THESES

EÖTVÖS LORÁND UNIVERSITY OF SCIENCES
FACULTY OF LAW
DOCTORAL SCHOOL

Viola Vincze LL.M. (Westminster)

THE LEGALITY OF THE USE OF LETHAL AUTONOMOUS WEAPON SYSTEMS IN THE CONDUCT OF HOSTILITIES

PHD DISSERTATION

THESES

Supervisor: Dr. Gábor Kajtár LL.M. Ph.D

Budapest, 2019

TABLE OF CONTENTS

I. A brief summary of the research task

There are numerous ongoing armed conflicts in the world and a perceptible trend of militarizing foreign affairs. These conflicts become more and more complex due to the applied and deployed technologies. Both the employed weaponry and the relevant legal framework became multidimensional in the last decades. There is an intricate and dense web of applicable international law (including the law of armed conflict, international human rights law, international criminal law) and national laws, as well as political, ethical, and moral expectations from both governments and civil society as to how armed conflicts should be fought.

Cyber assets, remote-controlled drones and non-lethal tactics and techniques are introduced to all current armed conflicts irrespective of conflicts' international or non-international nature. In this regard, respecting and observing the law of armed conflicts (LOAC) looks challenging as it is discernible from the most recent reports of non-governmental organizations, as well as from the cases of international criminal tribunals and national courts. The impact of technological development on the transformation of conflicts has recently captured the attention of lawyers and academics, too. Interest and research in the fields of cyber and military technology in addition to robotics has thrust *autonomous weapon technologies* into the limelight.

The development and deployment of *lethal* autonomous weapon systems (AWS) presents multifaceted (legal, military, political and moral) challenges. Governmental research, UN experts as well as scholars and military lawyers are all bound by the endeavour to find a common denominator regarding the basic concept and characteristics of AWS. In order to understand what capabilities, opportunities, advantages and risks AWS may represent as well as the related legal issues, the underlying concept of autonomy (as the most distinctive feature of AWS, distinguishing them from any other (traditional) type of weapons) shall be introduced. An AWS can be defined from various perspectives; one can assess its objective technical characteristics, its relationship with humans (based on a human-machine interface enabling their 'interaction') as well as its level or degree of autonomy. The outcome of any legal analysis will boil down to how we define AWS according to these angles.

The general purpose of this dissertation is to assess the legality of AWS through a *legal positivist* view, cleared as much as possible from moral and political influences. The research intends to show where the law stands today and assess whether it is adequate to answer the challenges posed by autonomous weapon technologies. One of the animating purposes of LOAC is to protect life and therefore, it cannot be fully separated from ethical questions and principles expressing moral values. These are however generally interiorized and articulated by the fundamental principles of LOAC (the principle of military necessity, the principle of distinction, the principle of proportionality, the principle of humanity and the principle of precautions).

The contemporary discussion of autonomous technologies is characterized by a rift between consequentialists and deontologists – between technological optimism and pessimism. The consequentialist approach is very much result-oriented, whereby the pre-established objectives necessitate the matching means including the use of autonomous systems. Deontologists on the other hand emphasize the possible breach of the right to dignity and the disregard towards ethical norms. Among these contradicting accounts, objective (data-based, descriptive) and subjective (personal and biased) perceptions and accounts often become blurred and tangled.

The interdisciplinary and complex nature of the subject matter requires the assessment of the relevant legal instruments of different but closely related regimes (LOAC, international human rights law and international criminal law), but the bulk of the dissertation is a *LOAC and targeting based analysis.*

The most important quest of the dissertation is to examine the legality of the use of AWS in the conduct of hostilities. This analysis has to be divided into the study of legality under the Law of Attack (as the segment of LOAC regulating the conduct of hostilities with regards to the means and methods that can be employed) and within the practical context of targeting. Concerning the legality of AWS under the Law of Attack, the central questions are related to the principles of LOAC. Can combatants, weapon operators, and military commanders comply with LOAC principles when deciding to employ weapon systems with autonomy in critical functions including the discriminating between civilians and lawful military objectives (principle of distinction)? Are AWS legally and technically able to effectively distinguish between lawful and unlawful targets? Can adherence to the principle of proportionality be ensured when the respective decision is delegated to a machine? Can a mathematical algorithm make adequate calculation with regards to accounting for collateral damage in relation to military advantage? What precautionary measures do military personnel have to take during the planning and execution of operations when selecting AWS as a means of attack under the principle of precaution?

Based on the above legal challenges, the second goal of the research is to analyse the employment of AWS in the targeting process. This can be regarded as the *'operationalization'* of the LOAC principles. In this regard, the dissertation is focusing on the different aspects of translating general LOAC principles into the specific language of operations.

For the purposes of the dissertation, my basic premise is to regard AWS *as weapons* (i.e. a means of warfare) in the legal sense as opposed to treating it as an entity capable of independent decisions. Either as a tangible, inanimate object or as an algorithm, AWS lack legal personality, i.e. weapon systems cannot be regarded as addressees of obligations under LOAC. AWS are not human, they cannot be considered as combatants, units or forces (comprising of persons), and, apart from certain exceptions, they cannot be regarded as methods of warfare either, since methods cover certain manners of how to use means of warfare.

To follow the application of AWS on the battlefield from the planning until the conduct of operations, the dissertation also examines state responsibility, as well as individual and command responsibility in case of LOAC violations involving AWS. Regarding the latter, accountability will require the existence of *the action (violation) and the intent* of perpetrators, which in most case may be difficult to prove. This hurdle may nevertheless be offset by the existence of situational awareness and adequate overview over the operations, as well as by sufficient experience and knowledge on the side of the prospective perpetrators.

