



Meaningful Human Control of Autonomous Weapon Systems

Definitions and Key Elements in the Light of
International Humanitarian Law and International Human Rights Law

AMANDA MUSCO EKLUND

Amanda Musco Eklund

Meaningful Human Control of Autonomous Weapon Systems

Definitions and Key Elements in the Light of International
Humanitarian Law and International Human Rights Law

Title	Meaningful Human Control of Autonomous Weapon Systems Definitions and Key Elements in the Light of International Humanitarian Law and International Human Rights Law
Titel	Meningsfull mänsklig kontroll av autonoma vapensystem Definitioner och huvudkomponenter enligt internationell humanitär rätt och internationella mänskliga rättigheter
Rapportnr/Report no	FOI-R-- 4928 --SE
Månad/Month	Februari/February
Utgivningsår/Year	2020
Antal sidor/Pages	59
ISSN	1650-1942
Kund/Customer	Försvarsmakten
Forskningsområde	Vapen, Skydd och säkerhet
FoT-område	FoT Samverkansprogrammet
Projektnr/Project no	E72876
Godkänd av/Approved by	Torgny Carlsson
Ansvarig avdelning	Försvars- och säkerhetssystem

Detta verk är skyddat enligt lagen (1960:729) om upphovsrätt till litterära och konstnärliga verk, vilket bl.a. innebär att citering är tillåten i enlighet med vad som anges i 22 § i nämnd lag. För att använda verket på ett sätt som inte medges direkt av svensk lag krävs särskild överenskommelse.

This work is protected by the Swedish Act on Copyright in Literary and Artistic Works (1960:729). Citation is permitted in accordance with article 22 in said act. Any form of use that goes beyond what is permitted by Swedish copyright law, requires the written permission of FOI.

Summary

There is widespread agreement amongst states and civil society that the concept “meaningful human control” is useful in discussions on autonomous weapon systems. The concept has been defined differently in the debate, but there are common factors which can be systemised into the following key elements of meaningful human control: *context-control, understanding the weapon system, understanding the environment, predictability and reliability, human supervision and the ability to intervene, accountability and ethical considerations.*

The discussions on autonomous weapon systems mainly focus on international humanitarian law aspects. However, a perspective on international human rights law is necessary when developing the concept of meaningful human control, for it to be useful also when human rights law is the governing legal framework. Analysing which requirements the right to life of the *European Convention on Human Rights* would place on a concept of meaningful human control reveals additional key elements such as *expressing meaningful human control through national regulations, necessity- and proportionality assessments and procedural obligations.* Since humanitarian law and human rights law differ regarding what use of force is lawful, an additional key element is the *ability to assess what legal framework governs the use of force.*

There is disagreement on how the precise meaning of the key elements should be interpreted in specific contexts. Terms like “predictability” and “understanding” do not have a definitive meaning and must be interpreted in each specific application. Central actors do not agree on several questions. Can weapon systems that autonomously select and engage targets be allowed? Must operators always be able to cancel an operation? A challenge for the future debate is to precise what meaningful human control means in relation to specific questions like these, to ensure that the concept is rooted in the operational reality and that it does not become contradictory in relation to existing non-controversial weapon systems. The conclusion of this report is that the concept of meaningful human control is still not sufficiently defined to form the basis of any regulation on autonomous weapon systems. However, the key elements are a useful starting point for continued discussions on what technology and use comply with humanitarian law and human rights law.

Keywords: meaningful human control, human judgement, autonomous weapon systems, LAWS, international humanitarian law, international human rights law, human-machine interaction, artificial intelligence, United Nations Convention on Certain Conventional Weapons, CCW.

Sammanfattning

Det råder bred enighet mellan stater och civilsamhälle om att begreppet "meningsfull mänsklig kontroll" är användbart i diskussioner om autonoma vapensystem. Begreppet har definierats på olika sätt i debatten, med det förekommer dock vissa gemensamma och återkommande krav som kan brytas ner till möjliga huvudkomponenter av begreppet meningsfull mänsklig kontroll: kontroll av sammanhanget, förståelse för vapensystemet, förståelse för sammanhanget vapnet används i, krav på förutsebarhet och tillförlitlighet, möjligheten att övervaka vapenanvändningen samt att ingripa, möjligheten till ansvarstillskrivelse och etiska överväganden.

Diskussionerna om autonoma vapensystem har huvudsakligen skett ur ett humanitärrättsligt perspektiv. Att även undersöka frågan utifrån mänskliga rättigheter är emellertid nödvändigt när begreppet meningsfull mänsklig kontroll utvecklas. Detta för att begreppet ska kunna användas även när de mänskliga rättigheterna utgör det ytterst gränssättande regelverket. Genom att analysera vilka krav som rätten till liv enligt Europakonventionen om skydd för de mänskliga rättigheterna skulle ställa på begreppet meningsfull mänsklig kontroll, uppenbaras ytterligare huvudkomponenter av begreppet. Exempelvis krav på att uttrycka meningsfull mänsklig kontroll genom nationella bestämmelser, nödvändighets- och proportionalitetsbedömningar, samt processuella skyldigheter. Eftersom humanitär rätt och mänskliga rättigheter skiljer sig åt gällande vilken våldsanvändning som är tillåten, kan ytterligare en huvudkomponent anses vara förmågan att avgöra vilket regelverk som är ytterst gränssättande för användandet av våld i den specifika situationen.

Det kvarstår många frågor relaterade till hur varje huvudkomponent ska tolkas i specifika situationer, då begrepp som "förutsebarhet" och "förståelse" är ord som saknar en bestämd betydelse och måste tolkas vid varje konkret tillämpning. Centrala aktörer inom debatten är oense om flera frågor. Kan vapensystem med förmåga att autonomt göra målval och bekämpa mål tillåtas? Måste operatörer alltid ha möjlighet att avbryta en operation? En utmaning för den framtida debatten är att precisera vad meningsfull mänsklig kontroll innebär i förhållande till specifika frågor som dessa, och att tillse att begreppet förankras i den operativa verkligheten samt inte blir motsägelsefullt i relation till existerande icke-kontroversiella vapensystem. Slutsatsen av denna rapport är att begreppet meningsfull mänsklig kontroll fortfarande inte är tillräckligt väldefinierat för att ligga till grund för en reglering av autonoma vapensystem. Begreppets huvudkomponenter utgör emellertid en användbar utgångspunkt för fortsatta diskussioner om vilken teknik och användning som är förenlig med humanitär rätt och mänskliga rättigheter.

Nyckelord: meningsfull mänsklig kontroll, autonoma vapensystem, LAWS, internationell humanitär rätt, internationella mänskliga rättigheter, människa-maskininteraktion, artificiell intelligens, FN:s konvention om vissa konventionella vapen, CCW.

Table of contents

1	Introduction	8
1.1	The emergence of the concept “meaningful human control”	8
1.2	Purpose and scope	10
1.3	Method and material	10
1.3.1	The challenge of considering many perspectives	10
1.3.2	Method and materials used to map and identify definitions	10
1.3.3	Method and materials used to identify key elements	11
1.3.4	Method and materials used to analyse key elements in light of human rights law	11
1.4	Disposition	12
2	Conflicting definitions of meaningful human control	13
2.1	Selection of actors and disposition	13
2.2	Definitions of civil society actors	13
2.2.1	The International Committee of the Red Cross	14
2.2.2	The International Panel on the Regulation of Autonomous Weapons	17
2.2.3	Article 36	19
2.2.4	The Campaign to Stop Killer Robots	21
2.3	Definitions of states and associations of states	23
2.3.1	United States	23
2.3.2	Russian Federation	24
2.3.3	Non-Aligned Movement	25
2.3.4	European Union	25
2.3.5	Sweden	26
2.4	Concluding comments on definitions	27
3	Key elements of meaningful human control	28
3.1	Overview table of recurring key elements	28
3.2	Context-control	29
3.3	Understanding the weapon system	30
3.4	Understanding the environment	30
3.5	Predictability and reliability	30
3.6	Human supervision and the ability to intervene	31
3.7	Accountability	31
3.8	Ethical considerations and the principle of human dignity	32
3.9	Differences and conflicts	32
3.10	Concluding comments on key elements	34
4	Meaningful human control according to the right to life	35
4.1	Why human rights law?	35
4.2	The interplay between humanitarian law and human rights law	36
4.3	The right to life of the European Convention on Human Rights	37

4.4	Key elements of meaningful human control according to the right to life	38
4.4.1	Expressing meaningful human control through national regulations	38
4.4.2	Absolute necessity assessments	39
4.4.3	Proportionality assessments	41
4.4.4	Procedural obligations	41
4.4.5	Human supervision and the ability to intervene	43
4.4.6	Context-control	43
4.4.7	Understanding the weapon system	44
4.4.8	Ability to assess what rules govern the use of force	45
4.5	Concluding comments on a human rights perspective	45
5	Final conclusions and reflections	47
5.1	Everyone wants control – But defining it is difficult	47
5.2	Emerging key elements are challenging to specify	48
5.3	A human rights perspective for a lasting concept of control	48
5.4	The future development of meaningful human control	49
	Bibliography	51
	International law instruments (treaties, statutes, etc.)	51
	European Union	51
	Case law	52
	European Court of Human Rights	52
	Judgements	52
	Decisions	53
	International Court of Justice	53
	Material from the CCW framework	53
	Reports	53
	Statements	54
	Working papers	54
	Interventions	55
	General exchanges	55
	Presentations	55
	Briefing papers	55
	Literature	56
	Articles published online and blogposts	57
	Institutional reports, policy papers, etc.	57
	Websites	58
	Other sources	59

Abbreviations

AI	Artificial intelligence
CCW	Convention on Certain Conventional Weapons
ECHR	European Convention on Human Rights
ECtHR	European Court of Human Rights
FOI	Swedish Defence Research Agency
GC	Grand Chamber (of the European Court of Human Rights)
ICJ	International Court of Justice
ICRC	International Committee of the Red Cross
iPRAW	International Panel on the Regulation of Autonomous Weapons
LAWS	Lethal Autonomous Weapons Systems
NGO	Non-governmental organisation
UN	United Nations

1 Introduction

This report is the master's degree project of the author and is written at the *Swedish Defence Research Agency (FOI)*. The writing of the report is supervised by Martin Hagström, Jessica Appelgren and Tam Beran from *FOI* and Per Bergling from *Umeå University*.

1.1 The emergence of the concept “meaningful human control”

As a response to the growing unease regarding autonomous weapon systems, the dystopian video “Slaughterbots” portrays a near-future scenario where swarms of cheap, weaponized drones using artificial intelligence (AI), shaped charge warheads and facial recognition are used to assassinate political opponents.¹ This fictional video, together with news on the development and use of increasingly autonomous weapons, raises several questions regarding the emerging technology's compliance with international humanitarian law and international human rights law (hereinafter humanitarian law and human rights law) and demonstrates the urgency of discussing legal aspects of autonomous weapon systems.²

The definition of autonomous weapon systems is debated and it is difficult to technically define what constitutes such a system.³ This report adopts a wide definition of autonomous weapon systems as a weapon system with a high level of autonomy, or many automated functions, which the system can perform independently without direct human involvement.⁴ Weapons with autonomous functions have been used for decades, and the list of existing weapons where human control is limited due to autonomous functions is long.⁵ But recently the unease regarding autonomous weapon systems has increased due to technological advancements in the field of AI, which cause worries over a potential decrease in human control over the use of force.⁶ The main fear is that this development would lead to “killer robots” making decisions on life or death.⁷

¹ Stop Autonomous Weapons [www.youtube.com/watch?v=9CO6M2HsoIA] 2019-12-03. The video was commissioned by a coalition of organisations and AI researchers, and released on YouTube by the Future of Life Institute.

² C.f. Hambling (2019 a); Hambling (2019 b); Tucker (2019), regarding news on the development, use and export of increasingly autonomous weapons by states such as the United States, Turkey and China.

³ See Brehm (2017), pp. 13–15; See more regarding the difficulties in defining autonomous weapon systems in chapter 2.2.2.

⁴ Cf. Hagström (ed.) (2016 a), pp. 4–5; Williams (2015); Heyns (2014), p. 2; A frequently used definition of autonomous weapon systems is the one of the United States Department of Defense: “A weapon system that, once activated, can select and engage targets without further intervention by a human operator.” United States Department of Defense (2012), pp. 13–14.

⁵ Regarding the historical use of weapons with autonomous functions, see Rantakokko (2019); Regarding existing weapons, these are e.g. active protection weapon systems, missiles and loitering weapons with autonomous modes or functions. For additional examples, see Group of Governmental Experts report (2019), annex III, § 4; See also Boulanin & Verbruggen (2017), chapter 3, regarding the state of autonomy in deployed weapon systems.

⁶ Cf. Amnesty (2015), p. 5, regarding worries about losing “effective human control” due to technical developments in the field of AI.

⁷ See, e.g., Campaign to Stop Killer Robots (2018), p. 1.

These fears are mainly addressed within the framework of the *United Nations (UN) Convention on Certain Conventional Weapons (CCW)*.⁸ The framework held informal *Meetings of Experts* since 2014 and established a *Group of Governmental Experts on Lethal Autonomous Weapons Systems (LAWS)* in 2017 to discuss questions related to emerging technologies in the area of LAWS.⁹ There are currently no norms of international law which specifically address autonomy in weapon systems.¹⁰ Within the *CCW* framework, state parties, non-governmental organisations (NGOs) and other actors are discussing if there is need for a new normative framework regulating autonomous weapon systems, their use and development beyond existing rules of international law.¹¹ So far, autonomous weapon systems have mainly been discussed from a humanitarian law perspective, although human rights law is also relevant for the questions raised by these systems.¹²

Within the discussion on autonomous weapon systems, the concept of “meaningful human control” has emerged as the key to defining where to draw the line between lawful and unlawful use of autonomous weapon systems. The concept has also been used as an expression of what requirements should be placed on technical systems, their use and the user to ensure compliance with international law.¹³ There is widespread agreement that the concept of meaningful human control is useful,¹⁴ even if the terminology used to refer to these human control aspects might differ.¹⁵ However, widespread agreement comes at the legislative cost that defining the concept has been more challenging.¹⁶

The terms “meaningful” and “human control” are relative, and not surprisingly different actors with different interest have different views on what the concept should entail.¹⁷ That is why reaching consensus in the *CCW* framework discussions is not as easy as deciding that the use of autonomous weapon systems is lawful as long as there is meaningful human control over the use of force. Instead, the real challenge is answering this question: what

⁸ *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 as amended on 21 December 2001 (CCW)*. The *CCW* only contains general provisions, whereas all prohibitions and restrictions regarding specific weapon systems are made in annexed protocols.

⁹ Regarding the background on LAWS in the *CCW*, see UNOG [www.unog.ch/80256EE600585943/(httpPages)/8FA3C2562A60FF81C1257CE600393DF6?OpenDocument] 2019-12-03.

¹⁰ See Brehm (2017), p. 8; Crootof (2015), p. 102.

¹¹ Possible normative frameworks and outcomes discussed are e.g. a legally binding instrument, political declaration, guidelines, principles, codes of conduct and improving implementation of existing legal requirements, including legal reviews of weapons. See Group of Governmental Experts report (2019), annex III, § 5.

¹² See, e.g. the guiding principles of the Group of Governmental Experts report (2019), annex IV, introduction paragraph & § c which state that the work of the Group of Governmental Experts should be guided by humanitarian law in particular; Cf. with Article 36 (2019 a), pp. 1–2, which criticises the guidelines in the report for not recognising human rights law; See more regarding the applicability of human rights law in chapter 4.1–2.

¹³ See, e.g., Campaign to Stop Killer Robots (2019 b), which proposes that a treaty on fully autonomous weapons should be based on a general obligation to maintain meaningful human control over the use of force.

¹⁴ See Crootof (2016), p. 53; Canellas & Haga (2015), p. 1; Ekelhof (2019), p. 343.

¹⁵ See, e.g., the Group of Governmental Experts report (2019), § 17(e) & annex IV, § c, where the terms “human judgement” and “human-machine interaction” are used for human control aspects; See list of different terms used during the Group of Governmental Experts meetings (such as “appropriate”, “substantive” and “effective human control”) in the Group of Governmental Experts report (2018), annex III, § 22.

¹⁶ See Ekelhof (2019), p. 343; Group of Governmental Experts report (2019), § 22(a), stating that although there is agreement on the importance of the human element, further work is needed to develop a shared understanding; The expression that widespread agreement comes at a “legislative cost” is inspired by Crootof (2016), pp. 53–54.

¹⁷ C.f. UNIDIR (2014), p. 3, regarding the terms “meaningful” and “human control” being inherently subjective concepts.

does “meaningful human control” mean? The answer to that question will depend on how states and the civil society will define the emerging concept in relation to existing law. This calls for an analysis of these actors’ definitions and a clarification of the concept from both a humanitarian law and human rights law perspective.

1.2 Purpose and scope

The definition of the central concept “meaningful human control” regarding the use of autonomous weapon systems is disputed. The purpose of this report is to clarify and problematise the concept and its future development in the light of humanitarian law and human rights law. To fulfil this purpose, the report:

1. maps and identifies how central actors define meaningful human control,
2. systematises and compares these definitions in order to identify key elements of meaningful human control,
3. problematises these key elements of meaningful human control and identifies additional ones in light of human rights law, with a focus on the right to life according to the *European Convention on Human Rights (ECHR)*.

The scope is delimited to primarily discussing military applications of autonomous weapon systems. However, the conclusions of the report are still of relevance for areas such as law enforcement, in particular chapter 4 that focuses on human rights law.

1.3 Method and material

1.3.1 The challenge of considering many perspectives

The question of autonomous weapon systems is not only a legal one. For an accurate legal analysis, it is necessary to consider other perspectives such as how operations are conducted and how the technology actually work. The writer’s competence is within the field of law, and the purpose of this report is to clarify and problematise meaningful human control from a legal perspective. Still, efforts are made to integrate an operational and technical perspective.

1.3.2 Method and materials used to map and identify definitions

The general purpose to clarify and problematise the concept of meaningful human control is achieved through three research questions. Each question is answered using a different method, which are all described below. To map and identify how central actors define meaningful human control, a selection of actors was made. This selection is described in each sub-chapter of chapter 2 and in relation to each actor. To map and identify these central actors’ definitions of meaningful human control, their material related to the concept is analysed by reading it with the objective to identify all aspects which relate to how autonomy in weapons should be controlled. These aspects are then summarised as the actor’s definition of meaningful human control.

