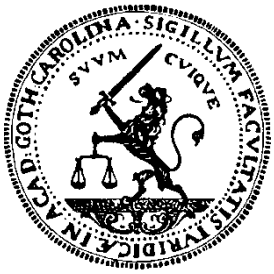


FACULTY OF LAW
Lund University

Cecilia Andersson

Killer Robots

– Autonomous Weapons and Their Compliance with IHL

JURM02 Graduate Thesis

Graduate Thesis, Master of Laws program
30 higher education credits

Supervisor: Markus Gunneflo

Semester: VT 2014

Contents

Abstract

The pursuit of weapons which distance the soldier from the actual battlefield has been going on ever since the transition from the waging of war using short blades, to the waging of war using bow and arrow. Today, that ambition has reached an almost completion with the ever-increasing number of unmanned, remote-controlled vehicles that are rapidly becoming the most common and prominent method of waging wars. Political incentives of cutting costs of warfare and sparing the lives of soldiers create the last push towards full autonomy. The emergence of increasingly autonomous weapons (AWs) has already generated a heated debate on the legality of these weapons, and two very polarized sides can be easily discerned.

The purpose of this thesis is to examine and analyze this debate, to look into the arguments put forth regarding the legality or illegality of autonomous weapons, and examine where the positions are in the debate. Focus is on the three fundamental principles in International Humanitarian Law (IHL): distinction, proportionality and precaution, and I discuss the arguments in both directions. Proponents often claim the ability of AWs to comply with IHL, with the development of sensors, algorithms, software and artificial intelligence (AI), which would allow the machine to satisfactorily distinguish between civilians and combatants, carry out proportionality assessments and to take the required precautions in its actions. Opponents instead argue that the development of AI has overpromised before, that sensors could never be able to distinguish between civilians and combatants in a contemporary battlefield and that proportionality and precaution assessments require a contextual understanding that only humans are capable of. The fundamental disagreement seems to lie in the uncertainty of the development of the software and technology, and the capability of machines to perform as well, or better than, humans. The issue of accountability is also examined in terms of what happens with the responsibility for breaches of IHL when we have assigned the task of targeting and firing, essentially, the life-and-death decision, to a machine. Different propositions such as placing the accountability onto the commander, programmer, manufacturer or even the machine itself are discussed. Issues relating to the moral and ethical aspects of changing the agents of war from humans to robots are also examined, and the possible consequences this might entail – both from a separate moral perspective and as part of the legality assessment, in terms of what would happen with the applicability of IHL if we would change the agents in war.

After having examined the debate on legality of AWs, some concluding remarks are drawn on what we are to do with the debate in the near future, where I present some of the more prominently discussed ways forward in terms of handling the emergence of these weapons. Finally, I end with some of my own reflections on what I have found in my analysis of the current debate, and what I believe are the more important aspects to continue discussing in the ongoing debate on the legality of autonomous weapons.

Sammanfattning

Jakten på vapen som distanserar soldaten från själva slagfältet har pågått ända sedan övergången från krigsföring med knivar till krigsföring med pil och båge. Idag har denna ambition närapå nått fullständighet med det ständigt växande antal obemannade, fjärrstyrda farkoster som snabbt håller på att bli den vanligaste och mest framstående metoden att föra krig. Politiska incitament såsom att kapa kostnader av krig och att spara soldaters liv innebär den sista knuffen mot full autonomi. Framväxten av alltmer autonoma vapensystem har redan genererat en passionerad debatt om lagligheten av dessa vapen, och två väldigt polariserade sidor är enkelt urskiljbara.

Syftet med det här arbetet är att undersöka och analysera den här debatten, att titta på de argument som förs fram gällande lagligheten eller olagligheten av autonoma vapen, och att undersöka var positionerna står i debatten. Fokus ligger på de tre grundläggande principerna i internationell humanitärrätt (IHL): distinktion, proportionalitet och försiktighet, och jag diskuterar argumenten i båda riktningarna. Förespråkarna framhäver ofta förmågan hos autonoma vapen att efterleva reglerna i IHL, genom utvecklingen av sensorer, algoritmer, mjukvara och artificiell intelligens (AI), vilket skulle göra det möjligt för maskinen att på ett tillfredsställande sätt skilja mellan civila och kombattanter, genomföra proportionalitets-bedömningar samt att företa nödvändiga försiktighetsåtgärder i sina aktiviteter. Motståndarna menar istället att utvecklingen av AI har lovat för mycket förut, att sensorer aldrig skulle kunna skilja mellan civila och kombattanter i ett nutida krigsfält och att bedömningar av proportionalitet och försiktighetsåtgärder kräver en kontextuell förståelse som endast människor kan klara av. Den grundläggande meningsskiljaktigheten verkar ligga i ovetskapen om utvecklingen av mjukvara och teknologi, och förmågan hos maskinerna att utföra uppgifter lika bra som, eller bättre än, människor. Frågan om ansvar undersöks också gällande vad som händer med ansvaret för överträdelser av IHL när vi överlåter uppgiften av att sikta och avfyra, i allt väsentligt, liv och död-beslut, till en maskin. Olika förslag om var ansvaret ska placeras, såsom på befälhavaren, programmeraren, tillverkaren eller till och med på maskinen själv, diskuteras. Frågor som relaterar till de moraliska och etiska aspekterna av att byta ut agenterna i krig från människor till robotar undersöks också, och de möjliga konsekvenser detta innebär – både från ett separat moraliskt perspektiv, men också som del av laglighetsbedömningen, beträffande vad som händer med tillämpligheten av IHL om vi byter agenterna i krig.

Efter att ha undersökt debatten om laglighet av autonoma vapen drar jag några slutsatser om hur vi ska fortsätta debatten i den nära förestående framtiden, där jag presenterar några av de mest diskuterade möjliga vägarna framåt när det gäller att hantera framväxten av dessa vapen. Slutligen avslutar jag med några egna reflektioner om vad jag har kommit fram till i min analys av debatten, och vad jag tror är de viktigaste aspekterna att bära med sig i den fortsatta debatten om lagligheten av autonoma vapen.

Acknowledgments

Sex år, två examen, en miljard minnen – så skulle min Lundatid snabbt kunna sammanfattas. Tiden här har, som många förvarnade, gått otroligt fort, och även om det känns som att dessa sex år försvunnit på ett ögonblick, har de innehållit upplevelser som skulle räcka en livstid. Författandet av denna masteruppsats i allmänhet, och detta förord i synnerhet, har dock på ett väldigt definitivt sätt påmint mig om att Lundatiden, ofrånkomligen, faktiskt går mot sitt slut.

Världens största tack, för ständigt stöd och kärlek, går till min underbara, galna, älskade familj – mamma, pappa, Elisabeth och Emil. Mitt liv stannar utan er.

Tack också till alla mina fantastiska fenomenala vänner – min andra familj – som på ett eller annat sätt förgyller min vardag och tillvaro. Ni är ovärderliga!

Tack till alla härliga JF-rävar jag haft turen att lära känna och jobba med i Juridiska Föreningen, ni fick mig att känna mig hemma från första terminen i Lund. Ett särskilt tack till styrelsen 2010/2011 för ett verkligt epic verksamhetsår!

Tack också till min handledare Markus Gunneflo för värdefullt stöd och insiktsfulla tips under uppsatsens gång, utan vilka jag nog fortfarande hade skrivit på mitt avgränsningskapitel.

Slutligen, tack till alla som befolkat Rummet, för att ni lyckats förvandla exjobbsskrivandet från vad som kunde blivit en outhärdlig mardröm till det som istället visade sig vara en av de bästa terminerna.

Lund, 25 maj 2014

Abbreviations

AI	Artificial Intelligence
AP I-II	Protocol Additional to the Geneva Conventions I-II of 8 June 1977
AW	Autonomous Weapon
AWS	Autonomous Weapon System
CCW	Convention on Certain Conventional Weapons of October 10 1980
CIHL	Customary International Humanitarian Law
DARPA	Defense Advanced Research Projects Agency
DoD	US Department of Defense
GC I-IV	The Geneva Conventions of 12 August 1949 I-IV
GPS	Global Positioning System
HRW	Human Rights Watch
ICC	International Criminal Court
ICJ	International Court of Justice
ICRAC	International Committee for Robot Arms Control
ICRC	International Committee of the Red Cross
IHL	International Humanitarian Law
IHRC	International Human Rights Clinic at Harvard Law School
MoD	UK Ministry of Defense
NGO	Non-Governmental Organization
SCI	Strategic Computing Initiative
UAV	Unmanned Aerial Vehicle
UGV	Unmanned Ground Vehicle
UK	United Kingdom
UN	United Nations
UNSC	United Nations Security Council
US	United States of America
WMD	Weapons of Mass Destruction
WWI	World War I
WWII	World War II

1 Introduction

1.1 Background

Wars are changing. This is hardly anything new, but something that has been prevalent throughout history. During the 20th century, the development of wars saw an unprecedented revolution – from the trenches of World War I, to the development of aerial warfare during World War II, from the development of nuclear wars at the end of WWII, to the rise of guerrilla warfare in the 1970s and 80s, from battlefields protecting civilian lives, to battlefields using civilians as human shields, from traditional interstate conflicts, to a post-9/11 context with a multitude of actors and battlefields. Our changing world has led to our changing wars, and where the rapid development of our world is going, so will the rapid development of wars follow.

One particular aspect of the evolution of warfare, has been the pursuit of distancing soldiers from the actual heat of battle as far as possible. Warfare has moved from short blades to long spears, from long spears to bow and arrow, from bow and arrow to cannons, from cannons to aerial bombing, from aerial bombing to cruise missiles.[1] Most notably, the development of air-power came to change the battlefield tremendously, and with this, military research came to focus mainly on how to further develop the use of air-power. During the 20th century, efforts to remote control aircrafts surged and several prototypes were flown both by the United States (US), British and German armies.[2] Since then, the automation of weapons and weapon systems has escalated, due to pressure to cost-cut from an increasingly diminished military budget, and a motivation of not risking the lives of soldiers which would lead to a more politically accepted war-waging.

However, the rise of military robotics is not without obstacles – sceptics are rapidly becoming more and more vocal in their concerns about both the legality and ethicality of autonomous weapons (AWs)[3] and the implications that they could have on future warfare. In the pursuit of more cost-effective weapons, weapons that save more lives, weapons that will not risk the lives of the soldiers – weapons that will be more politically accepted – it seems that the proponents are skipping important parts of the discussion of these weapons; not only the legal implications of today, but also the bigger impact and bigger picture of these weapons of tomorrow. The balance between what

[1] Sharkey, Noel "Saying 'No'! to Lethal Autonomous Targeting" in *Journal of Military Ethics*, Vol. 9, No. 4, 2010, p. 369.
[2] Krishnan, Armin *Killer Robots: Legality and Ethicality of Autonomous Weapons*, Ashgate Publishing, 2009, pp. 15-19.
[3] In short, AWs can be described as a weapon system that, once deployed, can carry out its mission autonomously, make decisions and address issues that arise along the way, and on its own, take targeting and firing decisions.

is scary in this development and what is militarily advantageous, is far from being struck and this unresolved issue will continue to influence the debate.

In general, the debate can be summed up in saying that the proponents of AWs emphasizes the advantage of machines over humans in their campaigns, and the superiority of machines in many aspects, such as human fallibilities in processing information, lower response times, fatigue etc.[4] – they are more desirable from a military point of view. The opponents instead emphasizes the adherence to international humanitarian law and the inability of the weapons to comply with the fundamental principles therein, as well as the ethicality and morality of using robots in war and taking the human completely out of the loop, and most importantly, letting machines make life-or-death decisions.[5]

1.2 Purpose and Research Questions

The purpose of this thesis is to examine the legality of autonomous weapons, and more specifically, the current debate on the legality of autonomous weapons. The main question in the debate is if and how autonomous weapons can comply with the fundamental principles of International Humanitarian Law (IHL); the principles of distinction, proportionality and precaution. After examining the arguments regarding the legality, I will move on to the sub-question, if these weapons are not able to comply with current IHL, or if information and detail on the compliance is insufficient, how are we to treat these new weapons and their legal status?

The point of departure will be in the current situation, where there are two strong sides to the use of AWs – the proponents and the opponents. Both sides are strong advocates for their cause, but both sides lack an objective argumentation in order to convince the reader of the claims of each respective side. Both sides tend to get overly subjective in their arguments, and tend to miss out on addressing the opposing side's arguments, thus giving the impression that each side is only interested in giving its own perspective, and not providing the full image.

One of the motivations for this thesis is to provide a more objective overlook of the debate and compare and value the arguments of each side in relation to the core issues that arise with AWs, in order to understand the debate and ultimately, where the disagreement lies. The purpose is thus rather to provide an overlook and not necessarily to come to an actual conclusion on legality. I will examine both sides of the debate and aim to keep the analysis as objective as possible. Nevertheless, in a polarized debate it is inevitable to stay completely objective, and as the debate leans back and forth, so will my reflections and conclusions. Furthermore, I will examine the situation where,

[4] Krishnan (2009), pp. 40-42.
[5] Asaro, Peter "On banning autonomous weapon systems: human rights, automation, and the dehumanization of lethal decision-making" in *International Review of the Red Cross*, Vol. 94, No. 886, Summer 2012, pp. 699-701.

if it is not possible to determine the legality of AWs (because of the limited amount of information and specificities of AWs today), what are we going to do with them and how are we to treat them? Perhaps a regulatory framework would be the best way forward in this situation? To leave a weapon system of this magnitude unregulated, even during its development phase, could possibly prove disastrous, as it could lead to unlimited proliferation and use of these weapons, in situations that were not even imaginable in the development.

Throughout the thesis, the underlying notion will at all times be the protection of civilians, and this is the reason why focus will be on these parts of IHL. Conversely, I will disregard other aspects that AWs will pose, such as the relation between robot and commander (e.g. the case of when a robot does not obey orders from the commander and the issues that may arise), and more technological aspects and issues of AWs that will not be relevant for the examination of AWs from a legal perspective.

1.3 Method and Material

In this thesis, I will first and foremost be performing an analysis of the argumentation, keeping in mind the purposes of such an analysis – first, to describe the arguments and the debate, and second, to value the extent of how well the arguments put forward really support the position taken. In the analysis, the method of "pro-et-contra" will be used, in order to carry out the second purpose, of examining the value of the arguments.[6] In this debate on AWs, the positions are very clear, thus the pro-side and the contra-side are easily identified, and the analysis there will consist of examining the sustainability of the argument, and its credence and relevance in the debate.[7]

As mentioned above, the purpose is perhaps not so much to come to an actual conclusion regarding the legality in itself, as it is to come to a conclusion regarding the arguments put forward in this heated debate, and how we should be viewing this and what lessons and conclusions we can draw from the result of the debate, and how this can help us in the further discussion on AWs. In performing this overview and analysis of the debate and the relevant arguments put forth, I will however allow myself to sidestep the structure of the debate in one aspect, and discuss the aspect of morality and ethicality of AWs – an aspect often overlooked or even ignored in the debate. When an analysis of a debate is being done, it is important to bring up and analyze what is being discussed, but I also think it is important to bring up and analyze what is *not* being discussed – in the present case, the morality and ethicality.

The materials used are mainly articles published in relevant scientific journals and books published in the field. Since the debate on the legality of AWs and the area of research in general is exceptionally new in character, the material

[6] Bergström, Göran & Kristina Boréus (eds.) *Textens Mening och Makt – Metodbok i Samhällsvetenskaplig Text- och Diskursanalys*, 3rd ed., Studentlitteratur 2012, pp. 93-94.
[7] Ibid., p. 127.

used is mainly based on the articles, as books tend to become dated fast within this field, whereas articles generally are more updated and are being published at a rate incomparable to that of books. To a limited extent, reports and studies from various organizations will also be used. With regards to the temporal aspect, the material I have used for the substantive parts of the thesis (i.e. publications) dates back no longer than to 2007, for the reasons mentioned above.

I have chosen to focus on the publications by certain distinguished authors and scientists, such as Noel Sharkey, Professor of Artificial Intelligence (AI) and Robotics; Peter Asaro, Philosopher of Technology with a background in AI and robot vision research; Robert Sparrow, Professor of Political Philosophy and Applied Ethics; Michael N. Schmitt, Professor at the US Naval War College, Jeffrey S. Thurnher, Lt. Col. US Army and Military Professor at the US Naval War College; Matthew C. Waxman, Professor of Law and member of the Hoover Institution's Task Force on National Security and Law; Kenneth Anderson, Professor of International Law and member of the Hoover Institution's Task Force, and Peter W. Singer, Director of the Center for 21st Century Security and Intelligence with expertise on military technology and contemporary warfare. The reason for focusing on these authors is that they are all prominent in their respective field, as well as prominent internationally in the debate on AWs.[8]

Research carried out in this field is of an inconclusive character. It is clear from the number of articles being produced, debates held and lobbying from Non-Governmental Organizations (NGO) that it is a very current subject and a lot of studies and research on the development of military technology are being done. However, these studies are mostly carried out by the military branches and are therefore often (inherently) secret and confidential. At the same time, it is also important to note that the research has to be approached carefully and with caution, as it is often speculative and uncertain.

Initially, the debate has been surrounding the issue of whether or not AWs will be able to comply with IHL, especially the principles of distinction and proportionality, as well as the sensitive area of accountability. However, the debate now seems to be shifting focus slightly, and with the evermore vocal calls for a preemptive ban,[9] the focus of the international community today seems to be moving more towards the question of whether or not there should

[8] For example, the Geneva Academy hosted a debate on "Autonomous Weapon Systems: Dangerous Killer Robots or Smarter and Less Harmful Warfare?" on 20 November 2013, between Matthew Waxman and Peter Asaro and moderated by Noam Lubell. The debate can be accessed at http://icrac.net/2013/11/geneva-academy-debate-matthew-waxman-vs-icracs-peter-asaro/.

[9] Cf. e.g. Opinion – Viewpoint "UK roboticist Prof Noel Sharkey calls for a pre-emptive ban on the deployment of autonomous weapons" in *The Engineer*, 11 March 2013; Human Rights Watch and International Human Rights Clinic *Losing Humanity: The Case against Killer Robots*, November 2012, p. 1 (hereinafter HRW Report); Asaro (2012), p. 687.

be an international convention set in place that prohibits AWs altogether, or at least that regulates the development and use of them.[10]

Furthermore, in order to understand the relevant material laws and regulations, I will for these parts of the thesis be employing the traditional legal method, i.e. examining material relevant for the field, analyzing it and drawing conclusions from my findings. This includes both current treaties and their commentary, case law (albeit limited) from the International Court of Justice, customary law and scientific research.

Finally, it is important to note that the thesis will have a mainly theoretical approach with regards to the discussion on AWs, because the amount of practical experience of these fully autonomous weapons is virtually non-existent. I will however draw some experience from e.g. the use of drones and systems with limited autonomy, but the main parts of the thesis will remain theoretical.

