the GGE on LAWS has held seven meetings, each lasting over one, two or three weeks. The last meeting of the GGE took place in Geneva in March and May 2023.

The 2023 session culminated in a final report, summarizing the positions adopted during the GGE meeting. The report of the GGE is consensus-based; therefore, it does not integrate all the topics discussed during the session. In regard to human control, the final report only notes that "Control with regard to weapon systems based on emerging technologies in the area of LAWS is needed

to uphold compliance with international law, in particular IHL, including the principles and requirements of distinction, proportionality and precautions in attack" (GGE on LAWS, 2023[b]: §21c). In this way, the final report does not reflect the richness of the discussions on human control, both because of the lack of consensus and due to the lack of time to reach this consensus during the sessions. This thesis will analyse the transcription of the meetings, in order to see in which ways states addressed human control.

Data set, the units of analysis and coding

As noted in the methodological section, data will be analysed following explanatory design: through quantitative content analysis first, and through qualitative content analysis second. The first approach will allow to depict the general timeline of the sessions, looking into:

1. How many times mentions to human control appear;
2. What qualifier or adjective, if any, accompanies the references to human control.

The second, qualitative, approach will focus on the message and the author of the message. This will be done by reorganizing the sample into new units of analysis, each of which will correspond to an aggregate intervention of a state. This section will explain the following:

1. Which states took position on the issue;
2. What positions have been expressed in regard to human control?

Altogether, this approach aims to reply to the first sub-question: "What do states understand under the notion of human control?". Following this structure, two samples based on the original data set were created. The first sample consists of 20 units of analysis, corresponding to each of the sessions that took place in March and May 2023. Each session lasted up to three hours, thus producing a transcription document of up to 20000 words. The second sample consists of 51 units of analysis, corresponding to every state that intervened at the GGE and the European Union. 14 non-state actors that

intervened at the GGE were excluded from the scope of this analysis. In both cases, coding was carried out inductively. This led to the development of 12 codes for quantitative analysis and 21 codes for qualitative analysis. For the visualisation of data, tables related to content analysis will be sited in the text, while figures and illustration can be found in the annex. The codebooks are also available in the annex (Tables 1 and 2).

Analysis

Linear analysis

Once the 20 units of analysis corresponding to each session were uploaded to Atlas.ti, conceptual quantitative content analysis allowed to trace every instance in which the phrase "human control" appears. Codification took place in a combination of automated and manual tracking. This allowed to establish that states made references to human control 504 times. Most of these references took place in March (Table 3).

	March sessions 10 591	May sessions 10 118	Totals
● Mention "human control" 504	424	80	504
Totals	424	80	504

Table 3: number of references to "human control". Source: own elaboration using Atlas.ti.

Quantitative analysis also shows how discussions involving human control evolved throughout the sessions (Figure 2 in the annex). 41% of the references to human control took place during the fifths and the sixths sessions. In May, mentions to human control decayed progressively, consistently with the idea expressed by the chairperson and a number of states that the GGE needed to focus on the final report and consensual positions.

Over five hundred references to human control denote the importance of this topic for the states. For comparison purposes, the term "autonomous weapons" has been used 413 times within the same sample, even though this number should be seen in light of the many denominations that LAWS receive.

In the word count of the entire sample, the words 'human' and 'control', taken separately, occupy the fifth and the tenth places, respectively, together with the words 'system', 'weapon', 'delegation', 'law', 'autonomous', 'paragraph', 'international' and 'proposal' (Figure 3 in the annex).

Inductive coding also allows to understand what qualifiers states tend to use when speaking about human control in relation to LAWS. Qualifiers and adjectives allow to better understand what states mean when they speak of human control, and what kind of human control they pursuit in relation to LAWS. 11 codes were developed inductively to cover all the references to 'human control' that were preceded by a qualifier. The eleven qualifiers identified through inductive coding and their occurrence are displayed in the table 4 below.

		March sessions 10 591	May sessions 10 118	Totals
◇ Appropriate human control	22	14	8	22
◇ Appropriate sufficient human control	1	1		1
◇ Diminished human control	1	1		1
◇ Due human control	1	1		1
◇ Full human control	2	2		2
◇ Meaningful and effective human control	1	1		1
◇ Meaningful human control	167	138	29	167
◇ Necessary human control	2	2		2
◇ Rigorous human control	1	1		1
◇ Significant human control	2	1	1	2
◇ Sufficient human control	5	5		5
Totals		167	38	205

Table 4: qualifiers used to refer to human control. Source: own elaboration, using Atlas.ti.

Qualitative analysis shows that, by far, the most common qualifier related to human control is the adjective 'meaningful'. The phrase 'meaningful human control' has been pronounced 167 times during the March and May sessions of the GGE. The second and the third positions are occupied by the words 'appropriate' (22) and 'sufficient' (5). In 299 cases, the phrase 'human control' was employed without any qualifier. In the section that follows, qualitative content analysis will allow to understand what states mean when they speak about meaningful, appropriate, sufficient, or other kinds of human

control; what expectations they place on this term, and what positions are the most supported within the GGE.

State by state analysis

The qualitative analysis adopted in this dissertation is based on a renewed set of units of analysis, which were developed manually, creating a unit every time a state intervenes, and aggregating all the interventions of the same state to the unit. The resulting 51 textual units (50 states and the European Union) show which states made the most references to the notion of 'human control'. As a prior quantitative step, the analysis shows that out of the 51 intervening states, 44 mentioned 'human control' at least once during the twenty sessions of the GGE. Most of the times, human control was mentioned without any qualifier. 34% of the mentions to human control referred specifically to 'meaningful human control', while 5% referred to 'appropriate human control'. The states that spoke about human control the most are the Philippines (53 mentions), the United States (35), Mexico (23), Palestine (22), Austria (21), Ireland (20) and Norway (20) (Figure 4 in the annex).

The codes, developed inductively, allow to outline elements of human control, retrieved from the positions of states. In addition to this, they show states' approaches to human control, and what type of considerations guide their reasoning.

Definitions and role of HC

Proposals on the definition of HC point out a series of secondary debates related to its role and the key elements. Some of those elements are common for the members of the GGE. For instance, states tend to agree that human control is a notion that can have varying degrees of intensity. One of the most wide-ranging definitions is proposed by the Philippines, where HC is understood

> "as a threshold of application of human judgment and intervention necessary to ensure the maintenance of human agency, respect for

42

responsibility, proportionality and accountability in undertaking decisions, regarding the use of any weapon, as well as the ability of human operators to effectively supervise any weapon, undertake the necessary interaction that could either be directive or preventive and to the activate, terminate or abort the operation of the weapon altogether".

Outlined in these terms, the notion of human control is compatible with weapons in-the-loop and on-the-loop.

Delegations supporting the development of the notion of human control tend to explain its substance by describing what HC must ensure. The Philippines complements its definition by outlining the need for predictability, that encompasses, in their interpretation, understandability, explicability, and traceability. The delegation of Palestine holds a similar view, expecting LAWS to be "predictable, reliable, understandable, and explainable and traceable". This means, *inter alia*, that the operators must be able to anticipate the performance of the weapon and understand its decision-making processes. The weapon must perform as intended and expected. Germany and Denmark reinstate that reliability is a key element of human control. In order to ensure it, Denmark outlines the need to develop specific trainings for the operators of LAWS. Mexico adds to this that, for a weapon to be reliable, it cannot autonomously redefine the parameters of its mission without human validation.

Also in regard to the functionality of HC, Argentina defends that human operators must be able "to define and validate the use of the systems and take critical decisions on the use of force". This position is supported by Ireland and Cuba. The latter refers to the selection and engagement with the target as the critical functions in which human control is indispensable. The standard of sufficient human control proposed by France indicates that operators must be "able to take a decision to deploy a weapon". This idea coincides with the directive control explained by the Philippines. Human-in-the-loop weapons can comply with this standard. Moreover, France suggests ensuring human control by setting temporary and regulatory limits to the development and the use of LAWS. Rather than human control, the United Kingdom advocates for fostering

context-appropriate human involvement. Similarly to France's proposal, human involvement must be ensured by pre-set temporal and also geographic limits for the deployment of LAWS. To these limits, Switzerland adds a third type of restriction, that would confine the use of LAWS to specific targets. Software and hardware specifications should ensure that the performance of LAWS is restricted in compliance with IHL, as presented by Tunisia.

Another way of understanding HC is pointing out what it is not. For this end, Palestine instrumentalised the concept of nominal human input (NHI), defined in the working paper as "an input performed by a human that does not materially affect the autonomous process" (State of Palestine, 2023: 3). An example of NHI is a button that an operator presses without exercising sufficient human judgement. A risk associated with the NHI is the automation bias, where operators overly relay on the weapons. Palestine defends that it is necessary to make sure that the required level of HC goes beyond the NHI. This position is supported by a number of states, including the Netherlands, Mexico, Switzerland, and, to some extent, the United States.

Finally, it is possible to illustrate some states' understanding of human control through a debate that arose between the Philippines, the United States, Uruguay, and Norway. The four states referred to a bullet in order to understand if HC exists in widely and currently used weapons. For the United States, the bullet analogy shows that "almost every weapon or munition at some point during its use is outside of human control: a bullet cannot be controlled after its fired". To this, the Philippines responds that a bullet, on the contrary, represents an example of MHC inasmuch it is triggered by the action of a human operator, who controls the effects of the use of the bullet by determining in which space and moment it will be fired. This leads the Philippines to incorporate another element to their definition of MHC: "meaningful human control is contextualised in ensuring limits over the use of force". Norway points out that "the intent of the use of a bullet" is precisely an illustration of the exercise of human control. Uruguay converges with this view, but notices that LAWS and

bullets are hardly comparable in regard to HC, because they involve completely different decision-making processes.

Human control and the qualifiers

As demonstrated in the quantitative analysis, mentions of human control are sometimes accompanied by qualifiers. Qualitative content analysis shows that states tend to be consistent in their way of referring to human control. At the same time, states do not tend to explain why they prefer one qualifier over others. France, Germany, Denmark, Australia, Bulgaria, and Norway are the states that show preference to the concept 'appropriate human control'. While the semantic implications of the term 'appropriate' were not discussed by any of these states, it is mentioned in combination with the idea that HC "should be retained during the whole cycle of the weapon system". France also employs the term 'sufficient human control', explaining that the human operator must "remain the only one able to take a decision to deploy a weapon" and continue "to define and validate any mission". The qualifier 'meaningful' is the most employed by states. States that, in one or another way, manifested the need to ensure meaningful human control are: Palestine, Philippines, Pakistan, Venezuela, Qatar, Algeria, Sri Lanka, Costa Rica, Austria, Belgium, France, El Salvador, Argentina, Cuba, Canada, Switzerland, Ireland, Uruguay, Ecuador, Mexico, Norway, Turkey, Kazakhstan, Colombia and Guatemala. Characterisations of meaningful human control include "predictability, reliability, understandability and explainability and traceability", together with the "human intervention and human judgement", and also "during the entire cycle of the weapon".

By reviewing the data state by state, it is possible to establish convergences. Content analysis did not encounter semantic differences between the three most used qualifiers. On the contrary, the adjectives meaningful, appropriate, and sufficient seem to refer to the same idea that human control must be qualitatively different from other types of human involvement. This can

also be seen in the fact that some countries, such as France and Norway, use the different qualifiers indistinctively. The semantic convergence was noticed by the states. The delegation of Argentina notices that the three qualifiers refer to the same need " to retain human control throughout the life cycle ensuring that operators have sufficient understanding of the systems and that they are in a position to define and validate the use of the systems and take critical decisions on the use of force". Qualitative content analysis confirms this view.

The object of human control

States offer different views on what the object of human control should be. The most common position is that human control should be exercised over the weapon or the machine. However, some states choose to use the formulations "human control over the use of force" and "human control over the effects" (Figure 5 in the Annex). For some delegations, this choice is intentional, because it leaves no space for confusion around the compatibility of autonomy in weapons, on the one hand, and a degree of human control over the same weapon, on the other hand. However, despite the fact that the three potential objects of human control have been identified, states did not engage in any discussion in this respect, nor provided explanatory remarks in their individual statements. Moreover, it has been noticed that some delegations use the different formulations indistinctively. For example, the Philippines, among the main proponents of MHC as a legal standard, referred six time to 'human control over the weapon', six times to 'human control over the use of force', and one time to 'human control over the effects' of a weapon.

The need for human control

37 out of the 51 intervening states verbalised the idea that human control is necessary, essential, or central to the regulation of autonomous weapons. The need for human control derives from three main types of reasoning: IHL considerations, ethical considerations, and security concerns (Figure 6 in the

annex). Intrinsic to the GGE settings, IHL considerations prevail significantly. The Philippines, Venezuela, Algeria, Palestine, Sri Lanka, Canada, Jordan, Denmark, Ireland, Bulgaria, and Colombia specifically expressed that HC is a requirement for a weapon to comply with IHL. Some other states raise questions instead of statements. For example, Belgium and Mexico inquire whether weapons that operate outside of human control can actually comply with IHL. The positions of the abovementioned states are grounded on the fundamental IHL principles, including distinction, proportionality, precautions in attack, humanity, and military necessity. They also refer to responsibility and accountability, elements intrinsic to IHL.