The material used for this analysis are not traditional sources of international law (international conventions, customary law, general principles, judicial decisions and academic writings), since autonomy in weapons is not regulated specifically in any of these sources.¹⁸ Documents submitted to the *CCW Group of Governmental Experts* are the

¹⁸ Regarding the traditional sources of international law, see art. 38(1), *Statute of the International Court of Justice*; Shaw, pp. 51–95.

most relevant material since the *Group of Governmental Experts* is the central forum of discussion regarding meaningful human control where actors express their suggested definitions. The *Group of Governmental Experts* is a group of experts from each state, which do not have the authority to sign a potential instrument on autonomous weapon systems.¹⁹ Even if the material from these meetings is the best material currently available, it does not necessarily represent the definite and official positions of the respective states. Other material used to analyse the actors' definitions are reports, national directives and other published material in which the actors express their definitions of meaningful human control. In addition, articles of scholars, mainly with a legal, military or technical perspective, are analysed to problematise the definitions.

Since the debate constantly develops, the most recent documents are prioritised when analysing definitions and key elements. As the debate proceeds, definitions might become outdated and this report considers developments and published material up until November 2019.

1.3.3 Method and materials used to identify key elements

To identify key elements of meaningful human control, the results from the analysis on definitions are used. The report systematises and compares these results with the aim to identify common aspects. These common aspects are interpreted into key elements of meaningful human control, which represent frequent themes in the debate. To clarify the similarities of these themes, this chapter also identifies differences in the definitions and potential reasons for these differences. Since the base of this analysis are the conclusions from the analysis on definitions, the references to the sources are not repeated when systematising the results since they are already accounted for.

1.3.4 Method and materials used to analyse key elements in light of human rights law

A large part of the material used in the chapters that analyse definitions and identify key elements is related to the *Group of Governmental Experts*. Since the scope of *CCW* is limited to armed conflict, the material from the *Group of Governmental Experts* is focused on humanitarian law.²⁰ This material is not suitable to analyse human rights aspects, and therefore other material is used to analyse the research question which focus on human rights law.

Since the aim is to analyse key elements and identify additional ones in light of the right to life according to the *ECHR*, the case law of the *European Court of Human Rights* (*ECtHR*), as the highest interpreter of the Convention, is analysed. As opposed to previous chapters, this one mainly analyses existing law, which is done by analysing the legal source with the highest authority in the field of the Convention – the *ECtHR* case law.

Since autonomous weapon systems are an emerging technology there is not yet any case law regarding these systems in particular. Therefore, as necessary when analysing a novel issue, existing general requirements are analysed to provide guidance on what might be required of a concept of meaningful human control. The cases selected for analysis are those that, within the scope of the report, involve aspects relatable to the question of

¹⁹ The decision to draft and sign a potential instrument would be made at the *CCW* Meeting of High Contracting Parties.

²⁰ Regarding the scope of the *CCW*, see art. 1(1–2), *CCW*; Human rights law is not excluded from the Group of Governmental Experts discussions, but usually only referred to as “other international law”. Regarding the focus on humanitarian law, see, e.g., Group of Governmental Experts report (2019), annex IV, introduction paragraph & §§ c, h.

autonomy in the use of force and meaningful human control. Cases from the Grand Chamber (GC) are prioritised, together with the more recent case law due to the evolutionary method of the *ECtHR*. According to this method, the Convention must be interpreted in the light of present-day conditions since it is considered a living instrument.²¹

When analysing an emerging technology that has not yet been dealt with in case law, parallels to existing case law is necessary. However, there are risks of analysing autonomous weapon systems through case law regarding for example other weapons or human combatants and state agents, since the novel aspects of autonomous weapon systems might be misrepresented. While parallels to other non-autonomous weapons underestimate the independent traits of autonomous weapon systems, parallels to human combatants and state agents would overestimate it.²² However, these parallels are necessary since this is the only existing case law which clarifies what requirements a concept of meaningful human control must relate to before there is case law on autonomous weapon systems. Therefore, parallels to existing case law is made with caution to avoid: 1) undermining what is new and particular about autonomous weapon systems, 2) equating these systems with either its human user or other weapons, and 3) assuming that the *ECtHR* will consider previous case law identically applicable to autonomous weapon systems.

1.4 Disposition

Each of the three research questions are analysed in a separate chapter. Chapter 2 maps and identifies how central actors define meaningful human control. In chapter 3 the definitions identified in the previous chapter are systematised and compared to identify key elements of meaningful human control. Chapter 4 problematises these key elements and identifies additional ones in the light of human rights law, with a focus on the right to life according to the *ECHR*. Conclusions are integrated throughout the entire report, but final conclusions and reflections are presented in chapter 5.

²¹ Regarding the ECtHR's evolutive interpretation, see case of *Tyrer v. the United Kingdom*, 1978-04-25, ECtHR, § 31; Case of *Kress v. France* [GC], 2001-06-07, ECtHR, § 70; Schabas (2015), pp. 47–48.

²² See Crootof (2018), who argues that there are limitations to making analogies in relation to autonomous weapon systems. Crootof considers that analogies to e.g. other weapons, combatants, child soldiers or animal combatants are misrepresenting and fail to address the legal issues raised by autonomous weapon systems.

2 Conflicting definitions of meaningful human control

This chapter explains how a number of states and civil society actors have defined meaningful human control, which will lay the foundation of discussion for the analysis in subsequent sections. One of the guiding principles agreed upon by the *Group of Governmental Experts* is that “in determining the quality and extent of human-machine interaction, a range of factors should be considered including the operational context and the characteristics and capabilities of the weapons system as a whole” (emphasis added).²³ This principle is a step on the way towards defining meaningful human control, but it is still unspecific.²⁴ Therefore this chapter will analyse what these “range of factors” are.

2.1 Selection of actors and disposition

The selection of actors is guided by an ambition to represent the three main positions on autonomous weapon systems: 1) opposition to any regulation on grounds that existing law is sufficient, 2) support of the idea of some form of regulation, but not necessarily legally binding, and 3) advocates of a ban arguing the use of autonomous weapon systems would not be able to comply with international law.²⁵

The choice to discuss civil society actors before state actors motivates a few words on disposition. In international law, it is logical to discuss states first since they are the main legal subjects and protagonists of international law.²⁶ However, in diplomatic negotiations such as the *Group of Governmental Experts*, states tend not to be expressive on their positions for strategical and political reasons, and their expressed views on meaningful human control are not the most specific.²⁷ Therefore, for presentational reasons civil society actors, which provide the most thorough discussion on meaningful human control, are discussed first since their detailed statements serve as an integrated introduction to relevant concepts. Even if the definitions of states are positioned later in the disposition, they are valued higher as a source than the definitions of civil society actors since states positions on meaningful human control are decisive for the development of international law.

2.2 Definitions of civil society actors

The selection of civil society actors has been determined by a combination of factors: they should either have a central role in the debate on meaningful human control, represent an

²³ Group of Governmental Experts report (2019), annex IV, § c; The guiding principles drafted by the Group of Governmental Experts were endorsed by the High Contracting Parties of the *CCW* in November 2019. See the *CCW* High Contracting Parties report (2019), § 31 & annex III.

²⁴ C.f. Group of Governmental Experts Report (2019), § 22(b) which states as a conclusion that “[f]urther clarification is needed on the type and degree of human-machine interaction required, including elements of control and judgement [...]”

²⁵ Cf. Scharre (2018), p. 266, regarding the idea of three main positions on autonomous weapon systems.

²⁶ See Shaw (2017), pp. 156–157; Cassese (2005), pp. 3–4, 71–72; States being a legal subject of international law implies that they have full legal capacity, rights, powers and obligations; But see Cassese (2005), pp. 134–150 and Shaw (2017), pp. 204–209, regarding that the traditional notion of states as the only legal subject of international law is being challenged.

²⁷ This could be explained by the strong interrelation between law and politics in international law, which means the creation of new norms is basically a negotiation between sovereign states. C.f. Shaw (2017), pp. 8–10; See generally regarding the concept of “sovereignty” in Cassese (2005), pp. 49–53.

authority on the subject, or have made thorough attempts of defining meaningful human control. It is important to underline that some actors are more influential than others. For example, the *International Committee of the Red Cross (ICRC)* which is a large international NGO considered having international legal personality under international law in some situations, has more power to influence other actors or define the issues related to meaningful human control than the smaller NGO *Article 36*.²⁸

2.2.1 The International Committee of the Red Cross

The *ICRC*'s unique role and status as a guardian, monitor and promoter of humanitarian law makes its definition of meaningful human control highly relevant.²⁹ The most notable features of the *ICRC*'s definition is its emphasis on the importance of ensuring that humans are able to make "context-specific legal judgements", its focus on humanitarian law aspects and that it is among the more detailed definitions. For a better understanding of the *ICRC*'s definition of meaningful human control it is necessary to put it in the context of the organisation's view on what limits on autonomy that can be derived from humanitarian law.

A fundamental legal basis for any theory of human control according to the *ICRC* is that humans and not machines must comply with humanitarian law, which inevitably means combatants must have human control over the use of force to be able to make the "context-specific legal judgement" that humanitarian law demands in specific attacks. This demand would in turn limit the extent of autonomy allowed by humanitarian law. Examples of central humanitarian law rules that require context-specific legal judgement are the rules of distinction, proportionality and precaution.³⁰ According to the *ICRC* these rules require context-based assessments, based on knowledge of the environment and expected effects of the weapon, which must be made by humans. When humans make these context-based assessments they must be reasonably close in time to the attack. When forming a plan, based on assumptions, these assumptions must continue to be valid until the execution of the attack, which according to the *ICRC* is a requirement for maintaining a meaningful level of human control.³¹

In defining meaningful human control, the *ICRC* begins with identifying three different stages of human control and thereafter three key elements of the concept. The different stages of human control are: 1) the development stage (during development and testing), 2) the activation stage (regarding the decision to activate the weapon system) and 3) the operation stage (the operation of the weapon system during which it independently selects and attacks targets). Human control must be retained at all stages for humanitarian law compliance according to the *ICRC*.³² The three key elements of human control are: 1) human supervision and the ability to intervene and deactivate, 2) predictability and reliability and 3) operational constraints.³³

²⁸ The *ICRC* is able to enter into international treaties under international law with States and international organisations. Shaw (2017), p. 207; See also Cassese (2005), pp. 133–134.

²⁹ Regarding the *ICRC*'s unique role and status, see art. 5, *Statutes of the International Red Cross and Red Crescent Movement*; Sandoz (1998); Cassese (2005), pp. 133–134.

³⁰ *ICRC* (2019 a), pp. 1–3; Regarding the sources of these rules, see the rule of distinction in art. 48, 51(2), 52(2), *AP I to the Geneva Convention* & art. 13(2), *AP II to the Geneva Convention*; Rule of proportionality in art. 51(5)(b) & art. 57(2)(b), *AP I to the Geneva Convention*; Rule of precaution in art. 58(c), *AP I to the Geneva Convention*.

³¹ *ICRC* (2019 a). The *ICRC* also recognises using the terms "appropriate levels of human judgment", "substantive", "appropriate", or "effective" human control, instead of "meaningful" human control.

³² *Ibid.*

³³ *ICRC* (2019 b).

The first key element of “human supervision and the ability to intervene and deactivate” expresses human control through supervisory control, which according to *ICRC* equals having “human-on-the-loop” supervision.³⁴ This element requires that the operator has the ability to adapt to changing circumstances, which require “situational awareness”. To make the context-based legal judgement that humanitarian law requires, the operator must have “sufficient information and understanding about the operation of the weapon system, the environment of use, and the interaction of the two, over the given time period and geographical area.”³⁵

Another necessary requirement is a “physical and/or communication link that permits adjustment of the engagement criteria and the ability to cancel the attack, as well as sufficient time for such intervention.”³⁶ Regarding what constitutes sufficient time for meaningful supervisory control, the *ICRC* considers there must be sufficient time for the operator to select and approve one of several options proposed by the system, to override and take back control, or to deactivate the system, before the weapon fires at a target.³⁷ This requirement can be related to the humanitarian law rules mentioned above, such as the rule of precaution, which requires that the warring parties take all feasible precautions to protect civilians and civilian objects under their control against the dangers resulting from military operations.³⁸ This obligation might involve being able to cancel an attack due to changing circumstances. Practically this means the supervision must be of a character that the user can, where feasible, cancel, suspend or modify attacks, up until the execution of the attack.³⁹

The second key element of “predictability and reliability” addresses the weapon system design, the task it is used for, the nature of the environment of use and the interaction between system and environment.⁴⁰ The *ICRC* is specific about defining the difference between predictability, as the knowledge of the consequences of use, and reliability, as the likelihood of failure.⁴¹ To ensure humanitarian law compliance regarding this key element the *ICRC* prescribes a few guidelines. Generally, the more complex the environment and the task are, the need for human control increases while the tolerance for autonomy decreases. Those planning or deciding on an attack that involves the use of autonomous weapon systems “must understand its capabilities and limitations in the given circumstances, in order to determine whether it will perform lawfully in the given circumstances.” This includes situational awareness of the environment over time as discussed above regarding the key element on human supervision.⁴² As a logical

³⁴ *ICRC* (2019 c), § 4; *ICRC* (2019 d), p. 4; Different references to humans’ position in “the loop” are common when discussing human control over autonomous systems. As defined by Scharre (2015), p. 10, “human *in* the loop” refers to machines that perform a function for some period of time, then stop and wait for human input before continuing”. “Human *on* the loop” refers to machines that perform a function entirely on their own but have a human in a monitoring role, who can intervene if the machine fails or malfunctions. “Human *out* of the loop” refers to machines that perform a function entirely on their own and humans are unable to intervene.

³⁵ *ICRC* (2019 b).

³⁶ *Ibid.*

³⁷ *ICRC* (2019 d), p. 10.

³⁸ See the rule of precaution in art. 58(c), *AP I to the Geneva Convention*.

³⁹ *ICRC* (2019 b).

⁴⁰ *Ibid.*

⁴¹ Predictability is defined as “the ability to “say or estimate that a specified thing will happen in the future or will be a consequence of something”, in other words knowledge of how the weapon system will function in any given circumstances of use, including the effects that will result”. Reliability is defined as “the quality of being trustworthy or performing consistently well”, in other words how consistently the weapon system will function as intended, without failures (malfunctions) or unintended effects.” *ICRC* (2019 d), pp. 10–12.

⁴² *ICRC* (2019 b).

consequence of the above, an autonomous weapon system that would be unpredictable by design would according to the *ICRC* be unlawful by nature. In the view of the *ICRC*, an autonomous weapon system that can change or adapt its functioning after deployment (for example by setting new goals) would be beyond human control, inherently unpredictable and therefore unlawful since any previous assessments would become invalid after deployment.⁴³ On these grounds, the *ICRC* has fundamental concerns about autonomous weapon systems controlled by machine learning algorithms, due to their “black box” manner making it difficult to understand how such a system reaches its output, and problems connected to bias.⁴⁴

The third key element of “operational constraints” is connected to the element of “predictability and reliability” since, according to the *ICRC*, all autonomous weapon systems will display some level of unpredictability due to their interaction with the environment. Operational constraints on the operation of an autonomous weapon system would be a way to increase the predictability.⁴⁵ Factors that, according to the *ICRC*, could be constrained to increase human control are: the task the weapon is used for, the types of targets it attacks, the type of force and effects it employs, the operating environment, the timespan between the human decision to activate the weapon and the eventual use of force self-initiated by the autonomous weapon system and the mobility (meaning the scope of movement over an area).⁴⁶

These are the three key elements of human control according to the *ICRC*. Yet a careful analysis of the *ICRC*'s material on autonomous weapon systems reveals a possible fourth element of meaningful human control. This factor is how ethical considerations influence the concept of human control. This report has a legal perspective and does not analyse ethical aspects. However, law and ethics are inherently related, particularly within humanitarian law and human rights law whose purpose is to protect persons.⁴⁷ According to the *ICRC*, ethical considerations are clearly connected to the discussions on requirements of human control over autonomous weapon systems and compliance with international law. Meaningful human control from an ethical viewpoint would imply a degree of control that “preserves human agency and upholds moral responsibility in decisions to use force.”⁴⁸ In the view of the *ICRC*, ethical considerations might lead to the development of new rules of humanitarian law and weapon bans through the Martens Clause.⁴⁹ The Martens Clause could be described as an ethical back up clause which provides that, in cases not covered by existing law, civilians and combatants remain under the protection of the “principles of humanity” and “dictates of public conscience”.⁵⁰

⁴³ *Ibid.*; Cf. with Article 36 (2019 b), p. 7, regarding systems that could change its target-profile after activation.

⁴⁴ *ICRC* (2019 c), § 10; See generally on different forms of machine learning systems and risks connected to these in *ICRC* (2019 d), pp. 14–19.

⁴⁵ *ICRC* (2019 c), §§ 8, 14.

⁴⁶ *ICRC* (2019 b).

⁴⁷ See *ICRC* (2018), §§ 5–7; Regarding the relationship between ethics and law, see also Sparrow (2017).

⁴⁸ *ICRC* (2018), §§ 82–85.

⁴⁹ *Ibid.*, §§ 8–9.

⁵⁰ Art. 1(2), *AP I to the Geneva Convention*; See also the preamble of *AP II to the Geneva Convention*.

2.2.2 The International Panel on the Regulation of Autonomous Weapons

The *International Panel on the Regulation of Autonomous Weapons (iPRAW)* is an independent group of experts from different states and scientific backgrounds. Its concept of human control (iPRAW generally use the term “human control” and not “meaningful human control”) is relevant since the panel’s mission is to provide an independent source of information and consultation to the *CCW Group of Governmental Experts*. This is done by suggesting possible approaches to autonomy and human control.⁵¹ The most notable features of *iPRAW*’s definition of human control is how it includes a military operational perspective, problematises the relation between the term “autonomy” and human control and underlines the legal importance of ethical considerations and human dignity.

iPRAW is not keen on making a specific definition of autonomous weapon systems. Autonomy is a notoriously difficult concept to define, together with defining the difference between autonomy and automation.⁵² According to *iPRAW* a focus on the concept human control is more important than defining autonomy, which is illustrated by the following example. If two weapon systems possessing the same autonomous function (such as selecting and engaging targets without human intervention) would operate in different contexts, they would require different levels of situational understanding, and consequently different types of legal judgements and human control would be required, even if the autonomous function is the same. Therefore, to conceptually grasp the problem, only a focus on human control is constructive.⁵³

iPRAW makes a useful approach to clarify the relation between autonomy and human control, stating that machine autonomy is intertwined with the concept of human control which exists on the same conceptual scale – more machine autonomy equals less human control, and vice versa. Since it is not necessary to define the term autonomy it is more productive to focus on the human control-side of the scale.⁵⁴ This report assumes a similar approach. Since it is difficult to define autonomous weapon systems from a technical point of view, because the definition will depend on the specific situation and the system in question, meaningful human control is used as a conceptual framework in this report.⁵⁵

iPRAW has used the military targeting cycle as a foundation when analysing aspects of human control from an operational perspective. A targeting cycle illustrates what steps and what decisions are taken during military targeting. There are different versions, but six steps are common: find, fix, track, target, engage and assess. Since a targeting cycle does not only focus on the steps “target” and “engage” it illustrates that these two steps (particularly the decision to engage) are the final steps of a long process of previous decisions.⁵⁶

iPRAW’s definition of human control is divided into two analytical categories. The first category refers to two necessary components of human control: 1) situational understanding and 2) possibility of human intervention. The second category assumes that each of these two necessary components can be divided into two different modes of

⁵¹ iPRAW is funded by the German Federal Foreign Office, but states that its views and findings do not reflect the position of the German government nor any other government. iPRAW (2018 c), p. 3.