II. Methodological approach of the dissertation

The objective of the dissertation is to strip the substance from the surrounding misperceptions and misinterpretations and introduce the question from a mainly *positivist legal point of view* with only touching upon ethical questions.

Based on Article 38 of the Statute of the International Court of Justice, the most important sources of present research are international conventions (most notably, Protocol I additional to the 1949 Geneva Conventions) and international customary law regarding the LOAC. This is naturally complemented by other (subsidiary) sources of international law including scholarly publications and judicial decisions although with regards to the latter, it should be noted that, thus far there exist only a few decisions concerning the use of autonomous technologies.

Apart from the sources of international law, the assessment of NATO documents and national military manuals is also necessary in order to study the applicable rules and regulations on targeting. Regrettably, targeting documents are often missed by authors deliberating over the subject of AWS, although these documents are in fact the most important reflections on the interpretation of LOAC, bridging black-letter law and practice. These military manuals and different training materials for state armed forces are especially relevant because they can be considered a part of state practice or evidence of *opinio juris* and thereby may indicate the evolution of customary norms. While it is common to refer to the general principles of attack (distinction, proportionality, precautions), their translation into the strategic, operational and tactical level of targeting through NATO and national military manuals should also be quoted and referenced. The subject of AWS can be assessed comprehensively only when one appreciates the way these weapon systems fit into – and maybe one day even replace parts of – the targeting process.

Commentary or *travaux préparatoires* to LOAC documents, if any, shall also be carefully read as they may contain valuable references as to the direction of negotiations, intent of parties and the different interpretations considered during the discussions. Further valuable sources include LOAC-related books and scholarly articles as well as the contribution that the International Committee of the Red Cross and different non-governmental organizations (for example the Human Rights Watch, Amnesty International, or Article 36 NGO) are making through their comprehensive reports, accounts, interpretations, and recommendations.

This dissertation is informed by and created from an IL and LOAC point of view. The examination of autonomy in weapon systems could touch upon different disciplines: ethics, robotics, military science, philosophy, and legal studies. Although the subject's interdisciplinary nature is acknowledged, this paper is written from a *legal* point of view, introducing only those technicalities (basic technical concepts and terminology regarding AWS) that are necessary for the legal appreciation of the subject.

Based on positivist methodology, the dissertation is built on the content of conventional law and customary rules, where conventions and treaties are assessed in light of the Vienna Convention on the Law of Treaties, according to which, treaties shall be interpreted together with other relevant rules of international law applicable in the relation between the parties.

The dissertation intends to identify and assess the applicable law and to highlight possible legal uncertainties as the basic methodological approach which is complemented with the examination of a possible demand for new or amended legislation. This methodology includes the introduction and explanation of the competing interpretations and concepts, as well as providing (mostly NATO and US) examples to highlight how LOAC conventional rules and customary principles are rendered operational by Allied and national documents.

The dissertation intends to analyse LOAC, the available NATO documents, national military manuals and other relevant documents in order to:

- assess AWS' legality under international law,
- see where and how the choice and employment of AWS can fit into the targeting process, and
- identify and highlight the existing guarantees in the process that ensure the exclusion of unnecessary, disproportional or for any other reason unlawful use of such weapons.

The applied methodology will identify if there are any uncertainties in law, in order to decide whether LOAC and NATO documents as well as selected national documents provide sufficient protection to avoid the use of unlawful AWS or the unlawful use of otherwise lawful AWS or alternatively, whether an AWS-specific regulation might be needed in the future. The assessment of the possibility of a future regulation covers the study of the prospect and reality of a ban, too.

The most important purposes of LOAC are to protect those who are not, or no longer participating directly in hostilities and to regulate the conduct of hostilities. This examination will stay within the framework of *jus in bello,* i.e. it will focus on how and in what circumstances AWS can be used in the conduct of hostilities following the commencement of an armed conflict. *Jus in bello* (LOAC) is to be distinguished from *jus ad bellum* which is the legal framework regulating the right of states to use force against other states or non-state actors.

From the legal point of view, the subject of the dissertation is limited to the examination of those norms of LOAC that regulate the conduct of hostilities. Lethal AWS may have serious implications regarding several human rights (e.g. right to life, right to human dignity, right to security, right to respect for family life, freedom of expression or freedom of assembly), but this writing will touch upon only the right to life, and other conducts (e.g. detention, interrogation or belligerent occupation) concerning human rights that are outside the battlefield context (e.g. right to dignity, prohibition of torture) are not covered.

III. Outcome of the research and its possible applications

General remarks – autonomy

LOAC, as part of international law, is a flexible, contextual, and intrinsically human decision-based regime. The dissertation examines AWS and autonomy through the prism of LOAC and determine if AWS, as a novel technology, can be used in compliance with the principles and rules of LOAC. It is also subject of the assessment whether resorting to AWS would limit or at least make the concerning LOAC rules seem rigid, since a human being can interpret the rules in light of a particular situation and, as a result, no two assessments can be the same, however, this flexibility may be lost when decisions are made by algorithms. This may result in stiff, although predictable outcomes that may not fit the specific circumstances of the situation.