States that articulated the requirement of human control based on ethical grounds referred to the need of human judgement in life and death decisions. Austria and Germany particularly emphasised that human control is an ethical requirement in critical functions of LAWS. They also mentioned values and moral imperatives. 14 states expressed, in one or another way, that weapons that target and engage autonomously are incompatible with human dignity. Finally, security considerations mentioned in relation to human control include the risk of unintended consequences, the "black-box problem" caused by unpredictable self-learning weapons, proliferation, cyberattacks, and acquisition by terrorist groups.

Conversely, some states sustain that HC is not the necessary nor the correct angle for deciding the regulations of LAWS (Figure 7 in the annex). Moreover, five states expressed the idea that HC is a distracting term, that only obfuscates the discussions of the GGE. For instance, the delegation of India urges the states not to "get distracted by terms such as human control, appropriate human control, sufficient human control, meaningful human control, et cetera, et cetera, which may lead to us swerving off the road as a group". Another pejorative view on HC, expressed by the Russian Federation and China is that HC is a politicised concept, rather related to political agendas than technical standards.

Human control as a criterion for prohibition

Among others, the Philippines specifically advocates for human control to become a criterion for prohibition. HC should become a legal threshold, in such a way that the compliance with MHC determines the legality of the use of a weapon. In their view, MHC acts as both a legal and a moral regulatory standard. It would also guarantee that there is a chain of human accountability and responsibility behind the performance of a weapon. Similarly, during the 2023 CCW, 18 states specifically stated that human control is or should become a legal threshold. However, this is a nuanced position, which leads to three types of debates.

The first debate centres on the relationship between HC and IHL. Three positions were identified. HC either is already a criterion for the prohibition of a weapon under IHL, or it is not yet but it should become such a criterion, or HC is not the right threshold for determining the legality of LAWS. Representing the first position, the delegation of Canada expressed their belief that "fully autonomous weapon systems that can operate without meaningful human control are already banned under international humanitarian law". Switzerland, Germany, and Norway maintain a similar point of view.

Rather than sharing the position that is already a criterion of prohibition, some states, including Argentina, maintain that human control "should be one of the criteria of prohibition". The delegation of the European Union adds to this that, if enforced, the requirement of human control "could strengthen human machine interaction compliance with IHL". Palestine specifically notices that the requirement of HC "is not currently codified in existing IHL", and advocates for the development of a legally binding regulation explicitly requiring HC. These positions express the need of a further development of the regulatory framework, by positivizing the requirement of HC.

The United States is one of the states that oppose the development of HC or MHC as a new legal standard. They recall that "international humanitarian law does not use the term human control nor contain a specific requirement that

weapons must be subject to human control". They also reinstate that "human control is not always resolved in compliance with IHL". Instead, they propose preserving an instrumental level of human-machine interaction and introducing precaution measures in the development and deployment phases.

The second debate focuses on the scope of prohibition. Within the sliding balance between autonomy and human control, some states propose concrete elements of human control that could underpin the normative requirement. Some of these elements, explored in the sections above, are the predictability of the weapon, the capacity of the operators to understand and predict its performance, human judgement, and the uninterrupted possibility of human intervention. Alternatively, instead of suggesting a concrete degree of HC that would determine the legality of the weapon, some states propose a binary approach, focusing only on the presence and the absence of HC. France and other states advocate for the prohibition of fully autonomous weapons that operate without any kind of human control. They promote a two-tier approach to LAWS, where fully autonomous weapons are prohibited, and regulations or positive obligations are developed for semiautonomous weapons. The Philippines, Austria, El Salvador, Cuba, Switzerland, and Ireland propose to establish HC as the criterion for prohibition under the first tier. However, this position is only one of the different proposals. Even though a two-tier approach has been discussed during the 2023 CCW and supported by 26 states and the European Union (Figure 8 in the annex), there is no consensus on where the limits of permissibility should lay, and whether HC should be the key criterion.

The third debate focuses on the GGE report and the format in which new regulations could be developed. The Philippines, Palestine, Austria, Cuba, Denmark, Norway, and Guatemala expressed their preference to include considerations on human control in the 2023 GGE final report. The opposite position was expressed by the Russian Federation, the United States, Australia, Belgium, and Israel. In addition to this, Philippines, Palestine, Austria, El Salvador, Cuba, and Ecuador advocate for the development of a legally binding

instrument that would regulate LAWS on the basis of human control. The coding scheme here was restricted to the statements in which the development of a new legally binding instrument was suggested specifically in relation to human control. The legally binding instrument could be either complemented or substituted by a code of conduct and good practices (as suggested by Turkey, Australia, Ireland, and the delegation of the European Union), confidence-building measures (United States, Cuba, Ireland, and Bulgaria) or a non-binding declaration (India, Venezuela, and the Netherlands). As an alternative to an international regulation, the Russian Federation and Pakistan expressed the preference to regulate LAWS internally, in national legislations.

Discussion

Content analysis shows that, despite its prominence at the 2023 GGE debates, HC remains a contentious issue. It is not possible to explain unambiguously what states understand as human control. The debate is fragmented, and it can be seen that the states that rely the most on the notion of HC are the ones that dedicate less percentage of their GDP to military spendings (World Bank, 2023). On the contrary, the United States, the Russian Federation and China— states that discard the notion of HC — are currently developing weapons with high levels of autonomy (Bode et. al., 2023).

In spite of these differences, the dialogue is open: in 38 instances, the delegations referred to other states' interventions when discussing human control. This demonstrates engagement and willingness to look for common grounds. Importantly, qualitative analysis did not encounter significant semantic differences between the notions of HC, MHC, and other qualifiers. During the 2023 CCW, Argentina and Switzerland noticed and appraised this convergence. Another convergence is related to the object of HC: some of the main proponents of HC as a legal standard use the expressions of "HC over the weapon" and "HC over the use of force" indistinctively. On the other hand, when explaining what HC is, states tend to start from the question of what HC

is necessary for. They accordingly define it through its role and functions, and those functions generally coincide. Altogether, these convergences indicate that common grounds are being built on a semantic level, even though the states employ different vocabulary.

The states' convergence on the two-tier approach is celebrated by the observers of the GGE (Bode & Watts, 2023). The two-tier approach does not necessarily centre on human control as a criterion for prohibition. Yet, depending on the evolution of the debates at the GGE in the upcoming years, it could. However, this would not necessarily lead to an effective regulatory framework. The Philippines warns against the risk of setting such a high threshold of HC that the prohibition becomes vane. In a context where LAWS are not defined on the international level, the prohibition of fully autonomous weapons that operate without any kind of human intervention would remain future-oriented in the best case, or simply ungrounded. This is one of the reasons why common understandings and definitions are of utmost relevance in regulating LAWS.

By combining the elements that states suggested during the GGE meeting in 2023, it is possible to propose a characterisation of HC. Human control is a notion that derives from the requirements of international law and particularly from IHL. It is also a moral imperative and a security guarantee. HC involves a human judgement in monitoring the weapon's critical functions. HC can be exercised in weapons where the operator is in or on the loop. HC is not a binary concept but a degree, and a degree of HC is required for a weapon to comply with IHL. In particular, and as manifested in the final report (GGE on LAWS, 2023[b]: §21c), HC is needed for complying with the principles of distinction, proportionality, and precautions in attack. HC also acts as a cohesive element in the chain of responsibility and accountability.

The analysis of states' interventions at the 2023 meeting of the GGE on LAWS shows what is understood under the notion of HC. The second chapter

of this thesis will provide a legal analysis in order to understand the relationship between this concept and IHL.

Chapter 2: Autonomous weapons, International Humanitarian Law, and the requirement of a human element

As observed in the previous chapter, IHL is the main legal framework that states employ to discuss the legality of LAWS. IHL is applicable to LAWS when they are used in conflicts. Analysis of the express wording of existing IHL treaty sources does not provide an answer to the question on whether IHL requires any sort of human control. IHL does not expressly require human control, and the expression 'human control' is not codified. This notwithstanding, this chapter will analyse whether IHL requires a comparable human element under other denominations, or as an implicit or embedded concept.[2]

Article 38 of the Statute of the International Court of Justice establishes that the main sources of international law are international conventions, international custom, and general principles of law.[3] In the case of IHL, the main treaty sources are the Hague Regulations of 1899 and 1907, the Geneva Conventions of 1949, the Additional Protocols to the Geneva Conventions of 1977 and the Convention on Certain Conventional Weapons of 1980. Moreover, customary law plays a fundamental role in IHL because it fills the void that is left by the lack of ratifications of some treaty sources, in particular the Additional Protocols. In 2005, the International Committee of the Red Cross published a study on customary international humanitarian law, which identifies 161 IHL customary rules that apply both in international and non-international conflicts, regardless of state parties' ratification of Conventions and the Protocols. For the purpose of this thesis, the analysis of customary law will be confined to the ICRC study.

---

[2] As explained in the chapter on methodology, and following the usual practice of normative research, legal sources employed in this chapter will be referenced in footnotes.
[3] Statute of the International Court of Justice (1945). Article 38.

States' interventions at the 2023 CCW outline several key IHL principles that are challenged by LAWS. The final reports specifically urge to ensure LAWS' compliance with the principles of distinction, proportionality, and precautions in the attack (GGE on LAWS, 2023b: §21c). These principles represent the spirit of IHL and reflect the main treaty and customary sources. In addition to this, the possibility to determine the responsibility of states and the accountability of individuals are essential for the compliance with IHL. This chapter will consecutively analyse each principle outlined above in order to understand what they entail and whether HC is required for their respect. After this, the requirements of HC will be contrasted with the degree of autonomy in weapons. For this, the models human in-the-loop, on-the-loop, and out-of-the-loop will be used.

Analysis

The application of the IHL to autonomous weapons

As stated above, IHL is applicable to LAWS when they are used in international armed conflicts (IACs) and non-international armed conflicts (NIACs). However, "only human beings are subject to legal rules" (Sassòli, 2014: 323) and "only humans are addressees of IHL" (Sassòli, 2019: 519). Therefore, it is not the autonomous weapons but the humans who are bound by legal obligations. Despite their growing autonomous functions, weapons are means of combat, not subjects. It is in this capacity that they must be compatible with IHL.

Distinction

The principle of distinction is contained in the articles 48, 51 and 52 of the Additional Protocol I of 1977. It establishes that only military objectives can be attacked in the conduct of hostilities. The civilian population and civilian objects must not become an object of attack. Civilians and combatants are two mutually

excluding categories: anyone who is not a combatant is a civilian, and in case of a doubt, the person must be considered a civilian.[4] This should be read in conjunction[5] with the prohibition to attack a combatant who is *hors de combat*, in situations related to captivity, surrender or incapacity.[6]

The principle of distinction also sets the prohibition of indiscriminate attacks, understood, *inter alia*, as "those which employ a method or means of combat the effects of which cannot be limited as required by [Additional Protocol I]"[7]. The customary interpretation of this principle associates indiscriminate attacks with "weapons whose effects are uncontrollable in time and space"[8]. For LAWS, this sets a preliminary requirement of control, related to the predictability of the weapon's performance, and the temporal and geographical confines of its deployment. Compliance with the principle of distinction must be ensured "at all times".[9] This means that the necessary assessment must take place before the attack but also during the attack. Once deployed, weapons with autonomous selection and engagement functions must be capable to correctly recognize and identify military targets. On the contrary, "weapons that are incapable of distinguishing between civilian and military targets" must be never used.[10]

In autonomous weapons, distinction between civilian and military targets includes identification, recognition, and assessment of the environment, which are enabled through sensors. However, sensors do not suffice in order to correctly identify a person as a combatant. For instance, the identification and recognition of combatants cannot rely solely on their clothing and emblems.

---

[4] Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts (AP I) (1977). Article 50(1).
[5] ICRC, *Study on Customary International Humanitarian Law (Study on CIHL)* (2005). Rule 1.
[6] AP I (1977). Article 41.
[7] AP I (1977). Article 51(4)(c).
[8] ICRC, *Study on CIHL* (2005). Rule 12.
[9] AP I (1977). Article 51(2). ICTY, *Prosecutor v. Stanislav Galić,* Judgement and Opinion, Case No. IT-98-29-T, 5 December 2003, §45.
[10] ICJ, *The Legality of the Threat or Use of Nuclear Weapons*, Advisory Opinion, 8 July 1996, §78.

IHL requires combatants to distinguish themselves from civilians but recognizes the practical impossibility to do so in some situations related to the natural conduct of hostilities.[11] Moreover, requirements related to the emblems, uniforms and identifiable insignia are not internationally unified (Winter, 2020), thus setting a constraint for machine-enabled recognition. Interpretation and common sense (Sassòli, 2019) are necessary in order to correctly identify a lawful target. Status assessment involves multifaceted elements that a machine might not assess in the same way as a human. Some of those elements are the environment, the behaviour, and the intent. Scholars frequently employ the scenario of a child's play: it is unclear whether an autonomous weapon would rightfully assess the situation where a child displays toy weapons (Wagner, 2014: 1392). As noted in the literature review, LAWS might also fail to correctly identify a person with a disability. This bears implications for the principle of distinction, because a walking stick or a metallic wheelchair could be mistakenly identified as a weapon (Quinn, 2021).