1.4 Delimitations

In the discussion on compliance with IHL, I have limited my focus to three cardinal principles of IHL – distinction, proportionality and precaution. This decision is made partly due to their importance in the entire body of IHL and partly on the prevalence of which they are being discussed in the international debate on AWs. With regards to the discussion on accountability, I will not be discussing the matters of state responsibility for internationally wrongful acts, even if these wrongful acts are committed by autonomous systems belonging to that state, since the legal complexities of e.g. attributability is a separate matter from the aims of this thesis. Finally, related areas such as neuroscience in relation to weapons ("neuroweapons") and performance-enhancing technology such as e.g. exoskeletons (worn in movies such as *Robocop* and *Iron Man*) will also be disregarded in the thesis, while these areas relate to the question of the development of warfare and technology, they do not relate to autonomous weapons as such.[11]

In the discussion on autonomous weapons, the question of drones will naturally arise, however, in this thesis I will not be addressing the matter of drones[12]. Today, current warfare is to a large extent already existing of remote-controlled weapon systems, or drones. They are considered the future of almost every air-force, and we are already at a place where not only states have access to drones. With the technology spreading ever faster, armed groups and insurgents have their own drones, and using easy instructions

[10] Altmann, Jürgen et al. "Armed Military Robots: Editorial" in *Ethics and Information Technology*, Vol. 15, No. 2, 20 June 2013, p. 74; Asaro (2012), pp. 688-689.
[11] See e.g. Dahm, Werner J.A., US Air Force presentation "Technology Horizons: Vision for Air Force 2010-2030, Capabilities Enabled by Science & Technology", Washington D.C., 25 June 2011, p. 15.
[12] I will however present a brief overview of remote-controlled weapons below, in chapter 3.2.

found online, there is even the possibility to make your own drone.[13] But we are already there, this is what happens today – what I am going to focus on is the next step, what happens *after* the drones? Therefore, the thesis will focus on the issue of autonomous weapons, where I will be discussing several different levels of autonomy, but primarily the development of what is being called full, or complete, autonomy.

1.5 Structure

This thesis will start with the second chapter where I will present a short overview of the vital parts of IHL that I will be discussing in relation to AWs and chapter three will provide the historic background as well as definitions of different autonomous weapons. These two chapters will serve as necessary backdrop for the examination that will follow in chapter four. Chapter four is by and large the core chapter of this thesis, and will offer an examination of the arguments on legality of these new weapons and their special features.[14] This will be the focal point where I examine the arguments put forward in the debate and show where the core disagreement on the legality of AWs lie. Chapter five will draw on chapter four for conclusions on the question posed in this first chapter. It will also examine some of the proposed ways forward in the debate on AWs, i.e. how we are to treat these weapons if we are not able to say anything for certain on the legality. Chapter six will to a certain extent summarize the findings from the previous chapters and offer some own reflections and concluding remarks. Reflections and my concluding remarks will primarily be found in the chapters assigned for this, however, some reflections will occur throughout the text as well.

[13] Noam Lubell Speech on "Robot Warriors, Terrorists & Private Contractors: What Future for the Laws of War?" in *Professorial Inaugural Lecture Series*, University of Essex, 13 December 2013.

[14] While I will present different kinds of autonomous weapons and discuss different levels of autonomy, the examination of legality will be on weapons with full, or complete autonomy.

2 Relevant International Humanitarian Law

2.1 The Fundamental Principles of IHL

The protection of civilian population and civilian objects is a principle that dates back to the very beginning of the waging of war, and the laws of this protection are highly developed and extensively codified, both in international and non-international armed conflicts. The purpose of waging a war to begin with is to defeat the enemy. In line with this, war-waging parties must respect the fundamental principle of military necessity – a combatant may only use the amount and kind of force necessary to defeat the enemy, to incapacitate the adversary and get them *hors de combat*, nothing else. Hence, the purpose is not in itself to kill the enemy, and any use of force more than what is required to partially or completely immobilize the enemy, is prohibited.[15]

If the most fundamental principle in regards to combatants is to only use the amount of force necessary to render them defeated, the most fundamental principle in regards to civilians is that they should be protected at all times. For the purpose of protecting the civilian population in times of war, several fundamental principles exist. With regards to the examination of autonomous weapons that will follow, I will focus on the three principles of distinction, proportionality and precaution, as these three have specific relevance for the scope of the use of AWs. The purpose of this chapter is thus to provide a brief overview of the rules and norms of IHL that are relevant for understanding the debate and the arguments put forth on the legality of AWs.

2.1.1 The Principle of Distinction

Perhaps the most essential expression of the protection of civilians in times of war, is the principle of distinction, which has twofold requirements: firstly, to distinguish between civilians and combatants as well as civilian objects and military objectives, and secondly, to only direct attacks on the latter. The principle of distinction can be traced back to the St. Petersburg Declaration of 1868, which in its preamble states that the only legitimate object of warfare is to weaken the military forces of the enemy, implicitly illegitimizing attacks on civilians.[16] In the Hague Regulation of 1907, Article 25 specifies that

[15] Greenwood, Christopher "Humanitarian Requirements and Military Necessity" in *The Handbook of Humanitarian Law in Armed Conflict*, Dieter Fleck (ed.), 2nd ed., Oxford University Press, 2008, p. 35.
[16] St. Petersburg Declaration Renouncing the Use, in Time of War, of Explosive Projectiles Under 400 Grammes Weight, Saint Petersburg of 11 December 1868.

attacks on towns, villages, dwellings or buildings that are undefended are prohibited – implicitly protecting civilian lives and objects.[17]

With regards to current treaty law, the principle is foremost found in Article 48 in Additional Protocol (AP) I, which states that parties to the conflict at all times must distinguish between civilians and combatants as well as between civilian objects and military objectives, and only direct operations against the latter. This basic rule is set out in order to ensure respect for and protection of the civilian population and civilian objects in times of war. In addition to Article 48, the operational Articles 50 and 52 are necessary in defining the scope of civilian population and civilian objects, as well as regulating more specifically under which circumstances military objectives may be attacked. In Article 51(3) the rule creates an exception to the prohibition, and states that civilians who take a direct part in the hostilities are exempted from the protection and may therefore be lawfully attacked during such time as they directly participate – an exception that applies to all kinds of weapons.

The importance of the principle of distinction also in customary law is first and foremost marked by the fact that the International Committee of the Red Cross (ICRC) has established the principle as customary law through its study on Customary International Humanitarian Law (CIHL), and sets it out as rule number 1.[18] Moreover, the International Court of Justice (ICJ), in its Advisory Opinion on the Legality of the Threat or Use of Nuclear Weapons, stated once and for all that the principle of distinction is one of the cardinal principles of humanitarian law, and that it constitutes an intransgressible principle of CIHL.[19] Furthermore, in the Rome Statute of the International Criminal Court (ICC), the deliberate attacking of civilians or civilian objects constitutes a war crime in both international and non-international armed conflicts.[20]

2.1.2 The Principle of Proportionality

Similar to the principle of distinction, the principle of proportionality is of vital importance in IHL. It demands the parties in conflict to ensure that their attacks are proportional, i.e. that the attacks that cause incidental loss of civilian life are not excessive in relation to the anticipated military advantage. Proportionality is a logical extension of the principle of distinction, and it is an expression of accepting that, while distinction should be upheld, it is impossible in wartime to avoid the loss of civilian life altogether. To this end, the principle of proportionality allows for a certain "collateral damage" in an

[17] Convention (IV) respecting the Laws and Customs of War on Land and its annex: Regulations concerning the Laws and Customs of War on Land, The Hague 18 October 1907.
[18] Henckaerts, Jean-Marie & Louise Doswald-Beck *Customary International Humanitarian Law*, Cambridge University Press, 2005, p. 25.
[19] International Court of Justice Advisory Opinion on the Legality of the Threat or Use of Nuclear Weapons (hereinafter Nuclear Weapons Advisory Opinion), 8 July 1996, paras. 78-79.
[20] Rome Statute of the International Criminal Court (hereinafter ICC Statute) Articles 8(2)(b)(i)-(ii) and 8(2)(e)(i)-(ii).

attack, if it is proportional to the anticipated military advantage to be gained from the attack. Evidently, the scope of what is an excessive loss of life and what is an anticipated advantage is impossible to define in general terms, and the balancing act it generates leaves an inevitable and considerable margin of appreciation in the assessment.[21]

In the context of the Geneva Conventions (GCs), the principle is expressed in AP I Article 51(5)(b) where attacks that "may be expected to cause incidental loss of civilian life, injury to civilians, damage to civilian objects, or a combination thereof, which would be excessive in relation to the concrete and direct military advantage anticipated" are considered as indiscriminate and as such, prohibited.[22] In determining the scope of "excessive", the AP I does not provide any further guidance, and the 1987 commentary to the protocol only gives reference to the commentary of Article 57, as Article 57(2)(a)(iii) also reiterates the principle[23], and merely states that the issue of proportionality poses a delicate problem. In some situations there might be no doubt, while in other situations hesitation might arise – the commentary finishes in noting that where such situations occur, the interest of the civilian population should prevail.[24]

Therefore, in relation to the wording in AP I, the scope of the provision remains unclear which creates difficulties in the application, as it is subject to subjective assessment and margins of appreciation. Objective standards are non-existent, and the assessment is necessarily based on future effects of the attack, an assessment that will be made by the military which has to act and make decisions only on the basis of the information available at that specific time. In the ratification of the protocol, several states filed declarations saying that the decision taken by the responsible person must be judged on the information available at the time, and not on information in hindsight.[25] In addition, the ICC Statute criminalizes disproportionate attacks in international conflicts, however not in non-international ones.[26]

In customary law, the principle corresponds to rule 14 of the ICRC CIHL study, which concludes that state practice establishes this principle as a rule of CIHL and that it is applicable in international and non-international conflicts alike.[27] The rule prohibits the launching of an attack that would cause excessive collateral damage, in the same wording as the provision set out in AP I Article 51, as mentioned above.

[21] Oeter, Stefan "Methods of Combat" in *The Handbook of Humanitarian Law in Armed Conflict*, Dieter Fleck (ed.), 2008, p. 198.
[22] See AP I Article 51(4) "Indiscriminate attacks are prohibited".
[23] Article 57 will be treated in chapter 2.1.3.
[24] Sandoz, Yvez, Christine Swinarski & Bruno Zimmerman (eds.), *Commentary on the Additional Protocols I and II of 8 June 1977 to the Geneva Conventions of 12 August 1949*, ICRC 1987 (hereinafter Commentary AP I), comment no. 1979.
[25] Oeter (2008), p. 205.
[26] ICC Statute Articles 8(2)(b)(iv).
[27] Henckaerts & Doswald-Beck (2005), p. 46.

2.1.3 The Principle of Precaution

The third of the fundamental principles of IHL in the pursuit of protecting the civilian population and avoiding or minimizing civilian loss or damage is the principle of precaution. It widens the scope of considerations that need to be made before an attack, it requires constant care to be taken to spare the civilian population, and specifically, it requires the parties to take all feasible precaution when conducting military attacks, in order to not cause unnecessary or superfluous injury or suffering. Precaution is not least important in attacks by remote-controlled weapons as the danger of military excessive collateral damage is exceptionally high with these weapons, even if recent technology development has improved the accuracy of targeting in these weapons.[28]

In treaty law, the principle of precaution first appeared in the 1907 Hague Convention (IX) Article 2(3), which required the commander to take all due measures to cause as little harm as possible.[29] Within the scope of the GCs, the principle is found in AP I Article 57. In the commentary to Article 57, special attention is drawn to the requirement in sub-paragraph 2(a)(i) of Article 57, that of properly identifying the objectives to be attacked as military objectives, and states that this identification should be carried out with the greatest of care, especially when the objective is located at a great distance. It also mentions that in those cases where there is the slightest doubt, the call for additional information and/or further reconnaissance is necessary to fulfill the requirements set out.[30]

One important aspect mentioned in the commentary is that in the cases of objectives located at a distance, reconnaissance is often carried out by aerial units and that the evaluation of this information must include a serious check of accuracy, noting the risk of the enemy setting up fake military objectives or concealing the true ones.[31] The extent of "everything feasible" was not further explained through the commentary, and it seems as if the interpretation is "everything practicable or practically possible", and that interpretation and application has to be a matter of common sense and good faith.[32]

The other important aspect codified in Article 57 is that of choosing a means of attack "with a view to avoiding, and in any event minimizing" collateral damage[33], as well as where there is a choice in the attack, to select the military objective that would generate the least collateral damage.[34] In general, this

[28] Oeter (2008), p. 190.
[29] Convention (IX) concerning Bombardment by Naval Forces in Time of War, The Hague 18 October 1907.
[30] Commentary AP I, comment no. 2195
[31] Ibid., comment no. 2195.
[32] Ibid., comment no. 2198.
[33] Additional Protocol I, Article 57(2)(a)(ii).
[34] Ibid., Article 57(3).

means that the weapon with the most accurate delivery parameters should be the one used in the attack where civilians are present or in the vicinity.[35] It also encompasses the choice between several military objectives, where the one which could achieve the same military advantage but that would constitute a smaller risk of harming civilians should be chosen. An example of this could be to instead of choosing a railway station inside a city, choosing a strategic point of the railway junction outside of the urban area – an alternative that would generate the same military advantage but cause less collateral damage.[36]

The temporal aspect in choosing military objectives is also important, such as planning the attacks at times when the presence of civilians is reduced.[37] In the choice between means and methods of combat in order to prevent collateral damage, the precision and range of the weapons at hand should be taken into account.[38] According to the ICRC CIHL study, the principle of precaution constitutes a rule of CIHL in both international and non-international conflicts.[39] With regards to the requirement to take all feasible precaution in the choice of means and methods of warfare, rule 17 of the ICRC study states that such precautions include considerations of the timing of the attack, avoiding combat in populated areas, the selection of means and methods of warfare proportionate to the target, the use of precision weapons and target selection.[40] The target selection comes back in rule 21 and re-iterates what Article 57(3) of Additional Protocol I says, that wherever a choice is possible, the least damaging one must be selected.

2.1.4 Martens Clause

At the Hague Peace Conferences in 1899 and 1907, the Russian delegate Friedrich von Martens introduced the clause that now bears his name, which was originally devised to handle a disagreement between the parties at the conference regarding the status of resistance movements in occupied territory. In the absence of an agreement, the clause was included to remind the parties that just because something was not explicitly regulated in the treaty, it did not mean that the interpretation of the status of the resistance movements was free. Instead, it had to be resolved by reference to the principles of humanity and the public conscience.[41]

In the creation of the APs in 1977, the Martens Clause was included in Article 1(2)[42] and the clause was intended to remind of the binding character of

[35] Oeter (2008), p. 210.
[36] Ibid., p. 211.
[37] Commentary AP I, comment no. 2200.
[38] Ibid.
[39] Henckaerts & Doswald-Beck (2005), p. 51.
[40] Ibid., p. 58.
[41] Greenwood (2008), pp. 33-34.
[42] Article 1(2) reads:

> In cases not covered by this Protocol or by other international agreements, civilians and combatants remain under the protection and authority of the

CIHL, to prevent interpretations *e contrario* of provisions of IHL that would be contrary to the very spirit of IHL itself, and also, to the extent that it is not covered by Article 36, to cover the development of new means and methods of warfare.[43] The exact scope and significance of the clause is still disputed[44] but it can be asserted that it means that the mere omission of a certain matter in international law, does not mean that international law necessarily is silent on the matter. The exact application of the clause with respect to requiring that all means and methods of warfare are to be measured against the standard of "the public conscience" is however not clear, as the concept of "the public conscience" is too vague to be interpreted in this manner.[45]

2.2 Article 36: Regulation of New Weapons

In the development of the APs, the drafters also saw the need to include a provision regulating the creation of new weapons, in order to ensure the future effective implementation of the basic rules prohibiting certain means or methods of combat, something that resulted in part III of AP I. At first, an internationalized control mechanism was envisioned, but due to the secrecy surrounding the development of new weapons in each state, it was deemed impracticable.[46] Instead, the drafters drew inspiration from the several states which already had such procedures implemented nationally, and established a mandatory system of national control procedure of legality, resulting in Article 36 of AP I.[47] The Article requires states to, in the study, development, acquisition, or adoption of a new weapon, means, or method of warfare, determine whether its employment would, in all or some circumstances, be prohibited by AP I or any other rule in international law applicable to the state.

The commentary to the Article specifies that the determination shall be based on the normal use of the weapon as anticipated at the time of the evaluation, and that the issue of state responsibility arises if the state fails to respect this obligation.[48] The commentary also addresses the difficulties in the assessment of legality, and the scope of "any other rule of international law applicable to the High Contracting Party". The rules set forth in AP I is quite self-explanatory, but the concept of "other rules of international law" implies any agreement related to disarmament, prohibition, limitation and restriction on the use of a certain weapon or a certain type of weapon, concluded by the

principles of international law derived from established custom, from the principles of humanity and from the dictates of public conscience.

[43] Fleck, Dieter "The Law of Non-International Armed Conflicts" in *The Handbook of Humanitarian Law in Armed Conflict*, Dieter Fleck (ed.), 2008, pp. 619-620.

[44] Which will also affect the applicability of the clause in the matter of autonomous weapons, see below chapter 4.3.4.

[45] Greenwood (2008), pp. 34-35.

[46] Commentary AP I, comment no. 1463-1464.

[47] Ibid., comment no. 1467.

[48] Ibid., comment no. 1466.

state, and perhaps most importantly, rules that form part of international customary law.[49]

The commentators in 1987 even envisaged the development of long distance and remote-controlled weapons, as well as the increased automation of the battlefield, as matters that would come to be affected by the provision in Article 36 and be object for such an evaluation of legality.[50] Also important to note is that if a weapon is found to be illegal in one state, it does not automatically create a rule of international law vis-à-vis third states, nor is there any obligation for the state to make its findings public or to reveal anything regarding new weapons during the process of development or manufacturing.[51] The status of Article 36 as customary law is not entirely cemented, and some argue that it is seen as best practice instead of binding law.[52] The ICRC contends that it is an obligation under Article 36 that several non-party states have implemented,[53] including the US which is at the forefront of research in military robots. However, the exact application and interpretation of the scope of Article 36 is not further developed in the Article or the rest of the AP, and therefore merits a brief examination.

2.2.1 When to Review?

In order to establish a more harmonized practice of Article 36, the ICRC issued a guide on the implementation of the Article, intended to provide some guidance to member states on vague or unclear matters. To begin with, the temporal aspect of the review is, according to the ICRC guide, set to take place at the earliest possible stage, i.e. at the conception or design of the weapon, and continue throughout the following technological development, such as developing prototypes and testing, and should at the very latest be done before entering into a production contract.[54] The guide also states that at each stage of the review, the intended or expected use of the weapon should be taken into consideration. One of the reasons for requiring the review to take place as early as possible is to avoid huge costs in the, often prolonged, development process for a weapon that in the legal review might end up being considered as illegal and therefore unusable. This also applies for the requirement to continuously review the weapon after it has left the design stage, and even the production stage, as the technological characteristics and its expected use may change during the continued development phase, i.e. the life cycle of the weapon.[55]

[49] Ibid., comment no. 1472.
[50] Ibid., comment no. 1476.
[51] Ibid., comment no. 1481.
[52] HRW Report (2012), p. 21.
[53] Henckaerts & Doswald-Beck (2005), p. 250.
[54] ICRC *A Guide to the Legal Review of New Weapons, Means and Methods of Warfare: Measures to Implement Article 36 of Additional Protocol I of 1977* (hereinafter ICRC Guide), Geneva, 2006, p. 23.
[55] Ibid., p. 24.

The review should also encompass future alterations and modifications to a certain weapon, since a weapon or weapon system may at a first review be deemed lawful in accordance with Article 36, but after alterations or modifications, may not at all times be considered as such.[56] With regards to AWs, this means that a lot of autonomous systems that are deployed today might require a re-review if they after a while develop the possibility to be weaponized, as well as weaponized systems with limited autonomy today, that have the possibility to be fully autonomous in the future.[57]

2.2.2 What Rules to Consider?

The second part in the legal review is to determine what rules of international law to consider. Article 36 mentions prohibitions "by this Protocol or by any other rule of international law", which includes specific prohibitions in international treaty law, such as e.g. the St Petersburg Declaration, the 1907 Hague Conventions, the 1925 Geneva Protocol for the Prohibition of the Use in War of Asphyxiating, Poisonous or Other Gases, the 1993 Chemical Weapons Convention and the 1997 Convention on the Prohibition of Anti-Personnel Mines. In addition to considerations of treaty law, the legal review must also consider relevant rules under customary law as well. The next step is to consider the legality of the weapon with regards to its ability to comply with the rules and principles set forth in the AP itself, most notably, the principles of distinction and proportionality.