From the practical point of view, weapon systems are a combination of weapons and the related computer technology, platform and the personnel necessary for its operation. At the very heart of AWS lies the control system (software) as the precondition to the autonomous capabilities of the weapon system (e.g. navigation, movement, obstacle avoidance, mapping the environment, identifying, selecting and tracking targets, engaging targets). Autonomy is probably the most distinct feature of the control system of an AWS, the *'cognitive engine'* that powers machines. With regards to AWS however, the main question is not whether to have autonomy or not. Rather, the *degree of autonomy* should be emphasized which is determined by the sophistication of the related computer technology (sensing software and control system). Autonomy can be defined as the capability of an AWS to perform a task without human intervention, according to its programming. An autonomous system is using probabilistic reasoning, i.e. based on the data collected by the sensors, it makes guesses about the best possible courses of action. What separates an AWS from automatic and automated systems is its capability to decide a course of action, from a number of alternatives, without depending on human control, although that may still be present. Unquestionably, the most important research field with regards to AWS will be that of autonomy and the basic research areas will include *inter alia* machine learning, natural language processing, computer vision, problem solving, logical reasoning, or human-machine interaction.

So far, there are no reports on developing (and employing) weapons systems that are capable of learning and adapting to their environment, i.e. using artificial intelligence. According to scientific research and media accounts, only partial artificial intelligence capabilities have been developed and true artificial intelligence probably will not be available in the coming years. Yet, technology has reached a point where the deployment of such systems will be practically *feasible within a short period of time.* Some military and robotics experts predict that fully autonomous weapons that could select and engage targets without human intervention could be developed within 30 years. The employment of such system would definitely mark a *paradigm shift* in the conduct of hostilities.

The theory (LOAC principles)

Although the Geneva Conventions and their Additional Protocols are the products of a different era that aimed to solve the puzzling questions relevant at the time, they had also anticipated the technological challenges parties to a conflict may face in future. If we add to the equation the exponentially accelerating pace of developing unanticipated technological solutions, the question still arises if legal concerns regarding AWS can be addressed by the existing body of LOAC. Can AWS be considered legal *per se* and can be used legally under LOAC? Can we treat AWS as we would treat any other weapon and thereby disregarding the high autonomy it may have in certain critical functions? Can we trust decision-making by an algorithm written for an AWS that is probably too complex for any end-user (combatant, operator or military commander) to comprehend in its totality? With regards to the above questions, the dissertation explores whether general LOAC principles are flexible enough to deal with the challenges posed by AWS including the serious ramifications deriving from delegated decision-making. The basic premise of the dissertation for studying the principles of LOAC is that LOAC is applicable to AWS as a means of warfare.

The first studied principle (*the principle of military necessity*) limits the right of the parties to the conflict to choose any means and methods. The modern concept of military necessity requires using only that kind and degree of force that is required to achieve the anticipated military advantage. At the heart of the concept lies the criterion that no defence shall be provided in the event of unlawful actions; on the contrary: a balanced principle of military necessity fosters gaining military advantage while also manifesting the humanitarian requirements of law. Own force protection (including both military personnel and weaponry) shall be considered a military advantage, therefore, the lawful use of autonomous technology on the battlefield under the principle can be more conveniently decided in light of the anticipated military advantage. In this case, the requirement that military objectives yield some military advantage would make any separate condition for military necessity unnecessary. The employment of AWS shall not be considered unlawful under the principle, for the reason that unlike manned systems, they can attack the enemy without placing an operator at risk (own force protection). Based on the above, the concept of military necessity does not in itself seem to render AWS unlawful, since there exist situations in which they are valuable militarily. The immaterial nature of the principle of military necessity seems too elusive to give us an indication as to AWS can be regarded lawful or unlawful under its provisions.

The Martens Clause as the transposition of *the principle of humanity* cannot be called a LOAC principle as such. It can rather be regarded as a guidance for positioning LOAC (including the rules of the conduct of hostilities) in the system of international law. It ensures that LOAC does not become a self-contained regime as it opens up the possibility to interpret its norms and mechanisms in the context of other norms of IL. Therefore, the Martens Clause does not provide us with a principle directly applicable in the conduct of hostilities, but it sets the norms of LOAC into the larger framework of IL. The Martens Clause implies that principles of general IL apply during armed conflicts even if there is no particular provision in the concerning treaty law. With LOAC being silent on a certain matter, the Martens Clause may serve as a backdoor that ensures that other sources of IL may provide protection to persons intentionally involved in the hostilities (e.g. combatants) and those who unintentionally found themselves in crossfire.

Apart from being prohibited by a special convention, to call any weapon (including AWS) unlawful *per se*, it must be non-compliant with the rules of Weapons Law as contained in LOAC. Regarding non-compliance with the concerning rules of Weapons Law, not having any rule *a contrario,* my fundamental premise is still that those rules that apply to conventional weapons (understood as widely used weapons such as small arms and light weapons, mines, bombs, shells, missiles, cluster munitions, etc. that are not weapons of mass destruction) equally apply to AWS as a new, unconventional type of weapon system. (Apart from the Martens Clause, the principle of humanity also requires the study of those rules of LOAC that prohibit *inter alia* the deployment of weapons causing superfluous injury or unnecessary suffering, the use of uncontrollable means and effects or weapons without the capability to discriminate between civilian and military objects but these refer to specific effects of weapons and the autonomy of any weapon system cannot be interpreted in light of these rules.)

As previously highlighted, neither the existing AWS, nor the possible truly autonomous systems of the future shall be seen as actors under LOAC. This premise will have a significant role to play when assessing *the principle of distinction*. The first and foremost objective of the principle is the prohibition of indiscriminate attacks. It generally requires military commanders and combatants to distinguish between the following categories in the area of operation:

- civilians – combatants
- civilians – civilians taking a direct part in hostilities
- combatants *hors de combat* – combatants
- civilian objects – military objects.

Applying this second principle to AWS, the central question will be whether AWS used against lawful military objectives can be used in such a way that personnel involved in the decision-making regarding its use or involved in its use can comply with the principle of distinction, i.e. whether AWS is capable of discriminating between lawful military objectives and unlawful targets. In this regard, the AWS's ability to effectively and reliably distinguish between military and civilian objectives will depend on the sophistication of its control system (algorithms), but the operational environment, the circumstances of the attack, the attacked military objectives and the length of deployment will also play an important role in the assessment.