Spatial or geographical restrictions contribute to ensuring that LAWS comply with the principle of distinction: if the performance of the weapon is restricted to a perimeter where there are no civilians, the evaluation of the personal statuses is simplified (Cherry & Johnson, 2020: 15). However, this measure is not necessarily sufficient. For example, LAWS might fail to recognize a combatant's wish to surrender due to the limitations in their assessment of circumstantial factors. It is unclear whether the weapon can always identify correctly a combatant *hors de combat*. In NIACs, where the status of combatants is imprecise,[12] the judgement required to identify the status of the persons might as well exceed the technical capacity of an autonomous weapon.

Similarly, the distinction between civilian objects and military objectives is not always self-evident (Sassòli, 2003). The military nature of an object is

---

[11] AP I (1977). Article 44(3).
[12] ICRC, *Study on CIHL* (2005). Rule 3.

determined by four elements, namely its "nature, location, purpose or use", which must be assessed in relation to their "effective contribution to military action"; in addition to this it is necessary to make sure that their" total or partial destruction, capture or neutralisation, in the circumstances ruling at the time, offers a definite military advantage".[13] Objects can have dual civilian and military use. Treaty law offers examples of objects that tend to have a civilian use, such as a school or a place of worship.[14] In case of doubt, they must not be attacked. The evaluation of the purpose and the use of an object is necessary for every specific case, and it requires complex context understanding and judgement. For example, a television and broadcasting stations have been in past considered to be military objectives, insofar the attack was "aimed at disrupting the communications network".[15] The legal debates surrounding this kind of attacks (Benvenuti, 2001) only confirm the fact that reasoning exercises required for the determination of a military target inherently involve a complex level of interpretation (Sassòli, 2003), hardly accessible to a machine.

Proportionality

The principle of proportionality is contained in the articles 51(5) and 57 of the AP I of 1977. It establishes the obligation to ponder the expected "concrete and direct military advantages" and the anticipated "incidental loss of civilian life, injury to civilians, damage to civilian objects, or a combination thereof"[16]. An attack conducted with autonomous weapons that disrespects the principle of proportionality is unlawful. In order to comply with the principle of proportionality, LAWS must be able to: comprehend the military advantage, evaluate the incidental civilian losses and determine whether the two elements are proportionate.

---

[13] AP I (1977). Article 52(2). ICRC, *Study on CIHL* (2005). Rule 8.
[14] AP I (1977). Article 52(3).
[15] ICTY, *Final Report to the Prosecutor by the Committee Established to Review the NATO Bombing Campaign against the Federal Republic of Yugoslavia*, 8 June 2000, §75.
[16] AP I (1977). Article 51(5).

In the military domain, proportionality is not established through a mathematic formula (Sassòli, 2019: 522). Anticipated civilian losses must be evaluated from the intersection of the principles of proportionality and distinction. As discussed above, sensor-enabled recognition might not be sufficient in order to correctly identify civilians nor foresee the incidental damage. In addition to this, proportionality assessment must consider complex reverberating effects of an attack, such as, for instance, the risk of attacking a dual-use object that provides services to the civilian population. On the other side of the proportionality balance, the evaluation of the expected military advantages is not confined to a specific part of the attack, but to the attack as a whole.[17] This means that a broader awareness of the military strategy underpinning the attack is necessary in order to assess the expected military advantages. Having assessed the two elements detailed above, the final evaluation of proportionality must be carried out with consideration to the principles of humanity and military necessity (Fleck, 2020). This sets another human-oriented element, because the principles of humanity represent the intersection between the legal and the ethical frameworks. In addition to this, and as noted by the Special Rapporteur on extrajudicial, summary or arbitrary executions, "Armed conflict and IHL often require human judgement, common sense, appreciation of the larger picture, understanding of the intentions behind people's actions, and understanding of values and anticipation of the direction in which events are unfolding. Decisions over life and death in armed conflict may require compassion and intuition".[18] Common sense, ethics, morality, compassion and intuition — some of the elements required for compliance with IHL and the assessment of proportionality — are inaccessible to machines.

---

[17] Rome Statute of the International Criminal Court (Rome Statute) (1998). Article 8(2).
ICRC, *Study on CIHL* (2005). Rule 14.
[18] Heyns, C. (2013). *Report of the Special Rapporteur on extrajudicial, summary or arbitrary executions, Christof Heyns*. A/HRC/23/47, §55.

Article 57(1)(a) of the AP I directs the requirement of proportionality assessment to "those who plan or decide upon an attack".[19] This wording implies a human element because IHL does not foresee any subject of decision making different than a human. It can also be inferred from the use of the relative pronoun "who", which mainly refers to human subjects. Van Den Boogaard establishes, based on ICTY judgements, that the human element lays in the figure of a "reasonable military commander", who must assess the proportionality of an attack in every specific case (2015: 22). For this end, human judgement is required because: "In determining whether an attack was proportionate it is necessary to examine whether a reasonably well-informed person in the circumstances of the actual perpetrator, making reasonable use of the information available to him or her, could have expected excessive civilian casualties to result from the attack".[20]

Article 57(1)(b) of the AP I establishes the obligation to cancel or suspend an attack if it fails to respect the principle of proportionality. This requirement entails a certain degree of human involvement, necessary in order to abort an attack at any point of time, even after it is initiated. For autonomous weapons, this sets a requirement of a continuous human monitoring, for an attack to be suspended whenever it does not comply with the principle of proportionality.

Precautions in attack

The obligation to take the necessary precautions is codified in the articles 57 and 58 of the AP I. This principle requires, *inter alia*, taking "all feasible precautions in the choice of means and methods of attack with a view to avoiding, and in any event, minimizing incidental loss of civilian life, injury to

---

[19] AP I (1977). Article 57(1)(a)(iii)
[20] ICTY, *Prosecutor v. Stanislav Galić,* Judgement and Opinion, Case No. IT-98-29-T, 5 December 2003, §58.

civilians and damage to civilian objects".[21] Once again, this instruction is directed to "those who plan or decide upon an attack".[22]

The principle of precautions in attack requires a choice of means and methods that minimize the damage inflicted to civilians. As Berrendorf and Bontridder (2023) recall, this choice must be made in good faith, an intrinsically human quality. In order to evaluate the possible damage, human commanders and operators must be able to understand and anticipate the performance of the weapon. However, sophisticated levels of machine-enabled decision making are inherently associated with a degree of unpredictability. In LAWS, this is frequently described as the black box problem, related to "the opaqueness of modern machine learning algorithms, which prevents any human from understanding their decision-making process" (Christie *et al*., 2023: 3). The use of unpredictable means of warfare is unlawful because it violates the principle of precautions in attack. Weapons whose performance cannot be anticipated by a commander must not be used. In addition to this, in the case of LAWS, the possibility of software malfunctioning and the vulnerability to cyberattacks (Solovyeva & Hynek, 2018: 174-175) must be taken into account in ensuring the necessary precautions in attack.

The responsibility gap

Common Article 1 to the Geneva Conventions commits states "to respect and ensure respect" for the provisions of the Conventions.[23] Grave breaches of the Conventions must be repressed by states.[24] Sates are responsible for "all acts committed by persons forming part of its armed forces".[25] The duty to prevent, suppress and report such breaches extends to the commanders.[26] In

---

[21] AP I (1977). Article 57(1)(a)(ii)
[22] AP I (1977). Article 57(1)(a)(iii)
[23] Geneva Convention for the Amelioration of the Condition of the Wounded and Sick in Armed Forces in the Field (GC I) (1949). Article 1.
[24] AP I (1977). Article 86(1).
[25] AP I (1977). Article 91.
[26] AP I (1977). Article 87(1).

circumstances detailed in the AP I, superiors are subject to responsibility for breaches committed by subordinates.[27] These provisions are also reflected in customary IHL.[28] In addition to this, "IHL rules criminalised as war crimes" entail criminal responsibility under International Criminal Law (ICL). ICL foresees individual criminal responsibility. Violation of the principles of distinction and proportionality, explored above, can constitute a war crime.[29] In this way, IHL and ICL establish two forms of international responsibility, that refer to states, on the one hand, and individuals, on the other hand (Gaeta & George Jain, 2019). Within the spectrum of individual criminal responsibility, a chain of responsibility can be established implicating the individuals involved in the commitment of a prohibited act.

The development and use of autonomous weapons do not *per se* alter the chain of responsibility. As LAWS are not addressees of IHL nor ICL, they cannot be held responsible for wrongful acts. However, the complexity involving LAWS is related to the determination of responsibility in case of a breach of IHL. Autonomous functions can lead to a responsibility gap, where it is unclear who is responsible for the violation of IHL during the development and deployment phases (Davison, 2018).

Awareness and knowledge, on the one hand, and intent, on the other hand, constitute necessary elements for the determination of criminal responsibility.[30] For instance, for a commander to be held accountable for a wrongful act of a subordinate, the former must have known or have had information sufficient to conclude that the wrongful act would or could take place.[31] ICC requires criminal intent or *mens rea* for the determination of criminal responsibility. Criminal intent is compound of two elements, namely

---

[27] AP I (1977). Article 86(2).
[28] ICRC, *Study on CIHL* (2005). Rules 149, 151, 152, 153.
[29] Rome Statute (1998). Articles 8(2)(b)(i), 8(2)(b)(ii), 8(2)(b)(iv), 8(2)(b)(vi), 8(2)(e)(i), 8(2)(e)(ii).
[30] AP I (1977). Articles 86(2), 87(3). Rome Statute (1998). Article 30(1).
[31] AP I (1977). Article 86(2).

the intention "to engage in the conduct" and the intention "to cause that consequence" or the awareness "that it will occur in the ordinary course of events".[32]

If the deployment of LAWS leads to an IHL violation, it is unclear who holds awareness of the potential breach within the responsibility chain. This means that, if a war crime is committed as a consequence of deployment of LAWS, it will not be punishable unless criminal intent can be identified in humans responsible for the weapon. The three types of roles which are typically associated with responsibility over LAWS take place during the development stage (e.g. military engineers and programmers), activation stage and operation stage (Davison, 2018). Due to the technical complexity of LAWS, it is possible that the technical knowledge required in order to be aware of their performance is distributed between the three stages, or not accessible to the operators during the operation stage. Unawareness of the consequences of the deployment of LAWS excludes the intent to cause such consequences, required for the determination of *mens rea*.

Research has attempted to address the responsibility gap related to autonomy in weapons. Bo (2021) identifies doctrine and jurisprudence[33] where criminal intent was rather determined based on the awareness that the consequence "will occur in the ordinary course of events". This broader interpretation refers to *dolus eventualis*, understood as *mens rea* where the perpetrator of a prohibited act "foresees its occurrence as a necessary or certain or highly probable consequence of the achievement of his main purpose and nevertheless engages in the conduct" (Bo, 2021: 288). This interpretation means that recklessness can also lead to criminal responsibility. Bo develops this argument in specific regard to LAWS: the use of unpredictable weapons could be equated with reckless behaviour. However, she concludes that even the

---

[32] Rome Statute (1998). Article 30(2).
[33] ICTY, *Prosecutor v. Jadranko Prlić and others,* Judgement, Case No. IT-04-74-T, 29 May 2013, §192.

extensive interpretation of *mens rea* does not suffice in order to mend the responsibility gap related to autonomy in weapons. It is evident, therefore, that autonomy in weapons represents a challenge for the determination of criminal responsibility.

Discussion

The analysis of legal sources presented in this chapter leads to the conclusion that IHL, taken independently or in combination with ICL, requires uninterrupted human involvement in the conduct of hostilities. Human judgement, common sense and understanding of the context, use, purpose, and intent are necessary in order to distinguish a military target, being this a combatant or a military object. Complex assessments that include elements of morality and a general awareness of the military objectives allow to determine proportionality of an attack. The principles of distinction and proportionality require human participation preceding the attack, but also while it is in progress, because an attack must be aborted whenever compliance with IHL fails to be ensured. The principle of precautions in attacks outlaws the use of weapons with unpredictable effects.

Therefore, the human element is necessary in IHL. Undeniably, humans do not always comply with the rules of war. Compassion, common sense and moral values are many times absent in battlefields, while rage and vengeance take their place. This argument, identified in the literature review, aims to show the superiority of LAWS over combatants, for example, in the application of the principle of distinction. However, what should matter is the fact that humans have the capacity to be guided by morally grounded judgements. If a combatant fails to do so and thereby commits a violation of IHL, they can be subject to individual criminal responsibility. In this way, human elements such as compassion and morality are essential in IHL, while misconduct incompatible with IHL, even if grounded on intrinsically human behaviour, is criminalized. In relation to LAWS, this means that the 'rage and revenge argument' mentioned

by scholars such as Horowitz (2016: 29) is misplaced. In relation to LAWS, this means that if a violation of the principles explored in this thesis leads to a grave breach of IHL, the human element is necessary in order to determine criminal responsibility under ICL.

The arguments presented above demonstrate that unconstrained autonomy in weapons is incompatible with the provisions and the spirit of IHL. When LAWS are deployed, the human element is required in the activation phase and during the attack, in order to ensure compliance with fundamental IHL principles. The requirement of HC derived from existing IHL can be contrasted with the models HITL, HOTL and HOOTL. LAWS where human operator remains out-of-the-loop are incompatible with IHL. The HOOTL model implies that the critical functions of the weapon are uncontrolled. Hence it follows that if the autonomous weapon errs in selecting a target or fails to perform a proportionality judgement, its performance cannot be stopped by a human operator. The very use of this weapon is incompatible with the principle of precautions in attack, because its effects cannot be aborted or, in other words, controlled. If the use of HOOTL LAWS leads to a grave breach of IHL, the determination of criminal responsibility would be at the very least problematic, because the opaqueness of complex decision-making processes makes awareness of the consequences inaccessible.