Lastly, the review should consider the weapon in light of the principles of humanity and dictates of public conscience, also known as the Martens Clause.[58] This last clause was found to constitute a rule of customary law by the ICJ which noted that it "had proved to be an effective means of addressing rapid evolution of military technology".[59]

2.3 Accountability

While accountability is not a necessary requirement for the determination of legality of a weapon, it is one of the main issues in the debate on autonomous weapons and therefore merits an examination. The issue of accountability is crucial in the enforcement of IHL and it arises in relation to questions of violations of IHL, i.e. when a provision, rule or norm of IHL has been violated, who can be held accountable for that violation? In terms of regular combatants it is often an easy matter to resolve – if a soldier did not respect the principle of distinction and intentionally or by negligence directed his attack against a civilian population, that would constitute a breach of IHL and the soldier would consequently be held accountable for that breach. One

[56] Ibid., p. 10.
[57] HRW Report (2012), p. 23.
[58] ICRC Guide (2006), pp. 10-14.
[59] Nuclear Weapons Advisory Opinion, paras. 78 & 84.

further argument is the necessity of accountability to ensure that a war is considered a "just" war, under the principle of *jus in bello*.[60]

Accountability ensures that persons committing war crimes are brought to justice and prevents impunity. Enforcing the laws and rules of IHL is however subject to ongoing discussion and the fact that the legal rules are not enforced through a central body is still one of the fundamental weaknesses of IHL.[61] In the United Nations (UN) Charter, the means of ensuring compliance lies with the Security Council (UNSC) in its responsibility of maintaining international peace and security.[62] This gives the UNSC the competence to take coercive measures against a threat to or breach of the peace. However, it is primarily directed against the misconduct of states, and not the misconduct of e.g. individual soldiers, which primarily lies under the jurisdiction of domestic law. For grave breaches of international law however, the ICC Statute has jurisdiction also over individuals.[63] The crucial issue of assigning accountability to someone, being able to hold someone responsible for breaches of international law, is one of the more disputed matters with regards to AWs, and will therefore be one of the more important obstacles to overcome in the discussion on legality of AWs.

[60] Sparrow, Peter "Killer Robots" in *Journal of Applied of Philosophy*, Vol. 24, No. 1, 2007, p. 67.
[61] Wolfrum, Rüdiger & Dieter Fleck "Enforcement of International Humanitarian Law" in *The Handbook of Humanitarian Law in Armed Conflict*, Dieter Fleck (ed.), 2008, p. 675.
[62] See e.g. the Charter of the United Nations, 26 June 1946, Articles 2(7), 24(1), 50 and 53.
[63] ICC Statute Articles 1, 5 and 8.

3 Autonomous Weapons

3.1 From Mechanical Knights to Remote-Controlled Flights: History of AWs

Scientists as early as the 15th century imagined automated machines, e.g. Leonardo da Vinci designed a mechanical knight that carried out complex movements through a system of strings and pulleys. However, it was not until the late 19th century before the more sophisticated automated machines started to see the light, where Nikola Tesla built a version of a remote-controlled electric boat that could also be designed to carry warheads, in other words, an early variation of the modern torpedo. In WWI, the German Navy built upon the ideas of Tesla and started to experiment with remote-controlled torpedoes, and the British Air Force as well demonstrated a prototype of an aerial torpedo in 1917. None of these attempts proved any distinct success however, and when the war ended, they were shut down.

When WWII hit Europe, the efforts of developing more automated weapons recommenced, and Germany, among others, managed to launch the *V-1*, an aerodynamic robot weapon, a flying bomb, that could be used on a massive scale and cause considerable damage to the enemy. While the *V-1* successfully attacked the United Kingdom (UK) and caused severe damages to the population and buildings, the guidance system of the *V-1* was far from accurate, and it was only possible to target large areas, and even then, the average miss distance was about five miles. Germany also designed a prototype of a remote-controlled ground vehicle, called *Goliath*, which were powered by an electric motor and could carry explosives more than one mile. In spite of this groundbreaking technology, the *Goliath* was not considered a success either, as it moved very slowly, required the operator to be in line of sight (because the machine was not equipped with a camera system), and the control cable that connected the machine to the joystick operating it, could easily become entangled or even severed by the enemy. As no apparent success had been reached with the attempts of automated weapons, researchers instead focused on weapons of mass-destruction, and with the detonation of the two nuclear bombs in Japan the war ended and the development of automated weapons came to an abrupt halt. The military research instead entered the nuclear and missile age.[64]

During the Vietnam War, the US military (primarily) saw the emerge of precision munitions and smart weapons that were autonomous in the sense that they could find and attack targets once they were launched by a human operator, with the targeting often based on radar or other sensor data.[65] Apart from this development, the nuclear weapon stand-off and the space race were predominant during the Cold War era, and progress in the field of automated

[64] Krishnan (2009), pp. 14-19.
[65] Ibid., p. 21.

weapons was slow, not to say non-existent. In 1983, Defense Advanced Research Projects Agency (DARPA) launched its Strategic Computing Initiative (SCI), aiming to achieve artificial intelligence within a decade, and spent over one billion dollars in trying to achieve this as well as developing intelligent machines that could wage war all by themselves, virtually removing humans from the battlefield altogether. This soon proved to be an overly ambitious goal, and while some evolution was seen in the creation of e.g. automated defense systems[66], the end of the Cold War in the 1990s once more halted the ambitions of autonomous weapons. This was partly due to cuts in the military budgets, and partly due to public expectations to instead focus research and budgets on welfare now that the military threat was finally gone. It was partly also because of the slow progress that had been achieved in AI and computer perception. The evolution in autonomous weapons and military robotics was also not believed to be useful in the featured operations of the 1990s: peacekeeping.[67]

In the early 2000s however, military technology in terms of automation finally started to receive attention, and in the wars in Afghanistan in 2001 and Iraq in 2003, the US started leading the development with a growing number and role, of unmanned systems. At the time, they were primarily intended for surveillance and reconnaissance, and secondly intended as a means of clearing the way for attacks by manned systems.[68] With the development of Unmanned Aerial Vehicles (UAVs), the many advantages of unmanned systems became clearer and clearer – decreased risk for the own forces probably being the biggest benefit as of yet.

Naturally, with all new systems of warfare there are disadvantages as well. With unmanned systems, the major factors that push the military towards an increasing autonomy in weapon systems are, firstly, the risk of technical glitches in remote-controlled systems and secondly, the interest of reducing manpower. The risk of electronic malfunction, cyber-attacks and hijacking of robots is an issue that arises when the machine is still connected to a human operator and relies on the operator for the actual targeting and firing of the weapon. Therefore, automating the weapon system and letting the machine take care of these matters on its own, eliminates the need for a human operator and facilitates the complete severing of the communication link and removes the risk of hacking.[69] The interest of reducing manpower in the military is due both to the decreasing interest and suitability of persons joining the military forces, and to cuts in the defense budgets. With remote-controlled systems there still needs to be a human operator in the picture, and as current unmanned systems require up to three human operators each, plus extra persons for support, it is easy to see the call for autonomous systems where the number of personnel needed for operating would be dramatically cut, as

[66] See below, chapter 3.2.2.1.
[67] Krishnan (2009), pp. 24-25.
[68] Ibid., p. 27.
[69] Ibid., pp. 38-39.

the system would in effect operate itself.[70] With these two major factors as traction, the intensified focus on automating weapons is understandable and the emergence of the different kinds of AWs a natural consequence.

3.2 Definition of Autonomous Weapons

As the field of autonomous weapons is an emerging one, the definitions tend to be different depending on the author, and no complete consensus has been reached.[71] As mentioned in the delimitation chapter, I will not be discussing remote-controlled weapon systems in the upcoming subsections, however, a basic distinction between these and autonomous systems is necessary for the further understanding of AWs. Remote-controlled systems are those that require the operation of a human being in the selection of targets, as well as activating, directing and firing of the weapons it carries. These are in the literature often referred to as UAVs or Unmanned Ground Vehicles (UGVs)[72] – in the public debate, they are most commonly known as "drones".[73] In the cases of these systems, the human, while often located elsewhere and many times far away, is still "behind the wheel". These unmanned systems are just that – unmanned – no man on board the actual vehicle or device, while autonomous systems once activated, have completely severed the tie with the human factor.

The word autonomy comes from the Greek words auto ("self") and nomos ("law"), meaning self-rule or self-governing.[74] Applying this to the term of autonomous weapons, it signifies a weapon that is self-governing, i.e. that does not require a human connection with the weapon. Even within this category there are different levels of autonomy and similarly, different definitions. The three levels of autonomy are, among others, defined in the report issued by Human Rights Watch (HRW), *Losing Humanity*, by the US Department of Defense (DoD) in its 2012 *Directive on Autonomy in Weapon Systems*, as well as by Armin Krishnan in his comprehensive book *Killer*

[70] Ibid., pp. 35-36. The numbers of flying hours by the *Predator* drone increased from 250 000 in June 2007 to more than a million in 2010. With this increase, the number of operators required will of course also surge, and the push towards full autonomy will intensify, see Sharkey (2010), p. 371.
[71] Some authors also discusses and differentiate between the concepts of "automated" and "autonomous", however, this is not an issue I will be discussing for the scope of this thesis. Cf. Asaro, p. 690, at note 5.
[72] The prevalence in the field (as well as the literature) of UAVs far exceeds that of UGVs, due to the higher degree of complexity of the environment in which UGVs operate. Whereas UAVs easily navigate in the airspace with the help of GPS, not having to avoid any obstacles, UGVs must be able to traverse through a much more complex terrain and identify and appropriately respond to obstacles on its way, cf. Krishnan, p. 55.
[73] ICRC *Autonomous weapons: States must address major humanitarian, ethical challenges*, FAQ, 2 September 2013.
[74] Krishnan (2009), p. 43.

Robots: Legality and Ethicality of Autonomous Weapons. The distinction they provide are as follows[75]:

Level of Autonomy	HRW	DoD	Krishnan
1	Human-in-the-loop	Semi-autonomous	Pre-programmed autonomy
2	Human-on-the-loop	Human-supervised autonomous	Limited or Supervised autonomy
3	Human-out-of-the-loop	Fully autonomous	Complete autonomy

In the following examination, the terminology of the definitions put forward by Krishnan will be the ones applied.

3.2.1 Pre-programmed Autonomy

The first level of autonomy indicates that the machine executes a specific function that has been pre-programmed into the system of the machine. Generally, weapons with pre-programmed autonomy have no or very limited capacity to diverge from the pre-set instructions and subsequently operate within very narrow parameters. There are also machines that will be able to operate with a structured control which allows for a somewhat greater autonomy, which follows a more organized "if-this-then-that"-algorithm.[76] This often consist of a decision process based on sensors that tells the robot "obstacle at left – move right".[77] Examples of machines with pre-programmed autonomy are robots designed for clearing mines[78], robots with tasks such as bomb disposal and cave clearance on the ground, as well as robots only used for surveillance from the air.[79]

3.2.2 Limited or Supervised Autonomy

With the second level of autonomy, the machine operates almost entirely on its own, which means that the variation in its behavior is much greater than in the pre-programmed systems, e.g. allowing the machine to find its own

[75] HRW Report (2012), p. 2; United States of America Department of Defense Directive 3000.09 on subject "Autonomy in Weapon Systems", 21 November 2012, pp. 13-14; Krishnan (2009), p. 45. The term "Lethal Autonomous Robots" or LARs, is increasingly used in the public debate, but will not be used in this thesis.
[76] Krishnan (2009), pp. 43-44.
[77] Sharkey (2010), p. 377.
[78] Krishnan (2009), p. 33.
[79] Sharkey (2010), p. 370.

way without requiring continuous human intervention. These systems would require human intervention when it comes to the more complex functions such as, first and foremost targeting, but sometimes also triggering of the weapon. As they are less capable of dealing with unforeseen situations and circumstances, the human operator would therefore also function as supervisor for the machine, where the machine at the event of an unforeseen situation would report back to the operator who then decides how to proceed. This category is probably the most common today and it encompasses a variety of different weapon systems that operate within a limited or supervised autonomy.[80] Today, there are both stationary and increasingly mobile machines that are able to e.g. patrol camps, military bases and even larger areas, with supervised autonomy.[81]

3.2.2.1 Automatic Weapons Defense Systems

One example of supervised autonomy in machines today are the automatic weapons defense systems, i.e. a system devised to sense an incoming threat, e.g. a missile or rocket, and then automatically respond in order to neutralize the threat. The human supervision is confined to accepting or overriding the machine's plan of action, something that must happen almost instantaneously and that is therefore not very common. The earliest version of this defensive system was the *MK 15 Phalanx*, installed on US Navy ships already in 1980. It is now being used by both the US and its allies.[82] Since then, the US has also developed a land-based version, the Counter Rocket, Artillery and Mortar System (*C-RAM*), which was first operated in Iraq in 2005 and functions in the same way as the *Phalanx*. When the *C-RAM* has detected a threat, the human supervisor certifies the target, however this has to happen within seconds in order for the system to be able to destroy the incoming threat in time.[83]

It is not only the US that has developed these kind of defense systems, Israel has set up its system *Iron Dome* on the Gaza border which uses a radar to detect incoming rockets and responds to those threats automatically. Between April 2011 and August 2012 it had a reported success rate of 80%. With this defense system as well, the response to a threat is being sent to a human operator, who instantaneously has to decide to fire or not, for the machine to be effective.[84] Additionally, Germany has developed the *NBS Mantis*, a defense system used to protect its operation bases in Afghanistan, which detects, tracks and shoots down incoming threats automatically, with human

[80] Krishnan (2009), p. 44.
[81] Lubell Speech (2013).
[82] Federation of American Scientists, "MK 15 Phalanx Close-In Weapons Systems (CIWS)", 9 January 2003; HRW Report (2012), p. 9.
[83] "Land-Based Phalanx Weapon System Completes Mission in Iraq" in *Naval Sea Systems Command (NAVSEA) newswire*, 16 February 2012; "C-RAM Transform Defense Tactics", *US Army news release*, 26 April 2012.
[84] Garamone, Jim "Iron Dome System Demonstrates US-Israeli Partnership" in *American Forces Press Service*, 1 August 2012; "Iron Dome Battle Management Demonstrated" in *Defense Update*, 2009.

intervention currently being limited to a monitoring role.[85] Also the UK is currently designing a supersonic aircraft known as the *Taranis* which is capable to fly autonomously but cannot engage targets without the approval of a human operator.[86]

The important difference to note between these and weapons with complete autonomy is that these weapons only have a defense function, and have no capabilities to take initiatives to attack targets on their own. They will engage automatically when they detect a threat, as a defensive mechanism, while a completely autonomous weapon would have a more offensive strategy of going out and looking for the targets on its own.

3.2.2.2 Sentry Robots

Along the lines of limited or supervised autonomy are also the development and deployment of sentry robots that operate on the ground. These robots have already been deployed in both South Korea and Israel. In 2010 the *SGR-1* sentry robot was installed along the demilitarized zone between South and North Korea which detects people in the zone using heat and motion sensors and then reports a warning back to a command center. At the command center, a human soldier can communicate with the identified person and decide whether or not to engage the weapon against the individual. Presently, the *SGR-1* only have autonomous surveillance capabilities and require a human to command the firing of a weapon.[87] While at present, it cannot fire without a human command, the machine has the possibility for an automatic mode, in which it has the capability to perform this firing decision on its own.[88]

A similar sentry system has been deployed along the Israeli border with Gaza, which senses movements and reports those signals back to a distanced facility where human soldiers evaluate the data and decides whether or not to fire at the object of the movement. The main targets are people trying to cross the border and sniper and rocket attacks, and while the sentry for the moment requires a human in the loop, the implication is that it will be able to operate completely autonomously in the future.[89]

3.2.3 Complete Autonomy

First and foremost, it is important to note that these types of machines only exist at an experimental level today, and that they are being built for research

[85] "Germany Orders MANTIS C-RAM Base Defense Systems" in *Defense Industry Daily*, 17 January 2011.
[86] Schmitt, Michael N., & Jeffrey S. Thurnher "'Out of the Loop': Autonomous Weapon Systems and the Law of Armed Conflict" in *Harvard National Security Journal*, Vol. 4, 2013, p. 239.
[87] Rabiroff, Jon "Machine Gun-Toting Robots Deployed on DMZ" in *Stars and Stripes*, 12 July 2010.
[88] Kumagi, Jean "A Robotic Sentry for Korea's Demilitarized Zone" in *Institute of Electrical and Electronics Engineers Spectrum*, 1 March 2007.
[89] "Sentry Tech: Long Distance Stationary Remote Controlled Weapon Station" in *Rafael Advanced Defense Systems Ltd.*, p. 1.

purposes only. Machines that are capable of complete autonomy are machines that can operate completely by themselves without any human intervention whatsoever, and to some extent also are capable to learn and adapt their behavior based on previous experiences, and in that sense, built upon an artificial intelligence designed to resemble human intelligence and thought-capacity.

Complete autonomy, and accordingly, the complete lack of human intervention, means that the operator only programs the machine with the objective of the mission, and the machine itself will find a solution to it and address the many problems that arise on the mission, on its own. However, the technology of today has not yet reached a level of intelligence in these machines to make them sufficiently predictable and controllable in order for them to be safe and useful for military purposes, although it is the long-term goal of DARPA to develop these kind of truly autonomous robots with cognitive, self-learning abilities.[90] Most authors stress the fact that there are currently no fully autonomous weapon systems,[91] and some, that there are no current intentions of developing such either.[92] However, the technology is rapidly moving towards full autonomy, and fully autonomous weapon systems are expected to be developed already within the next few decades.[93]

3.2.3.1 Artificial Intelligence

For a weapon system to be considered as completely autonomous, there is the requirement of AI in the system: the more developed the AI, the smarter the AW and therefore, the more tasks will it be able to carry out and subsequently, the more useful it will be in military operation. It seems therefore, that the research in AI and the development of it is a key aspect to the development of AWs, and the success of AWs will depend on it. Additionally, with particular importance for this thesis, the development of AI will be a key aspect in determining the possible adherence to the rules and principles of IHL. Ever since the launch of the SCI by DARPA in 1983, the development of AI has been a priority in the pursuit of better military technology.[94] While

[90] Krishnan (2009), p. 44.
[91] Cf. e.g. Krishnan (2009), p.1; Anderson, Kenneth & Matthew C. Waxman "Law and Ethics for Autonomous Weapon Systems: Why a Ban Won't Work and How the Laws of War Can" American University Washington College of Law Research Paper No. 2013-11, 2013, p. 3 in fine; HRW Report (2012), p. 3; Thurnher, Jeffrey S. "Examining Autonomous Weapon Systems from a Law of Armed Conflict Perspective" in *New Technologies and the Law of Armed Conflict*, Hitoshi Nagu & Robert McLaughlin (eds.), T.M.C. Asser Press, 2014, p. 214; Schmitt, Michael N. "Autonomous Weapon Systems and International Humanitarian Law: A Reply to the Critics" in *Harvard National Security Journal Features*, 2013, p. 3; Sharkey, Noel "The evitability of autonomous robot warfare" in *International Review of the Red Cross*, Vol. 94, No. 886, Summer 2012, p. 788.
[92] The UK Ministry of Defense (MoD) asserted in 2011 that there are currently no intentions of developing autonomous weapons, *Joint Doctrine Note 2/11 The UK Approach to Unmanned Aircraft Systems*, March 2011, paras. 507-508.
[93] HRW Report (2012), p. 46; Krishnan (2009), p. 169, DARPA estimates that human-level AI will be developed around the year 2030.
[94] Krishnan (2009), p. 46.

the success of AI in the field of military technology has been limited, it is quite established in computer science and programs based on AI already exist online, on platforms such as Narrative Science and Automated Insights. These platforms offer programs that processes large amounts of data and produces the outcome in neat reports or journal articles, with the purpose of replacing humans and therefore allowing companies to cut the costs of staff.[95]

The same idea applies to the idea of using AI in the military – replacing humans in order to cut the costs, but also to minimize or eliminate the risk of putting human soldiers in harm's way, a political factor as important as any. While the general field of AI has been around for decades, there is still no consensus on the definition of AI, and the field lacks a unifying theory that connects the vast variety of sub-disciplines. When AI is discussed in regards to research, the division most generally made is between weak and strong AI, where the former is designed to solve narrowly defined problems, and the latter designed to be able to tackle problems of a greater complexity, a machine that could match or even possibly exceed the intelligence of humans.[96]

Speaking in general terms, AI aims to use computers to simulate the human brain in order to technically reproduce or mimic human intelligence and cognitive abilities. The crucial challenge seems to be the understanding of the mechanisms and elements that is the basis of human intelligence, such as the ability to understand natural language, to recognize patterns, to apply knowledge and to learn. If it became possible to include these abilities in a computer program, then the creation of full, strong AI – human-like intelligence – would theoretically be possible.