⁵² iPRAW (2018 a), p. 8; See Williams (2015); Regarding the difference between “autonomous” and “automatic”, see also Hagström (ed.) (2016 a), pp. 4–5; Scharre (2015).

⁵³ iPRAW (2019), p. 1, where the example of the two weapon systems is further developed.

⁵⁴ iPRAW (2018 a), p. 8.

⁵⁵ C.f. Hagström (2016 b), pp. 24–25, regarding the difficulties to make definitions from a technical perspective.

⁵⁶ See iPRAW (2018 c), pp. 8–9; Regarding targeting cycles, see generally Roorda (2015); Ekelhof (2019).

control 1) “control by design” and 2) “control in use”.⁵⁷ Below, the two components of “situational understanding” and “human intervention” are analysed from the aspect of both “control by design” and “control in use”.

According to *iPRAW*, “situational understanding” implies the “ability of the human to understand the situation and its context including the state of the weapon systems as well as the environment”, which is necessary for human control. As demonstrated above with the example of the two weapon systems operating in different contexts, situational understanding is a dynamic and context-dependent concept which may vary in quality and quantity depending on the weapon system in question. Like the *ICRC*, *iPRAW* regards predictability and reliability as essential elements to determine the level of human control necessary. Predictability in the use of autonomous weapon systems requires that the human operator has situational understanding in order to have effective control, “understanding” meaning the human operator must be aware of the state of the system and its environment.⁵⁸

“The human operator” does not necessarily imply one particular human operator or commander since the concept of human control can be expanded to distributed authority in the view of *iPRAW*.⁵⁹ The view that the human control can be distributed to the military organisation as such, aligns with a military operational perspective on human control. The human role in military decision-making and legal assessments is rarely limited to one human individual, such as an operator, having full situational awareness and complete human control over the use of force in an operation. Ekelhof argues that an analysis on human control focusing solely on the relationship between the operator and the weapon is too narrow, and that a more appropriate approach is to recognise the distributed nature of control in military decision-making.⁶⁰ In addition to “the operator” not necessarily being one person, “the weapon” is not limited to one single weapon system, but also includes systems of multiple units executing a mission.⁶¹

Situational understanding through “control by design” implies the ability to monitor information about environment and system. In general, “control by design” refers to technical control, such as the technical specifications of the system which could be hardware and software design allowing an operator to exercise control while operating a system.⁶² While the operator is not required to understand the system on a software level, the design of the system must enable the operator to understand why it produces a specific outcome. This requirement would prevent having systems of a black box manner, as discussed above in relation to the *ICRC*’s concerns regarding machine learning algorithms.⁶³

Situational understanding through “control in use” refers to the appropriate monitoring of the system and the operational environment. In general, “control in use” expresses the operational dimensions of control, which imply the “procedural requirements to maintain control over the systems during planning, tasking and operation.” To exemplify, this mode

⁵⁷ Cf. with the *ICRC*’s concept of “different stages of human control” discussed above.

⁵⁸ *iPRAW* (2018 a), pp. 12–13, 15; *iPRAW* (2019), p. 2.

⁵⁹ *iPRAW* (2019), p. 1.

⁶⁰ Ekelhof (2019), p. 347, who uses current conventional air operations as an example to argue that instead of asking if the operator has meaningful human control over the weapon, a more appropriate question is “how the military organization as such can or cannot ensure meaningful human control over important targeting decisions.”

⁶¹ *iPRAW* (2018 c), p. 14.

⁶² *iPRAW* (2018 c), p. 6; *iPRAW* (2018 a), pp. 14–15.

⁶³ *iPRAW* (2019), p. 2; See *iPRAW* (2018 a), p. 15.

of control is necessary when changing situations on the battlefield demand new legal assessments based on awareness regarding the environment.⁶⁴

According to *iPRAW*, the “possibility of human intervention” refers to the option to appropriately intervene if necessary. This means the human operator must have the possibility to override the system and manipulate its action at any point in time. Intervention through “control by design” requires modes of operation that allow human intervention and require them in specific steps of the targeting cycle, whereas intervention through “control in use” requires authority and accountability of human operators, teammates and commanders, which comply with humanitarian law.⁶⁵

Beyond the two analytical categories analysed above, there is an additional concept of importance for the *iPRAW* definition of human control. Like the *ICRC*, *iPRAW* underlines the legal importance of ethical considerations and human dignity when defining minimum requirements for human control. The panel emphasises the ethical considerations on human dignity and the legal principle of human dignity which is established in international law.⁶⁶ The legal principle of human dignity can be described as a fundamental core value within both humanitarian law and human rights law. The *Charter of the United Nations* and universal human rights treaties express human dignity as something inherent in every human person and as the foundation of equal and inalienable rights.⁶⁷ *iPRAW* considers that it would be difficult to exclude human dignity from discussions on autonomous weapon systems, and that respect for human dignity might constitute an additional legal requirement for a minimum level of human control.⁶⁸

2.2.3 Article 36

Article 36 is a non-profit organisation focused on reducing harm from weapons, and a founding member of the *Campaign to Stop Killer Robots* (described below in section 2.2.4).⁶⁹ Its definition of meaningful human control is relevant since *Article 36* was the first actor to influentially use the concept and it has been involved in its development.⁷⁰ The most unique features of *Article 36*'s definition of meaningful human control is that it focuses on *what* the control should be exercised over (individual attacks), includes accountability as a key element and the way in which the NGO has used the concept to sketch a conceptual framework on how to frame a legal instrument on autonomy in weapon systems.

As a basis for its concept of meaningful human control, *Article 36* considers it developed from two premises: 1) there is wide agreement that weapon systems operating with no human control at all would be unacceptable, and 2) only involving a human in “the loop” does not ensure meaningful human control. For example, human control is not meaningful

⁶⁴ *iPRAW* (2018 c), p. 6; *iPRAW* (2018 a), pp. 14–15.

⁶⁵ *iPRAW* (2018 a), pp. 15–16; *iPRAW* (2018 c), p. 6.

⁶⁶ See *ICRC* (2018) regarding the *ICRC*'s view on human dignity, which is mainly analysed from an ethical perspective.

⁶⁷ See, e.g., the preambles of the *Charter of the United Nations*, *Universal Declaration of Human Rights*, *International Covenant on Civil and Political Rights* and the *International Covenant on Economic, Social and Cultural Rights*.

⁶⁸ See *iPRAW* (2018 b), and particularly p. 19, where three minimum requirements for human dignity in the use of force are suggested: 1) recognize a human being as a human, 2) understand the value of life and the significance of its loss, and 3) reflect upon the reasons for taking life and reach a rational conclusion that killing is justified.

⁶⁹ *Article 36* [www.article36.org/about/] 2019-12-03.

⁷⁰ *Article 36* (2016 a), p. 1, coined the phrase “meaningful human control over individual attacks”, which was introduced through a briefing paper submitted to the 2014 *CCW* informal Meetings of Experts, see *Article 36* (2014).

if a human would press a fire-button whenever a light came on without having any other information on the context.⁷¹

The following central arguments are important for understanding *Article 36's* view on the relation between meaningful human control and humanitarian law. Firstly, legal obligations are addressed to humans and humans apply the law, not machines that only execute the function they were programmed to do.⁷² Secondly, regarding *what* the control should be exercised over, the NGO underlines that it must be exercised over each *individual attack* since existing humanitarian law applies in relation to individual attacks.⁷³

Article 36 has identified four key elements of meaningful human control, but underlines that instead of specifying the necessary or sufficient degree and form of control, the elements provide a framework from which to start defining a more precise normative understanding.⁷⁴ These four elements are related to *technology, information, human action* and *accountability*.

The *technology* must be predictable, reliable and transparent. Predictable meaning “it can be expected to respond in certain ways”. A reliable technology would not be “prone to failure and is designed to fail without causing outcomes that should be avoided”. And lastly, transparent means a user can understand how the technology works. For example, it must be designed in a way that makes it possible for the system to provide the user with information such as the reasoning and goals used when executing its function.⁷⁵

Regarding *information*, the user must have accurate information on the outcome sought, the technology and the context of use. Information on the outcome sought means understanding to what purpose the technology is being used to enable the user to evaluate the legality of a proposed attack and military objective. Information on the technology implies understanding the process that will be applied, such as what a weapon system would identify as a target. And finally, information on the context of use is essential to understand the environment where the technology will operate, for example to be aware of the possible presence of civilians. These three aspects of information (intent, technology and context) are necessary considering that the technology is a tool for translating the user intent into outcomes in a particular context.⁷⁶

The key element of *human action* requires the possibility of timely human judgment and action, and a potential for timely intervention. When having the information from the key element of information discussed above, humans must then apply their judgement, for example by deciding to activate an autonomous weapon system. The “timely” requirement is important for the information on the context not to lose its accuracy. According to *Article 36*, it must also be possible for a human to intervene by cancelling or suspending

⁷¹ Article 36 (2016 a), p. 2; Article 36 (2014), p. 2; Article 36 recognises that other terms than meaningful human control can express the same, or similar, concept, such as “significant”, “appropriate”, “proper”, or “necessary” “human judgement” or “human involvement”.

⁷² Article 36 (2016 a), pp. 2–3. The NGO mainly focus on humanitarian law when situating autonomous weapon systems in a legal framework, but underlines that humanitarian law is not the only applicable legal framework; Regarding accountability, see Article 36 (2019 a), p. 2, which criticises how some actors continuously slip into implying the attribution of human responsibility to machines.

⁷³ Article 36 (2014), p. 4; See generally regarding what constitutes “an attack”, and the NGO’s concerns on a possible expansion of the concept of an “attack” in Article 36 (2016 a), pp. 1–3.

⁷⁴ Article 36 (2016 a), p. 1.

⁷⁵ Article 36 (2016 a), p. 4; Article 36 (2016 b), p. 2.

⁷⁶ Article 36 (2016 a), p. 4; Article 36 (2016 b), p. 2; Regarding questions on what a weapon system identifies as a target (target profiles), see generally Article 36 (2019 b) where the NGO argues for prohibiting systems that could “change the conditions under which they will apply force, those where such conditions cannot effectively be understood, and those that would target people.”

an attack, which could be considered implicit in the rule of proportionality which prohibits attacks that cause incidental civilian harm excessive in relation to the concrete and direct military advantage anticipated.⁷⁷

The last key element of *accountability* would mainly have effect after the use of force, but according to *Article 36* the existence of an accountability framework to a certain standard can affect humans to adapt their behaviour to the standards of that framework. This element is connected to the key element of human action, since ensuring that humans can make judgements and intervene is necessary for accountability.⁷⁸

Beyond these four key elements, the NGO has also expressed, as a requirement for meaningful human control, that humans control the space, duration and time of operation, which is connected to the NGO's position that meaningful human control must be exercised over each individual attack. An additional approach suggested is to "apply certain obligations to control the actual conditions in that context of use" by controlling who or what is within the context of operation through active management efforts.⁷⁹ This element is similar to what the *ICRC* describes as "operational constraints" which later on in this report will be analysed as the key element of "context-control".

In November 2019, *Article 36* published a policy note which draws on its previous writings on meaningful human control and sketches a conceptual framework on how to frame a legal instrument on autonomy in weapon systems. The aim of the suggested regulatory structure is not to suggest specific content of an instrument but developing a shared conceptual structure for future discussions. The regulation would cover "systems where force is applied on the basis of sensor data, without human evaluation of that data, and without a human setting the time and place of that application of force" which would be regulated through a combination of prohibitions and positive obligations. When these are exemplified, the key elements of meaningful human control discussed above, such as requirements on the predictability of the technology, are recurring. Prohibitions should for example cover systems that are so complex that they cannot be meaningfully controlled and positive obligations should ensure understanding of the systems used.⁸⁰

2.2.4 The Campaign to Stop Killer Robots

The Campaign to Stop Killer Robots (hereinafter *the Campaign*) is a coalition of NGOs working to "ensure that adequate levels of human control are retained in the use of force by banning the development, production, and use of fully autonomous weapons",⁸¹ and advocates a legally-binding instrument prohibiting weapon systems that can select and engage targets without meaningful human control.⁸² *The Campaign* does not have the same authority as the *ICRC*, but instead its definition is relevant since it represents many NGOs, is very active in the discussion on autonomous weapon systems and has the capacity to influence public opinion, which in turn affects states. In general, *the Campaign's* definition of meaningful human control is similar to the definitions of the

⁷⁷ Article 36 (2016 a), pp. 3–4; What is "timely" may vary between different technologies and contexts, Article 36 (2016 b), p. 3; Regarding the rule of proportionality, see art. 51(5)(b) & art. 57(2)(b), *AP I to the Geneva Convention*.

⁷⁸ Article 36 (2016 b), p. 3.

⁷⁹ See Article 36 (2019 b), pp. 9–10.

⁸⁰ Article 36 (2019 c).

⁸¹ Campaign to Stop Killer Robots (2018), p. 1; The Campaign's steering committee is composed of NGOs such as Human Rights Watch, Amnesty International and Article 36. Campaign to Stop Killer Robots [www.stopkillerrobots.org/about/] 2019-12-03.

⁸² Campaign to Stop Killer Robots (2019 a), p. 1; Campaign to Stop Killer Robots (2019 b).

ICRC and *iPRAW*, but a unique feature is how *the Campaign* has divided the concept into three categories of components: decision-making, technological and operational components, which will be discussed below.

The Campaign's concerns regarding fully autonomous weapons are based on the view that they are incapable of complying with the fundamental principles of both humanitarian law and human rights law.⁸³ However, it could be considered misleading to argue for a ban on the assumption that *the weapons* are incapable of complying with the law. Humans, and not weapons, are obliged to comply with the law and make legal assessments and the rhetoric of *the Campaign* incorrectly implies accountability of weapons.⁸⁴ It would be more appropriate to argue that *the use* of autonomous weapon systems could not comply with the law.

The Campaign underlines that the decision to take human life should not be delegated to machines, meaning not allowing machines to determine whom to target.⁸⁵ Regarding expressions that machines “determine” or “decide”, these could be considered anthropomorphic, meaning attributing human traits to non-human entities. It has been discussed if an anthropomorphic discourse confuses the discussion on autonomous weapon systems by prescribing technology characteristics it currently does not possess.⁸⁶

In a paper from November 2019 *the Campaign* presents key elements of a treaty on fully autonomous weapons, to which the concept of meaningful human control would be central. It suggests three types of core obligations: a general obligation to maintain meaningful human control over the use of force and a combination of negative and positive obligations based on different components of meaningful human control. As mentioned above, *the Campaign* divides the concept of meaningful human control into three categories of components, which include similar key elements as those discussed by the *ICRC* and *Article 36*. *Decision-making components* include requirements on the operator’s understanding of the operational environment and how the system functions, including what the system might identify as a target and sufficient time for deliberation. *Technological components* include requirements on predictability and reliability, the ability of the system to relay relevant information to the operator and the ability of human intervention after activation. *Operational components* include temporal and spatial limitations of the system’s operation and the permissible types of targets (e.g. personnel or material). *The Campaign* underlines that these components are not exhaustive and further analysis in negotiations may determine which of these or other components should be codified as prerequisites for meaningful human control.⁸⁷ An example of when there is not meaningful human control, according to *the Campaign's* view, is when there is no human in the decision-making loop during the selection and engagement of a target and when human control is only applied during initial deployment.⁸⁸

⁸³ See Campaign to Stop Killer Robots (2019 a), § 1.

⁸⁴ C.f. Article 36 (2019 a), p. 2; See, e.g., Human Rights Watch (2019), p. 1, where Human Rights Watch uses a similar rhetoric which implies that it is the autonomous weapon system which must comply with rules of humanitarian law.

⁸⁵ Campaign to Stop Killer Robots (2018), p. 1.

⁸⁶ One of the suggested guiding principles from the Group of Governmental Experts states: “In crafting potential policy measures, emerging technologies in the area of lethal autonomous weapons systems should not be anthropomorphized”. Group of Governmental Experts report (2019), annex IV, § i; Regarding the risks of an anthropomorphic discourse, see also Williams (2015), p. 54; Roorda (2015), pp. 153–154; Hagström (ed.) (2016 a), p. 2.

⁸⁷ Campaign to Stop Killer Robots (2019 b).

⁸⁸ Campaign to Stop Killer Robots (2018), pp. 1–3.

2.3 Definitions of states and associations of states

The following section will consider states' and associations of states' positions on autonomous weapon systems and definitions of meaningful human control. It is not possible to analyse all states within the scope of this report.⁸⁹ Instead, the ambition is to include states with different military capabilities, with some geographical diversity and to represent the three main positions on autonomous weapon systems mentioned above. Formally, there is sovereign equality between all states, but naturally some states are more influential than others due to differing levels of military and economic power.⁹⁰ For example, states like the *United States* and *Russian Federation* have a large influence on the discussion on autonomous weapon systems. Due to the evident bond between law and politics which characterises international law, it must be kept in mind that the official position states chose to communicate in international forums is affected by many political and security-related considerations and might not be completely representative of a state's actual position.⁹¹

2.3.1 United States

Unique features of the *United States*' position on the concept "meaningful human control" is that the state instead uses terms such as "appropriate levels of human judgement", that it has made the most in-depth analysis of human control aspects out of all the state actors analysed and how it focuses on that a system must effectuate human intent. The *United States* considers attempts of setting new international standards on autonomous weapon systems as impractical due to the uncertainties related to fast technological development.⁹² The *United States* presents three main arguments: 1) existing humanitarian law is sufficient to govern the use of autonomous weapon systems, 2) internal procedures for weapon reviews are essential and 3) the development of autonomous weapon systems have humanitarian benefits.⁹³ The position of the *United States* on meaningful human control differs significantly from some of the definitions above, as the concept of "human control" is discarded and instead the *United States* mainly uses the terms "human-machine interaction" and "appropriate levels of human judgement". A focus on "human control" is considered obscuring, as opposed to focusing on compliance with fundamental international law principles. The *United States* seems to assume that these principles require focus on the specific military effects.⁹⁴

Unlike the civil society actors that are concerned with autonomy in functions such as targeting and engagement and that the decision to kill would be delegated to machines, the *United States* does not consider that autonomy in functions such as target-selection and engagement amounts to delegating decision-making from humans to machines. The *United States* allows autonomy in targeting since, in the state's view, there is no legal

⁸⁹ See the official website of the *CCW* framework discussions on LAWS where submissions from all states involved are available. UNOG
[[www.unog.ch/80256EE600585943/\(httpPages\)/8FA3C2562A60FF81C1257CE600393DF6?OpenDocument](http://www.unog.ch/80256EE600585943/(httpPages)/8FA3C2562A60FF81C1257CE600393DF6?OpenDocument)] 2019-12-03.