HOTL and HITL models present a different and more disputed case. Findings of this chapter show that HOTL weapons, where a human operator has the possibility to stop the weapon if deemed necessary, do not necessarily comply with IHL either, for three reasons. First, the speed at which machine-enabled decision making operates cannot compare to the pace of human decision-making. This means that human supervision over the performance of the weapon can hardly suffice in order to timely detain an attack in progress, for instance, in case of reassessment of proportionality. Secondly, even if the human has the technical capacity to prevent the weapons' performance, the risks associated with automation bias mean that necessary human involvement can

lapse into a mere reactive response, that will not involve necessary human judgement nor context assessment by the operator. Thirdly, the model HOTL advances in the remedy of the responsibility gap but does not do so thoroughly. Even if a human operator has the role to supervise LAWS' critical functions, this level of involvement does not amount to the determination of criminal intent.

Finally, the HITL model, where the human participates in the selection and engagement with a target, are easier reconcilable with the IHL principles. In this model, judgements related to distinction and proportionality are reserved to the operator. This implies a greater level of predictability and understandability of the weapons' performance, in compliance with the principle of precautions in the attack. Awareness derived from the information accessible to the operator facilitates the determination of criminal responsibility in cases of grave IHL violations.

Indubitably, requirements related to a human element set a constraining framework for the development and the deployment of LAWS. This explains some states' reticence to adopt this or another regulatory framework in relation to the evolving autonomy in weapons. However, the assessment of the legality of new weapons is a legal requirement,[34] and this chapter has presented some of the considerations related to the human element that would determine the legality of contemporary and future autonomous weapons. In the pages that follow, IHL requirements will be contrasted with the notion of human control developed in the first chapter.

---

[34] AP I (1977). Article 36.

Chapter 3: Findings

Having assessed states' positions on human control and the IHL requirements related to a human element in lethal autonomous weapons systems, this chapter will contrast and combine these findings.

It is possible to observe alignment between the understanding of HC inferred from the 2023 meeting of the GGE and the requirement of a human element embedded in IHL. Coinciding with some states' positions, the need for human involvement does derive from IHL, as demonstrated in the chapter 2. This human involvement, explored in relation to the fundamental principles of IHL and individual responsibility, concurs with the categorisations that states provide in express mention to human control. For instance, HC involves human judgement and a constant human monitoring of the performance of the weapon. This is consistent with the findings of chapter 2. The normative analysis confirms that what some states define as nominal human input does not amount to a sufficient level of human control. NPI, closely associated with the automation bias, does not, for instance, ensure compliance with the principle of distinction because sufficient human judgement is not exercised in determining the status of a person. As states mention throughout the 2023 GGE meeting and conclude in their final report, human control is particularly relevant in regard to the IHL principles of distinction, proportionality, and precautions in the attack. The normative analysis confirms this finding. Overall, it is possible to conclude that the human element required by IHL coincides with the definition of HC drawn over the positions of states.

This is not a coincidence. Definitions and characterisations tailored by states under the CCW settings are grounded on IHL. This explains why findings of the first and the second chapters align. The three key contributions of this thesis are related to this alignment and will be explored in the following pages.

First finding: the requirement of human control derives from IHL

The analysis of states' positions acknowledges discussions on whether HC is required by IHL, should be required by IHL, or is not required by IHL. This thesis finds that, while HC is not explicitly required by IHL, this requirement is implicit. In particular, the principles of distinction, proportionality and precautions in the attack compel a degree of human involvement that coincides with states' understanding of human control. Compliance with IHL is incompatible with the lack of human control, based on the normative analysis, and this coincides with the positions expressed by a group of states.

Having this determined, this thesis advances in explaining how the notion of human control contributes to the understanding of the relationship between LAWS and IHL. HC participates in the regulation of LAWS by IHL, as an implicit requisite. The analysis of fundamental IHL principles allows to specify what HC is necessary for, thereby outlining the definition and the expected degree of human control. For the 2024 GGE meeting, this thesis suggests that states accept that HC is an implicit legal requirement. Once consensus is established on this ground, discussions should be channelled to the development of this notion. The rise of autonomy in weapons offers the opportunity to positivize the concept of human control, converting it in an explicit requirement oriented towards LAWS, but also towards other categories of weapons.

Second finding: human control can become the key criterion in the regulation of LAWS

Content analysis demonstrates the prevalence of HC as an element of the interstate discussions on LAWS. Normative analysis confirms the pertinence of this prevalence. Content analysis also shows the alignment over the characterisation of HC both among the supporters and the opponents of the

development of this notion as an explicit legal requirement. In view of the numerous convergences, demonstrated by previous research and by this thesis, it can be concluded that HC is a promising angle for future discussions on LAWS, that can lead to the development of new regulations.

Moreover, by placing the focus on human control, the GGE can achieve the intended technology-neutral approach to arms regulations, the pursuit of which was announced in the 2023 final report (GGE on LAWS, 2023[b]: §18). Instead of a specific regulation centred on LAWS, the requirement of HC allows to open the scope to all the categories of weapons. Most importantly, this would include weapons that will be developed in future. Some argue that regulations must not target future weapons, because "declaring these weapons illegal now would prevent States from researching and developing technologies that may lead to greater LOAC compliance" (Jensen, 2020: 53). On the contrary, this thesis holds that the GGE has the opportunity to guide future technological development pre-emptively, by ensuring that the warfare of tomorrow remains under the necessary degree of human control. At the same time, the technology-neutral approach allows to regulate LAWS while bypassing a series of complex discussions in which the GGE has not established consensus in the past years: for example, what the definition of LAWS should be.

As suggested by a number of states, human control could become the basis of a two-tier approach to LAWS. In this way, HC would determine the legality of an autonomous weapon, by setting both prohibitions and regulations. LAWS operating outside of human control would be prohibited expressly. The requirement of human control would function as a regulatory framework for LAWS and other categories of weapons. HC is a promising basis for a two-tier approach in comparison to other proposals expressed at the GGE. For instance, an alternative option formulated by some states suggested establishing the two-tier approach over the 'compliance with IHL'. This approach is, nonetheless, not concrete: states do not specify how compliance with IHL would be evaluated and certified. At the same time, human control does fulfil the role of ensuring

compliance with IHL. This thesis has advanced in suggesting a concrete understanding of human control, that could determine the legality of a weapon.

In addition to this, HC is not only a prospective but also a necessary focus. As long as HC is not enforced, states and the private sector can benefit from the legal vacuum. The lack of a specific international regulation centred on HC intensifies, for instance, the risk of attacks that result in civilian losses for which no individual can be held accountable (Wagner, 2014: 1404-1408). HC can provide guarantees that other approaches to LAWS lack. For example, temporal and geographic restrictions are not a panacea to the risks associated with LAWS, as demonstrated in this thesis. In turn, HC ensures that human judgement and situational context assessment remain compulsory in the selection and engagement with a target.

Third finding: the GGE on LAWS has the opportunity to emphasise the centrality of the human element in IHL

Normative analysis presented in this thesis reiterates that a core element of IHL lies in the fact that it is directed to humans. Humans, and not machines, are clearly the addressees of IHL rules and principles (Sassòli, 2019: 519). In addition to this, only humans are capable to comply with certain principles of IHL. This is not only related to the limitations of the autonomous capacities of LAWS in their current state of development. It is true that sensors present challenges to the identification and recognition of lawful targets. This sort of limitations can theoretically be addressed by further research and future technological developments. However, it is also true that, independently of the sophistication of the technology employed, what LAWS can never do is to make legal judgements required by IHL.

IHL does not contemplate any other subject but humans to exercise judgements related, for example, to the assessment of proportionality. While treaty sources do not tend to expressly formulate the rules as oriented to humans exclusively, some explicit requirements of a human element have been

identified throughout the normative analysis. This complements the implicit requirements, equally identified in the second chapter. The fundamental principles of IHL that have been here analysed are grounded on other considerations and principles, including the principles of humanity. It has been shown that compassion, morality and common sense, among others, must be part of the judgements and evaluations that take place during the targeting processes. Such qualities are accessible to humans, but not to the machines.

The increasing development of autonomous features in weapons leads to the decrease of human involvement in some aspects of warfare. In view of its mandate, the GGE on LAWS has the opportunity to reiterate that the human element is a legal requirement under existing IHL. First, humans are the only addressees of IHL. Second, human judgement is required inasmuch IHL does not contemplate an algorithmic decision-making to be the only process that takes place in relation to targeting. It also requires a complex evaluation that goes beyond machine capacities, not only due to the current state of the technological development, but also due to the very nature of the machines. LAWS cannot nor should not be expected to be able to reproduce moral assessments or common sense, required for the compliance with the fundamental IHL principles. In addition to this, responsibility for IHL violations can only lie with humans. Contrarily to what some suggest (Jensen, 2020: 37), the potential responsibility gap should not be disregarded only because the use of LAWS has not yet led to violations of IHL. States need to and have the opportunity to act pre-emptively. Content analysis outlined the claim that human control does not necessarily mean that IHL will be respected. While this is true, the requirement of human control would allow the determination of individual criminal responsibility in case of an IHL violation.

The CCW framework of the GGE on LAWS offers the settings where the need for human control can be reiterated and, potentially, positivised. By doing so, states would also reinforce the centrality of the human role in IHL,

which is utmost relevant in view of the rapid path of the technological developments in relation to weaponry and warfare.

Discussion

In replying to the research question, this analysis has shown that IHL requires LAWS to be compatible with a certain extent of human control. This coincides with the positions expressed by a number of states. As seen at the 2023 meeting of the GGE, states tend to use different terminologies and definitions. This thesis has progressed in pointing out convergences between the definitions and the elements of human control proposed by the states. In addition to this, normative analysis addressed the need for human control by analysing the fundamental IHL principles that can be challenged by the development of LAWS. The resulting notion of human control is grounded both on the combination of states' understandings of human control and on the normative requirements related to the human element in IHL. The parallelisms identified between the two approaches allow to conclude that what some states identify as human control coincides with the existing IHL requirements. This finding can be used to align states' positions during future meetings of the GGE on LAWS.

It has also been demonstrated that the need for human control derives implicitly from the existing IHL. This finding advances in bridging one of the debates identified during the 2023 meeting of the GGE. Having this settled, states can channel the forthcoming discussions towards more concrete issues related to human control. For instance, the 2024 GGE could explore ways to align the positions on what should be the object of control: whether the weapon, the use of force or the effects of the use of force. In doing so, it is necessary to adopt a technology-neutral approach: the requirements should focus on human control rather than on LAWS. Therefore, they must also be applicable to currently existing weapons. In 2024, the GGE can also review states' understandings of human control in order to specify the concrete requirements associated with it. This would set the necessary limits for future technological

developments and deployment of LAWS because, as demonstrated by this thesis, weapons that operate outside of human control are incompatible with the existing legal framework.

In addition to this, the notion of human control can be explicitly codified as a legal requirement. The development of a legally binding treaty is possible under the CCW framework. The two-tier approach, progressively accepted among the states, would address two significant gaps if human control is chosen as the criteria. First, the two-tier approach would positivize the prohibition of weapons that operate outside of human control. However, if the GGE only outlines this prohibition, it will run the risk of drawing an insufficient legal framework. While the definition of LAWS remains disputed, the scope of prohibition will allow broad interpretations, that eventually do not sufficiently restrain the development and use of autonomous weapons.

This is why the second, necessary, tier is related to the regulation: a concrete and specific requirement of human control would contribute to the regulation of LAWS and other weapons. Based on the analysis of states' positions and the current legal framework, this thesis suggests that human control must be exercised in two phases. On the one hand, human judgement should be a requisite during the selection of the target. On the other hand, constant human monitoring should be required during the engagement phase, in order to ensure that the operator can abort or detain the weapon at any point of time.

Conclusions

Autonomous features in weapons will be progressively developed and deployed. Lethal autonomous weapons systems are already used in conflicts. Autonomy offers numerous advantages to the military sphere, such as agility, risk reduction and support in decision-making. At the same time, autonomy inevitably leads to the decrease of human control. This thesis has shown that a degree of human control must be maintained, for LAWS to comply with IHL.

The necessary degree of human control has been formulated over the basis of states' positions and legal analysis. This methodological approach has allowed to develop the notion of human control within the existing legal framework and grounded on real and up-to-date positions of states. States play a central role in the development of international law, by creating and interpreting treaties and by participating in the formation of international custom through practices and the expression of *opinio juris*. This explains the importance of examining states' positions expressed in public and official fora.

As the first part of the analysis, the characterisation of human control has been outlined following the definitions and elements employed by states at the 2023 meeting of the Group of Governmental Experts on lethal autonomous weapon systems convened through the Conference on Certain Conventional Weapons. After this, the importance of the human element in IHL has been demonstrated by providing an analysis of three fundamental principles. It has been established that the notion of human control proposed by states coincides with the requirement of the human element embedded in IHL. Hence it follows that IHL implicitly requires human control.