However, the achievement of this has for a long time been a Holy Grail for researchers, and even if they are getting closer to systems that are capable of mimicking complex human thought processes, the triumph of a strong AI might never come.[97] In later years, the creation of more and more human-like – humanoid – robots have increased[98], and while they are mostly used as mechanized bartenders, the functions they carry out could possibly be easily transferred, from opening bottles and pouring drinks to pulling a trigger on a weapon.[99] In spite of this prediction, it is questionable to what extent robots with AI could actually behave like humans, and Krishnan notes that in the development of a strong AI, no matter the amount of troubleshooting and redesigning, it seems unlikely that the AI would develop anything that could

[95] See websites for Narrative Science, http://narrativescience.com/, and Automated Insights, http://automatedinsights.com/, respectively.
[96] Krishnan (2009), p. 47.
[97] Ibid., p. 48.
[98] One recent prototype of these humanoid robot, is the robot called *Atlas* developed by Boston Dynamics. *Atlas* is according to Boston Dynamics coordinated enough to climb using hands and feet, as well as handling tools designed for human use. While designed to negotiate rough terrain outdoors, the exact area of application for *Atlas* is however yet to be decided.
[99] Lubell speech (2013).

come close to the ability to instinctively find the best solution to a given situation, in a way that a human would.[100]

Peter Asaro makes the comparison between the argument that today's AI could meet the requirements that international law sets out, and the first kind of AI where computer developers predicted that within the decade, a computer could beat a Grand Master in chess. It took 40 years before the computer *Deep Blue* in 1997 finally beat the Grand Master Gary Kasparov. The important difference here though is that chess is a well-defined game based on rules and susceptible to computational analysis – it is not a matter of interpretation, and not a matter of social norms. Granted, international law is also based on rules, but other than that it should not be compared with a game like chess. International law requires interpretation and judgment in order for it to make sense and be applicable in real world situations. While precedents and established standards aid this interpretation, it is ultimately a matter of a case-by-case analysis, where considerations of innumerable factors needs to be made, considerations that conceivably only humans are capable of.[101]

Some authors argue on the other hand that recent leaps have been made in the research in artificial intelligence, and that this would facilitate taking the human out of the loop. It is even contended that the technology essentially exists today.[102] However, even some of the strong proponents of AWs admits that despite impressive advances in the field of AI, it is unlikely to achieve it in the near future[103], and that artificial intelligence has overpromised before.[104] This is likely also the reason why proponents generally tend to avoid the discussion on the progress of AI today, and stick to the more general claim of "strong AI is possible", in order to not overpromise the possible capacity and capability of AI once again.

[100] Krishnan (2009), p. 99.
[101] Asaro (2012), p. 705.
[102] Thurnher (2014), p. 215.
[103] Schmitt (2013), p. 20.
[104] Anderson & Waxman (2013), p. 14.

4 The Debate on Legality of AWs

To begin with, it is important to note the difference between two kinds of legality – legality of a weapon *per se*, and legality of a weapon with regards to the use of that specific weapon. I will examine the legality of autonomous weapons in light of both these aspects in turn, however, the first chapter which regards legality *per se* will only be discussed briefly and the different sides of proponents and opponents will not be discussed. This is due to the fact that the debate does not primarily revolve around the issue of if the weapons are illegal *per se*, but primarily around the issue of legality of the use of the weapons. Some discord does exist also around the first question, but not to an extent that merits an in-depth examination within the frames of the present thesis.

I will however start with an aspect of AWs that is *not* particularly prominent in the debate, but that nonetheless deserves some reflection, the moral and ethical aspect of using AWs.[105] In my view, the absence of this aspect from the growing debate on AWs is a bit disconcerting, because it seems as if advocates miss out on a fundamental discussion, before the legality should even be discussed. Indeed, the moral and ethical aspects can be seen as overarching principles that should be considered before the legality and compliance or non-compliance of the material rules of IHL is even discussed, but the moral and ethical aspects could also be a part of the legality assessment in itself.[106]

4.1 Morality and Ethicality of AWs

When parties discuss the legality of AWs, the arguments and the discussion is often limited to a discussion on the ability or inability for AWs to comply with the rules that IHL (primarily) sets out. Some authors touch upon the very important, but sometimes overlooked, issue of morality and ethics in relation to the deployment of AWs in battle, but the discussion on these moral and ethical issues often falls short of a proper review, and often leaves the impression that the author himself does not believe the argument is valid. In the debate, the issues presented below are discussed and considered as purely moral issues, but I would like to question this separation and wonder if the moral issues could be considered within the scope of legality as well. There are certain matters that may seem moral but that could have implications for the legality assessment, and in relation to the legality of AWs, I think that it

[105] While concepts of morality and ethicality are broad and could entail a number of different perspectives and points, it is outside the scope of this thesis to further explain or define these concepts. In the following chapter, the concepts will be used in their, if you will, generic sense.
[106] Most prominently, in the discussion of accountability, see below chapter 4.5.

is vital to consider the discussion on the fundamental idea that wars are fought by humans, and that it should be allowed more space in the current debate.

In all past and present regulations on how wars are to be fought, the agents of wars have always been humans. It has been taken for granted as a silent precondition, and not even considered worth mentioning in treaty texts etc., but the introduction of autonomous weapons in the arena of war begs this essential reconsideration – what happens to international humanitarian law if we change the agents of war? When the robot not only is capable of performing an order (to kill a combatant) but also is capable of making that decision on its own, perhaps the robot has been elevated morally from being a mere object to being a subject capable of morally meaningful actions (the decision to kill or not to kill)?[107] Are AWs still to be viewed as only a tool for soldiers to use, or is it elevated to being a soldier itself? The definition of a combatant as an operator of a weapon or weapon system, allows the potential interpretation of AWs as being combatants, at least in theoretical terms. Not only could this shift mean great difficulties in pure terminology of the laws and regulations, but in a moral and ethical sense as well.[108]

While IHL may be competent to change and adapt in order to encompass further developments in the means and methods of war, the current development where the agents of war may be changing and no longer be humans, is something that probably was not envisioned and therefore something that IHL does not consider. Therefore, if we enter into this discussion on new agents, in effect we leave the realm of IHL. Some argue that the discussion on personhood and status of robots could be compared to the legal status of children, and that robots therefore would be to consider as quasi-agents, with only a limited amount of responsibilities and duties.[109] This perspective might provide an explanation to the question of status, but it does not answer the problem of accountability, and more importantly, it does not at all resolve the moral and ethical issues tied to the matter of agents in war.

These issues of morality in relation to AWs are gradually starting to attract attention from scientists and authors on the subject, and concerns are starting to be raised: if our morality is what makes us humans, such as being able to have a gut instinct or a "sixth sense" that allows us to make decisions in war based on this, would robots who lack these qualities in essence be immoral?[110] And if we cannot know exactly what it is that makes us humans, that allows us to interpret complex situations and contexts[111], how are we to transfer this onto a machine? And perhaps more importantly, do we *want* to transfer this onto a machine? Would it be fundamentally immoral to delegate life-and-

[107] Krishnan (2009), p. 33.
[108] Liu, Hin-Yan "Categorization and legality of autonomous and remote weapons systems" in *International Review of the Red Cross*, Vol. 94, No. 886, Summer 2012, pp. 634-637.
[109] Lin, Patrick, George Bekey & Keith Abney "Autonomous Military Robotics: Risk, Ethics and Design" prepared for the US Department of Navy, Office of Naval Research by California Polytechnic State University, San Luis Obispo, 20 December 2008, pp. 58-60.
[110] Sharkey (2012), p. 790.
[111] Indeed, this is most often referred to as above, as a gut instinct or a sixth sense.

death decisions onto a machine that does not possess the moral instincts necessary to make a considered and informed decision about taking a human life?[112] Without human reason, judgment and compassion?[113] The "three laws of robotics" as set out by Isaac Asimov might seem appealing in theory, especially given its background in popular culture, but in practice, the application of a "robot morality" will prove more complicated than that.[114]

Some proponents offer solutions where morality and ethicality is programmed into the robots through complex algorithms or learning-based programs, such as Ronald Arkin who proposes a complex algorithm that could be developed and function as an "ethical governor" in AWs[115]. Others seem to already have taken a stand on the matter, and propose that it does not necessarily have to be human beings that are the agents of war, that machines are able to perform at least equally efficient as humans, and thus, that the moral imperative for using machines is already there.[116] This approach resembles the concept of "reduction to purpose" as explained by Martti Koskenniemi[117], where proponents often skip the discussion on morality of AWs, reducing the discussion to the purpose of saving as many (innocent) lives as possible.[118] If this purpose can be reached through the use of AWs, why take the often long and bothersome detour of discussing the moral aspects of letting AWs become the new agents of war? This shortcut could possibly lead to policymakers and military strategists to assume that technological advancement is the same as moral advancement.[119] Arkin also argues that the full moral capabilities of humans does not need to be reproduced in robots in order to attain an acceptable standard of moral in these

---

[112] Asaro (2012), p. 689.

[113] Ibid., p. 708.

[114] The Three Laws of Robotics was first mentioned in Asimov's short story "Runaround" in 1942. The laws are:

1. A robot may not injure a human being or, through inaction, allow a human being to come to harm.
2. A robot must obey the orders given to it by human beings, except where such orders would conflict with the First Law.
3. A robot must protect its own existence as long as such protection does not conflict with the First or Second Law.

[115] Arkin, Ronald *Governing Lethal Behavior in Autonomous Robots*, CRC Press Taylor & Francis Group, 2009, chapter 6 which discusses the creation of an "ethical governor".

[116] Arkin, Ronald "Lethal Autonomous Systems and the Plight of the Non-Combatant" in *AISB Quarterly*, No. 137, July 2013, p. 4; Cf. also Schmitt (2013), pp. 239-240; Anderson & Waxman p. 18.

[117] Cf. Koskenniemi, Martti "What Use For Sovereignty Today?" in *Asian Journal of International Law*, Vol. 1, Issue 01, 2011, p. 66. Koskenniemi mentions the concept of "reduction to purpose" in a discussion on the role and function of sovereignty of the world today, but the notion of the concept can nonetheless be transferrable.

[118] Arkin takes this position and states that "[…] if and how these new robotic systems can perform as well as, or better than, our soldiers with respect to adherence to the existing IHL. If achievable, this would result in a reduction in collateral damage, i.e. noncombatant casualties and damage to civilian property, *which translates into saving innocent lives*", in Arkin (2013), p. 2 (emphasis added).

[119] Kaag, John & Whitley Kaufman "Military frameworks: technological know-how and the legitimization of warfare" in *Cambridge Review of International Affairs*, Vol. 22, No. 4, December 2009, p. 587.

robots.[120] In addition, if we were to shift the moral responsibility onto the machine itself, we might be overlooking the discussion on what ethical judgment actually entails. Kaag and Kaufman argue on their side that warfare must be regarded as a strictly human activity and that we could never transfer moral responsibility onto a machine.[121]

As Peter Asaro puts it, the two most important questions in the consideration of moral and ethical aspects of AWs, are if a machine *could* make life-and-death decisions like this, and, perhaps the most important, if a machine *ought* to make these decisions at all. He begs the question that if we eliminate the human from the decision-making process, how do we ensure that the killing is not arbitrary? Overall, Asaro means that it would be morally and ethically wrong to relinquish these kinds of decisions to a machine, which is by no means developed or advanced enough to mimic a human and the human intellect, and that there is no automated process that should be accepted as a replacement for humans. Even if the technology could be made advanced enough to be able to satisfactorily perform the requirements for distinction and proportionality, and indeed perhaps be able to make fewer errors in a discrimination task, the decision of taking a human life should in any way not be transferred onto a machine.[122] Kaag and Kaufman points out that ethical judgments cannot be made by determinate rules and that it requires the flexibility and sensitivity that only humans possess.[123] Ethical decision-making is a human endeavor and occurs in unique and ever-changing circumstances, and the meaning of right and wrong in these circumstances cannot be determined by some sort of general metric that applies to all conceivable cases, but through a manner of unique interpretation on a case-by-case basis.[124] The ability to respond flexibly and contextually to ambiguous situations that present themselves in the battlefield is a reflection of the human capacity of such complex moral judgment.[125]

Krishnan points out the incapability of a machine to understand the finality of life, and that "[w]here there is no ability to die there is no true capability for ethical behavior." Therefore, AWs are not capable of ethical behavior, and not capable of making decisions that affects the life and death of humans.[126] Today, it seems that there is no reason to believe that, even with the emerging technology with the possible capacity to improve human life, technology would be able to solve ethical problems that have challenged humans for thousands of years.[127] Asaro asserts that the problem does not lie with the technology in itself, but on the contrary, that if the technology existed and was advanced enough, that technology ought to be used in order to assist the human soldiers to fulfill the requirements of distinction and proportionality,

[120] Arkin (2013), p. 4.
[121] Kaag & Kaufman (2009), p. 586.
[122] Asaro (2012), pp. 701 & 708.
[123] Kaag & Kaufman (2009), p. 587.
[124] Ibid., p. 590.
[125] Ibid., p. 600.
[126] Krishnan (2009), p. 133.
[127] Kaag & Kaufman (2009), p. 602.

but not to allow AWs to use that technology all on their own.[128] Also Krishnan makes this comment that technology might become superior to the performance of human soldiers, but that the crucial issue is that that specific technology is being used in the best and correct way, keeping in mind the ultimate goal of IHL – to protect those who are not, or no longer, taking a direct part in hostilities.[129]

When it comes to the intricate discussion on morality and ethicality in regards to warfare and AWs in particular, Anderson and Waxman hits the nail on the head in stating that "this is a difficult argument to address, since it stops with a moral principle that one either accepts or does not accept".[130]

Ultimately, the issues of morality in relation to AWs seem to remain in the shadow of the current debate, and while the question of the status of AWs as tools of combatants, or as combatants in themselves, could have grave consequences for matters such as accountability[131], it seems as the debate will continue to revolve around the technological capability to adhere to the laws and regulations.

4.2 Legality *Per Se*

The idea of some weapons being inherently inhumane, no matter the use of the weapon, has been recognized in treaty law since 1907[132] and is now part of customary law.[133] The legality of a weapon *per se* addresses weapons that by their very nature are unlawful, no matter what targets they are aimed at or no matter how it is being used. Legality of a weapon *per se* is governed primarily through two Articles of AP I: Article 51(4), indiscriminate weapons and Article 35(2), weapons that cause unnecessary suffering or superfluous injury.

4.2.1 Indiscriminate Weapons

Article 51(4) states that a weapon is considered illegal *per se* if it by its very nature is indiscriminate and cannot be aimed at a specific target and therefore is likely to cause harm to civilians as well as combatants in its use. The Article states that, among others, attacks which are indiscriminate are those that employ a means or method of combat that cannot be directed against a

[128] Asaro (2012), p. 702.
[129] Krishnan (2009), p. 30; Asaro (2012), p. 697.
[130] Anderson & Waxman (2013), p. 16.
[131] The nonsensical action of punishing a machine that is not capable of understanding the purpose of the punishment is discussed briefly under chapter 4.5, see also Sparrow (2007), p. 66, who claims that if no one can be justly held responsible for the actions of AWs, then the use of them in war would be unethical.
[132] Convention (IV) Respecting the Laws and Customs of War and Law and Its Annex: Regulation Concerning the Laws and Customs of War on Land, 18 October 1907, see Articles 22-23.
[133] Henckaerts & Doswald-Beck (2005), pp. 37-45, 237-250.

specific military objective, and attacks which may be expected to cause collateral damage that is excessive in relation to the military advantage anticipated.[134] This means that the assessment of distinction and proportionality already forms part of the determination of if a weapon is indiscriminate, as well as the determination of the lawful use of a weapon.[135]

Indiscriminate weapons that have been prohibited through international conventions over the years include anti-personnel mines (a mine that is designed to explode by the presence, proximity or contact of a person and that will injure or kill one or more persons)[136], and cluster munitions (a conventional munition designed to disperse or release explosive sub-munitions that each weigh less than 20 kilos).[137] Both anti-personnel mines and cluster munitions have been prohibited because they constitute weapons that by their very nature are incapable of distinguishing between civilians and combatants. Both weapons also often leave behind large amounts of unexploded ordnance which kills and injures both civilians and combatants for years and decades after its use. They therefore constitute indiscriminate weapons.

In the Nuclear Weapons Advisory Opinion, the ICJ discussed the opinion expressed by some states, that nuclear weapons would be unlawful in itself because it could never comply with the principles of IHL, including the prohibition of indiscriminate attacks, since "such weapons should kill and destroy in a necessarily indiscriminate manner, on account of the blast, heat and radiation occasioned by the nuclear explosion".[138] After establishing that "the use of such weapons in fact seems scarcely reconcilable" with the requirements of IHL, the Court did however state that it could not conclude with certainty that nuclear weapons were unlawful in all circumstances.[139] Even though the Court in this case did not conclude on the illegality of nuclear weapons, it did establish that the prohibition of indiscriminate weapons, along with the principles of distinction and proportionality, is of cardinal importance in IHL.[140] The idea the Court had, was that nuclear weapons could be lawfully used in the extreme circumstance of self-defense where the survival of the state was at stake.[141] Here, it is not difficult to draw an analogy between nuclear weapons and autonomous weapons, and, in fact, an armed robot would probably seem even less dangerous and cause less widespread injury than a nuclear weapon. Thus, with this analogy, allowing for the use of AWs in, at least, extreme circumstances.[142]

[134] AP I, Article 51(4)(b) and 54(5)(b).
[135] As discussed below in the chapter 4.3.
[136] Convention on the Prohibition of Use, Stockpiling, Production and Transfer of Anti-Personnel Mines and on their Destruction, Article 2(1).
[137] Convention on Cluster Munitions, Article 2(2).
[138] Nuclear Weapons Advisory Opinion, para. 92.
[139] Ibid., para. 95.
[140] Ibid., para. 78.
[141] Ibid., para. 97.
[142] Sharkey (2012), p. 797.

4.2.2 Unnecessary Suffering or Superfluous Injury

The second part of the inherent illegality is based on Article 35(2) which prohibits the use of weapons, projectiles and material and methods of warfare which are of a nature to cause unnecessary suffering or superfluous injury. This prohibition forms part of Article 35 which deals with basic rules of warfare, and which implies the fundamental principle of that the only legitimate object of acts of war, is to weaken the military forces of the enemy, or to disarm them, and that it is therefore prohibited to use means or methods of warfare which are excessive after having rendered the enemy *hors de combat.*[143] The prohibition of anti-personnel mines as mentioned above, was in part also based on this principle, as the explosion of such mines often cause severe injuries and disabilities on the victims.