⁹⁰ See Cassese (2005), p. 71; Regarding sovereign equality, see art. 2(1), *Charter of the United Nations*; Cassese (2005), pp. 48–53.

⁹¹ Regarding the bond between law and politics in international law, see Shaw (2017), pp. 8–10.

⁹² *United States* (2018 a), § 45.

⁹³ *United States* (2019), § 2 (a–c); Regarding the benefits of developing autonomous weapon systems, see *United States* (2018 b).

⁹⁴ See *United States* (2018 a), §§ 42–53.

requirement that the weapon system itself must be programmed to make humanitarian law assessments.⁹⁵ In the view of the *United States*, these humanitarian law assessments could be made by a human before activating an autonomous weapon system that would autonomously select and engage targets, provided that the system operates with sufficient reliability.⁹⁶

The key to lawful use of autonomous weapon systems regarding the “human-machine interaction” is ensuring that the system effectuate human intent when using force. The concrete dimensions on how the *United States* ensures this and “appropriate levels of human judgment over the use of force” is expressed in a *United States Department of Defense* directive.⁹⁷ The directive contains elements similar to those described previously, such as limitations on the time frame,⁹⁸ supervision and intervention,⁹⁹ demands on reliability and predictability,¹⁰⁰ and the requirements on understanding of the weapon systems and its effects.¹⁰¹ In more recent recommendations on AI ethical principles for adoption by the *Department of Defense*, the *United States Defense Innovation Board* recommends that AI is responsible, equitable, traceable, reliable, and governable.¹⁰²

2.3.2 Russian Federation

The submissions of the *Russian Federation* (hereinafter *Russia*) to the *Group of Governmental Experts* have not been as specific as those of the *United States* regarding the state’s view on meaningful human control. Similar to the *United States*, *Russia* states that existing international law is applicable to autonomous weapon systems and does not need to be adapted, underlines the potential benefits of developing such systems and highlights the importance of national weapon reviews.¹⁰³ The Russian view on meaningful human control, communicated to the *Group of Governmental Experts*, has been characterised by ambiguity throughout the years. In 2017 *Russia* stated that the concept of meaningful human control “though poorly developed, is a potential element of consent [...] However, it should be recognised that it will be very difficult to develop criteria for the “meaningfulness” of such control without politicising this issue.”¹⁰⁴ One year later *Russia* stated that trying to develop universal parameters for concepts such as maintaining “significant” human control could “hardly give practical results” since it doubted whether

⁹⁵ United States (2018 a), §§ 12–13, 42–45; See also United States Department of Defense (2016), p. 354.

⁹⁶ United States (2019), § 8.

⁹⁷ United States Department of Defense (2012); The term “appropriate” is considered flexible and a signal that human judgement is context-dependent. See United States (2018 a), § 9.

⁹⁸ United States Department of Defense (2012), 4.a.(1)(b), which requires that autonomous weapon systems “[c]omplete engagements in a timeframe consistent with commander and operator intentions and, if unable to do so, terminate engagements or seek additional human operator input before continuing the engagement.”

⁹⁹ *Ibid.*, 4.a.(3)(c), which requires that the interface between people and machines for autonomous weapon systems shall “[p]rovide clear procedures for trained operators to activate and deactivate system functions.”

¹⁰⁰ *Ibid.*, 4.a.(1)(a), which requires that autonomous weapon systems “[f]unction as anticipated in realistic operational environments against adaptive adversaries.”

¹⁰¹ *Ibid.*, 4.a (3), which requires that the interface between people and machines for autonomous weapon systems shall “[b]e readily understandable to trained operators.”; See also *ibid.*, enclosure 3, 1.b.(4), which requires that “[a]dequate training, TTPs, and doctrine are available, periodically reviewed, and used by system operators and commanders to understand the functioning, capabilities, and limitations of the system’s autonomy in realistic operational conditions.”

¹⁰² United States Defense Innovation Board (2019), pp. 27–41. The United States Defense Innovation Board is one of several independent federal advisory committees advising the Secretary of Defense.

¹⁰³ Russia (2019), §§ 6, 8–10.

¹⁰⁴ Russia (2017), § 12.

“criteria to determine a due level of "significance" of human control over the machine could be developed”, while in the same working paper stating “[w]e do not doubt the necessity of maintaining human control over the machine.”¹⁰⁵ In 2019 *Russia* expressed human control as being an “important limiting factor” and even specified some key elements, such as the possibility for human intervention, while adding that the specifics of human control should remain at the discretion of States.¹⁰⁶ To summarise, *Russia* has expressed concerns regarding the inclusion of existing military systems in any definition of what should be restricted, and disagrees with any definition based on autonomy in targeting and engagement since the human would set these functions to the system and therefore maintain human control.¹⁰⁷

2.3.3 Non-Aligned Movement

The focus of the *Non-Aligned Movement*, which represent a large part of the world’s developing countries, is different from the *United States’* and *Russia’s*, perhaps since these states are not in the forefront of developing autonomous weapon systems themselves.¹⁰⁸ The *Non-Aligned Movement* supports a legally binding instrument prohibiting LAWS, and call upon States to declare a moratorium on further development and use of LAWS.¹⁰⁹ Regarding defining meaningful human control, the *Non-Aligned Movement* has not been very specific, but mentions that LAWS must “remain under the direct control and supervision of humans at all times” and that this is a core element that must be an integral part of any legally binding instrument.¹¹⁰

2.3.4 European Union

The *European Union*, which has already made regulations of its own regarding autonomy,¹¹¹ supports a focus on human control within the *Group of Governmental Experts* discussions and delivers some key elements it considers crucial to ensure “sufficient human supervision”. These elements include 1) understanding of the system and its interaction with the environment, 2) that commanders at least retain ultimate command and responsibility for the decision to deploy the system, 3) humans must make the decisions on use of lethal force, exert control of lethal weapons systems they use and remain accountable for those decisions. The second element might seem like the *European Union* agrees to autonomy in functions like targeting and engagement, but the *European Union* also states that it is not acceptable to select and engage human targets without “some form or level of human control”. It is not clear if this human control can be exercised before the system is launched. Like the *ICRC*, the *European Union* underlines human control during the entire life cycle of the autonomous weapon system, such as

¹⁰⁵ Russia (2018), §§ 9, 11.

¹⁰⁶ See Russia (2019), § 7.

¹⁰⁷ See Russia (2019), § 4; Russia (2018), §§ 9, 11.

¹⁰⁸ As an example, the Non-Aligned Movement wants to direct discussions within the Group of Governmental Experts towards the risks of an arms race of fully autonomous weapons, the technology gap amongst states and the impacts on international and regional peace and security. Non-Aligned Movement (2018), § 4(e-f).

¹⁰⁹ *Ibid.*, §§ 5, 8–9. The term LAWS is used instead of autonomous weapon systems only when necessary to accurately describe the position of an actor who uses that term to define meaningful human control.

¹¹⁰ *Ibid.*, § 9.

¹¹¹ See, e.g., § 71, *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)*; § 4, *European Parliament resolution on the use of armed drones (2014/2567(RSP))*.

during design and deployment, and like *iPRAW* the *European Union* recognises the distributed nature of human control in military targeting practices.¹¹²

2.3.5 Sweden

Even if *Sweden* is a part of the *European Union*, it has expressed statements within the *Group of Governmental Experts* from a national perspective. The Swedish Government's website briefly mentions that regarding autonomous weapon systems (meaning weapons with no human control), the government considers it unacceptable that decisions of life or death are delegated to machines and that *Sweden* will participate in future discussions on such weapons.¹¹³

However, the Swedish position is that a legally binding instrument regarding LAWS would be premature due to the difficulty of defining LAWS, since it is a basic legal requirement for an instrument banning something to have a definition of what is prohibited.¹¹⁴ Instead, *Sweden* supports discussing the development of a political declaration to underline areas of common understanding. *Sweden* questions the *ICRC*'s working definition of autonomous weapon systems which is based on "autonomy in critical functions" such as "selecting and attacking targets", since it would include existing weapon systems that have been used by many states for decades without raising humanitarian concerns.¹¹⁵ The Swedish position is that fully autonomous weapon systems do not exist today.¹¹⁶

Regarding meaningful human control,¹¹⁷ *Sweden* considers that the development of the concept is of key importance in order to characterise what systems would be deemed to be LAWS.¹¹⁸ As a starting point, *Sweden* has stated that it believes humans should never be "out of the loop" regarding decisions on the use of force, which connects to its statement never to delegate life or death decisions to machines.¹¹⁹ However, it is important to keep in mind that this position depends on *Sweden*'s view on fully autonomous weapons as a non-existing technology which is defined based on its lack of human control – a technology that anyhow is not in line with what weapon systems the military seek to develop in *Sweden*'s view.¹²⁰

It follows from these starting points that *Sweden* supports the principle of meaningful human control and from the early discussions within the *CCW* it has, like many of the actors above, underlined that the necessary level of human control is context-dependent.¹²¹

¹¹² See European Union (2019). Except from human control, the European Union also uses the terms "human oversight", "human judgement" and "sufficient human supervision".

¹¹³ Regeringskansliet [www.regeringen.se/regeringens-politik/utrikes--och-sakerhetspolitik/nedrustning-och-ickespridning/] 2019-12-03; See also Sweden (2015), p. 1.

¹¹⁴ Sweden (2018 a), pp. 1–2; Cf. with ongoing political debate in Sweden, where some ministers argue for an "effective ban on lethal autonomous weapon systems, which are incompatible with international law" (my translation), Lövin & Linde (2019).

¹¹⁵ Regarding the definition of the *ICRC*, *ICRC* (2019 e), p. 1; Sweden (2018 b), pp. 1–2; In general, there are many existing weapons that have "autonomy in critical functions", such as all guided missiles.

¹¹⁶ Sweden (2018 c), p. 1.

¹¹⁷ Sweden considers that the exact name of the concept is not important and exemplifies using other terms than meaningful human control, such as "appropriate human involvement" and "sufficient human control". *Ibid.*, p. 1.

¹¹⁸ Sweden (2018 b), p. 3.

¹¹⁹ Sweden (2014), p. 1.

¹²⁰ See Sweden (2018 d), p. 1.

¹²¹ See Sweden (2015), pp. 1–2.

As many other states, *Sweden* has not clearly expressed its view on the components of human control, but it has provided some views on the concept. *Sweden* considers the demands that LAWS should remain under control at all times problematic. Depending on one's definition of LAWS many non-controversial existing weapons which do not remain under human control at all times and where there are no ways of calling off an attack after the weapon is launched could be defined as LAWS. This position is connected to *Sweden's* scepticism of defining LAWS based on autonomy in targeting and engagement, since the Swedish position is that meaningful human control can be exercised before the launch of a weapon system, as it the case with existing non-controversial weapon systems. This position would allow for autonomy in targeting and engagement, without the possibility of human intervention, as is the case with existing weapon systems.¹²²

To summarise, the Swedish position supports the further discussion on meaningful human control and that future definitions of LAWS and meaningful human control do not become contradictory to the use of existing non-controversial weapon system and is rooted in the operational reality. *Sweden's* current position is in the middle of a moving spectrum, with states opposing any regulation on grounds that existing humanitarian law is sufficient at one end, and on the other end states and NGOs advocating a ban arguing the use of autonomous weapon systems would not be able to comply with international law.

2.4 Concluding comments on definitions

The definition of meaningful human control is disputed, which this chapter has illustrated. None of these definitions claim to contain the final answer to the question of what meaningful human control should entail, and a contributing factor is probably that there is no consensus of what technology is discussed. Autonomous weapon systems are the reason for the discussion on meaningful human control, but the lack of a common definition of the systems which are the subject of discussion certainly complicates reaching a unanimous idea of what meaningful human control should entail. Perhaps that is why some actors advocate defining autonomous weapon systems based on the lack of meaningful human control, and thereby avoiding defining the technology. Actors' different motives and interests certainly also contribute to the wide range of definitions.

Still, many useful attempts of suggesting what meaningful human control should entail are analysed above. Most of the actors do not consider their suggestions an answer to the question of meaningful human control, but rather entry points for further discussions amongst the actors. There are many common key elements, but also differences, between the different definitions. The next chapter will systematise and compare these definitions in order to identify key elements of meaningful human control.

¹²² See Sweden (2018 c).

3 Key elements of meaningful human control

This chapter aims to systematise and compare the definitions in order to identify key elements of meaningful human control and problematise these from a legal perspective. The following recurring key elements identified will be discussed: *context-control*, *understanding the weapon system*, *understanding the environment*, *predictability and reliability*, *human supervision and ability to intervene*, *accountability* and *ethical considerations*. Lastly, some differences and conflicts among the different key elements and definitions are discussed. Due to the focus on humanitarian law in the debate which this analysis is based on, these key elements are mainly focused on meaningful human control from a humanitarian law perspective.

3.1 Overview table of recurring key elements

The table systematises a selection of the most frequent key elements of meaningful human control and provides an overview of phrases used by the actors to describe these elements.¹²³

	ICRC	iPRAW	Article 36	United States
How to control the context (“context-control”)	Operational parameters and constraints		Controlling the space, duration, time and conditions in the context of use	
How to understand the weapon	Understand capabilities and limitations in given circumstances	Situational understanding of the state of the weapon systems	The user(s) should understand the technology and the process that will be applied	Understand the functioning, capabilities, and limitations of the system’s autonomy in realistic operational conditions
Understanding of the environment of use	Situational awareness	Situational understanding of the environment	The user(s) should understand the context within which the technology will function	The user(s) should be properly trained to understand the context and the technology
Requirements on predictability and reliability	Predictability (as knowledge of the consequences of use) and reliability (as the likelihood of failure)	Predictability and reliability determine the necessary level of human control	The technology is expected to respond in certain ways and not be prone to failure	Function as anticipated in realistic operational environments against adaptive adversaries

¹²³ All actors are not included in the table, which aims to provide examples of different aspects of the definitions. The table includes the definitions of the ICRC, iPRAW, Article 36 and the United States that were amongst the more detailed. The construction of the table is inspired from a table in Ekelhof (2019), p. 344. Empty boxes do not exclude that an actor made statements on the topic in other material than the one analysed.

	ICRC	iPRAW	Article 36	United States
Requirements on the human supervision and the ability to intervene	Human supervision and the ability to intervene and deactivate	The option to appropriately intervene if necessary	Timely human judgment and action, potential for timely intervention	Clear procedures for trained operators to activate and deactivate system functions
Views on ensuring accountability		Authority and accountability of human operators, teammates and commanders	Accountability framework to a certain standard	
Considerations on ethics and the principle of human dignity	Preserve human agency and uphold moral responsibility in decisions to use force	Human dignity as legal requirement for human control		

3.2 Context-control

Many actors underline the context-dependent nature of meaningful human control and how it is related to the question of the lawfulness of using autonomous weapon systems.¹²⁴ Both humanitarian law and human rights law require context-specific legal assessments (for example assessments of proportionality) before and during the use of force. Therefore, a key element of meaningful human control is controlling the context in which an autonomous weapon system will operate. As stated by the *ICRC*, all autonomous weapon systems will have some level of unpredictability since they interact with an unpredictable environment. Therefore, a key element of increasing human control is to limit the unpredictability of the environment by operational constraints. The limitations discussed by the actors are: limiting the task, permissible types of targets, type of force and effects, the mobility of the autonomous weapon system, its time frame and duration of operation, spatial limitations and controlling the actual conditions in the context through active management efforts.

To understand how context-control can be used to increase meaningful human control it might be helpful to draw parallels to how meaningful human control is exercised through context-control in the use of existing non-controversial weapon systems.¹²⁵ An example of such systems which are deployed today are “sensor fuzed submunitions”. These weapon systems operate with no human control after launch and the meaningful human control is exercised before launch through context-control. A projectile is delivered at a predefined target area where two submunitions are released. The submunitions will self-initiate the detonation if their sensors detect an object matching their predefined target-profile, otherwise the submunitions will self-destruct.¹²⁶ Practically, the target recognition and engagement are made by the submunitions with no human control. Instead the human control has been provided before launch through for example spatial limitations, programming the target-profile (limiting the type of target), weather observations, calculations of the projectile's trajectory and deciding the position of the target.

¹²⁴ ICRC discuss context-specific judgements, Sweden states that the necessary level of human control is context-dependent and iPRAW underlines that situational understanding is a dynamic and context-dependent concept.

¹²⁵ Cf. Article 36 (2014), pp. 3–4.

¹²⁶ An example of such sensor fuzed submunitions are the BONUS-munitions. BAE Systems [www.baesystems.com/en/product/155-bonus] 2019-12-03.

3.3 Understanding the weapon system

The two key elements “understanding the weapon system” and “understanding the environment” both build on the vague concept of “understanding”. Simply requiring having “understanding” over something raises the question of who should have this understanding –the operator or the military organisation as such? Many of the actors write about the operator’s understanding, but some (*iPRAW*, *European Union*) have also recognised the distributed nature of military decision-making, which means the understanding can also be divided within the organisation.¹²⁷

Regarding what “understanding the weapon system” implies, the actors have formulated their requirements in different ways. Examples of requirements on understanding are: understanding the systems capabilities and limitations in given circumstances (*ICRC* and the *United States*), the state of the weapon, why the weapon produced a specific outcome and having the ability to monitor it (*iPRAW*), what the weapon might identify as targets (*Article 36* and *the Campaign*) and the functioning, capabilities, and limitations of the weapon’s autonomy in realistic operational conditions (*United States*). Requirements on understanding the weapon system can also be connected to concerns regarding the transparency of autonomous weapon systems, for example regarding the black box manner of machine learning discussed above. The issue of transparency, which has been mentioned by some of the actors, could be addressed through this key element since it requires that the user is able to understand how the autonomous weapon system works.