This research began with the question of how the notion of human control contributes to the understanding of the relationship between LAWS and IHL. It has been concluded that HC is already an implicit requirement of IHL and, as such, a regulatory criterion. This role can further be positivised if the work of the GGE leads to the development of a legally binding treaty that would

require HC. In order to ensure that current and future technological development maintains the necessary degree of human control, this thesis suggests the following:

FIRST, the convergences and alignments identified at the 2023 meeting of the GGE on LAWS indicate that, even when differing denominations are employed, the majority of states concur on the need to maintain human control in lethal autonomous weapons systems.

SECOND, the present normative analysis proposes a conceptualization of human control that includes human judgment, monitoring and capacity to intervene during the whole cycle of the performance of a weapon. In weapons with autonomous features, human control cannot be substituted by machine decision-making in the critical functions, namely the selection and engagement with a target.

THIRD, the preservation of human judgement in LAWS is necessary for the compliance with fundamental IHL principles. IHL is oriented to humans and requires human assessment, which is qualitatively different from the algorithm-enabled assessment of a machine. The correct identification of a target can entail, among others, situational awareness, understanding of the context and common sense, inaccessible to LAWS. For example, calculations related to the assessment of proportionality cannot be substituted by a formula or an algorithm.

FOURTH, a continuous human monitoring of the performance of LAWS and the human capacity to intervene are also necessary for the compliance with IHL. An operator must be able to abort or interrupt the attack at any point of time. This is necessary, for example, whenever there is a change in the assessment of proportionality or precautions in attack.

FIFTH, human control is necessary for the determination of individual criminal responsibility in cases where the use of lethal autonomous weapons systems leads to a grave breach of IHL.

SIXTH, the requirement of human control established in this thesis leads to the conclusion that the human must always remain in the loop. LAWS that do not allow for any kind of human intervention, and LAWS that select and engage a target without human control are not compatible with the existing IHL.

SEVENTH, if human control becomes the primary regulatory angle in the GGE deliberations on LAWS, it will allow for the development of a technology-neutral approach. In this way, the work of the Group of Governmental Experts on lethal autonomous weapon systems will result in a regulation that extends beyond LAWS, and is oriented to all the future technological advancements in the military field.

Bibliography

Academic literature

Acheson, R. (2022). "Autonomous weapons and patriarchy." *Women's International League for Peace and Freedom,* 1-24. Retrieved from https://reachingcriticalwill.org/images/documents/Publications/aws-and-patriarchy.pdf

Akkuş, B. (2023). "Legal Transplants: Applying Arms Control Frameworks to Autonomous Weapons." *Eskişehir Osmangazi Üniversitesi Sosyal Bilimler Dergisi, 24*(1), 128-148. Retrieved from https://dergipark.org.tr/en/download/article-file/2902170

Amoroso, D., & Tamburrini, G. (2020). "Autonomous weapons systems and meaningful human control: ethical and legal issues." *Current Robotics Reports, 1,* 187-194. Retrieved from https://link.springer.com/article/10.1007/s43154-020-00024-3

Balkin, J. M. (1988). "Footnote." *Northwestern University Law Review, 83*(1), 275-320. Retrieved from https://openyls.law.yale.edu/bitstream/handle/20.500.13051/2227/Footnote.pdf?sequence=2

Baxter, J. (2020). "Content analysis". *International Encyclopedia of Human Geography, 2*(2), 391-396. Retrieved from https://www.sciencedirect.com/science/article/pii/B9780081022955108054

Benvenuti, P. (2001). "The ICTY prosecutor and the review of the NATO bombing campaign against the Federal Republic of Yugoslavia." *European Journal of International Law, 12*(3), 503-530. Retrieved from https://academic.oup.com/ejil/article-pdf/12/3/503/1044467/120503.pdf

Berrendorf, A., & Bontridder, N. (2023). "Le développement, le déploiement et l'utilisation d'un système d'arme létale autonome dans un conflit armé:

légalité et responsabilité." *Association française pour l'Intelligence Artificielle, 120*, 96-100. Retrieved from https://www.researchgate.net/profile/Laurence-Vanin/publication/370769582_Norme_IA_-_Bulletin_AFIA/links/64622eda4353ba3b3b696a7e/Norme-IA-Bulletin-AFIA.pdf#page=97

Bo, M. (2021). "Autonomous Weapons and the Responsibility Gap in light of the *Mens Rea* of the War Crime of Attacking Civilians in the ICC Statute." *Journal of International Criminal Justice, 19*(2), 275-299. Retrieved from https://academic.oup.com/jicj/article-abstract/19/2/275/6181757

Bode, I. (2019). "Norm-making and the Global South: Attempts to Regulate Lethal Autonomous Weapons Systems." *Global Policy, 10*(3), 359-364. https://onlinelibrary.wiley.com/doi/abs/10.1111/1758-5899.12684?casa_token=RA8-Oj30SFkAAAAA:5GddlyGDQCSy-V5I6ceg7CHnI6KvBsaaqGVtBboTqKKJamWm4xVJ37Omaq-DxhRjnwTu8RvwanESXso9jQ

Bode, I., & Watts, T. F. (2023). "Loitering munitions: flagging an urgent need for legally binding rules for autonomy in weapon systems". *Humanitarian Law and Policy*. Retrieved from https://blogs.icrc.org/law-and-policy/2023/06/29/loitering-munitions-legally-binding-rules-autonomy-weapon-systems/?_hsmi=264450645&_hsenc=p2ANqtz-_J-gJvbeTjTLHXHqbLPjdZl6M183OzpOy0qnRDUJ9IZxo1o-g7M0IUNJaL6Nau87P915oq_1hs-_RiAIX44lw5PT8HIg

Bode, I., Huelss, H., Nadibaidze, A., Qiao-Franco, G., & Watts, T. F. (2023). "Prospects for the global governance of autonomous weapons: comparing Chinese, Russian, and US practices." *Ethics and Information Technology, 25*(1), 5. Retrieved from https://link.springer.com/article/10.1007/s10676-023-09678-x

Boulanin, V. (2019). *The impact of artificial intelligence on strategic stability and nuclear risk*. Stockholm: Stockholm International Peace Research Institute. Retrieved from https://policycommons.net/artifacts/2324888/the-impact-of-artificial-intelligence-on-strategic-stability-and-nuclear-risk/view/

Boulanin, V., & Verbruggem, M. (2017). *Mapping the development of autonomy in weapon systems*. Stockholm: Stockholm International Peace Research Institute. Retrieved from https://www.sipri.org/sites/default/files/2017-11/siprireport_mapping_the_development_of_autonomy_in_weapon_systems_1117_1.pdf

Cass, K. (2015). "Autonomous weapons and accountability: seeking solutions in the law of war." *Loyola of Los Angeles Law Review, 48*(3), 1017-1068. Retrieved from https://heinonline.org/HOL/Page?handle=hein.journals/lla48&div=37&g_sent=1&casa_token=mLXYnQdOdFMAAAAA:VybavDnO5_L1ywQvvbwuFMtf-I5UiPR5HaFhmq_9mJegRGRpGm5lDLwnKQLVidOHbmrwLGBa-sw&collection=journals

Chen, Y. Y., Shek, D. T., & Bu, F. F. (2011). "Applications of interpretive and constructionist research methods in adolescent research: philosophy, principles and examples." *International Journal of Adolescent Medicine and Health, 23*(2): 129-139. Retrieved from https://ira.lib.polyu.edu.hk/bitstream/10397/5065/1/C1.pdf

Chernyavsky, A., & Sibileva, O. (2020). "Автономное высокоточное оружие как вызов международному гуманитарному праву [Autonomous precision weapons as a challenge to international humanitarian law]". *Военное право [Military law]*, *6*(4), 229-238. Retrieved from http://www.voennoepravo.ru/files/4-2020.pdf#page=229

Cherry, J., & Johnson, D. (2021). "Maintaining command and control (C2) of lethal autonomous weapon systems: Legal and policy considerations."

Southwestern Journal of International Law, 27(1), 1-27. Retrieved from https://www.swlaw.edu/sites/default/files/2021-03/1.%20Cherry%20%5Bp.1-27%5D.pdf

Christie, E. H., Ertan, A., Adomaitis, L., & Klaus, M. (2023). "Regulating lethal autonomous weapon systems: exploring the challenges of explainability and traceability." *AI and Ethics*, 1-17. Retrieved from https://link.springer.com/article/10.1007/s43681-023-00261-0

Darlaston-Jones, D. (2007). "Making connections: The relationship between epistemology and research methods." *The Australian Community Psychologist, 19*, 19-27. Retrieved from https://www.researchgate.net/publication/284970190_Making_connections_The_relationship_between_epistemology_and_research_methods

Davison, N. (2018). "A legal perspective: Autonomous weapon systems under international humanitarian law." *Arms Unit, Legal Division of the International Committee of the Red Cross*, 1-18. Retrieved from https://www.icrc.org/en/document/autonomous-weapon-systems-under-international-humanitarian-law

Dremliuga, R. (2020). "General legal limits of the application of the lethal autonomous weapons systems within the purview of international humanitarian law." *Journal of Politics and Law, 13*(2), 115-121. Retrieved from https://heinonline.org/hol-cgi-bin/get_pdf.cgi?handle=hein.journals/jpola13&section=36&casa_token=_lu7xvDIzvUAAAAA:-s6O3ngulGtgwc6397GXNukwCmPaCeGqi4I39TS6D6183kzyOTqpcPeM43v3qSyA9AK_rhaSHF8

Duncan, N. J. & Hutchinson, T. (2012). "Defining and describing what we do: Doctrinal legal research." *Deakin Law Review, 17*(1), 83-119. Retrieved from https://openaccess.city.ac.uk/id/eprint/4335/1/

Egeland, K. (2016). "Lethal autonomous weapon systems under international humanitarian law." *Nordic Journal of International Law, 85*(2), 89-118.

Retrieved from https://brill.com/view/journals/nord/85/2/article-p89_1.xml

Ekelhof, M. A. (2017). "Complications of a common language: why it is so hard to talk about autonomous weapons." *Journal of Conflict and Security Law, 22*(2), 311-331. Retrieved from https://academic.oup.com/jcsl/article-abstract/22/2/311/3744969

Espada, C. G., & Hortal, M. J. C. (2013). "Autonomous Weapons Systems, Drones and International Law." *Revista del Instituto Español de Estudios Estratégicos, 2,* 1-30. Retrieved from https://www.researchgate.net/profile/Maria-Jose-Cervell/publication/324224637_Autonomous_weapon_systems_drones_and_International_Law/links/5ac6445eaca2720544d1db90/Autonomous-weapon-systems-drones-and-International-Law.pdf

Etzioni, A., & Etzioni, O. (2018). "Pros and cons of autonomous weapons systems." In Amitai Etzioni (Ed.) *Happiness is the Wrong Metric: A Liberal Communitarian Response to Populism* (pp. 253-263), Washington: Springer. Retrieved from https://link.springer.com/chapter/10.1007/978-3-319-69623-2_16

Figueroa, M. D., Orozco, A. H., Martinez, J., & Muñoz, W. J. (2023). "The risks of autonomous weapons: An analysis centred on the rights of persons with disabilities." *International Review of the Red Cross, 105*(922), 278-305. Retrieved from https://www.cambridge.org/core/services/aop-cambridge-core/content/view/54872E40DBC09DC95D5F1DFD6DB89F51/S1816383122000881a.pdf/risks_of_autonomous_weapons_an_analysis_centred_on_the_rights_of_persons_with_disabilities.pdf?casa_token=7glQRo0ErI8AAAAA:U_D28JnUJ-lIbIxuKdYnXV0yLg_45ELBPyl3gTDhvCDPy8FGsBQWA8Fh8CoisRxTYDMCrCOXblL-

Fleck, D. (2020). "The Martens Clause and environmental protection in relation to armed conflicts." *Goettingen Journal of International Law, 10*(1), 243-266. Retrieved from https://www.gojil.eu/issues/101/101_article_fleck.pdf

Gaeta, P., & George Jain, A. (2019). "Individualisation of IHL Rules Through Criminal Responsibility for War Crimes and Some (Un)Intended Consequences". In Dapo Akande, David Rodin and Jennifer Welsh (Eds.) *The Individualisation of War.* Oxford: Oxford University Press. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3853333

Homayounnejad, M. (2017). "Assessing the Sense and Scope of 'Autonomy' in Emerging Military Weapon Systems." *Transnational Law Institute Think*, paper 76/2017. Retrieved from https://ssrn.com/abstract=3027540

Horowitz, M. C. (2016). "Why Words Matter: The Real World Consequences of Defining Autonomous Weapons Systmes." *Temple International & Comparative Law Journal, 30*(1), 85-98. Retrieved from https://heinonline.org/HOL/Page?handle=hein.journals/tclj30&div=11 &g_sent=1&casa_token=HLAU80fXUQ0AAAAA:rkhaOvIvVklQHT h4AjY5wvAFOFsVxAOnYeuG_YS8Rrnjeo_TRb0BuBu8geen22RNje 4PWMNY3uk&collection=journals

Jacobson, B. R. (2017). "Lethal autonomous weapons systems: mapping the GGE debate." *DiploFoundation Policy Papers and Briefs*, *8*, 1-7. Retrieved from https://www.diplomacy.edu/wp-content/uploads/2021/09/Policy_papers_briefs_08_BRJ.pdf

Jensen, E. T. (2020). "The (Erroneous) Requirement for Human Judgment (and Error) in the Law of Armed Conflict." *International Law Studies, 96*(1), 577-602. Retrieved from https://digital-commons.usnwc.edu/cgi/viewcontent.cgi?article=2916&context=ils