Other weapons that have been considered to cause unnecessary suffering or superfluous injury include expanding bullets ("dum-dum-bullets"), exploding bullets, poisonous and asphyxiating gases, biological and chemical weapons, weapons that leave fragments not detectable by X-ray, incendiary weapons and blinding laser weapons. As there is no complete consensus on what kind of weapons constitute unnecessary suffering or superfluous injury, there is subsequently no absolute consensus on the unlawfulness of all of the above mentioned weapons, however, there is a general agreement that most of them are prohibited.[144] In the pursuit of developing more tenable criteria for this prohibition, the ICRC has launched the SIrUS-project which proposes the following criteria that would ban weapons if their use cause:

- A specific disease, specific abnormal physiological state, a specific and permanent disability or specific disfigurement; or
- Field mortality of more than 25% or a hospital mortality of more than 5%; or
- Grade 3 wounds as measured by the Red Cross wound classification scale; or
- Effects for which there is no well-recognized and proven treatment[145]

Given these criterion set forward, it seems implausible that AWs, by their very nature, would cause these kinds of suffering or injury and therefore be prohibited as such under this principle.

4.3 Legality of the Use of AWs

While no explicit consensus on the legality *per se* exists, the core disagreement seems to lie instead in the discussion on the lawfulness of the use of AWs which will be discussed in the present chapter. Starting from the

[143] Commentary AP I, comment no. 1411.
[144] Henckaerts & Doswald-Beck (2005), pp. 243-244.
[145] ICRC, The SIrUS Project: Towards a determination of which weapons cause "superfluous injury or unnecessary suffering", Robin M. Coupland (ed.), 1997, p. 8.

description of current fundamental principles of IHL as set out in chapter two, I will in this chapter examine the arguments set forth by the proponents and the opponents for AWs. I will discuss the principles in the same order as in chapter two, and finish each with some concluding remarks.

4.3.1 Distinction

4.3.1.1 Sufficient Sensors and Software: Arguments in Favor of AWs

As noted above, the principle of distinction is applicable in all armed conflicts, both international and non-international, and provides a protection for civilians in times of armed conflicts, unless and for such time as they directly participate in hostilities. The issue at hand is whether or not an AW would be able to adhere to this principle and fulfil it in a satisfactory way, if a machine would be able to distinguish between a civilian and a combatant, and especially between civilians who does or does not take a direct part in the hostilities.

Proponents for the use of AWs argue that algorithms that attribute values to sensor data are theoretically achievable, which then would make it possible for them to distinguish between civilians and combatants in the battlefield, as well as between civilian and military objectives and accordingly only direct its attacks on the latter.[146] These sensors would be equipped in a way that allowed the AW to e.g. recognize if the potential target is a child, is carrying a weapon or otherwise engaging in hostilities.[147] Another proposal consist of programming distinction into the AWs software through categories and samples of lawful targets (e.g. persons or weapons that fire at the AW), and incrementally develop this into inductive reasoning about certain characteristics of lawful targets that might not be on the list.[148] This method, as well as the method using sensors and recognition processes to identify combatants might be based on case-reasoning and simulations to improve the inductive learning process of the machine.[149] If these sensors and/or the programmed sample-technique would be developed enough, it could prove to be a good enough tool to distinguish between lawful and unlawful targets, and thus be considered to comply with the principle of distinction.

Some authors note that the surrounding context and environment is of crucial importance when determining the AWs capability to adhere to the principle of distinction, and that in the contemporary battlefield which is getting increasingly cluttered and less clear-cut, the requirement for distinction is much higher, and the challenge therefore much greater. In these new settings and often urban areas, the development of a finely calibrated sensor package and advanced recognition software is vital in order to comply with the

[146] Schmitt (2013), p. 17.
[147] Ibid.
[148] Anderson & Waxman (2013), p. 12.
[149] Ibid.

principle of distinction. Some also admit that even if such technology actually is set in place, there could be situations that would be so complex that AWs would simply be unable to meet the requirement, and the use of which therefore would be unlawful – the result being that the AW may only be used in situations and under circumstances where it would be able to distinguish satisfactorily.[150]

4.3.1.2 Civilians in Contemporary Conflicts: Arguments Against AWs

Opponents to AWs on the other hand, mean that AWs could never comply with the principle of distinction, particularly that they would not have the ability to sense or interpret the difference between civilians and combatants – especially in the context of contemporary armed conflicts, where everyone can be a civilian or a combatant, and where combatants often disguise themselves as civilians. Many authors point out the fact that the wars are changing, and that the conventional warfare has started to fade out and untraditional warfare emerge more and more, which makes it increasingly difficult, even for human soldiers, to distinguish between legitimate and illegitimate targets, in accordance with the principle of distinction.

It is more and more often the situation where combatants hides in cities or urban areas along with the civilian population, not to mention the fact that in these new conflicts, combatants rarely wear uniforms or other military insignia, thus making the only characteristic that allows them to be identified as combatants, their conduct, their "direct participation in hostilities".[151] In contemporary, urban battlefields with an increased chaos of combat, the task of distinguishing between civilians and combatants could prove to be beyond the capability of a machine, and even if the perception sensors were developed enough, it would be easy to trick the robots by concealing weapons or by exploiting the limitations the AW is bound by.[152]

Schmitt & Thurnher points out in their article that "not all battle spaces contain civilians or civilian objects", and this could be true for e.g. battles at sea. However, almost all land based as well as airborne battlefields will affect civilians to some extent, and thus the argument that AWs could be used in certain environments becomes virtually redundant, as these weapons would not be employable in the overwhelming part of contemporary battle spaces, and therefore have no practical use.[153] Noel Sharkey, one of the more vocal of the opponents to AWs, points out that if it was as simple as instructing a computer of "if a civilian, do not shoot", then the principle of distinction might be able to be fulfilled. However, there is no way to actually give the computer the information or a definition of what a civilian actually is, as IHL does not provide a sufficient enough definition that could be programmed into

[150] Thurnher (2014), pp. 220-221.
[151] HRW Report (2012), p. 30; Krishnan (2009), p. 29
[152] Krishnan (2009), p. 99.
[153] Schmitt &. Thurnher (2013), p. 246.

a machine, but only a definition that requires the use of common sense and deductive reasoning.[154] Kaag and Kaufman also stress the continuum of cases varying by the level of involvement or support provided in an attack, where the spectrum varies from a soldier firing his weapon on one end, to a civilian who does not play any role in the attack, on the other end. The determination of a legitimate target is therefore a matter of degree of involvement, and the creation of a set of rules that in advance prescribes the situations where lethal force is permissible is therefore unlikely.[155]

Some also point to the limited understanding an AW would have of human intention, an assessment that is a key aspect of distinguishing lawful targets from unlawful. Human Rights Watch poses the scenario of a mother running after her two children who is playing with toy guns near a soldier. Whereas a human soldier would identify the intention of the mother and the children as harmless, an AW might not perceive these intentions and instead see a person running toward it and two individuals with guns, which in the eyes of the AW would constitute lawful targets.[156] The same reasoning applies to persons, often children, which are forced to bear arms or carry weapons against their will. Without the subjective assessment made by a human soldier that can perceive things that a machine cannot, such as body language that would indicate this involuntary action, the machine would again consider these children as lawful targets and pull the trigger.[157]

Furthermore, to determine what "direct participation in hostilities" entails, the ICRC has established a set of guidelines which sets forth three requirements that when satisfied, conclude that a civilian is a legitimate target; firstly a threshold of harm, secondly a direct causation and finally a belligerent nexus.[158] This attempt of adopting guidelines is one means of determining who is a legitimate target and who is not, but it is in no way an exhaustive guide, as each of these assessments also requires a sophisticated understanding of the complex situation of each individual that might or might not be participating in the hostilities. Additionally, the guidelines, as with all other rules of IHL, require an immensely interpretative judgment in order to be correctly and appropriately applied in any given situation, something the AW could not fulfill.[159]

To argue in favor of a solution based on a fixed list of lawful targets may be technically correct, but is unrealistic with regards to the contemporary battlefields of today, and such a limitation to the use of AWs would mean that

[154] Sharkey, Noel, "Automating Warfare: lessons learned from the drones" in *Journal of Law, Information and Science*, Expert Commentary for ed. 21(2), 10 August 2011, p. 4.
[155] Kaag & Kaufman (2009), p. 599.
[156] HRW Report (2012), pp. 31-32.
[157] Sharkey (2010), p. 379.
[158] ICRC "Interpretive Guidance on the Notion of 'Direct Participation in Hostilities' Under International Humanitarian Law", Nils Melzer, 26 February 2009, p. 46.
[159] Asaro (2012), pp. 697-699.

they would be unfit in virtually all circumstances – an approach that hardly is realistic or desirable.[160]

4.3.1.3 Concluding Remarks

Clearly, the debate on whether or not autonomous weapons are or will be able to comply with this fundamental principle is heated, and each side of the discussion strongly advocates for their side. It seems as the discussion on distinction, and the capability of AWs to comply with this requirement, revolves around the issue of how the AW would be equipped (sensors, camera, pre-programmed list of targets etc.) and if it could ever be equipped in a way that satisfactorily and lawfully fulfills the requirements set out by the principle. Another core issue here as well is that if we were to employ autonomous weapons in the battlefield, the minimum standard should be that they were as good as or better than a human being. Again, if this distinction is hard enough for a human soldier to make, how is a machine going to make it as well or better? The proposition of combining the AWs with facial recognition software, is interesting, but the obvious limitation of the weapon only being applicable to certain, recognizable individuals, if – at all – it was able to actually identify an individual's face in real time moving circumstances.[161]

Even if AWs could be equipped with appropriate and sufficient sensors and recognition software, the problem of the contextual definition of civilians and combatants still remains, perhaps even more so in the more and more complex battlefields, where today the most distinguishing feature that separates combatants from civilians is not some form of easily defined outer characteristic, such as uniforms or insignias, but the vague concept of "direct participation in hostilities". While work has been done from the ICRC to create an interpretive guide on the notion of this direct participation,[162] the conclusions set out in this publication are not easily, if at all, transformed into algorithms and software codes, that could be programmed into an AW and as such, applied in combat in a way that would allow the AW to respect the principle of distinction.

As mentioned earlier, in the lines of the ideas of Asaro and Krishnan, one solution to the problems that would arise by letting AWs perform assessments on distinction, could be to instead use the machines in order to aid human soldiers to perform their assessment better and more accurately, and not allowing the machines to perform these assessments by themselves.

[160] Anderson, Kenneth & Matthew Waxman, "Law and Ethics for Robot Soldiers" American University Washington College of Law, Research Paper No. 2012-32, 2012, p. 10.
[161] Lubell Speech (2013); Sharkey (2010), p. 379.
[162] See the ICRC Interpretive Guidance on 'Direct Participation in Hostilities'.

4.3.2 Proportionality

4.3.2.1 Pre-programmed Probabilities: Arguments in Favor of AWs

In order for AWs to fully comply with the requirement of proportionality that IHL sets out, the machines would have to be able to estimate the expected amount of collateral damage in an attack, as well as the military advantage expected from that attack – and finally, weigh these two estimates together in order to determine if the collateral damage would be excessive and the attack therefore unlawful. In calculating the estimates of collateral damage, most militaries today have developed the procedure known as CDEM – Collateral Damage Estimation Methodology, which relies on objective and scientific criteria for its assessment.[163] According to this methodology, the attacking force considers factors such as the precision of a weapon, its blast effect, attack tactics and the probability of civilian presence in the proximity of the target.[164] As this methodology is already based on calculable algorithms, an AW would not have any problems in performing these calculations.

While the estimation of the anticipated military advantage is contextual, and often made on a case-by-case basis, some argue that it is conceivable to create a framework of pre-programmed values where the military operator pre-determines what constitutes excessive collateral damage in relation to a certain target. To comply with the principle of proportionality, these pre-determined values would have to be set at an utmost conservative level. Also geographical (depending e.g. on the surroundings and placement of the battlefield) or temporal (depending both e.g. the time of the day, i.e. when many civilians are out and about, and the time in the conflict, i.e. early on or late in the conflict) limits could be established to help the AW comport with the requirements of proportionality.[165] Attaching values to various targets, objects and categories of humans could also include an inductive element, where the machine learns from human examples and human judgment about proportionality, and carries out the probabilistic assessment based on this.[166] These, not yet developed but possible, algorithms that could estimate the military advantage, could be combined with thresholds for unacceptable collateral damage, and these thresholds would then mean that the AW is programmed not to fire upon a target if the estimation calculates an amount of collateral damage above the threshold.[167]

It could also be argued that AWs might be even better than humans to determine the proper amount of force to be deployed, as an AW could more quickly and precisely calculate blast effects and other weapon effects which would cause collateral damage. An AW could possibly perform hundreds of

[163] Thurnher (2014), p. 221.
[164] Schmitt & Thurnher (2013), p. 255.
[165] Thurnher (2014), pp. 221-222.
[166] Anderson & Waxman (2013), p. 13.
[167] Schmitt & Thurnher (2013), p. 255.

calculations at a time, which then would increase the lethality of the attack, while at the same time reduce the probability of civilian casualties, calculations that are far too complex for a human to perform equally well.[168]

In addition, the arming of the robot is an essential part of the proportionality assessment, and if the AW was to be armed with highly precise microprojectiles instead of larger missiles, the eventual damage caused by e.g. the projectile missing its target, would be relatively small. Furthermore, lacking a need or inherent reaction of self-defense, a robot would not risk responding with aggression and an excessive amount of force, something a human soldier would risk in a similar situation.[169]

4.3.2.2 Complex Contextual Calculations: Arguments Against AWs

Those opposing AWs, claim on the other hand that the principle of proportionality requires an assessment and a judgment inherent in humans, that a fully autonomous weapon could never replicate or improve. As even proponents like Michael Schmitt notes, the principle, while easily stated, is one of the most difficult ones in IHL to apply, particularly due to the difficulties in valuation.[170] Opponents primarily stress the fact that the weighing of the circumstances before taking a decision is ultimately a subjective matter. As with the arguments regarding distinction, they argue that the technology of determining which weapon might cause the least amount of collateral damage, instead should be used to assist commanders and human soldiers to do their estimations and assessment better – not for the machine to do this by itself.

The calculation of whether the minimum number of civilian casualties and damage to civilian property is proportional to the anticipated military advantage to be gained from that specific attack, and a decision based on this calculation, should be performed by humans.[171] If software and algorithms could be developed that makes it possible for a robot to calculate the estimated civilian harm – an imaginable version of the CDEM as mentioned above – all the robot would have to do is count the number of civilians present in the intended target area, decide which weapon to use and then calculate the blast radius of that weapon of both lethal and injurious effects. But once the robot has made this calculation, how is it to balance this against the military advantage, taking for granted that it is even capable of calculating what the military advantage is (which is far from certain) and come to a decision that respects the principle of proportionality? This balancing act is an inherently subjective test, something that human soldiers often have enough difficulties

[168] Krishnan (2009), p. 92; Arkin (2009), pp. 47-48.
[169] Krishnan (2009), p. 93.
[170] Schmitt, Michael N., *Essays on Law and War at the Fault Lines*, T.M.C. Asser Press, 2012, p. 190.
[171] Sharkey (2010), p. 380.

in considering, and that test could never be interpreted into calculations in a computer.[172]

While the criteria remain unclear as regards how to perform this balancing act, even for humans, there can be no alternative. This balancing process is so complicated and requires a vast amount of data and different factors to be taken into account, that an attempt to replicate this capability in a machine, designing a formula that would be both comprehensive and precise, simply seems futile. In the assessment of proportionality, common sense is irreplaceable and inimitable.[173]

To determine the proportionality of an attack or military operation depends to the largest part on context, and even if it might be possible to program an AW with pre-determined values, opponents mean that it is highly unlikely that it could be programmed to deal with the infinite number of scenarios it might face in the battlefield, and also to take into account every movement and to adapt to an ever-changing proportionality evaluation.[174] This means that AWs essentially could become unpredictable, with the increasing complexity of the software coding, consisting of millions of lines of codes combined from the work of several programmers. With the codes becoming more and more intricate and consist of such large amounts of information, there is the risk that no single individual is able to completely predict the effect of a given command, since portions of these large programs can come to interact in ways that are unexpected and untested.[175]

4.3.2.3 Concluding Remarks

With regards to the principle of proportionality, it is clear that the two-part balancing act that makes up the principle, is exceedingly hard to accomplish. The first part of the assessment, calculating the estimated collateral damage, could possibly be computerized, along the lines of the existing program CDEM. The second part proves much more of a challenge though, and there is strong disagreement on the possibility of a machine to actually calculate the anticipated military advantage satisfactorily, and as this is often based on a case-by-case analysis, the programming of this could prove too difficult for a computer to carry out. In addition, the principle also requires the weighing of these two assessments against each other, and an evaluation of whether the result of this weighing would be proportional or not.

In the steadily more complicated and complex algorithms required to perform these assessments, there is also the risk of the machine becoming too unpredictable in their behavior, a source of major concern, especially

[172] Lubell Speech (2013).
[173] Oeter (2008), p. 198
[174] HRW Report (2012), pp. 32-33.
[175] Marchant, Gary et al. "International Governance of Autonomous Military Robots" in *The Columbia Science and Technology Law Review*, Vol. XII, 2011, p. 284. This also affects the issue of accountability, discussed below in chapter 4.5.

considering that these robots may be employed in unstructured environments.[176] In the debate over AWs compliance with the principle of proportionality, the most pressing concern is whether or not a robot could be designed to calculated military advantage, a calculation that essentially is a matter of subjectivity that would be exceedingly hard for a robot to replicate. Ultimately, this would be dependent on what happens in the development of strong AI.[177]

4.3.3 Precaution

4.3.3.1 Capacity of Care and Caution: Arguments in Favor of AWs

The requirement of doing everything feasible to verify that a target is in fact a military objective, as set out in the principle of precaution, poses high demands on an AW, demands that nonetheless could be met, provided that the AW is equipped with sufficient on-board or external sensors that ensures the reliability of target identification, e.g. the ability to zoom in and narrow down the location of enemy forces, and efficient recognition of targets.[178] It is also asserted that in many cases, the recognition capabilities of the machine would be sufficiently advanced to meet this requirement of identifying a military objective in a reliable manner.[179] In addition, it is conceivable that in keeping with the requirement of identifying the objective with care, an AW with advanced telescopic sight and cameras with zooming capabilities, would be better equipped at doing this for objectives located at a great distance than a human soldier. One soldier expressed that with the advanced camera system, it is possible to read people's nametags at 300 meters, that same persons facial expressions, what weapon he is carrying and even if its selector is on fire or on safe.[180] Provided this technology develops, AWs will be able to comply with this part of the principle of precaution.

The second part of the principle proves more difficult, as this requires selecting the means of warfare likely to cause least collateral damage, without sacrificing the military advantage. As with the principle of distinction and proportionality, some proponents argue that the AWs would be superior to human soldiers in carrying out the necessary assessments with regards to precautions and avoiding collateral damage, because its sensor systems might be more precise or discriminatory than that of a human soldier, or because its ability to take decisions under certain circumstances, e.g. particularly dangerous ones, could exceed that of a human soldier.[181] One additional

[176] Ibid.
[177] See above, chapter 3.2.3.1.
[178] Schmitt (2013), p. 23.
[179] Thurnher (2014), p. 222.
[180] Singer, Peter W., "Military Robots and the Laws of War" adapted from *Wired for War: The Robotics Revolution and Conflict in the Twenty-First Century* by P.W. Singer, in *The New Atlantis*, Winter 2009, p. 35.
[181] Schmitt (2013), p. 25.

possibility to ensure that AWs comply with the principle of precaution could be that of arming them with non-lethal weapons, or requiring them to use non-lethal force as a first resort. One example of this is the PackBot developed by robotics company iRobot, which has the capability to be equipped with shotguns that can fire non-lethal rubber bullets.[182]

The alternative to only use non-lethal force as a first resort could also be a solution in order to take the necessary precautions, as this would allow the AW to initially attempt to only disarm the enemy, then take appropriate action depending on the response from the enemy, e.g. if the AW initially recognizes a child carrying a gun as a target, it could fire rubber bullets to disarm the child, instead of firing lethal bullets which would not be an appropriate response to the threat that child poses.[183] If the machine was programmed to use non-lethal force only, this would eliminate the risk of the AW not complying with the principles of IHL.[184] The option of using non-lethal force as a first resort is not really a plausible solution for humans, as they could be risking their lives if they chose to use non-lethal force in a similar situation – a risk that AWs could take, but not humans.