What complicates the assessment is that Additional Protocol I was drafted with a special attention to the *effects* of an indiscriminate attack (civilians and civilian objects being attacked) and not the implications and level of autonomy in the decision-making. AWS are means of warfare: weapon systems that can be employed by human decision-makers. These types of weapons require human beings (designer, programmer, military commander, operator, combatant, etc.) to design, program, activate and use them, therefore, they should not and cannot be seen as 'persons' (addressees of LOAC) aiming themselves. Based on their limited level of autonomy, my earlier starting point, that LOAC apply to AWS similarly as they apply to conventional warfare and conventional weapons (manned platforms) is still valid. Activation of an AWS is always a result of human decision-making and human judgment supposing that the AWS in question is capable of distinguishing thereby the persons responsible for its employment and use do not breach the principle of distinction (selecting and engaging only lawful targets).

One of the most important underlying questions the decision-maker has to deal with is whether the AWS to be selected has the sufficient capability to recognize and assess the nature and characteristics of targets (based on the circumstances) or not. The human judgment regarding the selection and activation of an AWS may well indicate that the level of AWS ability to discriminate is adequate and may also be decisive in addressing the concerns regarding responsibility. Assessing adherence to the principle require a twofold review: (1) investigating the possibility of positive identification of military objectives including combatants, civilians taking a direct part in hostilities as well as different military objects by AWS and (2) assessing whether an AWS can be aimed with an acceptable level of distinction. I believe, lawfulness will require positive answers to both questions. This presupposes the assessment of certain objective data and criteria (e.g. uniforms, distinctive clothing, insignia, in case of known persons, facial and body images) that can be measured as well as the AWS's ability to *ascertain intent* from human behaviour. AWS may comply with the first requirement, but the second one (interpreting intent) may prove to be challenging. Currently, softwares and algorithms are in most cases inadequate for understanding the context (circumstances) which is subject to change, yet comprehending its implications is a prerequisite for adhering to the principle of distinction.

In principle, certain AWS may possess the ability to comply with distinction in certain circumstances, but this compliance is dependent on the operational environment, the sophistication of the sensors and control system, as well as on the military objective to be identified (distinguished).

In certain cases, an AWS may be a better weaponeering solution (with a higher degree of precision and less chance for collateral damage) than other weapons but to make an informed decision, it is critical to determine first the type of environment where AWS are planned to be used, because the demand on the AWS to distinguish will depend on it. A higher degree of autonomy may be justified in a relatively stable, uncluttered environment against targets that can be identified without doubt (e.g. incoming missiles in desert warfare), whereas a lower degree of autonomy is needed in cluttered environment (e.g. in urban warfare), where distinction on the tactical level is not always possible. In the latter case employing AWS with high autonomy may prove to be unlawful.

The principle of proportionality enjoys a close relationship with the principle of military necessity and distinction (the latter is a prerequisite to the proportionality analysis as distinction is necessary for assessing collateral damage). The most important purpose of the principle is to protect civilians from excessive injuries and damages. The term 'excessive' is of crucial importance here as we will see, since it indicates that there will always be a certain risk of unintentional injury and death and collateral damage in civilian objects. The principle provides legal validation for the infliction of suffering that would be intolerable in peacetime. There is no rule in LOAC according to which all collateral damage is prohibited, yet, minimizing incidental injury and collateral damage limits potential international condemnation, supports proportionality and contributes to peace-making and rebuilding efforts. Under the principle, military personnel have to evaluate and measure two significantly different categories. On one side of the scale are the adverse effects (collateral damage), which are measured against the concrete and direct military advantage anticipated from the attack.

The most important question that arises regarding this principle is that if military advantage is subjective, contextual and changing as military operations are progressing (and thereby difficult to discern even for an experienced decision-maker), *how can an AWS assess the proportionality* of an attack? Regarding distinction, the crux of the problem is the inherent ability to recognize/identify lawful military objectives. As for proportionality, accepting that the assessment of collateral damage can be carried out by an AWS, the problematic issue is assessing the military advantage expected from the attack (as a whole) in a dynamic environment. Certain parameters can be written into an AWS's algorithm but positive identification of a lawful military target in itself is not enough – the military advantage expected from attacking it (at the time) necessitates an incredibly high degree of situational awareness, in-depth understanding of strategic, operational and tactical objectives (required end state) as well as an ability to assess and process all available information. I believe that it can only be achieved if the AWS has a sufficiently sophisticated algorithm and is able to receive and interpret continuous updates regarding the operation. Right now, the abstract thinking necessary for implementing distinction in a complex and rapidly changing operational environment seems to be possessed only by humans.

On the technical level, it would be feasible for an AWS to operate according to preprogrammed values. This effectively means that operators can set the values in light of the pre-determined excessive (unacceptable) collateral damage for the military objective to be attacked. In case constant update is not possible, these values shall be set at a conservative (low) level. The chances of the need to adjust the pre-set values can be significantly lowered if the use of AWS is restricted to a geographically limited, uncluttered environment for a shorter period of time. In this case, it is less likely that any adjustment is needed.

The technical possibility of pre-setting values for unacceptable collateral damage however does not mean that an AWS is able to address the qualitative judgment needed to comply with the principle of proportionality. In order to use AWS in populated area (consistent with the principle of proportionality), humans should set a value for the allowed civilian casualties, injuries and damage to civilian objectives for each type of military objective. In this case humans would do the military necessity and proportionality calculation and the AWS would call off the attack should the number of civilian casualties would exceed the predetermined allowable number.