Kayser, D. (2023). "Why a treaty on autonomous weapons is necessary and feasible." *Ethics and Information Technology, 25*(2), 1-5. Retrieved from https://link.springer.com/article/10.1007/s10676-023-09685-y

Kibiswa, N. K. (2019). "Directed qualitative content analysis (DQlCA): A tool for conflict analysis." *The Qualitative Report, 24*(8), 2059-2079. Retrieved from https://www.researchgate.net/profile/Naupess-Kibiswa/publication/350981784_Directed_Qualitative_Content_Analysis_DQlCA_A_Tool_for_Conflict_Analysis/links/62e5ed859d410c5ff37598f0/Directed-Qualitative-Content-Analysis-DQlCA-A-Tool-for-Conflict-Analysis.pdf

Krishnan, A. (2009). K*iller robots: legality and ethicality of autonomous weapons*. Farnham: Ashgate Publishing. Retrievable from: https://books.google.ch/books?id=klvlN9PgBeYC&dq=Killer+Robots:+Legality+and+Ethicality+of+Autonomous+Weapons+De+Armin+Krishnan&lr=&hl=es&source=gbs_navlinks_s

Leveringhaus, A. (2016). *Ethics and autonomous weapons*. Oxford: Palgrave MacMillan. Retrieved from https://books.google.ch/books?id=Cuw0DAAAQBAJ&dq=lethal+autonomous+weapons+without+human+operator+engage+target&lr=&hl=es&source=gbs_navlinks_s

Madill, A., Jordan, A., & Shirley, C. (2000). "Objectivity and reliability in qualitative analysis: Realist, contextualist and radical constructionist epistemologies." *British journal of psychology, 91*(1), 1-20. Retrieved from https://bpspsychub.onlinelibrary.wiley.com/doi/pdf/10.1348/000712600161646?casa_token=NZivr_sHkg0AAAAA:53gleRcqa5T2OdcwLKSSfQqgaMUqne8C6KDrddHMs2zLCKezPV8UXVQfrbxErdvgf-mnGLZadpH1yzTUDg

Massacci, F., & Vidor, S. (2022). "Building Principles for Lawful Cyber Lethal Autonomous Weapons." *IEEE Security & Privacy, 20*(2), 101-106.

Retrieved from https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9740711

McFarland, T., & Assaad, Z. (2023). "Legal reviews of in situ learning in autonomous weapons." *Ethics and Information Technology, 25*(1), 1-9. Retrieved from https://link.springer.com/article/10.1007/s10676-023-09688-9

Mitchell, M. (2023). "Analyzing the law qualitatively." *Qualitative research journal, 23*(1), 102-113. Retrieved from https://www.emerald.com/insight/content/doi/10.1108/QRJ-04-2022-0061/full/html?skipTracking=true

Nahavandi, S. (2017). "Trusted autonomy between humans and robots: Toward human-on-the-loop in robotics and autonomous systems." *IEEE Systems, Man, and Cybernetics Magazine, 3*(1), 10-17. Retrieved from https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=7815473&casa_token=ljMGmmPuU7MAAAAA:_cC-lixESwIbHmZCzfR9CBm7ufHMSX3NaNQEVyNu5fIMmz8op3qXYJtD4R6xXdxkpswb62yTmC1b&tag=1

Noone, G. P., & Noone, D. C. (2015). "The debate over autonomous weapons systems." *Case Western Reserve Journal of International Law, 47*(1), 25-36. Retrieved from https://heinonline.org/HOL/Page?collection=journals&handle=hein.journals/cwrint47&id=34&men_tab=srchresults

O'Reilly, M., Kiyimba, N., & Drewett, A. (2021). "Mixing qualitative methods versus methodologies: A critical reflection on communication and power in inpatient care." *Counselling and Psychotherapy Research, 21*(1), 66-76. Retrieved from https://onlinelibrary.wiley.com/doi/epdf/10.1002/capr.12365

Petman, J. M. (2017). *Autonomous Weapons Systems and International Humanitarian Law: 'Out of the Loop'?*. Helsinki: Unigrafia Oy. Retrieved from

https://um.fi/documents/35732/48132/autonomous_weapon_systems_a
n_international_humanitarian_law__out_of_the/c0fca818-3141-b690-
0337-7cfccbed3013?t=1525645981157

Riebe, T., Schmid, S., & Reuter, C. (2020). "Meaningful Human Control of
Lethal Autonomous Weapon Systems: The CCW-Debate and Its
Implications for VSD." *IEEE Technology and Society Magazine, 39*(4),
36-51. Retrieved from
https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9288988

Righetti, L., Pham, Q. C., Madhavan, R., & Chatila, R. (2018). "Lethal
autonomous weapon systems: ethical, legal, and societal issues". *IEEE
Robotics & Automation Magazine, 25*(1), 123-126: 124. Retrieved from
https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8314620

Roff, H. M. (2014). "The strategic robot problem: Lethal autonomous weapons
in war." *Journal of Military Ethics, 13*(3), 211-227. Retrieved from
https://www.tandfonline.com/doi/full/10.1080/15027570.2014.975010?
needAccess=true

Roff, H. M., & Danks, D. (2018). "Trust but Verify: The difficulty of trusting
autonomous weapons systems." *Journal of Military Ethics, 17*(1), 2-20.
Retrieved from https://www.daviddanks.org/s/Roff_Danks-Pub.pdf

Roff, H. M., & Moyes, R. (2016). "Meaningful human control, artificial
intelligence and autonomous weapons." *Briefing Paper Prepared for the
Informal Meeting of Experts on Lethal Autonomous Weapons Systems,
UN Convention on Certain Conventional Weapons*. Retrieved from
https://article36.org/wp-content/uploads/2016/04/MHC-AI-and-AWS-
FINAL.pdf

Rosert, E., & Sauer, F. (2018). "Perspectives for Regulating Lethal Autonomous
Weapons at the CCW: A Comparative Analysis of Blinding Lasers,
Landmines, and LAWS." Paper prepared for the workshop *New
Technologies of Warfare: Implications of Autonomous Weapons Systems
for International Relations*, 5th EISA European Workshops in

International Studies, Groningen, 6-9 June 2018, 1-39. Retrieved from https://www.academia.edu/download/56715793/rosert_sauer_ewis18.pdf

Saket, S. (2020). "Lethal Autonomous Weapons: Conundrum of Morality and Legality." *Supremo Amicus, 18*, 815-824. Retrieved from https://heinonline.org/HOL/Page?handle=hein.journals/supami18&div=95&g_sent=1&casa_token=snaQXVwT3JIAAAAA:Vu468vMCQC54G-xKo4tyvU-6XDz5OjiwgDCg001oOrpTNiSPt3g7exmGn9LPC2YBf__TeS2Inw&collection=journals

Sandvik, K., & Lohne, K. (2015) "Lethal Autonomous Weapons: Killing the 'Robots-don't-Rape' Argument." *IntLawGrrls, 5*. Retrieved from https://ilg2.org/2015/08/05/lethal-autonomous-weapons-killing-the-robots-dont-rape-argument/

Sassòli, M. (2003). "Legitimate targets of attacks under International Humanitarian Law". Background Paper prepared for the *Informal High-Level Expert Meeting on the Reaffirmation and Development of International Humanitarian Law*, Cambridge, January 27-29, 2003. Retrieved from https://www.hpcrresearch.org/sites/default/files/publications/Session1.pdf

Sassòli, M. (2014). "Autonomous weapons and international humanitarian law: Advantages, open technical questions and legal issues to be clarified." *International Law Studies, 90*(1), 308-340. Retrieved from https://digital-commons.usnwc.edu/ils/vol90/iss1/1/

Sassòli, M. (2019). *International humanitarian law: Rules, controversies, and solutions to problems arising in warfare*. Massachusetts: Edward Elgar Publishing.

Schmitt, M. N. (2012). "Autonomous weapon systems and international humanitarian law: a reply to the critics." *Harvard National Security*

Journal, 4, 1-37. Retrieved from https://centaur.reading.ac.uk/89864/1/Schmitt-Autonomous-Weapon-Systems-and-IHL-Final.pdf

Schuller, A. L. (2017). "At the crossroads of control: the intersection of artificial intelligence in autonomous weapon systems with international humanitarian law." *Harvard National Security Journal, 8*(2), 379-425. Retrieved from https://heinonline.org/HOL/Page?handle=hein.journals/harvardnsj8&div=11&g_sent=1&casa_token=GugQ_DfHXVQAAAAA:vH9RK9PCb0wpIbk6wBg9ZUy9oWj4y8DFxjhoM7U8-Uk8ASqMR27AXzUawz5_3ipfRK49v5DpSw&collection=journals

Seixas-Nunes, A. (2022). *The Legality and Accountability of Autonomous Weapon Systems: A Humanitarian Law Perspective.* Cambridge: Cambridge University Press. Retrieved from https://books.google.ch/books?id=MtduEAAAQBAJ&dq=autonomous+weapons+novelty&lr=&hl=es&source=gbs_navlinks_s

Sharkey, A. (2019). "Autonomous weapons systems, killer robots and human dignity." *Ethics and Information Technology, 21*(2), 75-87. Retrieved from https://link.springer.com/article/10.1007/s10676-018-9494-0

Sharkey, N. (2010). "Saying 'No!' to Lethal Autonomous Targeting." *Journal of Military Ethics, 9*(4), 369-383. Retrieved from https://www.tandfonline.com/doi/pdf/10.1080/15027570.2010.537903?casa_token=2_bINY834WQAAAAA:oNfE6xDe3ezjUzP9U6Ye8mtqB3GQCpxkJyLZ17t_3HHGR1-lr4Nt7wiE9U8PPYWjQ19WJ8UuLc-0fNg

Smith, P. T. (2019). "Just research into killer robots." *Ethics and information technology, 21,* 281-293. Retrieved from https://idp.springer.com/authorize/casa?redirect_uri=https://link.springer.com/article/10.1007/s10676-018-9472-6&casa_token=MD4YSl_7ibIAAAAA:HLciMAS-

FG2CEPIKkOZxDN0gle6ABXsYcwXf3Ot-
OaP9wZuDFmwsjplYeGLRMACHqXovqR9vT_4srY83olk

Snelson, C. L. (2016). "Qualitative and mixed methods social media research: A review of the literature." *International Journal of Qualitative Methods, 15*(1), 1-15. Retrieved from https://journals.sagepub.com/doi/epub/10.1177/1609406915624574

Solovyeva, A., & Hynek, N. (2018). "Going Beyond the" Killer Robots" Debate: Six Dilemmas Autonomous Weapon Systems Raise." *Central European Journal of International & Security Studies, 12*(3), 166-208. Retrieved from https://cejiss.org/images/issue_articles/2018-volume-12-issue-3/07-solovyeva-new-1.pdf

Sparrow, R. (2007). "Killer robots." *Journal of applied philosophy, 24*(1), 62-77. Retrievable from https://onlinelibrary.wiley.com/doi/full/10.1111/j.1468-5930.2007.00346.x?casa_token=19vHZV9ags4AAAAA:HQzBKLeHs PEHfSEAJyrAmoFRJRiU_bTkMjUSG2TH3fKR3oM69_RM_WXPaT t0OqpbmVodwUQs_3gbhTiVGw

Szpak, A. (2020). "Legality of use and challenges of new technologies in warfare–the use of autonomous weapons in contemporary or future wars." *European Review, 28*(1), 118-131. Retrieved from https://www.cambridge.org/core/services/aop-cambridge-core/content/view/1F2EDF5458CD2EDCF999DEAA9B188124/S1062 798719000310a.pdf/legality-of-use-and-challenges-of-new-technologies-in-warfare-the-use-of-autonomous-weapons-in-contemporary-or-future-wars.pdf

Taddeo, M., & Blanchard, A. (2022). "A comparative analysis of the definitions of autonomous weapons systems." *Science and engineering ethics, 28*(37), 1-22. Retrieved from https://link.springer.com/article/10.1007/s11948-022-00392-3

Toscano, C. P. (2015). "Friend of humans: an argument for developing autonomous weapons systems." *Journal of National Security Law and Policy, 8*(1), 189-246. Retrieved from https://heinonline.org/HOL/Page?handle=hein.journals/jnatselp8&div=13&g_sent=1&casa_token=ONZs1OzPUOUAAAAA:paAHXBnD4JWrWqO8_Fi9MaVXQYbcXtUZ1sqKHL_sqEX8J2zZ1AdJAYh7F3wDanuPzXIdryTsW8Y&collection=journals

Van Den Boogaard, J. (2015). "Proportionality and autonomous weapons systems." *Journal of International Humanitarian Legal Studies, 6*(2), 247-283. Retrieved from https://brill.com/view/journals/ihls/6/2/article-p247_3.xml

Wagner, M. (2014). "The dehumanization of international humanitarian law: legal, ethical, and political implications of autonomous weapon systems." *Vanderbilt Journal of Transnational Law, 47*, 1371-1424. Retrieved from https://scholarship.law.vanderbilt.edu/cgi/viewcontent.cgi?article=1243&context=vjtl

Winter, E. (2020). "The compatibility of autonomous weapons with the principle of distinction in the law of armed conflict." *International & Comparative Law Quarterly, 69*(4), 845-876. Retrieved from https://eprints.ncl.ac.uk/fulltext.aspx?url=268967%2F8F747129-DB0F-4D98-B40F-384E6273D1FA.pdf

Winter, E. (2022). "The compatibility of autonomous weapons with the principles of international humanitarian law." *Journal of Conflict and Security Law, 27*(1), 1-20. Retrieved from https://academic.oup.com/jcsl/article-abstract/27/1/1/6513626

Zhang, Y., & Wildemuth, B. M. (2009). "Qualitative analysis of content. Applications of Social Research Methods to Questions." *Human Brain Mapping, 30*(7), 2197-2206. Retrieved from https://philpapers.org/rec/ZHAQAO

Web pages, official publications, and news articles

Automated Decision Research (2023). *Report: Convergences in state positions on human control*. Retrieved from https://automatedresearch.org/news/report/convergences-in-state-positions-on-human-control/. Accessed 15 July 2023.