4.3.3.2 Meagre Measures by Machines: Arguments Against AWs

Again, as with the two previous principles, it is conceivable that if the technology becomes advanced enough, that technology could be used by human soldiers to better comply with the requirements set out by these principles – i.e. allowing human soldiers to benefit from e.g. the advanced camera system mounted on a remote-controlled machine in order to make better informed decisions with regards to precaution, but not allowing the machine to make these decisions on its own.[185]

The argument that AWs may be able to use non-lethal force is not unproblematic either, as even non-lethal weapons mounted on an autonomous weapon system (AWS) can cause indiscriminate injuries and even deaths, when using weapons such as rubber bullets, Tasers and new directed-energy weapons (weapons directing microwaves, lasers, sound etc.).[186] Ultimately, an AW is by definition a machine and no matter how technologically advanced it might become and how perfectly it may mimic us humans, it could never be truly like humans. In that sense, opponents doubt how we could trust that a machine will be able to take measures in order to take constant care of the civilian population, and to protect it, if it has no real relation to the civilian population, no actual awareness about the population, or anything else outside itself and its calculating programs for that matter.[187] The moral and ethical constrain, a result of the human capability of

[182] Singer (2009), p. 37.
[183] HRW Report (2012), p. 38.
[184] Singer (2009), p. 46.
[185] Ibid., p. 35.
[186] Krishnan (2009), p. 107.
[187] Ibid., p. 131.

empathizing with other humans, that may hinder a human soldier from an attack or taking an extra measure to ensure that precaution is taken, will not be able to be made by a machine. The commentary to Article 57 also emphasizes that the interpretation above all must be a question of common sense and good faith for military commanders, and that the weighing of humanitarian and military interests at stake must be carefully done in each and every attack. This begs the question if a machine could ever replicate the common sense and good faith required.[188] A human would in all circumstances have a superior comprehension on situations that arise and if that new situation justifies the use of lethal force. It would not simply pull the trigger blindly because a programmed algorithm said so.[189]

4.3.3.3 Concluding Remarks

The question of whether or not AWs will be able to fulfill the requirements that the principle of precaution sets out, is probably the strongest point for the proponents, and conversely the weakest for the opponents. This can even be inferred already from the fact that not all[190] opponents actually discuss the principle of precaution. The principle can, in its practical application, be boiled down to the two requirements of target selection and weapon selection. Both of these requirements are important when it comes to AWs but apply at different stages.

Weapon selection applies first of all to the commanders deciding at the beginning of an operation to select a weapon – a manned weapon or an autonomous weapon depending on which weapon would cause the least damage and suffering. It also applies to the autonomous weapon system when launched – in the situation where the AWS identifies a target, it is required of the robot to choose the weapon or way of using its weapon in the way that causes the least damage and suffering.

Target selection applies to the AW in the same sense, when it is deployed in the battlefield, the robot is required to choose its targets in accordance with the principle of precaution (as well as distinction and proportionality), and to refrain from engaging the weapon if it, in its assessment of the target, discovers that it would not fulfill the demands of precaution. This is where questions and disagreements still exists regarding the capability of an AW to be sufficiently aware of its surroundings to fulfill these requirements of weapon and target selection.

With regards to taking all feasible precautions to cause the least possible harm, the emerging thought of allowing AWs if the lethal decision-making is always done by a human has been proposed,[191] meaning that non-lethal weapons should be the principal rule but that AWs could be equipped with

[188] Commentary AP I, comment no. 2208.
[189] Krishnan (2009), p. 132.
[190] HRW, for example, does not mention the principle of precaution in its report.
[191] Altmann et al. (2013), p. 76.

lethal weapons if it was programmed in a way that required a human authorization for the use of the lethal force.[192]

Unfortunately, in the debate on precaution (precaution in particular, but to some extent also with regards to the principles of distinction and proportionality) it often seems as though the propositions put forward by the proponents, rather are efforts to appease the opponents than sincere propositions, as these propositions often would mean that the AWs would be unusable in practice (e.g. the suggestions that AWs would only operate in battlefields where no civilians are present, thus complying with the principle of distinction, or that the AWs could use non-lethal force in order to comply with the principle of precaution).

4.3.4 Martens Clause

The issue at hand here is whether or not the Martens Clause should be taken into account during the assessment of legality of the autonomous weapons, and the issue lies in the disagreement between proponents and opponents as to the role of the clause. While opponents like the HRW asserts that the legal review should take the clause into consideration, proponents like Schmitt and Thurnher mean on the other hand that the clause only applies in the absence of treaty law.

HRW cites the ICJ which recognizes the clause as part of customary law, and also observes that the clause has "proved to be an effective means of addressing rapid evolution of military technology".[193] As such, the HRW, along with the ICRC, asserts that the clause should be taken into account in the legal review.[194] Schmitt and Thurnher on the other hand are of the opinion that the clause only acts as a failsafe mechanism and is not meant to be considered an overarching principle that is required to be taken into consideration in every case. In their view, the legality of weapons is sufficiently covered in treaty law and therefore the clause does not merit consideration in the review process.[195] Even if the applicability or non-applicability of the Martens Clause is of no crucial importance in the debate on legality, the non-consensus on the matter nevertheless points to the deep disagreement that exists between the two sides.

4.4 Compliance with Article 36

The possible compliance or non-compliance is understandably an assessment that is not expected of the machine itself, but of the state developing and deploying it. Therefore, it is in regards to the compliance with Article 36, a matter of state responsibility in relation to international law and not the breach of international law by the weapon as such. Obviously, all states that are

[192] Singer (2009), p. 46.
[193] Nuclear Weapons Advisory Opinion, paras. 78 & 84.
[194] HRW Report (2012), p. 25; ICRC Guide (2006), p. 17.
[195] Schmitt & Thurnher (2013), pp. 275-276.

parties to the AP I are bound by this obligation, but even though the customary law status of this obligation is not entirely clear, and non-party states therefore technically only have to ensure that new weapons are lawful before its use, many non-party states have these measures in place for review of new weapons, such as the US. The obligation to perform the legal review in general is confirmed in a DoD directive, stating that the acquisition and procurement of weapons and weapon systems shall be consistent with all applicable domestic and international law.[196]

In the UK, this general obligation is found in their military manual, declaring that the weapon review process is conducted in a progressive manner as concepts for new means and methods of warfare are developed, and that it takes into account likely future developments in the law of armed conflict.[197] These are but two examples of states that in their military manuals or guidelines have included this obligation of performing the legal review of new weapons, and while practice of domestic regulations of the legal review of new weapons is not as widespread as desired, there are some good examples on how states have implemented the requirement of Article 36 in their domestic military regulation.[198]

With regards to AWs, explicit mentioning of reviews are scarcer. The US directive of 2012 concerning AWs devotes one of its enclosures entirely to the guidelines for review of certain autonomous or semi-autonomous weapon systems, thus fulfilling the obligation of performing the legal review.[199] The UK issued a Joint Doctrine Note on the UK Approach to Unmanned Aircraft Systems in March 2011, which states that legal review should continue to be part of the development cycle and must identify the legal, moral and ethical concerns in regards to the development of these new weapon systems.[200] The continued compliance with Article 36 in regards to the development of AWs will also be a matter to discuss in the debate and the way forward.

4.5 Accountability

Initially, it is important to note the difference between accountability due to intentional breaches and accountability due to mistakes – clearly, in the case of a software programmer or commander programming or launching the AW to engage in actions that would amount to war crimes, that person would be held accountable for this intentional breach. However, the much more likely scenario is that of the robot itself making mistakes.[201] In that case, in the

[196] US DoD Directive 5000.01 on subject "The Defense Acquisition System", 12 May 2003, Enclosure 1, para. E1.1.15.
[197] UK MoD, *The Joint Service Manual of the Law of Armed Conflict*, 2004, paras. 6.20 and 6.20.1.
[198] For more examples of domestic regulation of the legal review, see the ICRC Guide (2006), note 8.
[199] DoD Directive 3000.09, Enclosure 3, in particular paras. 1(a)(5) and 1(b)(6).
[200] MoD (2011), para. 708.
[201] Schmitt & Thurnher (2013), pp. 277-278.

determination and attribution of responsibility over actions committed by an AW, there are several options of where the responsibility should be placed; the commander that deployed the AW, the programmer or the manufacturer that designed and produced the AW, the operator that supervises the actions of the AW, or even the AW itself. In much the same way as a military commander is not held accountable for the actions of his subordinates (except in exceptional circumstances as part of the command responsibility doctrine), as the human soldiers are autonomous beings, the attribution of responsibility on a commander of a fully autonomous weapon seems like an inappropriate idea.[202] The application of the command responsibility doctrine would also prove insufficient, as this is based on the knowledge or possible knowledge[203] of the IHL-violation. In the case of a violation by an AW, this requirement of knowledge would fall on the fact that the commander is not the one responsible for programming the weapon.[204]

If the violation committed by the AW instead is viewed as a technical fault, the responsibility would be attributed to the programmer or manufacturer, but also this proposal is ineffective, owing to the complete autonomy after the weapon has been deployed. If a weapon is construed in way that allows it to "feel" its surroundings, especially in a complex battlefield, and make own decisions based on this information, technical malfunctions will not be where the main problem lies. In addition, the criminal responsibility under IHL would only arise if the programmer acted intentionally, and if we are to attribute violations by the AW as a technical mishap, the responsibility for the programmer would fall here.[205]

To attribute responsibility to the manufacturer could prove an even worse solution, as this could lead to the scenario where no manufacturer would be willing to produce these weapons at all, if they were to risk criminal liability for what the weapon might do in the battlefield.[206] The proposal of holding the civilian software writer accountable on a civilian liability level may not be reasonable either. Marchant et al. sees the possibility of a software writer failing to code the machine sufficiently to e.g. recognize a civilian, and the machine then attacks civilians in a battlefield. In this scenario, Marchant means that "it is conceivable that the software writer of the code might be responsible for the mistaken actions".[207] This solution may be questioned on the basis that it is not reasonable that that amount of responsibility should lie with the civilian software writer, for them to have to imagine every

[202] HRW Report (2012), p. 42.
[203] AP I, Article 86(2) "The fact that a breach of the Convention or of this Protocol was committed by a subordinate does not absolve his superiors from penal or disciplinary responsibilities, as the case may be, *if they knew, or had information which should have enabled them* to conclude in the circumstances at the time, that he was committing or was going to commit such a breach and *if they did not take all feasible measures* within their power to prevent or repress the breach." (emphasis added)
[204] HRW Report (2012), p. 43.
[205] Ibid., pp. 43-44; AP I Article 85(3).
[206] HRW Report (2012), p. 44.
[207] Marchant et al. (2011), pp. 282-283.

conceivable scenario of war in order to program the robot correctly, and escape individual liability.

Furthermore, the scenario of one single software writer is becoming less and less probable, since programs today consist of millions of lines of codes put together by a team of programmers, each responsible for parts of the program but none being familiar with the entire program. In addition, with the increasing complexity of the machine, the risk of unexpected, not programmed behavior may arise, as a result of the sheer complexity of the program.[208] Assuming that robots lack common sense or the contextual understanding required, even relatively sophisticated algorithms and software codes can be subject to failure when they face situations outside the design parameter they were intended for.[209] In addition, all of the proposed solutions that involve holding an individual accountable, whether it be the software programmer, the manufacturer or the commander, has one major flaw – does it make sense to hold the human behind the machine accountable, when the machine by definition is designed to and supposed to calculate and reach decisions faster and better than a human being? In that case, obviously the human behind it could always refer to the better judgment of the machine, which is one of the reasons the machine was created for in the first place, one of the very *raisons d'être* of the machine.[210]

Finally, assigning accountability to the machine itself is a possibility that has been discussed, although perhaps never as a serious solution. This is at the moment a dubious idea for several reasons – one being that it would be futile to punish the machine for a misconduct or breach of IHL, since it does not possess the capability to feel remorse or anything like it, nor any moral sensibility that would restrain it from repeating that misconduct.[211] Robots today are far too underdeveloped to understand any kind of accountability on their part, as well as understanding or influencing their behavior in the future.[212] Therefore, the point of attributing accountability to someone for an unwanted behavior, in order to prevent that behavior from recurring in the future, would be moot.[213] This solution also entails the moral and ethical aspects discussed earlier[214], and the very difficult question of whether or not a machine would have sufficient moral and ethical capacities in order to be considered a fully autonomous agent, that willingly and knowingly have performed this unwanted act or behavior.

[208] Ibid., p. 284.
[209] Scharre, Paul "Why Unmanned", in *Joint Force Quarterly*, Issue 61, 2nd Quarter, 2011, p. 92.
[210] Lubell Speech (2013); Sparrow (2007), p. 70.
[211] Sparrow (2007), pp. 71-72.
[212] HRW Report (2012), p. 45.
[213] As Sharkey puts it, it would be like "telling your washing machine that if it does not remove stains properly you will break its door off", Sharkey (2010), p. 380.
[214] See above, primarily chapter 4.1.

4.6 Concluding Remarks

First of all regarding the legality of AWs, it is true that AWs are not technically outlawed, since the terms "autonomous weapons" or "robots" do not appear anywhere in any prohibitive instruments. The problem lies perhaps not with the use of autonomous systems as such, such as patrolling or guarding systems, but when it comes to autonomous weapon systems, lethal autonomous robots capable of taking life-and-death decisions on their own – then we have a far more complicated situation.[215]

The biggest and most important obstacle in the current debate on AWs is the fact that the technology is not yet developed or elaborated enough to be able to say anything for certain on the capability of AWs to adhere to the principles of IHL. The advocates on all sides of the debate are left with basing their arguments on what might happen in the future, regardless of how close or distant that future might be. In this approach, authors naturally choose the future view that would support their cause and arguments best. For example, Schmitt states that "there is *no question* that autonomous weapon systems *could be* programmed"[216], while HRW emphasizes that "it is *highly unlikely* that a robot could be pre-programmed"[217] and that e.g. value judgments "cannot be boiled down to a simple algorithm."[218]

In general, claims that support the development of AWs clearly assume that the technology is possible, but as a result, AWs may be prematurely introduced to the battlefield before the robotics and AI experts are even certain that strong AI capabilities can be produced.[219] One of the major problems with this information-gap is that the international community is not in agreement of how AWs are to be treated until that information and knowledge is entirely clear. The question of if the laws of today are capable to deal with the changes in military technology (especially the question of the new actors on the battlefield, who actually is carrying out the fighting, if robots are considered tools of soldiers or soldiers in themselves)[220] is not clear. Furthermore, the process of changing international law is so long and cumbersome that even if negotiations are held today, the legislation finally emanating from that is a legislation for the past, not the present and certainly not the future.[221] One conceivable solution is that the international community could take command over these issues and start discussing AWs on an international arena, possibly through the forums of the UN or the Conference of Disarmament.

If we were to disregard the uncertainty of development of technology for a second, there are still several other concerns that affect the legality of AWs

[215] Krishnan (2009), pp. 112-113.
[216] Schmitt (2013), p. 20, emphasis added.
[217] HRW Report (2012), p. 32, emphasis added.
[218] Ibid., p. 34.
[219] Ibid.
[220] See above, chapter 4.1.
[221] Lubell Speech (2013).

and the way they are being viewed and discussed in the current debate. One important issue that is frequently brought up in the discussion is the role emotions play in the battlefield – an argument used both by proponents and opponents. Proponents vehemently argue the benefits of having weapons in war that are not affected by human emotions; an AW would not get hungry or tired, would not have any sense of feelings such as revenge if the soldier next to it got killed, and no self-preservation. They would not let their judgment be clouded by emotional aspects, but be able to keep the assessment and information-processing as objective and unsullied as ever – perhaps allowing them to also use force less often.[222] The lack of self-preservation would also mean that the AW could take more risks in the battlefield and therefore limit the amount of force being used.[223] With fewer decisions made in the heat of the battle and clouded by emotions such as revenge, fear, panic or anger, AWs could possibly act more humane in war and therefore commit fewer war crimes than humans.[224]

Arkin points out the dismal record in ethical behavior in the battlefield by human soldiers and the potential causes for war crimes; high losses in the own forces spurs revenge-seeking, high turnover in the chain of command leads to a weakened leadership, dehumanization of the enemy through using derogatory names, inexperienced, immature or poorly trained troops, external pressure of e.g. producing a high number of eliminated enemies, pleasure from power of killing or a sense of anger and frustration. He means that there is undoubtedly room for improvement and that AWs could be a help in addressing these problems.[225] In addition, the removal of human soldiers from the battlefield would mean less mental health issues due to less exposure to violence and traumatizing events in war. The statement by Gordon Johnson, member of the (now defunct) Pentagon's Joint Forces Command is very telling for this point of view:

> They don't get hungry. They're not afraid. They don't forget orders. They don't care if the guy next to them has just been shot. Will they do a better job than humans? Yes.[226]

But the use of weapons that does not have emotions – is that necessarily a good thing? No revenge, sure, but no compassion, empathy or intuition either? The importance of these human emotions and qualities *cannot* be exaggerated. HRW notes that even if human-like cognition in AWs became feasible, they would still lack certain human qualities such as emotions, compassion and the ability to understand other human beings, and therefore, human oversight of AWs would in any case be a necessity to ensure the protection of civilians as well as combatants in times of armed conflict.[227] In

[222] Arkin (2009), pp. 29-30.
[223] Anderson & Waxman (2013), p. 15; Arkin (2009), pp. 120-121.
[224] Anderson & Waxman (2013), p. 15.
[225] Arkin (2013), p. 2.
[226] Weiner, Tim "'GI Robot' Rolls Toward the Battlefield" in *The New York Times*, 17 February 2005.
[227] HRW Report (2012), p. 29.

addition, it could even be conceived that emotions are necessary in order to make ethical judgments, and that a machine that is not capable of feeling compassion for the sufferings of others, is not capable of making good moral or ethical decisions.[228] Of all human emotions robots lack, one of the more important is the skill of reasoning. Calculating and estimating is not the same as reasoning, and reasoning applies both to reasoning on their own, but also reasoning with others. The problem of unanticipated situations that arise is often mentioned in the debate and the capability of robots to cope with these situations is discussed. While unanticipated situations would be the same for human as well as robot soldiers, in the event of such a situation, the human soldier would be able to reason his way to the best solution, and to reevaluate and adapt the actions in the new situation in accordance with the changed conditions, a quality that robots, presumably, does not possess.

While proponents continue to argue that the absence of emotions will make the battlefield more humane and emphasize the shortcomings of humans in war (such as mentioned human fallibilities, lower response times and fatigue)[229], opponents instead stress the importance of human qualities such as showing kindness, mercy and compassion, factors that can restrain the use of force in the battlefield.[230] In general, it feels as though the proponents, not so much overestimates the capabilities of AWs, but underestimates the capabilities of humans! In their arguments on how and why machines will be able to outperform human soldiers, there seems to be very limited amount of consideration and more importantly, value, of how human soldiers in an instant can read the body language, facial expression, atmosphere and other relevant factors of a possible target, and in an instant is capable of deciding the status of that target as lawful or unlawful.