3.4 Understanding the environment

Many of the actors have identified understanding the operational environment as an important key element for compliance with international law. Some have referred to this element as “situational awareness” or “situational understanding”. This key element is connected to others, such as “human supervision and intervention”, “predictability” and “understanding the weapon system” since understanding the environment is essential to understand how the weapon system will interact with it. The core of this key element is that humanitarian law and human rights law require context-based assessments, which in turn require understanding the context. For example, it is necessary to have information and understanding of the environment in order to detect and adapt to changing situations in the operational environment (such as civilians entering the area) which would require making new legal assessments.

3.5 Predictability and reliability

Predictability and reliability appear as a key element in most of the actors’ discussions on meaningful human control and is connected to many of the other key elements. It could even be considered expressing a more overarching “goal-element” which the other key elements strive to achieve. For example, by using “context-control” the predictability increase, by “understanding the weapon” the use of it will be more predictable and by “understanding the environment” the autonomous weapon system’s interaction with it will be more predictable. The *United States* considers predictability and reliability as the key to lawful use of autonomous weapon systems since it ensures that the system will “effectuate human intent”, which in other words means the system functions predictably and reliably

¹²⁷ Cf. Ekelhof (2019) who considers it more accurate to speak about how the military organisation, not the operator, exercise meaningful human control over important targeting decisions. She exemplifies with how limited understanding the operator of a F-16 aircraft might have in a conventional air operation.

as anticipated by humans.¹²⁸ Other actors mention that predictability and reliability is connected to the design phase, for example by not designing an autonomous weapon system that could change its own function after deployment, which could render its use inherently unpredictable.

3.6 Human supervision and the ability to intervene

This key element is related to the discussions of humans *in, on* and *off* the loop, whereas this key element would imply a “human on the loop” situation. This element can be divided into two main components: 1) the human supervision over the operation of the autonomous weapon systems and 2) the ability to intervene through for example modifying or cancelling an attack. The different actors have characterised the necessary link between the autonomous weapon system and human in different ways. Regarding the aspect of supervision, the main requirement is that the system can relay relevant information to the human operator. Regarding the aspect of intervention, actors have mentioned different actions that should be possible for a human to perform, such as the ability to: cancel, suspend and modify attacks, overriding the system and the possibility of partial or complete deactivation.

Questions that arise in relation to requirements on supervision and intervention are when in time humans must have this control over the system, and for how long they must maintain the possibility to intervene? Some actors, such as the *ICRC, Article 36* and *the Campaign*, use more relative terms such as “sufficient time” for intervention, “potential for timely intervention” or “sufficient time for deliberation”.¹²⁹ Other actors use more absolute terms requiring the possibility to intervene “at any point in time” (*iPRAW*) or that the weapon must be under direct control and supervision “at all times” (*Non-Aligned Movement*). There are problematic aspects of the stricter approach. As *Sweden* has stated, requirements to keep autonomous weapon systems under control at all times would imply that there is no meaningful human control over some existing non-controversial weapon systems where there are no ways of calling off an attack after the weapon is launched. An example of such a weapon system is the sensor fuzed submunitions discussed above, and there are many other examples of existing weapons that lack this level of human control after launch.¹³⁰ This aspect of the key element “human supervision and intervention” might make it difficult to reach agreement between the actors on its specific meaning.

3.7 Accountability

One of the guiding principles agreed upon at the 2019 *Group of Governmental Experts* was that “human responsibility for decisions on the use of weapons systems must be retained since accountability cannot be transferred to machines.”¹³¹ Some of the actors (*Article 36, iPRAW, European Union*) have expressed accountability as one of the key

¹²⁸ See United States Department of Defense (2012), 4.a.(1)(a).

¹²⁹ Cf. with the United States definition of a “human-supervised autonomous weapon system” as an autonomous weapon system “that is designed to provide human operators with the ability to intervene and terminate engagements [...] *before unacceptable levels of damage occur*” (emphasis added). United States Department of Defense (2012), p. 14.

¹³⁰ E.g. active protection weapon systems, missiles, loitering weapons, naval or land mines and “sentry” weapons with autonomous modes or functions. See Group of Governmental Experts report (2019), annex III, § 4.

¹³¹ Group of Governmental Experts report (2019), annex IV, § b; Regarding accountability, see also annex IV, § d & § 17(b-c).

elements of meaningful human control. The *European Union* underlines that humans must make the decisions on use of force, exert control of weapons they use and remain accountable for those decisions. It is not surprising that questions on accountability are connected to the question of human control. For example, command responsibility according to international criminal law requires that the commander exercised *effective command and control*.¹³² The reason for accountability being framed as a key element by some actors might be the result of worries regarding the possible difficulties in assigning responsibility to humans for unlawful acts caused by the use of autonomous weapon systems, also referred to as the “accountability gap”.¹³³ This report does not address questions of accountability independently, but mentions them to underline why some actors have chosen to include accountability as a key element of meaningful human control.

3.8 Ethical considerations and the principle of human dignity

This report has a legal perspective and therefore the discussions on ethics is limited to discussing how actors have expressed ethical considerations as a key element that might have legal implications. The *ICRC* has expressed that ethical values might drive the legal development of the concept of meaningful human control and mentioned the Martens Clause of humanitarian law as an example, even if the legal status of the Martens Clause is debated.¹³⁴

Both the *ICRC* and *iPRAW* also underline the relevance of human dignity. While the Martens Clause is relevant for humanitarian law, the legal principle of human dignity also has a strong foundation within human rights law, and as *iPRAW* stated: human dignity might constitute an additional *legal* requirement for a minimum level of human control according to both humanitarian law and human rights law.¹³⁵ In sum, ethical considerations and the principle on human dignity could be considered an additional key element of meaningful human control.

3.9 Differences and conflicts

When using the concept meaningful human control, the users of the concept might have different interpretations over *what* this meaningful human control should be exercised. *Article 36* has stressed that meaningful human control must be maintained *over each individual attack* since existing humanitarian law rules apply in relation to individual attacks.¹³⁶ But other possible interpretations could be exercising meaningful human control over the use of force in general, the weapon system itself, the effects of the attack

¹³² See art. 28, *Rome Statute of the International Criminal Court*.

¹³³ Regarding the “accountability gap”, see Human Rights Watch (2015); Campaign to Stop Killer Robots [www.stopkillerrobots.org/learn/] 2019-12-03; Scharre (2018), pp. 261–263.

¹³⁴ See ICRC (2018), § 9, which underlines that it is debated whether the Martens Clause is a “legally-binding yardstick against which the lawfulness of a weapon must be measured, or rather an ethical guideline”; See also Scharre (2018), pp. 263–266; Sparrow (2017).

¹³⁵ See *iPRAW* (2018 b), pp. 9–10, 19–20.

¹³⁶ It might be challenging in the particular situation to assess and limit what constitutes an “individual attack”. See ICRC (1987), p. 1453, § 4783, regarding the definition of an attack, which states an attack “is a technical term relating to a specific military operation limited in time and space”; See also art. 49, *AP I to the Geneva Convention*, which states that “[a]ttacks” means acts of violence against the adversary, whether in offence or in defence.”

or the critical functions of the autonomous weapon system.¹³⁷ All actors do not clarify what they consider should be the object of meaningful human control, and it would increase the quality of discussions on the concept and reduce misunderstanding if there was more clarity regarding the intended object of control.¹³⁸

Some of the central actors use a discourse which gives the impression that decisions on life or death would be delegated to machines.¹³⁹ An anthropomorphic discourse which implies that machines would make legal assessments and decisions using the same reasoning attributes as humans might confuse the discussion since it ascribes human abilities to non-human entities. As technology stands today, machines can only compute (make algorithmic calculations) based on functions programmed by humans. Therefore, the decision-making of machines is not comparable to that of humans since a machine's decision-making is based on algorithmic calculation results based on quantified descriptions (models) of the world.¹⁴⁰

The discourse of not delegating life or death decisions to machines is used by *the Campaign* to argue for a ban on fully autonomous weapons. *The Campaign* argues that these weapons would be incapable of complying with international law since a machine could not make the legal assessments required.¹⁴¹ This argument is based on the view that a “fully autonomous weapon” is defined based on its lack of human control and therefore the weapon would make the decision to kill – which would be unacceptable since the weapon would be incapable of performing legal assessments.

The *United States* represents an opposing position since it considers that autonomy in for example target-selection does not amount to delegating decision-making from humans to machines, and that there is no legal requirement that the weapon itself should be able to perform legal assessments.¹⁴² As a consequence, the *United States* allows autonomy in targeting and engagement since it considers that the legal assessments necessary can be made by humans before activating an autonomous weapon system.¹⁴³ This position differs from the view of *the Campaign*, which considers autonomy during selection and engagement of a target to be an example of when there is no meaningful human control since human control would only be applied during the initial deployment. These conflicting views illustrate one of the biggest differences between the actors' positions – the different views on the lawfulness of autonomy in targeting and engagement.

The Campaign opposes autonomy in selection and engagement of targets. Also, the *ICRC* belongs to the group of actors that are sceptical of weapon systems with autonomy in its “critical functions” (which are selection and attack of targets according to the *ICRC*), since the user would not know the exact timing, location or nature of the attack which would be initiated by the weapon due to environmental factors. *iPRAW* must also be considered

¹³⁷ See, e.g., Campaign to Stop Killer Robots (2019 b), p. 5, proposing a focus on control over the use of force rather than control over specific technology since this would remove the need to predict technological developments.

¹³⁸ See UNIDIR (2014), p. 2.

¹³⁹ See, e.g., Regeringskansliet [www.regeringen.se/regeringens-politik/utrikes--och-sakerhetspolitik/nedrustning-och-ickespridning/] 2019-12-03; Campaign to Stop Killer Robots (2018), p. 1.

¹⁴⁰ See Hagström (ed.) (2016 a), p. 2; Brehm (2017), p. 21.

¹⁴¹ See Campaign to Stop Killer Robots (2019 a), § 1; Campaign to Stop Killer Robots (2018), p. 1; C.f. with Roorda (2015), according to whom “the debate on whether robots will ever be capable of fulfilling the distinction and proportionality assessments is both speculative and irrelevant.”

¹⁴² See United States (2018 a), § 53; Cf. Roorda (2015), who considers that there is no requirement that a weapon do e.g. a proportionality assessment, and that the human must only have control over the effect of the use of force.

¹⁴³ See Roorda (2015), p. 166 who considers that weapons with autonomy in e.g. targeting without human input may be lawful through the human making the critical decision about when and how to employ the system given the conditions ruling at the time.

belonging to the group critical to autonomy in certain functions, since its version of the key element of “human supervision and intervention” requires that the human operator has the possibility to intervene *at any point in time*.

The *United States* belongs to an opposing group of actors which does not consider that autonomy in functions like targeting and engagement leads to the conclusion that life or death decisions have been delegated to machines. According to this position, the decision to use lethal force is made by humans at the activation stage before launch and not by the machine during targeting and engagement, meaning the weapon is a tool to achieve a military goal set by humans. This approach also considers the other steps of the “targeting cycle” as relevant when ensuring meaningful human control. Other actors which could be considered close, or belonging to this group, is *Russia* due to its disagreement with any definition of LAWS based on autonomy in targeting and engagement since these functions would be designed and preset to the system by the human and therefore human control would be maintained. *Sweden* has also questioned the working definitions of LAWS based on “autonomy in critical functions”, since it would include existing weapon systems. Examples of such existing weapon systems are the sensor fuzed submunitions mentioned above and various guided missiles.

3.10 Concluding comments on key elements

A precise definition of meaningful human control becomes complicated once the concept and its key elements are applied to specific applications, such as autonomy in certain functions perceived as critical. As the previous analysis of definitions and their key elements illustrates, it is not possible (or perhaps even constructive) at this stage of the development of the concept “meaningful human control” to draw any final conclusions on what meaningful human control *is* and the form of control it requires. What the analysis above has strived to achieve is clarifying the complexity of the concept, but also systematising some of the main key elements in order to clarify the concept as much as possible. In the following chapter this clarification will serve as a foundation for further discussion with the aim to problematise the key elements of meaningful human control and identify additional ones in light of human rights law.

4 Meaningful human control according to the right to life

4.1 Why human rights law?

Previous chapters are based on central actors' discussions on meaningful human control, which mainly have been held within the framework of the *CCW*. These discussions have mainly focused on humanitarian law aspects of meaningful human control since the scope of the *CCW* and its protocols is limited to armed conflict.¹⁴⁴ However, humanitarian law is not the only legal framework governing armed conflict or the use of autonomous weapon systems in general since human rights law also is applicable.¹⁴⁵ Analysing meaningful human control through the lens of human rights law is necessary since it applies at all times, while the application of humanitarian law depends on the existence of armed conflict in which humanitarian law has precedence over human rights law as *lex specialis*. Human rights law might be the governing legal framework in many situations. For example, during military operations in situations that cannot be classified as an armed conflict, in situations of occupation or armed conflict in which humanitarian law and human rights law often overlap in practice, or when an autonomous weapon system is deployed within law enforcement.¹⁴⁶

This chapter focuses on the right to life according to the *European Convention on Human Rights (ECHR)*. Even if other human rights are relevant to discuss in relation to autonomous weapon systems, the right to life is central to the debate on meaningful human control which focuses on the lethal aspects of autonomous weapon systems.¹⁴⁷ The right to life appears in all major human rights instruments, but a focus on the regional *ECHR* is relevant for many states.¹⁴⁸ Obviously for its member states who are bound by the Convention, but also for non-member states involved in coalition military operations with *ECHR* member states. Another reason to focus on the *ECHR* is the innovative and particularly detailed case law of the *ECtHR*, and because other human right bodies are inspired by it.¹⁴⁹ The *ECHR* case law can also be considered an expression of fundamental European constitutional concepts and theories, which are of particular interest for European states developing their position on autonomous weapon systems.

¹⁴⁴ See art. 1(1–2) *CCW*. Art. 1(2) *CCW* excludes application in “situations of internal disturbances and tensions, such as riots, isolated and sporadic acts of violence, and other acts of a similar nature, as not being armed conflicts.”

¹⁴⁵ See Droege (2007), pp. 310–324; Gaggioli Gasteyger & Kolb (2007), pp. 116–118; Brehm (2017); Crootof (2015), pp. 98–99, 107–109.

¹⁴⁶ See Droege (2007), p. 310, regarding the overlap between humanitarian law and human rights law in situations of occupation and non-international armed conflict.

¹⁴⁷ Regarding other relevant human rights, see, e.g. the right to respect for private life, to an effective remedy, not to be discriminated against, not to be subjected to inhuman or degrading treatment, to liberty and security of person and freedom of movement; See generally a brief discussion on some of these rights in Brehm (2017).

¹⁴⁸ See, e.g. the right to life in universal human rights instruments such as art. 3, *Universal Declaration of Human Rights*; art. 6, *International Covenant on Civil and Political Rights*; Regarding regional human rights instruments, see, e.g. art. 2, *ECHR*; art. 4, *American Convention on Human Rights*; art. 4, *African Charter on Human and Peoples' Rights*.

¹⁴⁹ See Gaggioli Gasteyger & Kolb (2007), pp. 115–116.

4.2 The interplay between humanitarian law and human rights law

Since autonomous weapon systems might be deployed in situations where their use will be governed by both humanitarian law and human rights law, it is necessary to clarify the interrelationship between the two legal frameworks, and to underline the importance of analysing meaningful human control from both perspectives.

Human rights law applies at all times, while the application of humanitarian law depends on the existence of armed conflict, meaning the two legal framework will apply concurrently in armed conflict.¹⁵⁰ In practice, they often complement and mutually reinforce each other.¹⁵¹ The relationship is considered one of specialised and general law, where humanitarian law is *lex specialis* in armed conflict and displaces human rights law which is *lex generalis* when the two frameworks are in conflict.¹⁵² The most influential interpretation of the parallel application of human rights law and humanitarian law originates from the *International Court of Justice (ICJ)*,¹⁵³ according to which the right to life applies also in hostilities, but the assessment of what is an arbitrary deprivation of life will be determined by humanitarian law since it is *lex specialis*.¹⁵⁴ This means that even when humanitarian law is *lex specialis*, human rights law is still applicable. However, when a human rights court applies human rights law the more specific humanitarian law is used for interpretation.¹⁵⁵ The interplay between the two frameworks can result in three situations: 1) some rights are exclusively regulated by humanitarian law, or 2) exclusively by human rights law, or 3) a combination of both.¹⁵⁶

As a human rights court the *ECtHR* is not mandated to apply humanitarian law,¹⁵⁷ but it can still interpret human rights in light of humanitarian law which it has done regarding the right to life.¹⁵⁸ And vice versa, in some situations the *ECtHR* has ruled on military

¹⁵⁰ The ECtHR has clarified through its case law that the right to life applies during armed conflict. See case of *Al-Skeini and Others v. the United Kingdom* [GC], 2011-07-07, ECtHR, § 164; Case of *Hassan v. the United Kingdom* [GC], 2014-09-16, ECtHR, §§ 35–57, 77; Case of *Güleç v. Turkey*, 1998-07-27, ECtHR, § 81; See also Schabas (2015), p. 153; Bantekas & Oette (2016), p. 650; Droege (2007), p. 336; Gaggioli Gasteyger & Kolb (2007), pp. 116–118.

¹⁵¹ See Droege (2007), pp. 336–337, 340–344; But see Gaggioli Gasteyger & Kolb (2007), p. 115, regarding arguments that the two legal frameworks contradict each other regarding the right to life.

¹⁵² See Droege (2007), pp. 337–339, 344–348; Bantekas & Oette (2016), pp. 657–659; But see Gaggioli Gasteyger & Kolb (2007), pp. 118–124, questioning the interpretation of the *lex specialis* relationship as a mere conflict of rules mechanism which would derogate an entire legal regime by the other, since it is not realistic an entire field of law is always more specific.

¹⁵³ The ICJ is an important source of international law since it is the principal judicial organ of the UN, which settles international legal disputes submitted by states and gives advisory opinions on issues referred to it by the UN. See Shaw, pp. 803–804, 847–848; The ECtHR refer to the ICJ in its case law regarding the relationship between humanitarian law and human rights law, see, e.g. *Hassan*, §§ 35–57, 77, 104.

¹⁵⁴ *Advisory Opinion on Legality of the Threat or Use of Nuclear Weapons*, 1996, ICJ, § 25, regarding the right to life in art. 4 of the *International Covenant on Civil and Political Rights*.

¹⁵⁵ See *Hassan*, § 104; Case of *Varnava and Others v. Turkey* [GC], 2009-09-18, ECtHR, § 185; Case of *Loizidou v. Turkey* [GC], 1996-12-18, ECtHR, § 43; See also Bantekas & Oette (2016), p. 659.