Congress of the United States (2023). "Defense Primer: U.S. Policy on Lethal Autonomous Weapon Systems". *Congressional Research Service*, 1-3. Retrieved from https://crsreports.congress.gov/product/pdf/IF/IF11150. Accessed 10 June 2023.

Cramer, M. (2021). "A.I. Drone May Have Acted on Its Own in Attacking Fighters, U.N. Says". *The New York Times*. Retrieved from https://www.nytimes.com/2021/06/03/world/africa/libya-drone.html. Accessed 25 February 2023.

International Committee of the Red Cross (1995). *Vienna Diplomatic Conference Achieves New Prohibition on Blinding Laser Weapons and Deadlock on Landmines*. Retrieved from https://www.icrc.org/en/doc/resources/documents/misc/57jmlc.htm. Accessed 6 March 2023.

International Committee of the Red Cross (2022). *What you need to know about autonomous weapons*. Retrieved from https://www.icrc.org/en/document/what-you-need-know-about-autonomous-weapons#:~:text=When%20will%20these%20weapons%20come,anti%2Dpersonnel%20mines%20in%201997. Accessed 17 February 2023.

Ministry of Foreign Affairs and Religion of Costa Rica (2023). *Communiqué of the Latin American and the Caribbean Conference of Social and Humanitarian Impact of Autonomous Weapons*. Retrieved from https://www.rree.go.cr/files/includes/files.php?id=2261&tipo=documentos. Accessed 6 March 2023.

Spoljaric, M. (2023). "Risks from the unconstrained use of autonomous weapons in armed conflict are stark". *ICRC blog: Who we are*. Retrieved from https://blogs.icrc.org/ir/en/2023/04/risks-from-the-unconstrained-use-of-autonomous-weapons-in-armed-conflict-are-stark/. Accessed 10 June 2023.

Tavsan, S. (2021). "Turkish defense company says drone unable to go rogue in Libya". *Nikkei Asia*. Retrieved from https://asia.nikkei.com/Business/Aerospace-Defense-Industries/Turkish-defense-company-says-drone-unable-to-go-rogue-in-Libya. Accessed 26 February 2023.

UK Parliament (2023). *How should autonomous weapons be developed, used and regulated?* Retrieved from https://committees.parliament.uk/committee/646/ai-in-weapon-systems-committee/news/186511/how-should-autonomous-weapons-be-developed-used-and-regulated/. Accessed 25 February 2023.

UN Geneva (2023). *Digital Recordings Portal*. Retrieved from https://conf.unog.ch/digitalrecordings/. Accessed 20 February 2023.

United Nations Association – UK (2023). *33 states call for urgent negotiation of new international law to limit autonomous weapons*. Retrieved from https://una.org.uk/news/33-states-call-for-urgent-negotiation-of-new-international-law-to-limit-autonomy-in-weapons-systems. Accessed 2 Julz 2023.

United Nations Institute for Disarmament Research (2017). "The Weaponization of Increasingly Autonomous Technologies: Concerns, Characteristics and Definitional Approaches". *UNIDIR Resources*, *6*, 1-34. Retrieved from https://unidir.org/sites/default/files/publication/pdfs//the-weaponization-of-increasingly-autonomous-technologies-concerns-characteristics-and-definitional-approaches-en-689.pdf. Accessed 6 March 2023.

United Nations Office for Disarmament Affairs (2023). *The Convention on Certain Conventional Weapons*. Retrieved from https://disarmament.unoda.org/the-convention-on-certain-conventional-weapons. Accessed 12 June 2023.

United Nations Treaty Collection (2023). *Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons which may be deemed to be Excessively Injurious or to have Indiscriminate Effects*. Retrieved from https://treaties.un.org/Pages/ViewDetails.aspx?src=IND&mtdsg_no=XXVI-2&chapter=26&clang=_en. Accessed 21 July 2023.

Weapons Law Encyclopaedia (2023). *Weapons system*. Retrieved from https://www.weaponslaw.org/glossary/weapons-system. Accessed 25 February 2023. Accessed 2 March 2023.

World Bank (2023). *Military expenditure (% of the GDP)*. Retrieved from https://data.worldbank.org/indicator/MS.MIL.XPND.GD.ZS. Accessed 16 June 2023.

International law, jurisprudence, reports, and working papers

Treaty sources and statutes

Statute of the International Court of Justice (1945).

Geneva Convention for the Amelioration of the Condition of the Wounded and Sick in Armed Forces in the Field (1949).

Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts (1977).

Rome Statute of the International Criminal Court (1998).

Customary law

International Committee of the Red Cross, *Study on Customary International Humanitarian Law* (2005).

Jurisprudence

International Court of Justice, *The Legality of the Threat or Use of Nuclear Weapons*, Advisory Opinion, 8 July 1996.

International Criminal Tribunal for the former Yugoslavia, *Final Report to the Prosecutor by the Committee Established to Review the NATO Bombing Campaign against the Federal Republic of Yugoslavia*, 8 June 2000.

International Criminal Tribunal for the former Yugoslavia, *Prosecutor v. Stanislav Galić,* Judgement and Opinion, Case No. IT-98-29-T, 5 December 2003.

International Criminal Tribunal for the former Yugoslavia, *Prosecutor v. Jadranko Prlić and others,* Judgement, Case No. IT-04-74-T, 29 May 2013.

UN reports and working papers

Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons System (2023[a]). *Non-exhaustive compilation of definitions and characterizations submitted by the Chairperson*. CCW/GGE.1/2023/CRP.1. Retrieved from https://reachingcriticalwill.org/images/documents/Disarmament-fora/ccw/2023/gge/documents/characterisations.pdf

Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons System (2023[b]). *Report of the 2023 session of the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems*. CCW/GGE.1/2023/2. Retrieved from https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-_Group_of_Governmental_Experts_on_Lethal_Autonomous_Weapons_Systems_(2023)/CCW_GGE1_2023_2_Advance_version.pdf

Heyns, C. (2013). *Report of the Special Rapporteur on extrajudicial, summary or arbitrary executions, Christof Heyns*. A/HRC/23/47. Retrieved from

https://www.ohchr.org/sites/default/files/Documents/HRBodies/HRCouncil/RegularSession/Session23/A-HRC-23-47_en.pdf

Panel of Experts on Libya (2021). *Letter dated 8 March 2021 from the Panel of Experts on Libya established pursuant to resolution 1973 (2011) addressed to the President of the Security Council*. S/2021/229. Retrieved from https://undocs.org/Home/Mobile?FinalSymbol=S%2F2021%2F229&Language=E&DeviceType=Desktop&LangRequested=False

Quinn, G. (2021). *Report of the Special Rapporteur on the rights of persons with disabilities, Gerard Quinn*. A/76/146. Retrieved from https://documents-dds-ny.un.org/doc/UNDOC/GEN/N21/196/98/PDF/N2119698.pdf?OpenElement

State of Palestine (2023). *State of Palestine's Proposal for the Normative and Operational Framework on Autonomous Weapons Systems*. CCW/GGE.1/2023/WP.2/Rev.1. Retrieved from https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-Group_of_Governmental_Experts_on_Lethal_Autonomous_Weapons_Systems_(2023)/CCW_GGE1_2023_WP.2_Rev.1.pdf

Annex

Codes		Groups
● ◇ Mention "human control"	504	Mentions
○ ◇ Meaningful human control	167	Qualifiers
○ ◇ Appropriate human control	22	Qualifiers
○ ◇ Sufficient human control	5	Qualifiers
○ ◇ Full human control	2	Qualifiers
○ ◇ Necessary human control	2	Qualifiers
○ ◇ Significant human control	2	Qualifiers
○ ◇ Appropriate sufficient human control	1	Qualifiers
○ ◇ Diminished human control	1	Qualifiers
○ ◇ Due human control	1	Qualifiers
○ ◇ Meaningful and effective human con...	1	Qualifiers
○ ◇ Rigorous human control	1	Qualifiers

Table 1. The codebook developed inductively for the quantitative content analysis. The bars in the right column show the prevalence of the codes. Source: own elaboration, using Atlas.ti.

Codes	
○ ◇ Human control is necessary	148
○ ◇ References the two-tier approach	94
● ◇ IHL considerations	73
○ ◇ Human control as a criterion for prohibition	52
○ ◇ Definition, denomination and elements of human control	50
○ ◇ Crossreference to other's position or intervetion	38
● ◇ Human control over the weapon	23
● ◇ Ethical considerations	18
○ ◇ Human control is a degree	16
○ ◇ Human control is a distracting term	13
● ◇ Human control over the use of force	13
○ ◇ Human control should be part of the final report of the GGE	13
○ ◇ HC/MHC should not become a legal standard	10
○ ◇ Human control should be regulated through an internationally binding instrument	9
○ ◇ Human control should not be part of the final report of the GGE	8
○ ◇ Nominal human input	8
○ ◇ Human control should not be a criteria for prohibition	5
○ ◇ Human control is a politicized term	3
○ ◇ Human control should be regulated by the states internally	3
● ◇ Security considerations	3
● ◇ Human control over the effects	1

Table 2. The codebook developed inductively for the qualitative content analysis. The bars in the right column show the prevalence of the codes. Source: own elaboration, using Atlas.ti.
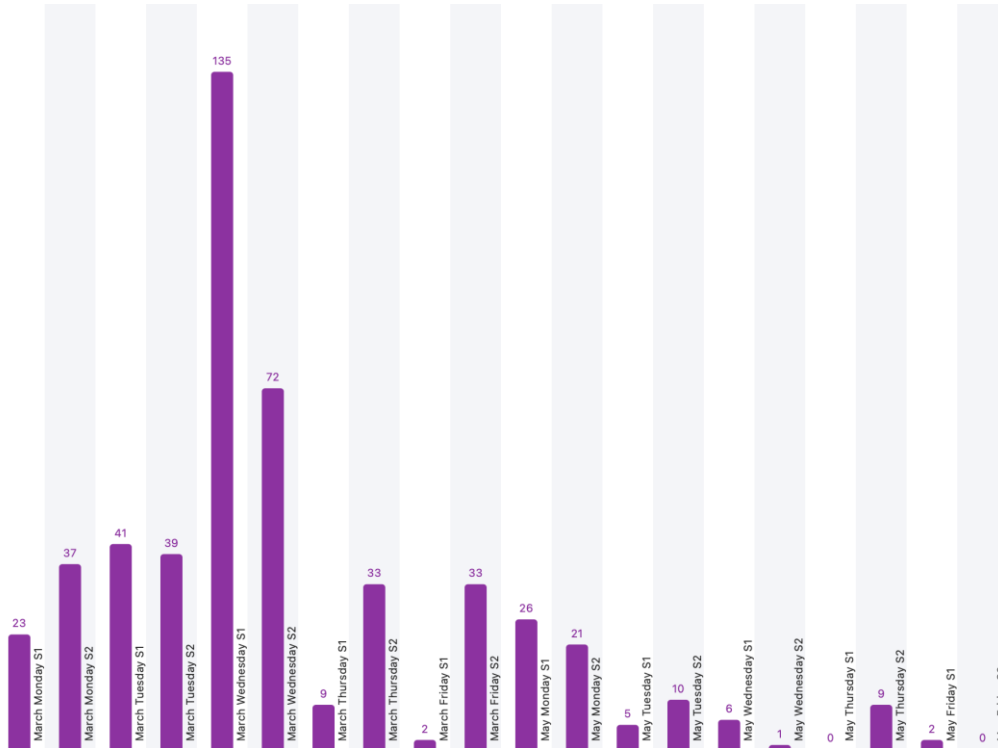
Figure 2. Evolution of the discussions involving the concept of "human control". Each vertical bar corresponds to one of the twenty sessions of the 2023 CCW on LAWS. Source: own elaboration, using Atlas.ti