For the time being, humans make the decisions regarding proportionality, thereby being responsible for compliance with the principle. Humans provide the highest level command regarding the deployment of an AWS (based on the operational environment) and guidance concerning the values to be set as maximum collateral damage. It is however not insurmountable that with technological advancement, algorithms would enable AWS to adjust these base levels in order to address the changes in the operational environment. To allow for these adjustments, continuous data feed has to be ensured. On the other hand, it may also be (at least technically) possible that an AWS can learn to apply the principle on the basis of feeding it with scenarios and the correct answers regarding proportionality. Nevertheless, until such time as this may come, I believe that the use of AWS in compliance with proportionality necessitates certain limitations and restrictions (e.g. temporal and geographic limitations and use restricted to situations where the risk of CD is non-existent or low).

The principle of precaution basically requires those who plan, decide and execute an attack to take constant care in order to avoid or at least minimize the risk of collateral damage. This will include *inter alia* to select a weapon or weapon system and a military objective that involves the least risk for collateral damage as well as to cancel or suspend the attack in case of unlawful (civilian) targets and the breach of proportionality. The precautionary principle would prohibit to use AWS in cases when other feasible weapons are available that pose less risk for collateral damage and promise the same military advantage. Lawful military objectives shall be attacked by weapons that expose civilians to the least possible harm (without forfeiting military advantage) – if this weapon is an AWS, then *it shall be selected as a matter of law*. In this regard, an outright ban on AWS would act against compliance with the principle of precaution and may have a counterproductive effect.

Precaution shall also be applied post-attack too, as long as the physical possibility exists to suspend or abort an attack if it is expected to cause excessive collateral damage, i.e. becomes disproportionate. The addressees of LOAC are human beings, and under the precautionary requirement, military personnel have to do everything feasible to act with the greatest care. With regards to AWS, this means the use of its sensors' capabilities to the largest extent possible in order to recognize targets.

The requirement also implies the optimal match of weaponeering solutions to the selected military objectives, i.e. choosing weapons that will cause the least collateral damage yet deliver the military advantage anticipated. An AWS can also operate with the greatest possible care – according its algorithm. If an AWS can be programmed to run regular checks (in order to confirm target identity and the proportionality of the attack), I believe that under the principle it has to be programmed to run theses checks in order to ensure that the military personnel is compliant with all LOAC principles.

Operationalizing LOAC principles

Following the theoretical analysis, the assessment of *how the above principles and rules are put to practice* is also necessary (i.e. how they are operationalized during the targeting process). Understanding how these rules guide targeting is insurmountable to the analysis of the adequacy of how LOAC today deliver its functions and to deciding on whether any new law is needed as an answer to the challenges posed by AWS. It may also serve as a reminder that stringent practices are for a long time in place to give effect to the principles and rules of LOAC.

The contemporary concept of targeting evolved only after the introduction of airpower in World War I and the targeting process has been shaped by technological developments, organizational structures as well as an inter-service competition for scarce resources.

Notwithstanding the growing number of studies and scholarly articles on AWS in light of LOAC, the targeting process (especially the targeting cycle followed by NATO and the US, being the most detailed available documents on targeting doctrine) is generally understated. Any employment of AWS on the battlefield today is effectively happening through the targeting process, therefore, it is vital to understand how it *'operationalizes'* distinction, proportionality and precautions, i.e. how the process ensures compliance with these principles on the strategic, operational, and tactical level.

Targeting is a process of selecting targets and choosing the appropriate weapon to attack them with an aim to achieve the desired operational effects in support of the commander's objectives. Adherence to the LOAC principles shall be ensured throughout the whole targeting process and by all military personnel involved. Precaution with regards to the selected means and methods of warfare shall be taken into consideration in all phases of targeting. The proportionality analysis can be carried out only if the lawful target is already developed (selected and vetted) and an AWS has been chosen as a weaponeering tool. Under LOAC, the lawfulness of attacks involving AWS must be ensured and assessed by the person launching the weapon or ordering the weapon to be launched. Not being an addressee of LOAC, the assessment of LOAC compliance (lawfulness of the attack) cannot be delegated to an AWS.

The phases of the targeting process are described in the targeting doctrines. The targeting process links strategic-level guidance (desired end state) with tactical targeting activities on the battlefield through the targeting cycle. As future autonomous technology developments unfold, what may become a question is what functions (or tasks under the targeting phases) AWS can take over from the activities covered by the phases of targeting.[1] With the improving imaging, target recognition, and data processing capabilities, the attention may turn towards carving a slice of ISR (intelligence, surveillance, reconnaissance) as this field is knowingly struggling with the extremely high data load. These data are collected not only from ISR activities but also available from open sources such as databases or social media.

[1] Generally, the phases of targeting are:

PHASE 1: political and strategic direction, establishing objectives and desired end state
PHASE 2: target development, collateral damage estimation
PHASE 3: capabilities analysis, collateral damage estimation
PHASE 4: force planning and assignment, approval of prioritized targets
PHASE 5: mission planning, force execution, positive identification, target validation
PHASE 6: targeting assessment (effects)

These capabilities could certainly facilitate the target development (target vetting, target validation) in the future.

Regarding the LOAC principles, on the strategic and operations level, distinction in targeting is adhered to within the target development, force planning and assignment, as well as force execution targeting phases. The requirement of taking feasible precautions permeates the whole targeting process from the planning phase until battle damage assessment.