¹⁵⁶ *Advisory Opinion on the Legal Consequences of the Construction of a Wall in the Occupied Palestinian Territory*, 2004, ICJ, § 106.

¹⁵⁷ See Bantekas & Oette (2016), p. 649; Gaggioli Gasteyger & Kolb (2007), p. 124.

¹⁵⁸ Cf. art. 31(3)(c) of the *Vienna Convention on the Law of Treaties*, according to which any relevant rules of international law applicable in the relations between the parties may be considered when interpreting a treaty. The ECtHR has referred to this article in e.g. *Hassan*, § 102; *Loizidou*, § 43; See also *Varnava*, § 185.

operations “against a normal legal background”, meaning examining the military lethal force used with the same criteria of law enforcement in peacetime.¹⁵⁹

It is important to note that for the *ECHR* to apply extra-territorially the state must have jurisdiction through “effective control”,¹⁶⁰ either over an area, or over a person subject to a state agent’s authority and control.¹⁶¹

4.3 The right to life of the European Convention on Human Rights

The right to life is protected differently in human rights law and humanitarian law, and before analysing key elements of meaningful human control according to the right to life, a brief introduction of article 2 *ECHR* is in its place.¹⁶² Article 2 reads:

“1. Everyone’s right to life shall be protected by law. No one shall be deprived of his life intentionally save in the execution of a sentence of a court following his conviction of a crime for which this penalty is provided by law.

2. Deprivation of life shall not be regarded as inflicted in contravention of this article when it results from the use of force which is no more than absolutely necessary: (a) in defence of any person from unlawful violence; (b) in order to effect a lawful arrest or to prevent the escape of a person lawfully detained; (c) in action lawfully taken for the purpose of quelling a riot or insurrection.”

Unlike humanitarian law, derogations can be made to human rights. The right to life is not absolute since derogations are allowed “in respect of deaths resulting from lawful acts of war”, which means killing lawful under humanitarian law.¹⁶³ An interference in the right to life might also be lawful if it 1) has a legitimate aim, which consists of the exhaustively pre-listed situations in sub-paragraph 2 a–c,¹⁶⁴ 2) the use of force must be no more than

¹⁵⁹ See, e.g., case of *Isayeva v. Russia*, 2005-02-24, ECtHR, § 191; Case of *Khamzayev and Others v. Russia*, 2011-05-03, ECtHR, § 185; Case of *Kerimova and Others v. Russia*, 2011-05-03, ECtHR, § 253; In the case of *Isayeva*, §§ 133, 191, this was motivated by the fact that no martial law and no state of emergency had been declared, and no derogation had been made under art. 15 *ECHR*; See also Schabas (2015), pp. 155–156.

¹⁶⁰ See art. 1 *ECHR*, which regulates the applicability of the Convention. Regarding the “effective control” test, see Grabenwarter (2014), pp. 7–10; Droege (2007), p. 325; In relation to military operations under the aegis of the UN it is relevant to consider case law where the Court did not focus on whether the respondent States exercised extra-territorial jurisdiction, but whether the Court was competent to examine those States’ contribution to a UN peacekeeping mission which exercised the relevant control of the area, since the UN is a distinct legal entity and not a contracting party to the *ECHR*. See, e.g., case of *Behrami and Behrami v. France and Saramati v. France, Germany and Norway* [GC], 2007-05-02, ECtHR.

¹⁶¹ Regarding control over an area, see *Al-Skeini*, §§ 138–140; Case of *Loizidou v. Turkey (preliminary objections)* [GC], 1995-03-23, ECtHR, § 62; Case of *Banković and Others v. Belgium and Others* [GC], 2001-12-12, ECtHR, § 70; Regarding control over a person, see case of *Öcalan v. Turkey* [GC], 2005-05-12, ECtHR, § 91; Case of *Issa and Others v. Turkey*, 2004-11-16, ECtHR, § 71.

¹⁶² Regarding the difference between the protection of life in humanitarian law and human rights law, see Droege (2007), pp. 344–345; Brehm (2017), p. 26; Gaggioli Gasteyger & Kolb (2007), pp. 134–137.

¹⁶³ Derogation to the *ECHR*, and under which conditions that is permitted, is regulated in art. 15 *ECHR*; Regarding derogations to the right to life in respect lawful acts of war, see art. 15(2) *ECHR*; Bantekas & Oette (2016), p. 343; Regarding the term “lawful acts of war”, see Schabas (2015), pp. 601–602; Another exception to the right to life according to art. 2(1) is the death penalty, which in turn is limited by protocols no. 6 and 13 which abolish the death penalty in times of peace and war for the states party to them.

¹⁶⁴ See case of *McCann and Others v. the United Kingdom* [GC], 1995-09-27, ECtHR, § 149; Grabenwarter (2014), pp. 18–20; Since derogations are allowed for killing lawful under humanitarian law, these legitimate aims are expanded in armed conflict to also include legitimate targets according to humanitarian law. See Park (2018), p. 42.

absolutely necessary for the achievement of the legitimate aim,¹⁶⁵ and 3) the force must be strictly proportionate to the achievement of the aim.¹⁶⁶

4.4 Key elements of meaningful human control according to the right to life

4.4.1 Expressing meaningful human control through national regulations

The analysis of meaningful human control according to the right to life begins with the requirement to clearly regulate the use of autonomous weapon systems. The right to life contains two substantive obligations, and one of them is the obligation to protect the right to life by law.¹⁶⁷ This means the state must put in place a legal framework which defines the limited circumstances when the use of force is allowed. In order to comply with the inherent principle of strict proportionality in article 2, this framework must fulfil a series of requirements described in case law.¹⁶⁸ These requirements indicate what standards both the use of autonomous weapon systems and the regulation of them would need to relate to. To exemplify, in the case of firearms the national regulation must “make recourse to firearms dependent on a careful assessment of the situation”, for example by basing it on an “evaluation of the nature of the offence committed by the person in question and of the threat he or she posed.”¹⁶⁹ The regulation must also “secure a system of adequate and effective safeguards [...] against avoidable accident”.¹⁷⁰ Regarding weapons in general, the Court has underlined that it is of primary importance that national regulations exclude the use of weapons that carry “unwarranted consequences”.¹⁷¹

These requirements can be connected to the key elements analysed above which aim to ensure that humans can make context-based assessments and that the technology will function reliably and predictably. Living up to these requirements might pose a challenge for a state which intends to regulate the use of autonomous weapon systems, which due to their autonomous functions probably will require stricter standards on the regulation than when regulating the use of uncontroversial weapons like firearms. The national regulation will most likely be required to ensure that the use of autonomous weapon systems will comply with the requirements of for example “unwarranted effects” and “safeguards against avoidable accidents” from previous case law. Even if previous case law concerns

¹⁶⁵ See *McCann*, §§ 148; Case of *Isayeva, Yusupova and Bazayeva v. Russia*, 2005-02-24, ECtHR, § 169; Case of *Andronicou and Constantinou v. Cyprus*, 1997-10-09, ECtHR, § 171; Case of *Giuliani and Gaggio v. Italy* [GC], 2011-03-24, ECtHR, § 175.

¹⁶⁶ See *McCann*, § 149; *Isayeva, Yusupova and Bazayeva*, § 169; *Andronicou and Constantinou*, § 171; *Giuliani and Gaggio*, § 176.

¹⁶⁷ See the first sentence of art. 2(1) *ECHR*; See also case of *Boso v. Italy*, 2002-09-05, ECtHR, § 1 under the heading “The Law”; The first sentence of art. 2(1) also implies a positive obligation in certain circumstances to take preventive operational measures to protect the right to life. See case of *L.C.B. v. the United Kingdom*, 1998-06-09, ECtHR, § 36; Case of *Osman v. the United Kingdom* [GC], 1998-10-28, ECtHR, § 115; Case of *Keenan v. the United Kingdom*, 2001-04-03, ECtHR, § 89.

¹⁶⁸ See *Giuliani and Gaggio*, § 209; Case of *Makaratzis v. Greece* [GC], 2004-12-20, ECtHR, §§ 57–59.

¹⁶⁹ *Giuliani and Gaggio*, § 209; Case of *Nachova and Others v. Bulgaria* [GC], 2005-07-06, ECtHR, § 96; Regarding the definition of “firearms”, these could consist of e.g. handguns, rifles, shotguns and machine guns.

¹⁷⁰ *Giuliani and Gaggio*, § 209; *Makaratzis*, § 58.

¹⁷¹ See case of *Tagayeva and Others v. Russia*, 2017-04-13, ECtHR, § 595.

other kinds of weapons, such as firearms, it seems reasonable that the Court would not place less strict standards on autonomous weapon systems.

Clear regulations also serve the purpose of reducing the autonomy of action when using force. The Court has held that when regulations governing the use of force are unclear, the absence of instructions lead to law enforcement officials enjoying “a greater autonomy of action and [...] more opportunities to take unconsidered initiatives [...]”.¹⁷² Even if this case law is related to the autonomy of humans while using firearms, it illustrates that in the use of force there is little room for autonomy not subjected to regulations ensuring strict proportionality.

4.4.2 Absolute necessity assessments

The second main substantive obligation of the right to life is the prohibition of deprivation of life, delimited by a list of exceptions.¹⁷³ A requirement for any of those exceptions to apply is that the use of force must be no more than absolutely necessary for the achievement of the legitimate aims.¹⁷⁴ These aims are: to defend a person from unlawful violence, effect a lawful arrest, prevent the escape of a person lawfully detained and lawful acts taken for the purpose of quelling a riot or insurrection.¹⁷⁵ The assessment of if the use of force was absolutely necessary is usually the main focus of the Court’s right to life case law. Therefore, ensuring that a necessity assessment is made and that the use of force is absolutely necessary would probably constitute a key element of meaningful human control according to the right to life.

Streletz, Kessler and Krenz v. Germany is a case regarding the border-policing regime of East Germany (GDR) resulting in the killing of East Germans attempting to escape to West Germany.¹⁷⁶ The case illustrates the need to make necessity assessments in the light of automated use of force. The weapons used in this case, anti-personnel mines and automatic-fire systems, were not autonomous in the sense of the autonomous weapon systems discussed in this report.¹⁷⁷ However, due to their “automatic and indiscriminate effect”, together with the categorical nature of the border guards’ orders to “annihilate border violators [...] and protect the border at all costs”, the Court considered that the automated killing flagrantly infringed the fundamental rights of the constitution and violated the right to life.¹⁷⁸ The use of force was considered in no sense absolutely necessary, and in clear disregard of “the need to preserve human life”.¹⁷⁹ This case is not about the autonomy of the weapon technology itself, but the organisation of the operation as such and the absence of a necessity assessment when automating the killing. As regards to autonomous weapon systems, this case illustrates that there must be control over the individuated use of the system (in the sense of making, and complying with, necessity

¹⁷² See *Makaratzis*, § 70. The case concerned the use of firearms in a police chase.

¹⁷³ See *Boso*, § 1 under the heading “The Law”.

¹⁷⁴ The term “absolutely necessary” indicates “that a stricter and more compelling test of necessity must be employed than that normally applicable when determining whether State action is “necessary in a democratic society” under paragraphs 2 of Articles 8 and 11”, see *McCann*, § 149; *Case of McKerr v. the United Kingdom*, 2001-05-04, ECtHR, § 110; *Giuliani and Gaggio*, § 176.

¹⁷⁵ Art. 2(2)(a–c) *ECHR*. To defend a *person* from unlawful violence implies that defending an autonomous weapon system would not be a legitimate aim.

¹⁷⁶ *Case of Streletz, Kessler and Krenz v. Germany* [GC], 2001-03-22, ECtHR, §§ 13–18.

¹⁷⁷ The “automated fire system” consisted of fence-mounted, directional SM-70 fragmentation mines, triggered by trip-wire. Brehm (2017), p. 43.

¹⁷⁸ *Streletz, Kessler and Krenz*, § 73.

¹⁷⁹ *Ibid.*, §§ 96–97, 102.

assessments), because otherwise the use of lethal force will probably be considered having automated and indiscriminate effects which flagrantly would violate the right to life.¹⁸⁰

In the case of *McCann and Others v. the United Kingdom* British soldiers shot to death suspected Irish Republican Army (IRA) operatives they believed were about to commit a bombing, but after they were killed it was discovered that the suspects were unarmed and that there was no bomb.¹⁸¹ However, it was not the actions of the soldiers in themselves which gave rise to a violation of the right to life, but the control and organisation of the operation as a whole.¹⁸² The case illustrates that the planning stage of an operation is connected to whether the use of force was absolutely necessary, and therefore constitutes another relevant aspect of meaningful human control.¹⁸³ The Court took into consideration surrounding circumstances like “the planning and control of the actions”, and established that failure to “exercise *strict control* over any operations which may involve the use of lethal force” (emphasis added) can make up a violation.¹⁸⁴ Since the planning and control were insufficient the use of lethal force was not absolutely necessary since the state did not “minimise, to the greatest extent possible, recourse to lethal force.”¹⁸⁵

The requirements on the planning stage of operations are of particular importance for the discussion on meaningful human control. When using autonomous weapon systems there might be a timespan between the human decision to launch the weapon system and the eventual use of force initiated by the system. The requirement to plan and exercise “strict control” over operations possibly involving the use of lethal force would probably place stricter demands on the planning stage before launching an autonomous weapon system which could self-initiate the use of force, than when engaging state agents. The legal assessments made during the planning stage must ensure that the use of the weapon system will comply with requirements of necessity and strict proportionality even after launch. Like the “military targeting cycle”, focusing on the planning stage takes into consideration that the steps of “targeting” and “engagement” are preceded by a chain of decisions which the state remains responsible for.¹⁸⁶ This approach demonstrates that it is important for compliance with human rights law that meaningful human control entails ensuring control over the use of lethal force early in the process.

Perhaps this aspect will be even more important in relation to autonomous weapon systems than in cases such as *McCann* regarding the shooting by human agents. The reason why the actions of the soldiers did not, in themselves, give rise to a violation in *McCann* was the soldiers’ “honest belief which [was] perceived, for good reasons, to be valid at the time but subsequently [turned] out to be mistaken.”¹⁸⁷ Justifying an infringement based on a mistaken honest belief will probably not be accepted when an autonomous weapon system kills someone by mistake. The concept of an “honest belief”

¹⁸⁰ Cf. *Ibid.*, § 73.

¹⁸¹ *McCann*, §§ 93, 96, 198.

¹⁸² *Ibid.*, §§ 199–201. The insufficient planning consisted of not arresting the suspects at an earlier stage, not making sufficient allowances for the possibility their information might be erroneous and the automatic recourse to lethal force when the soldiers opened fire, see *ibid.*, § 213.

¹⁸³ See, e.g., other cases where the Court also considered that the right to life had been violated due to inadequate planning, *Isayeva, Yusupova and Bazayeva*, §§ 195–199; *Isayeva*, § 200. Both cases involved military operations where bombings from airplanes were executed in an indiscriminate manner in respect of the civilian population.

¹⁸⁴ *McCann*, §§ 150–151.

¹⁸⁵ *McCann*, § 194; See also case of *Huohvanainen v. Finland*, 2007-03-13, ECtHR, § 94; *Andronicou and Constantinou*, § 186; Case of *Bubbins v. the United Kingdom*, 2005-03-17, ECtHR, § 136.

¹⁸⁶ Cf. Bantekas & Oette (2016), pp. 343–344.

¹⁸⁷ *McCann*, § 200; See also *Andronicou and Constantinou*, § 192; *Bubbins*, § 138; *Huohvanainen*, § 96.

would be difficult to apply to a machine, unless the Court would consider whether the human operator or military organisation had an honest belief that the use of force was necessary. Such an argument would most likely not be accepted since this belief must be subjectively reasonable with regards to the circumstances *at the relevant time*.¹⁸⁸ This requirement will not be met in the case of autonomous weapon systems with a timespan between the human decision to launch the weapon system and the eventual use of force initiated by the system, unless there are possibilities for human supervision and intervention providing sufficient environmental understanding for an operator to form an honest and genuine belief valid at the relevant time.

4.4.3 Proportionality assessments

Beyond necessity, another required assessment is the one of proportionality. This balancing act requires that the force used must be strictly proportionate to the achievement of the legitimate aim set out in 2(2)(a–c).¹⁸⁹ A proportionality assessment is already complex since it requires balancing values that are incomparable, such as the value of life and military advantage. To successfully code this assessment would most likely be an even more complex task. However, the case law does not require that the weapon makes the proportionality assessment, which is the responsibility of humans using the weapons. Therefore, another aspect of meaningful human control is the need to ensure that humans are able to assess the proportionality of the use of force.¹⁹⁰ Particularly when it comes to new technology, the Court has emphasised that states that take on a pioneer role in the development of new technologies have a special responsibility to strike the right balance in their proportionality assessments.¹⁹¹

4.4.4 Procedural obligations

On top of its two substantive obligations, article 2 also contains a distinct, implicit procedural obligation to effectively investigate alleged breaches of the substantive obligations.¹⁹² Several different requirements on an investigation are described in the Court's case law, but this section will focus on those particularly relevant to the discussion on autonomous weapon systems. In sum, the procedural obligation of the right to life requires that the use of autonomous weapon systems is explainable, predicable and reliable.

¹⁸⁸ See case of *Armani Da Silva v. the United Kingdom* [GC], 2016-03-30, ECtHR, § 248.

¹⁸⁹ See *McCann*, § 149; *Isayeva*, § 173; *Isayeva, Yusupova and Bazayeva*, § 160; See also *Giuliani and Gaggio*, §§ 176, 209, regarding the principle of strict proportionality being inherent in art. 2. The principle of proportionality is not spelled out in the article but established in case law.

¹⁹⁰ Cf. case of *Kakoulli v. Turkey*, 2005-11-22, ECtHR, regarding the shooting of a Greek Cypriot in the buffer zone between northern and southern Cyprus. The Court underlined that even in situations of unrest soldiers do not have the right to open fire automatically on persons only because they seem suspicious, since they cannot take recourse to lethal force without making an assessment whether there is an imminent risk of death or serious harm. Without such an assessment the use of force is neither proportionate nor absolutely necessary. *Ibid.*, §§ 119–121.

¹⁹¹ See case of *S. and Marper v. UK* [GC], 2008-12-04, ECtHR, § 112, which does not concern the right to life, but the respect for private life in relation to retention of DNA information. Still, the case illustrates the Court's view that states pioneering within the development of new technologies have a special responsibility to strike the right balance between the advantages of the new technology and the rights at stake.