Figure 3. The word cloud quantifying the most common words (except prepositions and similar) used during March and May meetings of the GGE. Highlighted are the words "control" and "human". Source: own elaboration, using Atlas.ti.
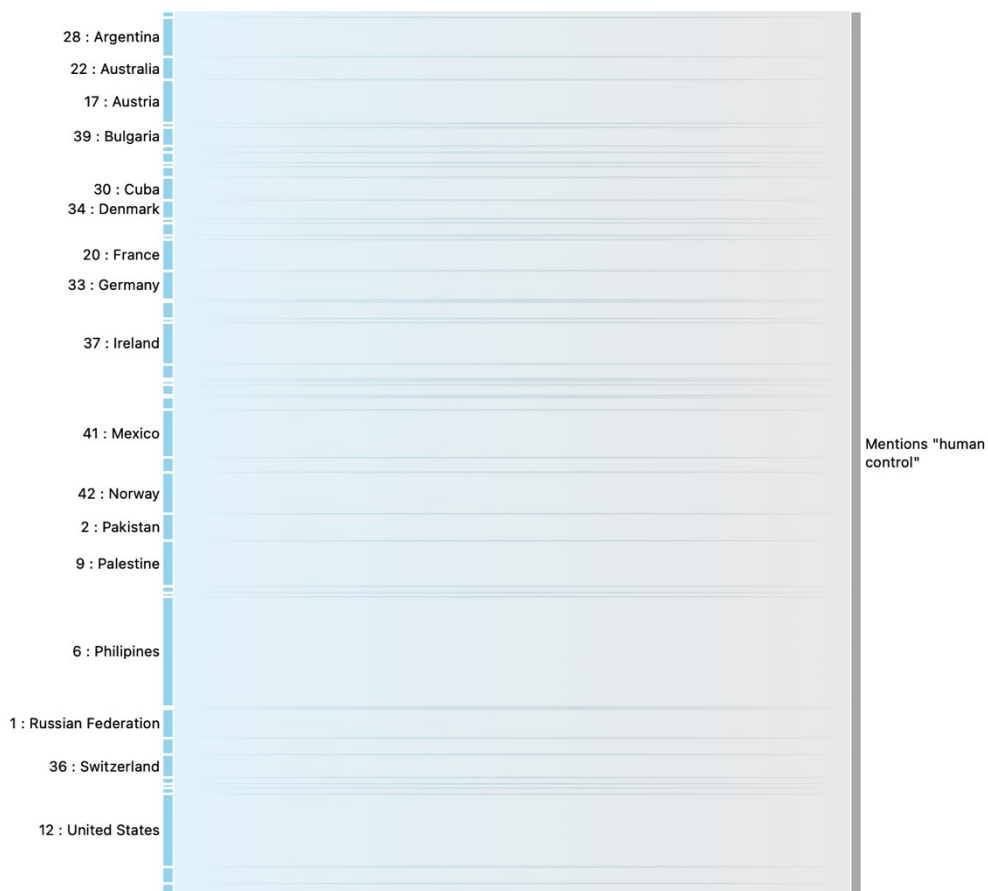
Figure 4. A Sankey diagram showing what states mentioned human control the most. The number of mentions is visualised through the width of each raw. The numbers preceding the name of the country correspond to the order in which countries intervened during the 2023 meeting of the GGE. Source: own elaboration, using Atlas.ti.
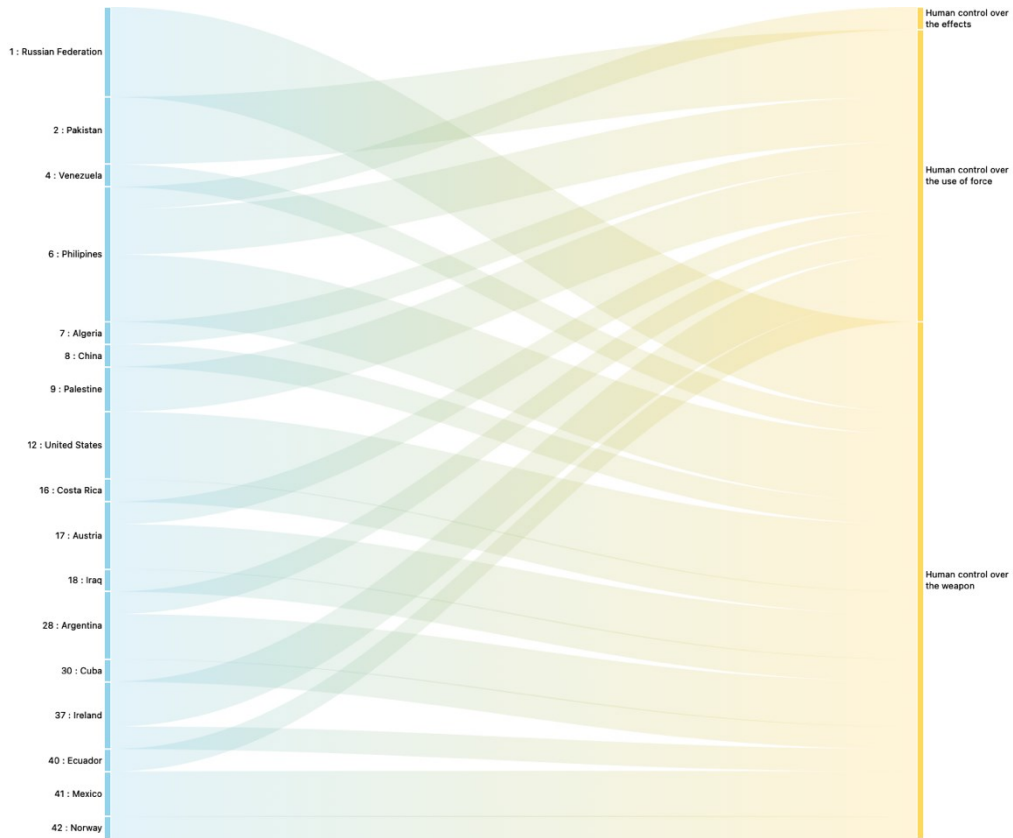
Figure 5. A Sankey diagram illustrating the debate over the object of human control. Source: own elaboration, using Atlas.ti.
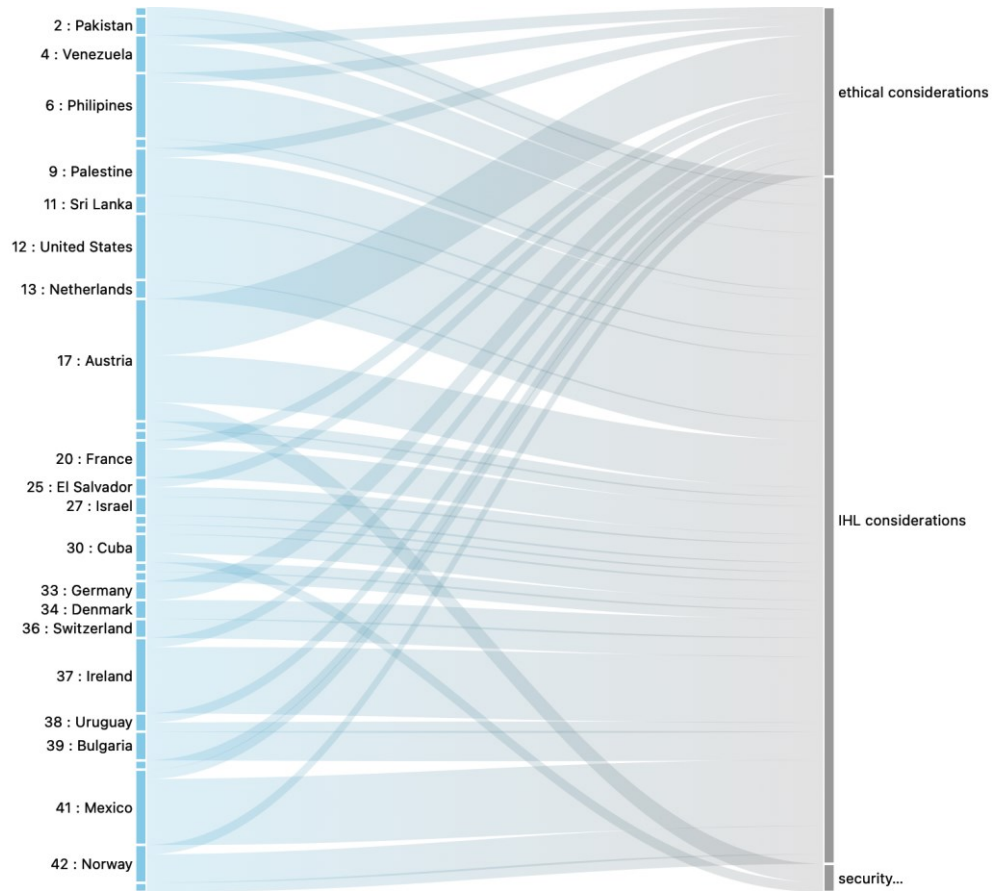
Figure 6. A Sankey diagram showing the prevalence of IHL-related considerations when exposing the need for human control. Source: own elaboration, using Atlas.ti.
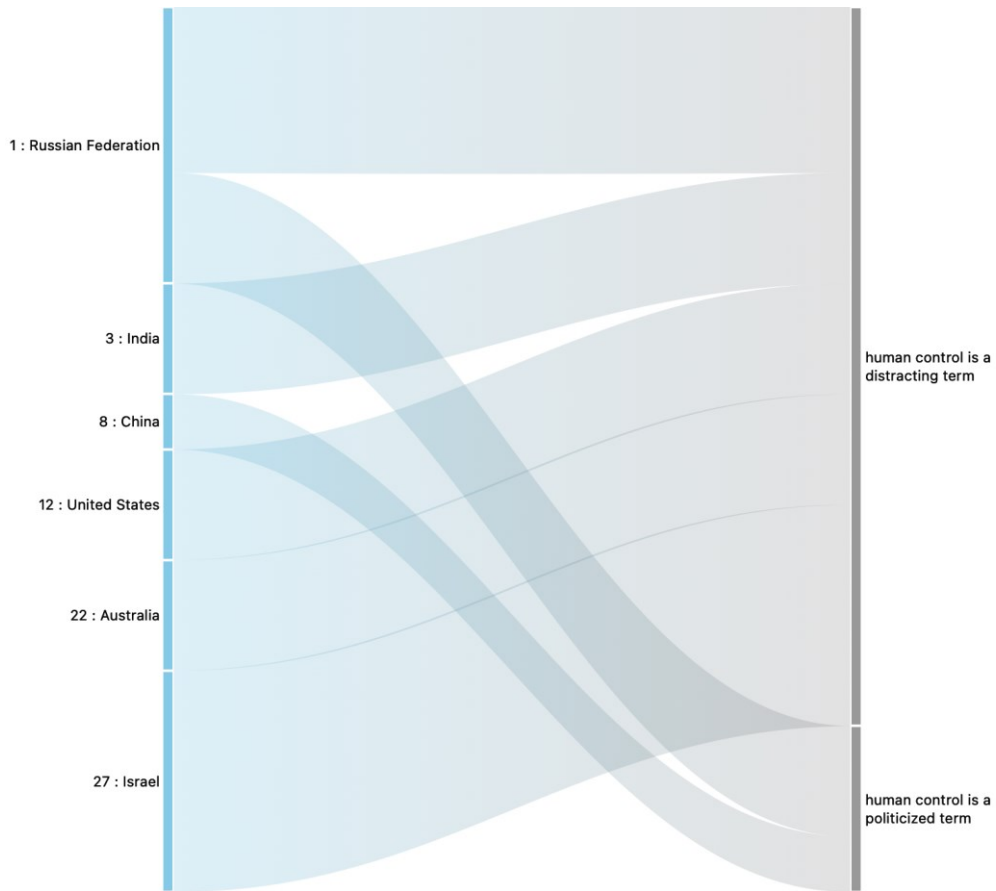
Figure 7: A Sankey diagram showing the ideas expressed by the states that argue that HC is not the right angle for developing regulations on LAWS. Source: own elaboration, using Atlas.ti.
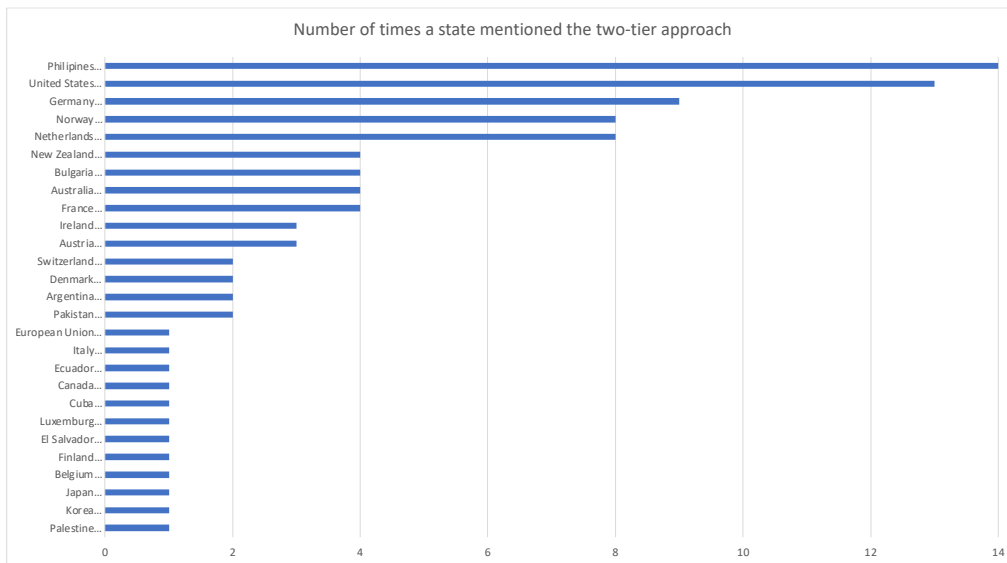
Figure 8: The number of times a state referred to the two-tier approach. Source: own elaboration.