It is imperative however to question *when* these feasible precautionary measures should be taken (based on the targeting phases, at (1) target development (selecting military objectives), (2) capabilities analysis (weaponeering: selecting the most adequate means and methods of warfare that match the targets), (3) force planning and assignment (assigning operative units), as well as (4) force execution (verifying the military nature of target, assessing proportionality, cancelling or aborting mission). The principle of proportionality covers the collateral damage estimation (CDE) and the assessment of the anticipated military advantage. The CDE is closely tied to the principle of distinction as estimating collateral damage is only possible if one is aware whom and what we can regard as civilian or civilian object. In CDE therefore, distinction and the proportionality test overlap.

Violating the right to life

The most important purposes of LOAC are to regulate the conduct of hostilities and to protect those who are not or no longer participating directly in hostilities (e.g. civilians, *hors de combat*, shipwrecked, PoW). While LOAC's roots go back hundreds of years, human rights are relatively young in comparison. The dissertation assesses only one human rights with regards to employing AWS on the battlefield: the right to life. In order to assess whether the right to life under international human rights law (IHRL) is applicable in armed conflicts, one shall examine the concepts existing under IHRL and LOAC.

The notion of arbitrary deprivation of life is not used in LOAC, instead the LOAC regime applies the unlawful killing concept. To connect the two concepts, the prohibition of arbitrary deprivation of the right to life under IHRL also encompasses unlawful killing in the conduct of hostilities, i.e., the killing of civilians and persons *hors de combat* not justified under the rules on the conduct of hostilities. This certainly includes the death of civilians as a result of violating distinction, proportionality and precaution. Generally, the determination of a possible violation of the right to life (guaranteed by IHRL) shall happen only with reference to the LOAC concept of arbitrary deprivation of life.

Notwithstanding the differences in the protected values, applicability, and terminology between LOAC and IHRL, there seems to be no inherent contradiction between their respective provisions regarding the right to life. There is a definite overlap between the provisions of IHRL and LOAC with the latter being more detailed and context-specific (as opposed to the general principles of IHRL). In order to ensure the maximum available protection afforded by both branches of international law, IHRL should be seen as supplementing LOAC provisions for the purposes of adequate application. For this to materialize however, hierarchy and interpretation should be agreed upon.

I believe the violation of the right to life shall be assessed under LOAC as *unlawful killing*. Unlawful killing can indeed be 'committed' by employing an AWS. In case of man in the loop weapons, an AWS is clearly instrumental, and target engagement is controlled by the operator. However, an AWS has to 'pass' multiple tests before being included in a state's weapon arsenal in order to ensure that the weapon is not unlawful *per se*.

Considering an AWS lawful as it is, only its application in a particular way may violate LOAC and result in unlawful killing (e.g. deploying an AWS designed to desert warfare in close-in urban warfare where it cannot comply with the principle of distinction or where its employment results in death or personal injury excessive compared to the anticipated military advantage). This only underlines the responsibility of the commander to ensure the AWS's adequate application (in light of the circumstances and the operational environment).

State and individual responsibility

There are two main types of responsibility: state and individual responsibility (including the individual responsibility of a person and command responsibility). The dissertation focuses primarily on command responsibility.

Under common Article 1 the Geneva Conventions, parties undertake to respect and to ensure respect for the Convention in all circumstances. This provision may also be regarded as basis for state responsibility although its scope of application is narrower: the obligation applies in respect of other states party to Geneva Conventions. Article 1 not only enables states to control the conduct of their bodies and persons acting on behalf of them, but it also makes the responsibility of the states to ensure that these bodies and persons respect the Conventions in all circumstances. Not adhering to this obligation may entail the responsibility of the state under international law. I believe that the concept of state responsibility cannot provide meaningful accountability. It is questionable how willing governments would be to enter into any discussion, negotiation or dispute settlement regarding claims arising from the use of AWS as the development and deployment of such weapons (not to mention their design and specification) may not be an information they want to disclose. Furthermore, in cases of covert operations, the lack of political acceptability may be another issue states could face. It has to be carefully measured by governments and military commanders alike whether anticipated benefits of resorting to AWS overweigh the possible loss of support in case things go wrong in the operational theatre. AWS are not addressees of LOAC and clearly, they are not organs of a state or state agents either, therefore, the persons whose actions will be attributable to a state is probably the commander deciding on its use and its operator. A state would also be held responsible when using an AWS that it has not, or has inadequately, tested or reviewed prior to deployment.

Individual responsibility has a different legal base than that of state responsibility as it is mostly based on international criminal law (ICL). It entails two different types of responsibility: that of an individual (for example a combatant or weapon system operator) and that of the commander (command responsibility). The latter can be indirect when an individual commits a violation of LOAC and the commander is not exercising sufficient oversight over the person responsible for carrying out his orders. In case of an AWS, this is definitely complicated by the fact that an AWS cannot be regarded as a person responsible for following the orders of a commander. Command responsibility can also be direct when the commander is ordering the violation.

One of the crucial questions regarding the use of AWS will be that of assigning criminal accountability in case of an unlawful killing resulting from the violation of LOAC. This require us to examine the actions and intent of perpetrators in order to find someone accountable.

For the purposes of present dissertation, alleged perpetrators will include only combatants, operators, and military commanders, and individual responsibility (the responsibility of military personnel operating an AWS, i.e. combatants, operators, or military commanders) and command responsibility (the indirect or direct responsibility of a military commander) is distinguished.

Assigning individual responsibility is easier in case there is only a low degree of autonomy in critical functions (human-in-the-loop AWS). In this case the causal relation between the activation or operation and the violation is rather close, facilitating decision with regards to responsibility for LOAC violations. For AWS with a high degree of autonomy, assigning responsibility may be problematic, and this may be further hindered by any time lapse between the decision to activate the weapon or activation and the unlawful act.