As mentioned earlier in regards to the development of artificial intelligence,[231] the capability to read contexts is conceivably something that only humans are capable of, and the human factor in assessments and application of law is possibly inimitable. The interpretation of international law is based on context, social norms and judgment, and qualities like that are extremely difficult for a robot to mimic, if possible at all. In addition, humans are far superior to machines or computers when it comes to managing information that is incomplete, contradictory or unformatted and to make decisions when it is hard to foresee the possible outcomes or consequences of a certain scenario of actions.[232]

Another issue that is often mentioned against AWs is the technological weaknesses associated with them, such as framing problems and weak software. The framing problem is related to the difficulties in AI and programming, specifically, how the robot interacts with its environment in

[228] Kaag & Kaufman (2009), p. 598.
[229] Krishnan (2009), pp. 39-41.
[230] Sharkey (2012), p. 793.
[231] See above, chapter 3.2.3.1.
[232] Kaag & Kaufman (2009), p. 602.

relation to how any given situation was represented in that robot. This problem arise from the desire of not having to write endless codes and formulas to describe the effects caused by certain actions which would render the robot too slow to be militarily useful. This would result in a robot that is prone to use force indiscriminately or disproportionately as it often would miss important information or misinterpret situations.[233]

The problem with weak software is where the software becomes more and more complex and therefore possibly less and less reliable, safe and trustworthy. If the existence of a reliable software, programmed into the AWs, is the safeguard for a robot to act in a way that would be in compliance with the mentioned principles of IHL, as well as moral and ethical considerations, then the risk of problems with weak software, or technical glitches in the software, is a risk not worth taking.[234] Could a machine actually look at the big picture, not just identifying if an individual is a threat, but understand the context and everything around it, and would it be able to read human emotions and understand them? Other suggestions put forth to justify the use of AWs are that the machines could be programmed to use force only when there is zero doubt. While this is a nice idea in theory, practice shows that in warfare, there is never zero doubt. This approach would therefore render the weapon impractical and unusable in the actual battlefield. Another idea is that AWs would be programmed to never fire first, to only return fire. Again, this is feasible in theory, but in a battlefield it would not be a practical solution. The same goes for the suggestion to never use AWs when civilians are around, because nowadays, civilians are rarely completely absent from the battlefields.

Clearly, a large amount of issues still remains regarding both the legality and morality of developing and using AWs, and the disagreements regarding these many aspects are not easily resolved. Despite of all uncertainties and concerns, as well as all arguments put forth both in support of and against AWs, the fact still remains that they are being developed and researched, and the issue then becomes how we are to move forward.

[233] Krishnan (2009), p. 99.
[234] Ibid., p. 100.

5 Regulate or Terminate: Regulating the Terminator

Having reached the conclusion that it might not be possible today to conclude anything decisive about the possibilities or capabilities of AWs to comply with the rules of IHL, the subsequent question then becomes how we are to treat them. As far as IHL is concerned, perhaps we need to think about a more creative and forward thinking approach to how we interpret and apply these principles if we want them to remain relevant. However, we also need to remember that the law might not always be the answer, and in certain areas (such as the present one with AWs), sometimes arguments from policy perspectives and ethical aspects are stronger than legal rules when it comes to dealing with lethal decisions made by robots.[235] The most pressing need now is to acknowledge that the problems surrounding AWs exist, thus legitimating it as an international concern, drawing attention of relevant experts and in doing so, demonstrate that the international community is taking the issue seriously.

While it is clear that not enough is known about the potential risks of AWs, waiting for the technology to develop could mean that it might be too late to undertake any meaningful regulations because the commercial drive behind the technology would be too strong and rooted.[236] Jürgen Altmann, military nanotechnology expert, expresses concerns over the proliferation of AWs and warns that unmanned, remote-controlled systems could be deployed in high numbers and could thereafter relatively simply and covertly be changed into autonomous systems.[237] The fact that military robotics technology is moving forward at an unprecedented pace is a point of concern when it comes to the implementation of an international regulation in order to both ensure the sufficient protection of civilians, but also to prevent the development of a dynamic that could possibly upset existing strategic balances, something that could result in the destabilization of the entire international system.[238]

5.1 An Outright Ban

The call for an outright ban on the development and use of autonomous weapons was made already in 2009 by the newly founded organization International Committee for Robot Arms Control (ICRAC). The organization promotes the prohibition of the development, deployment and use of armed AWSs and states that machines should not be allowed to make killing decisions.[239] Among the founders of ICRAC are Noel Sharkey, Peter Asaro

[235] Lubell Speech (2013).
[236] Marchant et al. (2011), p. 314.
[237] Krishnan (2009), p. 145.
[238] Ibid., p. 146.
[239] The mission statement reads:

and Robert Sparrow, and the members consists of experts in robotics technology, robot ethics, international relations, international security, arms control, international humanitarian law and international human rights law.[240] As of October 2013, the organization had gathered more than 270 computer scientists, engineers, AI experts, roboticists and others, representing 37 countries, in its call to ban AWs.[241]

In 2012, HRW and the International Human Rights Clinic at Harvard Law School (IHRC) issued the same call, and recommended all states to prohibit the development, production and use of fully autonomous weapons through an internationally binding legal instrument.[242] HRW was shortly after one of the leading forces behind the launch of the Campaign to Stop Killer Robots in April 2013, an international coalition of NGOs working to ban fully autonomous weapons.[243] These three represent the largest forces behind the call to ban AWs, and they all propose the same solution: the implementation a pre-emptive and comprehensive ban on the development and use of fully autonomous weapons, something that would be achieved through an international legally binding instrument, as well as implementation in domestic laws and other measures. The call for a ban is based on concerns about taking humans out of the loop when it comes to targeting and firing, essentially killing, decisions. Concerns also relate to the possible consequences of letting autonomous weapons develop and proliferate freely without any international control or overview.

The argument most vehemently put forward is that it would be irresponsible to simply wait and see, and the opponents stress the need to initiate an international ban before too many states develop the technology and we "venture down a path from which there is no return", as Sharkey puts it.[244] HRW in particular also stresses the threat that AWs would pose to civilians in times of war, and that a ban would ensure that the targeting and firing decisions are always carried out by a human soldier, capable of adhering to the principles of IHL.[245] Asaro points out several benefits of introducing a

---

> Given the rapid pace of development of military robotics and the pressing dangers that these pose to peace and international security and to civilians in war, we call upon the international community to urgently commence a discussion about an arms control regime to reduce the threat posed by these systems. We propose that this discussion should consider the following:
> Their potential to lower the threshold of armed conflict;
> The prohibition of the development, deployment and use of armed autonomous unmanned systems; machines should not be allowed to make the decision to kill people;
> Limitations on the range and weapons carried by "man in the loop" unmanned systems and on their deployment in postures threatening to other states;
> A ban on arming unmanned systems with nuclear weapons;
> The prohibition of the development, deployment and use of robot space weapons.

[240] ICRAC website, "Who we are", http://icrac.net/who/.
[241] Statement issued by the ICRAC, 16 October 2013, available at http://icrac.net/wp-content/uploads/2013/10/List-of-Signatories-ICRAC-calll.pdf.
[242] HRW Report (2012), p. 5.
[243] Stop Killer Robots website, "About Us", http://www.stopkillerrobots.org/about-us/.
[244] Sharkey (2013).
[245] HRW Report (2012), p. 46.

binding legal treaty (besides clarifying the international community's stand on AWs) including to avoid a slippery slope towards AWs, to shape future investments in technology development and to establish the legal principle that autonomous systems are not sufficiently morally capable to make life and death decisions about humans. According to this, a ban would be desirable in order to protect human rights, as well as other norms protecting individuals.[246]

However, there is the concern from proponents of AWs that a ban might disrupt the development of a weapon that is possibly more capable than humans of respecting legal and moral norms, and that it would disregard the potential for a weapon that could minimize harm to civilians.[247] Schmitt even argues that it would be irresponsible to ban AWs at this early stage of development, as this would disregard the possible advantages that AWs might have on warfare, such as the possibility to attack an enemy without risk for the attacker. He also contend that the development is not yet at a point where we can conclude if AWs are more or less harmful than human operated systems. The argument here is that it would be irresponsible *not* to wait and see what possible benefits the development of AWs might have for future warfare.[248]

In light of this, some authors propose alternative approaches, where AWs are regulated through a framework convention, rules of procedures, policies or directives. These proposals emphasizes the view that with incremental development of AWs, the regulation of acceptable use should also be incremental.[249] The biggest obstacle however, is the lack of political will, when one of the main motivations for developing AWs is just that, generating political will for warfare, by removing combatants from the heat of battle. While there are successful precedents of NGOs campaigning leading to a prohibitive convention,[250] the feasibility of such a convention in the case of AWs is probably significantly lower, judging by the difference in military advantage and effectiveness between landmines and AWs. States that have a lot to win by the development of AWs will surely not be prepared to sign a treaty banning the development or use of these weapons, so the proposition might fall on its own unreasonableness. Instead of imposing a prohibition of AWs altogether, more and more authors propose the solution of negotiating a sort of framework convention, an arms control agreement.

5.2 Adoption Within the Frames of CCW

While it has traditionally been weapons of mass destruction (WMD) that have received the overwhelming part of international arms control attention, conventional weapons are becoming increasingly sophisticated and military

[246] Asaro (2012), pp. 693-694.
[247] Schmitt & Thurnher (2013), p. 234.
[248] Schmitt (2013), pp. 36-37.
[249] Anderson & Waxman (2013), pp. 2-3.
[250] The International Campaign to Ban Landmines was a coalition of NGOs formed in 1992 which eventually led up to the adoption of the Mine Ban Treaty in 1997.

effective and could soon be no less important to control and regulate than WMDs. Therefore, it could be important to bring AWs, which are currently only indirectly (or possibly insufficiently) covered by treaties, onto the agenda of international arms control before things get out of hand.[251]

The United Nations Convention on Certain Conventional Weapons (CCW) was drafted in October 1980 and seeks to prohibit or restrict the use of certain conventional weapons which are considered excessively injurious or has indiscriminate effects. It is an annex to the GCs and contains five protocols, each of which regulates one type of weapon; weapons with non-detectable fragments, landmines and booby traps, incendiary weapons, blinding laser weapons and clearance of explosive remnants of war. The aim of the convention is to provide additional protection for civilians as well as combatants for injuries caused by weapons not sufficiently covered by the GCs and APs.[252]

One significant feature of the CCW is the possibility to expand it and adopt additional protocols to respond to the development of new weapons. This means that any state party may suggest an additional protocol to be added, and a conference may then be held to discuss the proposition.[253] This particular feature of the CCW is what has attracted recent considerable attention in relation to AWs, especially from NGOs such as ICRAC and the Stop Killer Robots campaign, and its coalition members. In November 2013, these two NGOs motioned to bring the issue on the agenda for the Conference of the State Parties for the CCW, and on 15 November, the Conference decided to convene a Meeting of Experts on 13-16 May 2014 for their first ever meeting discussing questions related to AWs.[254] The agenda for the Meeting of Experts includes discussions on technical issues, ethics and sociology and legal aspects. The legal aspects in particular, address issues relating to implications for the principle of humanity and the Martens Clause, the impacts on the principles of distinction, proportionality and precaution, the compatibility of AWs and Article 36 and finally the impact on issues of responsibility and accountability for violations of international law.[255] The outcome of that meeting will then be discussed at the next Conference of the State Parties in November 2014.[256] A possible outcome of this process could lead to a negotiation which in turn could lead to a new protocol for the CCW, which could then either ban or at the very least, regulate the development and deployment of AWs.

---

[251] Krishnan (2009), p. 157.
[252] CCW Preamble.
[253] CCW Article 8(2)(a)-(b).
[254] Final Report, Meeting of the High Contracting Parties to the 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 on 14-15 November 2013, Report published 16 December 2013 (hereinafter CCW Final Report), paras. 18 & 32.
[255] CCW Expert Meeting on Lethal Autonomous Weapon Systems, Agenda on 13th-16th May, 12 May 2014.
[256] CCW Final Report, annex I.

Krishnan notes that the international community could benefit immensely from adopting some kind of arms control which would prevent or at least slow down the arms race in military robotics and the proliferation of these weapons. With the need to contain the potentially very negative consequences of this increasingly advanced technology on international security, a regulation could prevent a situation in which self-evolving autonomous defense systems develops in a way that in the long run could possibly threaten the very existence of humanity.[257]

5.3 Soft Law Governance

While the idea of a hard law regulation, both in terms of a pre-emptive ban and a regulation within the CCW, might be a good idea in theory, again it might not be practicable due to the problem of political will. With all international treaties, there is always the problem of the state having to give something up on a national level, for the good of the international community, whether it is the state restricting its free trade in order to protect the environment, or making sure the state complies with its human rights obligations in order to receive aid. With this in mind, it is not hard to see the reluctance of states to sign and ratify international treaties that directly restrains the decisiveness of the state, *especially* when it comes to the capability to develop its military forces and the protection and security of the very state itself. In addition, limitations such as the resources and time needed for negotiations, difficulties in enforcement and compliance, and lack of flexibility in international instruments makes this traditional model less and less realistic. In that case, an alternative approach based on soft law might prove more appropriate.[258]

While soft law has the distinct disadvantage of not being formally binding and non-enforceable, instead it has the advantages of being more flexible, capable of a more speedily launch and more easily adaptable to changes in both technology and the political and social scene. Marchant et al. suggests that some of the methods applied in other areas of emerging technologies possibly could be applied to military technology as well,[259] including codes of conduct, transgovernmental dialogue, information sharing and confidence-building measures.

Codes of conduct – non-binding general guidelines that defines the responsible, ethical behavior – has the advantage of being able to be developed and implemented by various entities in society, including governmental agencies, industries, NGOs, scientific societies etc., and they may be important as a first step in providing some protection and governance until more formal instruments can be negotiated. While codes can be created rather quickly compared to more formal instruments, the biggest downside is the difficulty in application, and if there are multiple codes for one area, there

[257] Krishnan (2009), pp. 156-157.
[258] Marchant et al. (2011), p. 306.
[259] Ibid.

is no clear hierarchy of which code takes precedence over another. Recently, codes of conduct have been created for research projects on synthetic biology and nanotechnology, areas that might have potential for military application. Although a code of conduct might not be enough to completely regulate the area of AWs, it could certainly be a step forward in creating a common goal for such a regulation.[260]

Transgovernmental dialogues are more informal and flexible arrangements between government officials from different countries who discuss and coordinate policies and they have become increasingly common in areas that concerns international coordination, such as national security issues. This provides a forum where information and best practices can be shared in the pursuit of harmonization of policies and enforcement practices, and these dialogues can greatly enhance cooperation and influence policy outcomes in the international arena. As this method also proves effective in starting a discussion among policymakers from different countries, starting transgovernmental dialogues could also be an important step towards more concrete policy measures in the future.[261]

The concept of information sharing and confidence-building measures first arose in the area of international relations and have frequently been used in international conflicts as initial measures in order to reduce hostilities and to enhance stability, trust and security. Within the scope of AWs, these kind of measures could consist of either unilateral or multilateral initiatives, and nations could commit to a limited moratorium[262] on the deployment of AWS, as well as commit to share information on technical issues and agree to host international conferences to discuss the issue of AWs.[263]

[260] Ibid., pp. 307-310.
[261] Ibid., pp. 311-313.
[262] Arkin supports the idea of a moratorium to ensure that technology meets international standards, Arkin (2013), p. 5.
[263] Marchant et al. (2011), p. 313.

6 Reflections

In April 1978 an international panel of experts convened to discuss the issue of "should weapons of dubious legality be developed?", shortly after the adoption of the two APs in 1977.[264] The same questions that were discussed in that panel are the ones that are being discussed today regarding the development of autonomous weapons. If we cannot be sure of the legality of a weapon, should we really allow the development of that weapon to continue? In this very new but quickly polarized debate, the positions seem to have reached a stalemate. It is clear that we cannot conclude anything decisive about the legality of AWs, their ability to comply with the rules of IHL, before we know if the technology is actually possible, if we can actually develop a strong enough AI that will be able to respect the rules of IHL. While the debate between proponents and opponents continues, the international community must start to take a stand on how these weapons are to be treated.

It is safe to say that increasingly autonomous systems and robots are here to stay – their usefulness in many areas in our society are undeniable and far too great to be underestimated. Already, autonomous machines have already started to carry out meaningful tasks, such as sifting through the wreckage at Ground Zero after 9/11[265], delivering cargo to operating bases[266], or with the ability to function as Medbots or Searchbots[267], or even used in the civil society as advanced nursing or elder-care robots[268]. However, this use of autonomous systems are far from the use of autonomous weapon systems, weapon systems that in effect will make life-and-death decisions in the battlefield – a distinction that is often ignored or at least overlooked. Taking this decision-making power out of the hands of humans and placing it in a machine will have enormous moral and ethical consequences, and could even affect the way the legality of the weapon is being considered. Ultimately, a limit must be drawn where what is military advantageous, comes to a point where we lose humanity in warfare.

If the analysis of the legality of AWs would consist only of considerations of military effectiveness, it would be easy to see the appeal of developing more and more autonomous weapon systems which would replace humans in the battlefield and thus completing the seemingly unattainable goal of entirely distancing humans from the actual battle – machines would not get tired, hungry, wounded, affected by emotions, they would not need salary, pensions, insurances. The military advantages that AWs could pose in filling the roles of the three Ds – dull, dirty or dangerous – are strong arguments in

[264] "Should Weapons of Dubious Legality Be Developed", Proceedings of the 72nd Annual Meeting of the American Society of International Law, 1978.
[265] Singer (2009), p. 33.
[266] Thurnher (2014), pp. 216-217.
[267] Singer (2009), p. 37.
[268] Anderson & Waxman (2013), p. 2.

favor of the use of AWs.[269] But the development of more and more autonomous weapons could lead down a path that today only seems as science fiction or popular culture – the similarities to *Star Wars*, *Terminator*, *Robocop* and others are uncanny – but that could be reality in only a few decades. What if the development slips out of our hands, out of human control? A weapon that is instructed to learn and adapt through its mistakes could eventually evolve into an unpredictable, virtually indestructible, killer machine.[270] All of these, possible but more or less probable, scenarios for the future, might seem overly exaggerated today, but the seemingly beneficial and simple action of removing humans from the battlefield has the very real and disconcerting effect of making the wars easier to enter and easier to wage. Replacing humans with robots could have two important implications – the shift of the burden of war to civilians, a real risk especially in the context of today's asymmetric warfare,[271] and most importantly – lowering the threshold of war. The turn to AWs alters the political calculations of war, the prospect of being able to wage a war without the risk of casualties in the own forces would remove one of the greatest deterrents of war.[272] In removing this biggest disincentive to start wars, we could be seeing a future of warfare where nations are more and more prone to use armed force, because the costs are so much lower than they ever would be with human soldiers.[273]

If we were to witness, with the rise of autonomous weapons, the emergence of an entire industrialization of warfare with clean-killing battlefields where hi-tech countries fight each other without risk to their own forces – a battle of technology – a battle between robots – then this emerging technology and autonomous weapons could be accepted. But as this science fiction-utopian scenario is more than highly unlikely, we are left with the scenario where some hi-tech countries are in possession of these weapons and will use them against human soldiers, and will use them in contemporary battle-spaces where combatants blend in together with civilians. In this scenario, as well as the scenarios of war we have witnessed for centuries (and are witnessing today), civilians will be caught in between and come to harm. In these, far more likely scenarios, it is obvious that the protection of civilians must continue to be a priority and that considerations of this must be taken when discussing the legality of AWs.