¹⁹² See *Armani Da Silva*, § 229; Case of *Ergi v. Turkey*, 1998-07-28, ECtHR, § 82; *Giuliani and Gaggio*, § 298; *McCann*, § 161; Regarding the procedural obligation being a distinct obligation which can make up an independent violation of the right to life, see case of *Silih v. Slovenia* [GC], 2009-04-09, ECtHR, §§ 153–154.

Human rights law contains stronger procedural safeguards for the individual than humanitarian law,¹⁹³ and the Court reviews the procedural obligation of the right to life in both times of peace and armed conflict.¹⁹⁴ The right to life does not only cover intentional killing, but also situations where the deprivation of life is an unintended outcome.¹⁹⁵ Therefore, the use of unpredictable or unreliable technology where killing would be an unintended outcome is also covered by article 2.

The concerns regarding the lack of transparency of autonomous weapon systems have been discussed above, particularly the black box manner of machine learning. When it comes to the procedural obligations of the right to life there are standards of an effective investigation of an alleged violation which would require explainable and transparent technology. An investigation of an alleged breach must firstly be *adequate* to be effective, meaning the investigation is capable of establishing the facts, determining if the force used was justified and identifying those responsible.¹⁹⁶ It is not an obligation to perform results, but of having the means necessary.¹⁹⁷ In order to comply with these requirements, meaningful human control must entail at least having the means to effectively investigate an alleged violation. This requirement might be challenging for the use of systems with machine learning functions since they might not be transparent or predictable enough to allow effectively investigating the legality of its use.¹⁹⁸ A particular challenge would arise when states use autonomous weapon systems in those situations where the Court previously has placed the burden of proof on the state to provide a satisfactory and convincing explanation for injuries or death, such as when the situation in question is wholly, or in large part, within the exclusive knowledge of the authorities.¹⁹⁹

Another aspect of the procedural obligation of the right to life is connected to the prohibition of discrimination (article 14 *ECHR*) and the duty to “take all reasonable steps to unmask any racist motive and to establish whether or not ethnic hatred or prejudice might have played a role in the events.”²⁰⁰ Even if this case law concerns racist motives of human state officials, it is worth noting that the Court in the future might need to handle possible bias in the functioning of an autonomous weapon system due to for example biased input data. Even if a machine would not have racist motives of its own it will be interesting to see if, and how, the Court will uphold the obligation to unmask racist motives in relation to an autonomous weapon system allegedly affected by bias.

¹⁹³ See Droege (2007), pp. 348–355; See, e.g., the right to an effective remedy (art. 13 *ECHR*) and the right to a fair trial (art. 6 *ECHR*).

¹⁹⁴ See Gaggioli Gasteyger & Kolb (2007), pp. 152–157.

¹⁹⁵ See *McKerr* § 110; *McCann* § 148.

¹⁹⁶ See case of *Ramsahai and Others v. the Netherlands* [GC], 2007-05-15, ECtHR, § 324; *Armani Da Silva*, § 233; *Giuliani and Gaggio*, § 301.

¹⁹⁷ See case of *Tahsin Acar v. Turkey* [GC], 2004-04-08, ECtHR, § 223; *Al-Skeini*, § 166; Case of *Jaloud v. the Netherlands* [GC], 2014-11-20, ECtHR, § 186.

¹⁹⁸ See Brehm (2017), p. 6, according to whom humans must be involved in algorithmic targeting processes in a manner which maintains a possibility to explain the reasoning underlying algorithmic decisions in concrete circumstances.

¹⁹⁹ See *McKerr*, § 109; Case of *Salman v. Turkey* [GC], 2000-06-27, ECtHR, § 100; Case of *Çakıcı v. Turkey* [GC], 1999-07-08, ECtHR, §§ 85–87; Case of *Timurtaş v. Turkey*, 2000-06-13, ECtHR, § 82.

²⁰⁰ *Nachova*, § 160; Case of *Angelova and Iliev v. Bulgaria*, 2007-07-26, ECtHR, § 115.

4.4.5 Human supervision and the ability to intervene

The possibility for human supervision and intervention has been discussed above as a key element of meaningful human control. The right to life also seems to require supervision of a certain standard according to the *ECtHR*'s case law. Most of the case law regarding the right to life implicitly presuppose a possibility for human supervision and intervention since it demands assessments of both necessity and proportionality which must be valid during the prevailing circumstances.²⁰¹ The case law does not specifically analyse the particular aspects of how to supervise and intervene with autonomous weapon systems. However, it does deliver basic guidelines on the level of situational awareness required of state officials, which reasonably must also be guaranteed in the use of autonomous weapon systems.

A case regarding Russian aerial bombings during the Chechen War which resulted in civilian deaths, clarifies that there must be sufficient systems of information transmissions to pilots which are capable of briefing them on the presence of civilians.²⁰² In this case the forward air controller was not at the scene and unable to see the area or make any evaluation of the targets, and neither the air controller nor the pilot were aware of the announcement of a humanitarian corridor that day.²⁰³ According to the Court, the operation was not planned and executed with the requisite care for the lives of the civilian population and therefore amounted to a violation of the right to life.²⁰⁴ The Court's view on the operation and interaction between the air controller and the pilot can be compared to *Article 36*'s example of a human pressing a fire-button whenever a light came on, which would not constitute meaningful human control.

The case law has not yet revealed what requirements the Court would place on the human supervision and intervention of autonomous weapon systems. However, the existing case law illustrates basic guidelines which require some form of human situational awareness in the use of force, and a certain standard of the communication in military operations. It is difficult to say what view the Court would have on the supervision of existing weapons with autonomous functions, such as sensor fuzed submunitions, and where it would draw the limit on future autonomous weapon systems. However, most likely the element of "context-control" discussed below will have an impact on that question.

4.4.6 Context-control

The key element of context-control is also relevant to discuss through the lens of the right to life. There is case law on the use of anti-personnel mines, a weapon which lack many of the autonomous characteristics of the autonomous weapon systems discussed, but still involves aspects of questionable human control once the mine has been activated. Can these aspects be mitigated through context-control according to the right to life?

In two cases regarding anti-personnel mines, *Albekov and Others v. Russia* and *Paşa and Erkan Erol v. Turkey*, the Court illustrates what is not sufficient context-control and clarifies that the state has a duty to control the environment around the mines to ensure protection from the risk of death and injury.²⁰⁵ In *Albekov* the failure to locate and deactivate mines, mark and seal off the mined area and warn the civilians affected

²⁰¹ Cf. Brehm (2017), p. 48, who considers "a requirement for human agents involved in the use of an AWS to remain constantly and actively (personally) engaged in every individual application of force" inherent in the right to life.

²⁰² See *Isayeva, Yusupova and Bazayeva*, §§ 187–189; See also Gaggioli Gasteyger & Kolb (2007), p. 142.

²⁰³ See *Isayeva, Yusupova and Bazayeva*, §§ 187–189; See also Park (2018), p. 45.

²⁰⁴ See *Isayeva, Yusupova and Bazayeva*, §§ 199–200.

²⁰⁵ See case of *Albekov and Others v. Russia*, 2008-10-09, ECtHR; Case of *Paşa and Erkan Erol v. Turkey*, 2006-12-12, ECtHR.

constituted a violation of the right to life.²⁰⁶ The case of *Paşa and Erkan Erol* concerned anti-personnel mines laid around a military area. The context-control consisting of two rows of barbed wire, warning signs and informing the villagers was not considered sufficient since the area in question was used as pasture land by villagers and frequented by children, and the safety measures used did not prevent children from entering the area.²⁰⁷ In another case, employing aviation bombs in a populated area without prior evacuation of civilians was considered “impossible to reconcile with the degree of caution expected”.²⁰⁸

Context-control does not solve the issue of ensuring that the use of autonomous weapon systems is lawful (for example absolutely necessary and proportional) which the element of “human supervision and intervention” addresses. Instead, the element of context-control places a guiding limit on which circumstances autonomous weapon systems may be deployed in. It also presents requirements on which safeguards are required to reduce the risk of unlawful targets being at risk. The basic requirements of context-control which can be discerned from existing case law of other weapons are focused on ensuring civilians do not enter an area where they might be at risk. The focus on keeping civilians out of risk is probably why existing active protection systems with autonomous functions are not as controversial as the idea of offensive autonomous weapons. Active protection systems are mainly used to prevent incoming missiles or projectiles from acquiring and destroying a target, and usually operate in an environment where the chance of civilian presence is low (for example the Phalanx system which is used to protect warships at sea). Their use is also mainly limited to non-personnel targets (such as incoming missiles) which reduces the risk to violate the right to life. Keeping these factors in mind, it is probable that the lawful applications of autonomous weapon systems will be strictly limited and oblige a state to employ sufficient safeguards when deploying a system in areas with civilian presence. Especially if the system in question could be considered having indiscriminate effects, such as the case of anti-personnel mines and aviation bombs.

4.4.7 Understanding the weapon system

The *ECtHR* case law sets requirements which can be usefully added to the analysis of the key element “understanding the weapon system” discussed above. Particularly relevant is the right to life case law regarding the training of state agents. This case law mainly concerns the use of firearms and does not consider the particular question of what understanding of the weapon should be required of humans using an autonomous weapon system. However, the case law on firearms generally illustrates that technical training is required and what understanding that training must result in.²⁰⁹

In *Nachova and Others v. Bulgaria*, where two unarmed fugitives were shot during an attempted arrest, the Court stated that “law enforcement agents must be trained to assess whether or not there is an absolute necessity to use firearms, not only on the basis of the letter of the relevant regulations, but also with due regard to the pre-eminence of respect for human life as a fundamental value”.²¹⁰ State officials should also receive effective

²⁰⁶ *Albekov*, § 90; See also Schabas (2015), p. 148. In *Albekov* the state denied having laid the mines, §§ 7–10, but according to the Court deciding who laid the mines was not necessary to adjudicate the issue since the government in any event was aware of the presence of mines near the village which triggered their positive obligations under art. 2 to protect the residents from the risks involved, § 85.

²⁰⁷ *Paşa and Erkan Erol*, §§ 33–38.

²⁰⁸ *Isayeva*, § 191; See also Schabas (2015), p. 157.

²⁰⁹ See case of *Sašo Gorgiev v. the former Yugoslav Republic of Macedonia*, 2012-04-19, ECtHR, § 51, according to which the necessary technical training must be given when equipping police forces with firearms.

²¹⁰ *Nachova*, § 97; See also *Kakoulli*, § 110.

training with the objective of complying with international standards for human rights and policing, according to the case of *Şimşek and Others v. Turkey*.²¹¹

Since the case law is about the use of firearms it is difficult to draw specific conclusions on what understanding of autonomous weapon systems the Court would require. However, a reasonable assumption is that the Court most likely would not set lower requirements on the training and understanding of autonomous weapon systems than on firearms. At least regarding the use of firearms, it is possible to conclude that in order to comply with the right to life 1) the operators must receive training, and 2) this training must be effective, which requires that operators are able to assess necessity, relevant regulations, compliance with human rights and policing standards and taking into account the respect for human life. In sum, the training must enable the operators to evaluate the legality of the use of the weapon, which in turn would require that the technology is predictable and reliable.

4.4.8 Ability to assess what rules govern the use of force

Since the lawfulness of the use of force sometimes will differ depending on if humanitarian law or human rights law is the governing legal framework, the final element of meaningful human control which will be analysed is the ability to assess what rules govern the use of force in the particular case. To exemplify, proportionality assessments are different in human rights law, where the use of force must be proportional to the legitimate aim to protect life, and in humanitarian law, where incidental loss of civilian life must be proportional to the military advantage anticipated.²¹² The differences between the two legal frameworks, regarding what use of force is lawful, make it essential to ensure that humans are able to assess what rules govern the use of force in the specific situation and have the ability to adapt operations accordingly.²¹³

4.5 Concluding comments on a human rights perspective

A human rights perspective is necessary when developing the concept of meaningful human control for the concept to be useful also when human rights law is the governing legal framework. Different aspects from a human rights perspective have been analysed to clarify what requirements existing case law would place on a concept of meaningful human control of autonomous weapon systems for compliance with the right to life. This has resulted in an in-depth analysis of the key elements from the previous chapter, such as: *human supervision and the ability to intervene, understanding the weapon system and context-control*. The analysis has also revealed additional key elements from a human rights perspective, such as: *expressing meaningful human control through national regulations, absolute necessity and proportionality assessments, procedural obligations and the ability to assess what rules govern the use of force*.

Since this analysis is about an emerging technology and a new concept in development, the analysis is based on existing case law which is not entirely comparable to the technology and concept analysed. Therefore, the conclusions of this chapter do not claim to exhaustively answer what is, will or should be meaningful human control according to the right to life of the *ECHR*. Instead the conclusions contribute with useful basic

²¹¹ See case of *Şimşek and Others v. Turkey*, 2005-07-26, ECtHR, § 109.

²¹² See Droege (2007), pp. 345–346; But see Gaggioli Gasteyger & Kolb (2007), pp. 137–138, who consider that the concrete operation of the principle of proportionality seems broadly equivalent in both humanitarian law and human rights law.

²¹³ See Brehm (2017), pp. 6, 39.

guidelines and aspects that should be considered in the continued process to form the concept of meaningful human control.

The requirements to assess absolute necessity and proportionality narrow the scope for lawful autonomy in targeting and engagement due to the need to individuate the use of force to the circumstances at the prevailing time.²¹⁴ However, lawful autonomy in these functions are still not excluded. As long as these requirements are met humans could possibly exercise meaningful human control through key elements such as context-control, human supervision and intervention and by ensuring control over an operation at the planning stage.

Even if it is not possible to foretell how the *ECtHR* will construct its own view on meaningful human control in the future, this view will most likely be guided by the Court's previous methods of interpretation. In line with previous case law, the right to life will most likely be interpreted to make its safeguards practical and effective.²¹⁵ The Court might develop its jurisprudence in any direction to maintain the *effet utile* of those safeguards in line with the principle of the Convention as a living instrument which must be interpreted in the light of present day conditions.²¹⁶ Most likely the *ECtHR* will strictly construe the provisions of article 2 when adjudicating on the future use of autonomous weapon systems and step carefully not to dilute the safeguards of the right to life due to its fundamental character.²¹⁷

After focusing on meaningful human control according to human rights law a relevant question is: can the concept have different meanings depending on what legal framework governs the specific situation? From a practical perspective the answer is yes. Since human rights law and humanitarian law place different requirements on the use of force, what constitutes meaningful human control according to each legal framework will vary. However, from a more general conceptual perspective the concept will still essentially express the same core idea of maintaining meaningful human control over the use of force no matter what legal framework governs the situation.

²¹⁴ See Brehm (2017), pp. 48, 69, who considers that the scope for autonomous targeting is extremely limited under human rights law.

²¹⁵ See *McCann*, § 146; Case of *Soering v. the United Kingdom* [plenary], 1989-07-07, ECtHR, § 87; *Loizidou (preliminary objections)*, § 72; See also Schabas (2015), pp. 49–50, regarding the principle of effectiveness.

²¹⁶ See *Tyrer*, § 31; *Kress*, § 70; See also Schabas (2015), pp. 47–48 regarding the dynamic and evolutive interpretation of the Convention.

²¹⁷ See, e.g. the Court's statement in *McCann*, § 147, regarding that art. 2 must be strictly construed due to its fundamental character and since it enshrines one of the basic values of a democratic society.

5 Final conclusions and reflections

5.1 Everyone wants control – But defining it is difficult

The analysis of how central actors define meaningful human control illustrates some of the most frequent themes and definitions in the debate on the concept. After having analysed the different definitions it becomes clear that defining meaningful human control is difficult and that more discussion is needed before the actors might reach a compromise on what the concept should entail.

In those discussions the definitions of states will be more influential than the definitions of civil society actors since states decide if, and how, they will regulate autonomous weapon systems. Most likely states with more military and economic power, such as the *United States*, *Russia* and *China* will have a strong influence on the discussion about autonomous weapon systems.²¹⁸ Out of these states the *United States* has been the most open to share its view on meaningful human control, while for example *Russia* has been more reserved in communicating its view on the concept. Both of these states have clearly expressed that they do not support any new regulation on autonomous weapon systems based on meaningful human control at this point, on grounds that existing international law is sufficient. These positions stand in clear contrast to the recent suggestions from *The Campaign* and *Article 36* of a binding regulation based on meaningful human control. Considering the current position of the *United States* and *Russia*, it seems unlikely that these states would support such a regulation.

The definitions of the civil society actors represent the most detailed of the debate. These actors influence the development of the concept meaningful human control through lobbying activities, participating in the discussion and by providing suggestions. For example, *Article 36* was the first actor to suggest using the term “meaningful human control”, the authority of the *ICRC* on humanitarian law might influence how the issue of autonomous weapon systems is framed and *the Campaign* influence public opinion through campaigning which in turn might affect the position of states.

One thing that all actors seem to have in common is that no one seem to consider it neither desirable nor lawful to use autonomous weapon systems without any human control at all. Therefore, there is widespread agreement that the aspect of human control is useful to discuss. Nevertheless, when it comes to defining aspects of meaningful human control which would limit autonomy in weapons use more specifically, consensus is still far from reached. The technology discussed is still in development and states are cautious to limit themselves and their defence capabilities. At the same time, they might be interested in limiting others.

Agreeing on what technology is being discussed is not easy and it complicates the discussion that the object of discussion cannot be clearly defined. The difficulties in defining the technology and the context-dependent nature of the legal restraints on the use of force, motivate focusing on the concept meaningful human control instead of defining the technical aspects of autonomy. Such a technical definition would not be constructive

²¹⁸ The position of China is not independently analysed in this report. Regarding the position of China, see generally China (2018); Li & Xie (2019), regarding a Chinese perspective on legal regulation of autonomous weapon systems under humanitarian law; Kania (2018), who claims China uses a “strategic ambiguity and shifting approach” in its approach towards LAWS.

since the concept of meaningful human control is context-dependent and varies depending on the specific application and system in question.

5.2 Emerging key elements are challenging to specify

Analysing the definitions of meaningful human control shows that the discussion is progressing and that there are certain key elements of meaningful human control emerging from the debate. The key elements all represent different ways to limit autonomy and increase human control. From a legal perspective, a common factor of all these key elements is that they strive to ensure that humans are able to make the legal assessments necessary according to existing humanitarian law and human right law and have the capacity to adapt operations accordingly.²¹⁹

The existing legal requirements on the use of force are not the main disputed issue. For example, most actors would agree that weapons cannot be completely unpredictable and unreliable. Therefore, it might be easy to agree on a general key element of the vague notion of predictability. Instead, the challenge of clarifying the key elements of meaningful human control lies in reaching agreement on their meaning in specific applications. This includes specifying how much control each key element should require in different contexts and how they relate to each other when looking at specific questions. For example, can meaningful human control be exercised before launch of an autonomous weapon system and thereby allow autonomy in targeting and engagement? Can context-control increase the predictability up to a level that the possibility for human supervision and intervention is not necessary after launch? Is this only possible in certain operational environments? When in time must there be human supervision and possibility for intervention to ensure meaningful human control? At all times? This report has illustrated that when it comes to the specifics, the actors do not have the same answers to these questions. However, the key elements are a useful conceptual starting point for further discussions.