In case of a direct perpetration by a commander (e.g. intentionally deploying an AWS in circumstances where it is incapable to adhere to LOAC), we cannot talk about command responsibility, but *direct perpetration* with a direct perpetrator (therefore his responsibility cannot be indirect). In these cases, both *actus reus* (violation) and *mens rea* (mental element or intent) can be established (the commander making the decision regarding deployment is aware of the strong possibility of the violation). The lack of *mens rea* however raises the question whether a responsibility gap exists preventing liability to be imposed on a commander. I believe this question should be answered in the negative, taking into consideration that the military commander has indirect control over the action of the AWS and he exercises judgment when deciding (based on legal and targeteer advice) which means and methods of attack to choose in order to match the targeted military objective. This is further underlined by the precautionary requirement, under which the commander has to make sure that the most adequate means and methods are chosen with the least risk of collateral damage. If the commander is unfamiliar with a particular AWS's potential capabilities and limits, under the principle of precaution he has do everything *feasible* to get acquainted with its specifics in order to make a well-informed decision regarding employment.

I believe that the sole exception from liability would be if humans could be left outside the whole robotic cycle (design-manufacturing-procurement-programming-targeting). This would practically require that AWS design, manufacture and use AWS.

Commanders are responsible for preventing violations of the law and for taking necessary action. The fact that a breach of LOAC was committed by a subordinate does not absolve his superior from penal or disciplinary responsibility, as the case may be (if he knew, or had information which should have enabled him to conclude in the circumstances at the time) that the subordinate was committing or was going to commit a breach and his commander did not take all feasible measures within his power to prevent or repress the breach.

International Criminal Law (Rome Statute) is using the concept of 'knew, or owing to the circumstances should have known' which can be regarded as a stricter responsibility involving the commander's failure in exercising control over his subordinates. The most important question regarding this requirement is how the 'knew, or owing to the circumstances, should have known' expression can be translated and applied in cases of unlawful killing involving AWS and whether the criteria is in agreement with the concerning provision of Additional Protocol I on the intent and knowledge of the perpetrator.

I believe it would be more appropriate to talk about responsibility for implementing certain decisions. At the present state of technology, we cannot attribute free will to autonomous weapons but we should see them for what they are: a software programmed to carry out tasks according to their preprogrammed algorithm and preset percentage of autonomy in critical or non-critical functions. We should also recall here that the removal of humans from the final target engagement decision does not act against assigning responsibility as commanders play a substantial role in the targeting process. I believe military commanders are the best situated in the process to have situational awareness regarding deployment and under command responsibility to take all necessary and reasonable measures to prevent violations. A commander shall have an adequate overview of the operation; besides commanders do not belong to junior military personnel, therefore they have possibly long years of experience and knowledge – I believe these further attributes also point towards the possibility of establishing their responsibility for targeting decisions.

The future of warfare – the warfare of future

In the dissertation, I intended to touch upon those legal and other issues that shape the debate regarding the legality and use of AWS on the battlefield. Many of its opposers are fervently fighting to achieve a ban (or at least a moratorium) on its research and military use. It is questionable whether such an approach can be supported (promoting any ban is underlined by the argument (or premise) that LOAC – as it exists today – is not sufficient (enough) to regulate the field of AWS). I believe however that when we study the adequacy of LOAC, we have to take into account the targeting procedure as a whole, too, because any field application of AWS happens through either a national or NATO targeting process which is a complex procedure requiring in-depth assessment of the operational environment, military objectives, and the matching capabilities. It is therefore not an arbitrary decision of any military personnel to use AWS but the result of a sometimes rather lengthy procedure involving the expertise of the military commander, legal advisor, political advisor, targeteer, etc. where the decision-making authority rests at the highest level.

Notwithstanding the reasons and arguments supporting the call for a ban, I believe that it would be counterproductive. A ban would mean that no AWS are available even if they would be suitable for the circumstances and the targeted military objective (i.e. sufficiently fast and accurate thereby threatening no or less civilians). In this case, fielding traditional weapons instead of AWS may result in more or more extensive collateral damage (greater harm).

Logically, a preemptive ban is unnecessary for another reason: if an AWS cannot comply with LOAC, its development of fielding would be unlawful under LOAC which makes any ban redundant. This assumes the adequate application of the existing law, mostly through the targeting process.

In principle, LOAC and autonomy are not incompatible, and compliance can be facilitated certain legal and operational requirements. First, AWS can be lawfully employed under LOAC if its use will realize military objectives which cannot be attained by other weapons or weapon systems that would cause less collateral damage. During the targeting process, the available capabilities (weaponeering solutions, including AWS) are matched to the selected target sets. Second, the targeted military objective is also of crucial importance as the AWS must comply with the principle of distinction. For this, the required input (types of targets, e.g. ballistic missiles; dimensions; parameters; assigned values; facial features; identities, etc.) shall be preprogrammed based on objective parameters or traceable and sufficient intelligence. Third, using AWS in a primarily defensive role ensures a conservative approach, i.e. the system only attacks when encounters (recognizes) a threat, for example in case of missile defence systems where the system’s response is triggered by a signature, sound or emitted heat of objects (e.g. ballistic missiles). Ideally, intentionally targeting individuals shall only happen in circumstances where positive identification of lawful targets is possible, and no civilians are endangered; therefore, the chance of collateral damage is excluded. Fourth, limiting discrimination (and generally targeting) to a predefined combat zone would also lessen the chance of incidental injury and collateral damage and therefore could contribute to improving compliance with distinction. Based on the above, in principle, certain AWS may possess the ability to comply with distinction, but this compliance is dependent on the operational environment, the manner or use and on the target itself.