While current efforts to create these kinds of new legal regimes for weapons are increasingly led by states that have a low likelihood of ever using these weapon systems in combat or by NGOs,[274] it is also increasingly clear that

---

[269] Singer (2009), p. 31. Dull is a reference to boring tasks that often takes a lot of time (such as staring at a monitor for thirty hours without blinking); Dirty means battlefields that are filled with biological or chemical weapons, and Dangerous conditions could include space, rough seas or flights with very high gravitational pressure.
[270] Krishnan (2009), pp. 53 & 114. The idea of an evolving, adapting weapon was for example envisaged in the animated movie *The Incredibles*.
[271] HRW Report (2012), pp. 39-40.
[272] Asaro (2012), p. 692; Krishnan (2009), p. 114.
[273] Anderson & Waxman (2013), p. 17.
[274] Schmitt, Michael N., foreword to *New Technologies and the Law of Armed Conflict*, Hitoshi Nasu & Robert Mc Laughlin (eds.), 2014.

the introduction of autonomous weapons on the battlefield is something that engages and upsets the international community. Instead of allowing the unregulated development and deployment of AWs into the battlefields, a compromise could possibly be reached, where the technology that is being developed, instead of being programmed into self-governing machines, is being used as aid and assistance for human soldiers. That way, combatants and civilians would still be able to benefit from all the distinct advantages that the ever-improving technology provides, without relinquishing the essential factor of human decision-making. Using the machines as aids and tools instead, would also sidestep the moral and ethical obstacles that arises in changing the agents of war from humans to robots. Again, if we do not know exactly what it is that makes us humans, we cannot know exactly what it means for the application of the rules of IHL if we remove the human factor. This is an additional point of concern in the development of AWs – what happens to the actual application of IHL if we change the actors in the battlefield? Are the principles of distinction, proportionality and precaution even applicable? Or do we need a completely new set of rules that applies to the use of robots in the warfare, principles and rules that will regulate this new form of war? I personally do not believe it will come to this even with the rapid development of AWs, but the fact that we do not know how IHL applies if we are to change the agents in war should be considered in the current and coming debate.

So far, the debate has already engaged some of the world's most highly regarded scientists, authors and experts, on both sides of the debate. One interesting factor in the debate is that some of the world's most noteworthy professors and military experts have chosen the opposing side of the debate, the side that argues *against* the development and use of AWs, and have also gone to the lengths of creating an NGO that actively works for the prohibition of these weapons. Even if this circumstance does not say anything substantial about the legality of AWs, I think that this position does say something about the character of these weapons and of how we should be treating them. To my knowledge, since the call for banning landmines, no weapon has spawned this much international attention, a circumstance I find worth noting. In addition, on the other side of the debate, the proponents, we find that the most vocal advocates are scientists and experts are actually employed by the US military![275] If anything would call the arguments from the side of the proponents into question, it is the fact that these advocates actually have own interests in the debate.

The other thing that disturbs me when having analyzed the debate and the many articles, pleas, editorials and viewpoints expressed, is the fact that the

[275] Michael Schmitt, Jeffrey Thurnher, Ronald Arkin, Matthew Waxman and Kenneth Andersson are all representing the US government military branches in some way (Schmitt is professor at the US Naval War College; Thurnher is Lt. Col. in the US Army and professor at the US Naval War College; Arkin is professor at Georgia Tech and funded by *inter alia* DARPA and the US Army; Waxman is professor at Columbia Law School and has previously served in senior positions at the State Department, DoD and National Security Council and, together with Anderson (who is professor at Washington College of Law), member of the Hoover Institution's Task Force on National Security and Law.

debate seems to have boiled down into a clear-cut, plain discussion on legality, while at the same time, the parties to the discussion are well aware that we cannot yet say anything for certain on the legality, since we cannot yet see the full development of autonomous weapon systems. As a soon-to-be international lawyer, yes, legality is important and it is the focus of this thesis, but I do find it alarming that the debate seems to have come to the point of legality, and stopped there. It is disquieting that issues such as morality and ethicality of using robots as soldiers, of changing the very agents of war, of letting machines make life-and-death decisions, that these questions have so far been extremely overlooked in the debate. In several of the articles used for this thesis, the issue of morality pops up once every now and then, indicating that there are concerns on this point, but at the same time showing the unwillingness or inability to discuss these matters in a satisfactory manner. It is clearly difficult to discuss moral and ethics, especially in the context of war, but at the moment, the vital part is just to make sure that these issues are taken into consideration, and to widen the scope of the assessment to not just encompass the legality and military effectiveness of these possible weapons, but to also encompass moral and ethics. As argued above, I believe that there is room for the moral issues to become part of the legality evaluation as well, e.g. to the extent of where IHL is still applicable if we change the agents in war, and the issue of accountability, for the same reason.

However, the aspect of this debate that I find most important to bring into the coming considerations, conferences, debates or discussions on the legality of AWs, is the bigger picture of future warfare. While the balance between scary and advantageous must be maintained until we know more, the considerations of what consequences might come from introducing AWs onto the battlefield, cannot wait. This argument deserves some extra weight in the debate – by removing humans from war, we also remove the biggest disincentive for war altogether. In these discussions, we, the international community, must ask ourselves if this is the view of war and use of armed force we want to see for the future? All studies on humanitarian law and peacebuilding are based on the idea that wars are of evil and that we should do everything humanity can to refrain from it, but would the use of AWs really further this cause? Or would it in fact move in the opposite direction?

Yes, wars are changing. With regards to this particular change, perhaps it is too early for us to say anything definitively on the legality of autonomous weapons. It is however not one day too soon to be talking about how we are to treat the development of increasingly autonomous weapons, and with each week comes new developments and improvements in the field of military technology and with that, new legal and moral challenges. What now lies before us is for us to decide whether we are to direct the technology ahead of us or if we are to let the technology direct us. Us directing the technology could possibly lead to more humane wars with more precise attacks on combatants and less collateral damage on civilians. Technology directing us could instead lead to a decreased humanity in warfare and, ultimately, us losing control over the way future wars are waged.

Bibliography

Books

Arkin, Ronald *Governing Lethal Behavior in Autonomous Robots*, CRC Press, 2009.

Bergström, Göran & Kristina Boréus (eds.) *Textens Mening och Makt; Metodbok i Samhällsvetenskaplig Text- och Diskursanalys*, 3rd ed., Studentlitteratur Lund, 2012.

Fleck, Dieter (ed.) *The Handbook of Humanitarian Law in Armed Conflicts*, 2nd ed., Oxford University Press, 2008.

Henckaerts, Jean-Marie & Louise Doswald-Beck *Customary International Humanitarian Law, Volume I: Rules and Volume II: Practice*, ICRC, Cambridge University Press, 2005.

Krishnan, Armin *Killer Robots: Legality and Ethicality of Autonomous Weapons*, Ashgate Publishing, 2009.

Nasu, Hitoshi & Robert McLaughlin (eds.) *New Technologies and the Law of Armed Conflict*, T.M.C. Asser Press, The Hague, 2014.

Sandos, Yvez, Christine Swinarski & Bruno Zimmerman (eds.) *Commentary on the Additional Protocols of 8 June 1977 to the Geneva Conventions of 12 August 1949*, ICRC, Martinus Nijhoff Publishers, Geneva, 1987.

Schmitt, Michael N. *Essays on Law and War at the Fault Lines*, T.M.C. Asser Press, The Hague, 2012.

Periodicals

Altmann, Jürgen, Peter Asaro, Noel Sharkey & Robert Sparrow "Armed Military Robots: Editorial" in *Ethics and Information Technology*, Vol. 15, No. 2, published online 20 June 2013.

Anderson, Kenneth & Matthew C. Waxman "Law and Ethics for Robot Soldiers" American University Washington College of Law Research Paper No. 2012-32, 2012.

Anderson, Kenneth & Matthew C. Waxman "Law and Ethics for Autonomous Weapon Systems: Why a Ban Won't Work and How the Laws of War Can" American University Washington College of Law Research Paper No. 2013-11, 2013.

Arkin, Ronald "Lethal Autonomous Systems and the Plight of the Non-Combatant" in *AISB (The Society for the Study of Artificial Intelligence and Simulation of Behaviour) Quarterly*, No. 137, July 2013.

Asaro, Peter "On banning autonomous weapon systems: human rights, automation and the dehumanization of lethal decision-making" in *International Review of the Red Cross*, Vol. 94, No. 886, Summer 2012.

Kaag, John & Whitley Kaufman "Military frameworks: technological know-how and the legitimization of warfare" in *Cambridge Review of International Affairs*, Vol. 22, No. 4, December 2009.

Koskenniemi, Martti "What Use For Sovereignty Today?" in *Asian Journal of International Law*, Vol. 1, Issue 01, 2011.

Lin, Patrick, George Bekey & Keith Abney "Autonomous Military Robotics: Risk, Ethics and Design" prepared for the US Department of Navy, Office of Naval Research by California Polytechnic State University, San Luis Obispo, 20 December 2008.

Liu, Hin-Yan "Categorization and legality of autonomous and remote weapons systems" in *International Review of the Red Cross*, Vol. 94, No. 886, Summer 2012.

Marchant, Gary E., Braden Allenby, Ronald Arkin, Edward T. Barrett, Jason Borenstein, Lyn M. Gaudet, Orde Kittrie, Patrick Lin, George R. Lucas, Richard O'Meara & Jared Silberman "International Governance of Autonomous Military Robots" in *The Columbia Science and Technology Law Review*, Vol. XII, 2011.

Scharre, Paul "Why Unmanned" in *Joint Force Quarterly*, Issue 61, 2nd Quarter, 2011.

Schmitt, Michael N. "Autonomous Weapon Systems and International Humanitarian Law: A Reply to the Critics" in *Harvard National Security Journal Features*, 2013.

Schmitt, Michael N. & Jeffrey S. Thurnher "'Out of the Loop': Autonomous Weapon Systems and the Law of Armed Conflict" in *Harvard National Security Journal*, Vol. 4, 2013.

Sharkey, Noel "Automating Warfare: lessons learned from the drones" in *Journal of Law, Information and Science*, Expert Commentary for ed. 21(2), 10 August 2011.

Sharkey, Noel "Saying 'No!' to Lethal Autonomous Targeting" in *Journal of Military Ethics*, Vol. 9, No. 4, 2010.

Sharkey, Noel "The Evitability of Autonomous Robot Warfare" in *International Review of the Red Cross*, Vol. 94, No. 886, Summer 2012.

Singer, Peter W. "Military Robots and the Laws of War" in *The New Atlantis, A Journal of Technology and Society*, Winter 2009, adapted from *Wired for War: The Robotics Revolution and Conflict in the Twenty-First Century* by P.W. Singer, 2009.

Sparrow, Peter "Killer Robots" in *Journal of Applied of Philosophy*, Vol. 24, No. 1, 2007.

Electronic sources

Automated Insights website.
http://automatedinsights.com/ last accessed on 25 May 2014.

Boston Dynamics on the robot *Atlas*.
http://www.bostondynamics.com/robot_Atlas.html last accessed on 25 May 2014.

Campaign to Stop Killer Robots website.
http://www.stopkillerrobots.org/ last accessed on 25 May 2014.

CCW Expert Meeting on Lethal Autonomous Weapon Systems, Agenda on 13th-16th May, 12 May 2014.
http://www.unog.ch/80256EDD006B8954/(httpAssets)/CDB6FA89EAFFEFB3C1257CD3005314A2/$file/ProvisionalAgenda_CCWMXLAWS+12May14.pdf last accessed on 25 May 2014.

C-RAM Transform Defense Tactics", *US Army news release*, 26 April 2012.
http://www.army.mil/article/78724/C_RAM_transforms_defense_tactics/ last accessed on 25 May 2014.

Dahm, Werner J.A., US Air Force presentation "Technology Horizons: Vision for Air Force 2010-2030 Capabilities Enabled by Science & Technology", Washington D.C., 25 June 2011.
http://sdsi.asu.edu/wp-content/uploads/2012/01/Dahm-AFA-Southwest-Presentation.pdf last accessed 25 May 2014.

Debate: "Autonomous Weapon Systems: Dangerous Killer Robots or Smarter and Less Harmful Warfare?" Geneva Academy 20 November 2013.
http://icrac.net/2013/11/geneva-academy-debate-matthew-waxman-vs-icracs-peter-asaro/ last accessed on 25 May 2014.

Federation of American Scientists "MK 15 Phalanx Close-In Weapons Systems (CIWS), 9 January 2003.
http://www.fas.org/man/dod-101/sys/ship/weaps/mk-15.htm last accessed on 25 May 2014.

Final Report, Meeting of the High Contracting Parties to the 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, 2013 Session on 14-15 November 2013, Report Published 16 December 2013. http://daccess-dds-ny.un.org/doc/UNDOC/GEN/G13/646/33/PDF/G1364633.pdf?OpenElement last accessed on 25 May 2014.

Garamone, Jim "Iron Dome System Demonstrates US-Israeli Partnership" in *American Forces Press Service*, 1 August 2012. http://www.defense.gov/news/newsarticle.aspx?id=117354 last accessed on 25 May 2014.

Germany Orders MANTIS C-RAM Base Defense Systems" in *Defense Industry Daily*, 17 January 2011. http://www.defenseindustrydaily.com/Germany-Orders-Skyshield-C-RAM-Base-Defense-Systems-05418/#deutschland-nbs last accessed on 25 May 2014.

Human Rights Watch & International Human Rights Clinic, Human Rights Program at Harvard Law School, Report on "Losing Humanity: The Case against Killer Robots", November 2012. http://www.hrw.org/sites/default/files/reports/arms1112ForUpload_0_0.pdf last accessed on 25 May 2014.

International Committee of the Red Cross "A Guide to the Legal Review of New Weapons, Means and Methods of Warfare: Measures to Implement Article 36 of Additional Protocol I of 1977", January 2006. http://www.icrc.org/eng/assets/files/other/icrc_002_0902.pdf last accessed on 25 May 2014.

International Committee of the Red Cross "Autonomous weapons: States must address major humanitarian, ethical challenges" FAQ, 2 September 2013. http://www.icrc.org/eng/resources/documents/faq/q-and-a-autonomous-weapons.htm last accessed 25 May 2014.

International Committee of the Red Cross "Interpretive Guidance on the Notion of 'Direct Participation in Hostilities' Under International Humanitarian Law", Nils Melzer, 26 February 2009. http://www.icrc.org/eng/assets/files/other/icrc-002-0990.pdf last accessed on 25 May 2014.

International Committee of the Red Cross "The SIrUS Project: Towards a determination of which weapons cause 'superfluous injury or unnecessary suffering'", Robin M. Coupland (ed.), 1997. http://www.loc.gov/rr/frd/Military_Law/pdf/SIrUS-project.pdf last accessed on 25 May 2014.

International Committee for Robot Arms Control website. http://icrac.net/ last accessed on 25 May 2014.

International Committee for Robot Arms Control, Scientist's call for ban on AWs, 16 October 2014. http://icrac.net/wp-content/uploads/2013/10/List-of-Signatories-ICRAC-calll.pdf last accessed on 25 May 2014.

Iron Dome Battle Management Demonstrated" in *Defense Update*, 2009. http://defense-update.com/photos/iron_dome_bms.html last accessed on 25 May 2014.

Kumagi, Jean "A Robotic Sentry for Korea's Demilitarized Zone" in *Institute of Electrical and Electronics Engineers Spectrum*, 1 March 2007. http://spectrum.ieee.org/robotics/military-robots/a-robotic-sentry-for-koreas-demilitarized-zone last accessed on 25 May 2014.

Land-Based Phalanx Weapon System Completes Mission in Iraq" in *Naval Sea Systems Command (NAVSEA) newswire*, 16 February 2012. http://www.navsea.navy.mil/Lists/NewsWires/DispForm.aspx?ID=12 last accessed on 25 May 2014.

Narrative Science website. http://narrativescience.com/ last accessed on 25 May 2014.

Noam Lubell Speech on "Robot Warriors, Terrorists & Private Contractors: What Future for the Laws of War?" in *Professorial Inaugural Lecture Series*, University of Essex, 13 December 2013. http://vimeo.com/channels/642873 last accessed on 25 May 2014.

Opinion – Viewpoint "UK roboticist Prof Noel Sharkey calls for a pre-emptive ban on the deployment of autonomous weapons" in *The Engineer*, 11 March 2013. http://www.theengineer.co.uk/military-and-defence/opinion/say-no-to-killer-robots/1015720.article last accessed on 25 May 2014.

Panel on "Should Weapons of Dubious Legality Be Developed", Proceedings of the 72nd Annual Meeting of the American Society of International Law, 1978. http://heinonline.org/HOL/Page?handle=hein.journals/asilp72&div=7&g_sent=1&collection=journals#36 last accessed on 25 May 2014.

Rabiroff, Jon "Machine Gun-Toting Robots Deployed on DMZ" in *Stars and Stripes*, 12 July 2010. http://www.stripes.com/news/pacific/korea/machine-gun-toting-robots-deployed-on-dmz-1.110809 last accessed on 25 May 2014.

Sentry Tech: Long Distance Stationary Remote Controlled Weapon Station", in *Rafael Advanced Defense Systems Ltd.* http://www.rafael.co.il/marketing/SIP_STORAGE/FILES/4/1104.pdf last accessed on 25 May 2014.

Stop Killer Robots Campaign website. http://www.stopkillerrobots.org/ last accessed on 25 May 2014.

United Kingdom Ministry of Defense "The Joint Service Manual of the Law of Armed Conflict", 2004. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/27874/JSP3832004Edition.pdf last accessed on 25 May 2014.

United Kingdom Ministry of Defense "Joint Doctrine Note 2/11 The UK Approach to Unmanned Aircraft Systems", March 2011. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/33711/20110505JDN_211_UAS_v2U.pdf last accessed 25 May 2014.

United States of America Department of Defense Directive 3000.09 on subject "Autonomy in Weapon Systems", 21 November 2012. http://www.dtic.mil/whs/directives/corres/pdf/300009p.pdf last accessed on 25 May 2014.

United States of America Department of Defense Directive 5000.01 on subject "The Defense Acquisition System", 12 May 2003, Certified current as of 20 November 2007. http://www.dtic.mil/whs/directives/corres/pdf/500001p.pdf last accessed on 25 May 2014.

Weiner, Tim "'GI Robot' Rolls Toward the Battlefield" in *The New York Times*, 17 February 2005. http://www.nytimes.com/2005/02/16/technology/16iht-robot.html?pagewanted=all&_r=0 last accessed on 25 May 2014.

Treaties and Cases

Treaties

Convention on Cluster Munitions of 3 December 2008

Convention on the Prohibition of Use, Stockpiling, Production and Transfer of Anti-Personnel Mines and on their Destruction of 3 December 1997

Hague Convention (IV) respecting the Laws and Customs of War on Land and its annex: Regulations concerning the Laws and Customs of War on Land of 18 October 1907

Hague Convention (IX) concerning Bombardment by Naval Forces in Time of War of 18 October 1907

Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International and Non-International Armed Conflicts (Protocol I and II) of 8 June 1977

Saint Petersburg Declaration Renouncing the Use, in Time of War, of Explosive Projectiles Under 400 Grammes Weight of 11 December 1868

The Charter of the United Nations of 26 June 1945

The Geneva Conventions of 12 August 1949 I-IV

The Rome Statute of the International Criminal Court of 17 July 1998

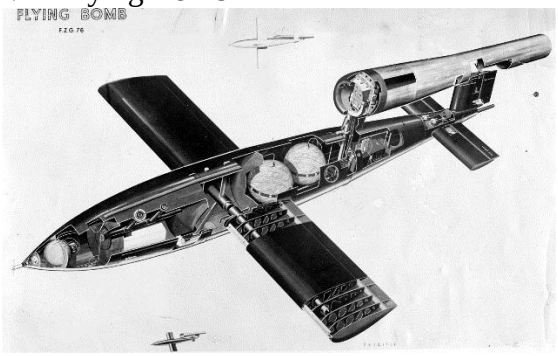
The United Nations Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to Be Excessively Injurious or to Have Indiscriminate Effects of 10 October 1980

Cases

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

Appendix – Images

V-1 Flying Bomb



Goliath UGV

MK-15 Phalanx CIWS

Iron Dome

SGR-1 Sentry Robot (image courtesy of www.hrw.org)

Predator Drone

Atlas Humanoid Robot