Regarding proportionality, the concerning decisions are made by humans. Collateral damage estimation could be technically carried out by an AWS yet calculating military advantage is problematic considering the present state of technology because of its highly contextual nature and exposure to rapid changes. I am confident that it will not be an insurmountable barrier in the far future but until the technology is available, that fielding of AWS in compliance with proportionality necessitates certain limitations and restrictions (e.g. temporal and geographic limitations and use restricted to situations where the risk of collateral damage is non-existent or low).

The principle of precautions requires military personnel involved in planning, decision-making, execution regarding operations to take constant care in order to avoid or at least minimize the risk of collateral damage. Challenging to program it as it is, if there are more military objectives promising similar military advantage, the one involving the least risk of collateral damage shall be chosen. The same requirement applies to AWS: the one exposing the least civilian to risk shall be selected, without forfeiting military advantage. I believe that this is probably the strongest argument against those supporting a ban: in case of a ban, the possibility of lessening the collateral damage from armed conflict will be taken away. Military personnel have the responsibility to act with the greatest care during targeting, yet, part of this obligation can be complied with by programming an AWS to regularly confirm target identity.

Based on the assessment of the dissertation, AWS are not unlawful *per se* (based on the fact that they incorporate autonomous features). However, AWS that can autonomously select and engage targets without direct human control or supervision can be used lawfully only in a fairly limited number of circumstances, mostly in simple, static, unchanging environments.

In cases falling outside these limited circumstances, human operators shall maintain some kind of oversight over the AWS (to remain a so-called red card holder with the ability to revoke or change any inadequate AWS decision).

Law is in the air?

Considering the exponential development of new technologies and weapons, further discussions and debate regarding the AWS is unavoidable; what seemed to be science fiction yesterday, is the reality of today's warfare. This process however cannot be limited to governments' intervention; representatives of civil society and NGOs, as well as roboticists, scientists, lawyers and ethicists also need to be involved in the process. This could ensure that members of the society will have a correct understanding of the implications of applying robotic technology in weaponry. Right now, there seem to be a rift between those who want to prevent the use (and proliferation) of AWS based on primarily ethical reasons and those (mostly military lawyers) who believe that in certain circumstances AWS can lawfully be deployed. I believe governments are somewhere in the crossfire between these two approaches: on the one hand they want to avoid unpopular decisions and therefore (influential) NGOs' and the civil society's concerns shall be adequately addressed by them; on the other hand, governments (based on their available resources) want to possess enhanced capabilities.

I believe that the existing LOAC framework is adequate to protect civilians, but in order to fulfill its function, adherence ought to be ensured. With regards to AWS, advocating a new law (a binding multilateral convention regarding ban or restriction) may not work considering the national interests and reluctance of those states who are largely involved in AWS development and use. Instead of a preemptive ban or a multilateral convention, a framework convention may be more successful in drawing the attention to the issue. This could circumvent the problem of committing states to a binding agreement before the full capabilities of AWS are known and could facilitate an open dialogue among states (and civil society). Alternatively, states may choose to agree on a set of operational guidelines as to how they expect AWS to interact in environments where there is no human oversight (e.g. if you shoot at a robot, expect that it will shoot back). This would be a more feasible solution to regulate their use than any convention. Or, similarly to the Tallinn Manual, a group of experts may be entrusted with collecting the applicable rules of LOAC and International Law in general. Should NATO decide to commit itself to the issue, then establishing a NATO accredited Centre of Excellence (COE) could reasonably study the new technology or an already existing COE (e.g. the Cooperative Cyber Defence COE, Modelling and Simulation COE or Military Engineering Centre of Excellence) could extend its research to AWS as well.

I believe that the recent technological developments in the field of autonomy and artificial intelligence and the need to address the different defence interests of states will inevitably lead to the fielding of a growing number of sophisticated AWS. To develop or possess AWS is ultimately a political decision that may be shaped by society's reluctance and negative opinion. Autonomous weapon systems can be the source of unthinkable military power. What is needed is to harvest the benefits their use offer, while keeping them under effective and tight control. This means finding the right balance between human and machine using both technical and legal tools.

IV. List of relevant publications

A katonai szükségesség elve az új fegyverrendszerek tükrében in: JOGI TANULMÁNYOK 1. (2016) pp 467-478, ELTE Állam- és Jogtudományi Doktori Iskola – Kari Doktorandusz Konferencia, Budapest

Military Necessity in: Tavaszi Szél 2016 – Spring Wind 2016 Tanulmánykötet, I. kötet (Gábor Keresztes ed.) pp 416-427, Doktoranduszok Országos Szövetsége

A békéhez való jog in: Themis, az ELTE Állam- és Jogtudományi Doktori Iskola Elektronikus Folyóirata (2016) pp 191-208

Application of the Right to Dignity during the Conduct of Hostilities in Light of the Recent Robotic Weapon Developments in: Tavaszi Szél 2017 – Spring Wind 2017 Tanulmánykötet, I. kötet (Gábor Keresztes ed.) pp 380-391, Doktoranduszok Országos Szövetsége

Taming the Untamable: The Role of Military Necessity in Constraining Violence in: ELTE LAW JOURNAL (2016/2), pp 93-118 (2018)

The Role of Customary Principles of International Humanitarian Law in Environmental Protection in: Pécs Journal of International and European Law (2017/II), pp 19-39 (2018)

A katonai szükségesség elvének angolszász értelmezése az USA és Nagy-Britannia hadijogi kézikönyvei alapján in: Védelmi alkotmányosság az új típusú biztonsági kihívások erőterében (Ádám Farkas ed.) pp 83-98 (2018) Magyar Katonai Jogi és Hadijogi Társaság