5.3 A human rights perspective for a lasting concept of control

A human rights law perspective is necessary for a constructive discussion on meaningful human control since there are many situations where human rights law would be the primary legal framework from which to evaluate the lawfulness of the use of autonomous weapon systems. Therefore, the key elements of meaningful human control were analysed in the light of human rights law, with a focus on the right to life according to *ECHR*. This analysis identified additional key elements of meaningful human control which are different from the humanitarian law centred key elements originating from the definitions of the central actors. Considering these results, it is important to include a human rights perspective in future discussions developing the concept of meaningful human control for the concept to be useful also when human rights law is the governing legal framework

²¹⁹ Cf. Brehm (2017), p. 8, regarding that meaningful human control seems to entail ensuring humans have the opportunity and capacity to make the necessary legal assessments and act in a way required to comply with existing law.

5.4 The future development of meaningful human control

When considering the different positions and definitions of meaningful human control – is it realistic to believe that the discussions within the *CCW* framework will result in a unanimous definition of meaningful human control containing key elements similar to those systematised in this report? The question on if, and how, to regulate autonomous weapon systems and meaningful human control boils down to that wide consensus comes at the legislative cost of an unspecific result open to many interpretations.

A possible future regulation on autonomous weapon systems could result in an instrument somewhere on a scale between ambiguity and clarity depending on if the priority is either that many states accede to it or to write an ambitious instrument. For instance, one possibility is a legally binding protocol from the *CCW* framework that would need to be ambiguous and unclear in order for a high number of states to sign and ratify it.²²⁰ States currently opposing new regulation, such as the *United States* and *Russia*, would most likely not accede to any legally binding instrument unless it was ambiguous. If negotiations in the *CCW* framework would not result in any regulation, another possibility is that civil society actors such as *the Campaign* might initiate drafting a more ambitious instrument outside the *CCW* framework together with a smaller number of interested state parties.²²¹

The conclusion of this report is that even if both these possible forms of regulation have their advantages and disadvantages, any regulation would be premature at this point. This conclusion does not exclude that new regulation will not be necessary in the future since this report analyses the concept of meaningful human control, and not whether new regulation is required to regulate autonomy in weapon systems. However, the analysis of the concept meaningful human control in this report shows that even if the concept is developing, it is not sufficiently defined to form the basis of a regulation on autonomous weapon systems at this point.²²²

There are examples of concepts that are not precisely defined but still functional, such as “unnecessary suffering” and “indiscriminate effects” in humanitarian law.²²³ These concepts exemplify that it might be necessary for legal terms to have a degree of imprecision to be applicable in different contexts. However, these examples are general principles of humanitarian law that are not restricted to the use of particular weapon systems, as opposed to the case of “meaningful human control” where the ambition is to use the concept to restrict the use of a particular category of weapon systems. Autonomous weapon systems are an emerging technology that has not been technically defined and according to some does not even yet exist. Therefore, arguing that meaningful human control might not need to be precisely defined on grounds that other functional legal concepts are not would overlook the complexity of regulating the use of particular weapon

²²⁰ It should be noted that such a protocol would not apply to situations outside armed conflict due to the limitations in the scope of the *CCW* in art. 1(2) *CCW*.

²²¹ Cf. Campaign to Stop Killer Robots (2019 c), in which the Campaign “[...] urges states to launch negotiations here at the *CCW* or elsewhere on a legal-binding instrument [...]”.

²²² But see Crotoof (2016), pp. 54–55, 58–59, who argues that there might be benefits in imprecision when regulating new technology through international consensus. Indefiniteness within international law has the strength that in a consensus-based system it is easier to get states to initially agree on progressive but vague principles which with time will be developed and specified.

²²³ See, e.g. the concept of “unnecessary suffering” in art. 35(2), *AP I to the Geneva Convention*; Regarding “indiscriminate effects”, see, e.g. the prohibition on indiscriminate attacks in art. 51(4), *AP I to the Geneva Convention*.

systems, which have not been technically defined, with a concept that is imprecisely defined.²²⁴

A vague regulation rushing ahead of a shared agreement on meaningful human control would risk having little effect. Most key elements of meaningful human control analysed in this report are characterised by the use of unspecific terms, for example “predictability” and “understanding”. Even the words of the concept “meaningful human control” are relative. If these key elements would be codified, for example by stating “humans must *understand* the weapon system”, the function of such a regulation as a restriction of international law on states’ behaviour could be questioned due to the unspecific meaning of the terms, which would be open for interpretation. The terms could certainly be interpreted and specified in for example case law, but states might not consider such developments legally binding if the states did not intend for the terms to have such a specific meaning when they acceded to the regulation in question. There is no unified system of sanctions in international law, and states will follow norms they perceive themselves obliged to obey.²²⁵ Another aspect is the questionable legality of basing a ban or prohibition on vague definitions that do not clarify what is prohibited. Undefined terms can easily be applied in an arbitrary manner, or simply lack effectiveness.²²⁶

It can also be discussed if it is desirable to develop an instrument with more specific definitions of meaningful human control, but with fewer state parties. The instrument would lack universality and risk undermining the universality of humanitarian law and human rights law. Explicitly banning a weapon, without great support, leaves opportunity for states that do not ratify the new instrument to still use the weapon in question (within the limits of existing international law). Instead, if the ground for the unlawfulness of the weapon originates from existing law as a living document all parties must relate to that – including those who would not ratify a ban – which would support the universality of humanitarian law and human rights law.

In conclusion, there might be widespread agreement that the concept of meaningful human control is useful when discussing autonomous weapon systems. However, the distance between different positions on the concept, when it comes to specific applications, implies that it will be challenging to further precise the concept, especially in a legally binding instrument. The overarching purpose of this report was to clarify and problematise the concept of meaningful human control and its future development. The ambition is that the conclusions made will be useful in future discussions on defining meaningful human control, its key elements and how these should be specified. Moreover, returning to the dystopian opening of this report, perhaps its conclusions might also be useful in addressing some of the concerns raised by the “Slaughterbots” video.

²²⁴ But see UNIDIR (2014), p. 4, which suggests that it is perhaps not necessary to precisely define “meaningful human control” since many other widely shared concepts are not defined in themselves.

²²⁵ C.f. Shaw (2017), pp. 3–4.

²²⁶ Cf. Article 36 (2016 a), p. 2, which states that without a normative requirement on human control “the legal framework itself is open to divergent and progressively broader interpretations that may render human legal application meaningless.”

Bibliography

All sources with an URL were accessed 2019-12-03.

International law instruments (treaties, statues, etc.)

Charter of the United Nations, 1945-10-24, 1 UNTS XVI.

Statute of the International Court of Justice, 1946-04-18, United Nations.

Universal Declaration of Human Rights, 1948-12-10, 217 A (III).

European Convention for the Protection of Human Rights and Fundamental Freedoms, 1950-11-04.

International Covenant on Civil and Political Rights, 1966-12-16, United Nations, Treaty Series, vol. 999, p. 171.

International Covenant on Economic, Social and Cultural Rights, 1966-12-16, United Nations, Treaty Series, vol. 993, p. 3.

Vienna Convention on the Law of Treaties, 1969-05-23, United Nations, Treaty Series, vol. 1155, p. 331.

American Convention on Human Rights, 1969-11-22.

Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts (AP I), 1977-06-08, 1125 UNTS 3.

Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of Non-International Armed Conflicts (AP II), 1977-06-08, 1125 UNTS 609.

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 as amended on 21 December 2001, 1980-10-10, United Nations, Treaty Series, vol. 1342, p. 137.

African Charter on Human and Peoples' Rights, 1981-06-27.

Rome Statute of the International Criminal Court, 1998-07-17, United Nations, Treaty Series, vol. 2187, No. 38544.

Statutes of the International Red Cross and Red Crescent Movement, in ICRC, *Handbook of the International Red Cross and Red Crescent Movement*, 14th ed., 2008, available at [www.icrc.org/en/publication/0962-handbook-international-red-cross-and-red-crescent-movement].

European Union

European Parliament resolution on the use of armed drones (2014/2567(RSP)), 2014.

Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), OJ L 119, 4.5.2016, pp. 1–88.

Case law

European Court of Human Rights

Judgements

Albekov and Others v. Russia (68216/01), 2008-10-09.

Al-Skeini and Others v. the United Kingdom [GC] (55721/07), 2011-07-07.

Andronicou and Constantinou v. Cyprus (25052/94), 1997-10-09.

Angelova and Iliev v. Bulgaria (55523/00), 2007-07-26.

Armani Da Silva v. the United Kingdom [GC] (5878/08), 2016-03-30.

Bubbins v. the United Kingdom (50196/99), 2005-03-17.

Çakıcı v. Turkey [GC] (23657/94), 1999-07-08.

Ergi v. Turkey (23818/94), 1998-07-28.

Giuliani and Gaggio v. Italy [GC] (23458/02), 2011-03-24.

Güleç v. Turkey (54/1997/838/1044), 1998-07-27.

Hassan v. the United Kingdom (29750/09), 2014-09-16.

Huohvanainen v. Finland (57389/00), 2007-03-13.

Isayeva v. Russia (57950/00), 2005-02-24.

Isayeva, Yusupova and Bazayeva v. Russia (57947/00, 57948/00, 57949/00), 2005-02-24.

Issa and Others v. Turkey (31821/96), 2004-11-16.

Jaloud v. the Netherlands [GC] (47708/08), 2014-11-20.

Kakoulli v. Turkey (38595/97), 2005-11-22.

Keenan v. the United Kingdom (27229/95), 2001-04-03.

Kerimova and Others v. Russia (17170/04, 20792/04, 22448/04, 23360/04, 5681/05, 5684/05), 2011-05-03.

Khamzayev and Others v. Russia (1503/02), 2011-05-03.

Kress v. France [GC] (39594/98), 2001-06-07.

L.C.B. v. the United Kingdom (23413/94), 1998-06-09.

Loizidou v. Turkey (preliminary objections) [GC] (15318/89), 1995-03-23.

Loizidou v. Turkey [GC] (15318/89), 1996-12-18.

Makaratzis v. Greece [GC] (50385/99), 2004-12-20.

McCann and Others v. the United Kingdom [GC] (18984/91), 1995-09-27.

McKerr v. the United Kingdom (28883/95), 2001-05-04.

Nachova and Others v. Bulgaria [GC] (43577/98, 43579/98), 2005-07-06.

Osman v. the United Kingdom [GC] (23452/94), 1998-10-28.

Paşa and Erkan Erol v. Turkey (51358/99), 2006-12-12.

Ramsahai and Others v. the Netherlands [GC] (52391/99), 2007-05-15.

S. and Marper v. UK [GC] (30562/04, 30566/04), 2008-12-04.

Salman v. Turkey [GC] (21986/93), 2000-06-27.

Sašo Gorgiev v. the former Yugoslav Republic of Macedonia (49382/06), 2012-04-19.

Šilih v. Slovenia [GC] (71463/01), 2009-04-09.

Şimşek and Others v. Turkey (35072/97, 37194/97), 2005-07-26.

Soering v. the United Kingdom [plenary] (14038/88), 1989-07-07.

Streletz, Kessler and Krenz v. Germany [GC] (34044/96, 35532/97, 44801/98), 2001-03-22.

Tagayeva and Others v. Russia (26562/07, 14755/08, 49339/08, 49380/08, 51313/08, 21294/11, 37096/11), 2017-04-13.

Tahsin Acar v. Turkey [GC] (26307/95), 2004-04-08.

Timurtaş v. Turkey (23531/94), 2000-06-13.

Tyrer v. the United Kingdom (5856/72), 1978-04-25.

Varnava and Others v. Turkey [GC] (16064/90, 16065/90, 16066/90, 16068/90, 16069/90, 16070/90, 16071/90, 16072/90, 16073/90), 2009-09-18.

Öcalan v. Turkey [GC] (46221/99), 2005-05-12.

Decisions

Banković and Others v. Belgium and Others [GC] (52207/99), 2001-12-12.

Behrami and Behrami v. France and Saramati v. France, Germany and Norway [GC] (71412/01, 78166/01), 2007-05-02.

Boso v. Italy (50490/99), 2002-09-05.

International Court of Justice

Advisory Opinion on Legality of the Threat or Use of Nuclear Weapons, 1996-07-08, I. C. J. Reports 1996, p. 226.

Advisory Opinion on the Legal Consequences of the Construction of a Wall in the Occupied Palestinian Territory, 2004-07-09, I. C. J. Reports 2004, p. 136.

Material from the CCW framework

URL is not attached to the references below since the material from the *CCW* framework discussions is available at the official webpage of the *CCW* [[www.unog.ch/80256EE600585943/\(httpPages\)/4F0DEF093B4860B4C1257180004B1B30?OpenDocument](http://www.unog.ch/80256EE600585943/(httpPages)/4F0DEF093B4860B4C1257180004B1B30?OpenDocument)].

Reports

Report of the 2018 session of the Group of Governmental Experts on Emerging Technologies in the Area of LAWS, CCW/GGE.1/2018/3, 2018-10-23.

Report of the 2019 session of the Group of Governmental Experts on Emerging Technologies in the Area of LAWS, CCW/GGE.1/2019/3, 2019-09-25.

Revised draft final report of the 2019 Meeting of the High Contracting Parties to the *CCW*, CCW/MSP/2019/CRP.2/Rev.1, 2019-11-15.

Statements

Campaign to Stop Killer Robots (2019 a), *Statement to the Convention on Conventional Weapons Group of Governmental Experts on lethal autonomous weapons systems*, statement for the Group of Governmental Experts (2019/03/25–29), 2019-03-27.

Campaign to Stop Killer Robots (2019 c), *Statement to the Convention on Conventional Weapons annual meeting*, statement for the Meeting of the High Contracting Parties to the CCW (2019-11-13–15), 2019-11-14, not published at the webpage of the CCW and available at [www.stopkillerrobots.org/publications/].

European Union (2019), *Further consideration of the human element in the use of lethal force; aspects of human-machine interaction in the development, deployment and use of emerging technologies in the area of lethal autonomous weapons systems*, statement for the Group of Governmental Experts (2019/03/25–29).

Human Rights Watch (2019), *Agenda Item 5(a) regarding Challenges Posed to International Humanitarian Law*, statement for the Group of Governmental Experts (2019/03/25–29), 2019-03-26.

ICRC (2019 a), *An exploration of the potential challenges posed by emerging technologies in the area of lethal autonomous weapon systems to international humanitarian law*, statement for the Group of Governmental Experts (2019/03/25–29).

ICRC (2019 b), *under agenda item 5(b) further consideration of the human element in the use of lethal force; aspects of human-machine interaction in the development, deployment and use of emerging technologies in the area of lethal autonomous weapon systems*, statement for the Group of Governmental Experts (2019/03/25–29).

ICRC (2019 e), *Agenda Item 5 (c): CHARACTERISATION. The importance of critical functions*, statement for the Group of Governmental Experts (2019-03-25–29), 2019-03-25.

iPRAW (2019), *Statement by iPRAW during the CCW GGE on LAWS: Human Control*, statement for the Group of Governmental Experts (2019-03-25–29), 2019-03-29.

Working papers

China (2018), *Position Paper*, working paper for the Group of Governmental Experts (2018-04-9–13), CCW/GGE.1/2018/WP.7, 2018-04-11.

ICRC (2018), *Ethics and autonomous weapon systems: An ethical basis for human control?*, working paper for the Group of Governmental Experts (2018-04-9–13), CCW/GGE.1/2018/WP.5, 2018-03-29.

ICRC (2019 c), *Autonomy, artificial intelligence and robotics: Technical aspects of human control*, working paper for the Group of Governmental Experts (2019-08-20–21), CCW/GGE.1/2019/WP.7, 2019-08-20.

Non-Aligned Movement (2018), *General principles on Lethal Autonomous Weapons Systems*, working paper for the Group of Governmental Experts (2018-04-09–13), CCW/GGE.1/2018/WP.1, 2018-03-28.

Russia (2017), *Examination of various dimensions of emerging technologies in the area of lethal autonomous weapons systems, in the context of the objectives and purposes of the Convention*, working paper for the Group of Governmental Experts (2017-11-13–17), CCW/GGE.1/2017/WP.8, 2017-11-10.

Russia (2018), *Russia's Approaches to the Elaboration of a Working Definition and Basic Functions of Lethal Autonomous Weapons Systems in the Context of the Purposes and Objectives of the Convention*, working paper for the Group of Governmental Experts (2018-04-09–13), CCW/GGE.1/2018/WP.6, 2018-04-04.

Russia (2019), *Potential opportunities and limitations of military uses of lethal autonomous weapons systems*, working paper for the Group of Governmental Experts (2019-03-25–29), CCW/GGE.1/2019/WP.1, 2019-03-15.

United States (2018 a), *Human-Machine Interaction in the Development, Deployment and Use of Emerging Technologies in the Area of Lethal Autonomous Weapons Systems*, working paper for the Group of Governmental Experts (2018-08-27–31), CCW/GGE.2/2018/WP.4, 2018-08-28.

United States (2018 b), *Humanitarian benefits of emerging technologies in the area of lethal autonomous weapon systems*, working paper for the Group of Governmental Experts (2018-04-09–13), CCW/GGE.1/2018/WP.4, 2018-03-28.

United States (2019), *Implementing International Humanitarian Law in the Use of Autonomy in Weapon Systems*, working paper for the Group of Governmental Experts (2019-03-25–29), CCW/GGE.1/2019/WP.5, 2019-03-28.

Interventions

Sweden (2018 a), *Intervention on Agenda item 6 (d) – Options for way forward*, 2018-08-29, not published at the webpage of the *CCW* and available with the author.

Sweden (2018 b), *Intervention on Agenda item 6 (a) – Characterization of the systems under consideration in order to promote a common understanding on concepts and characteristics relevant to the objectives and purposes of the Convention*, 2018-04-10, not published at the webpage of the *CCW* and available with the author.

Sweden (2018 c), *Intervention on Agenda item 6 (b) – consideration of the human element in the use of lethal force; aspects of human machine interaction in the development, deployment and use of emerging technologies in the area of LAWS*, 2018-04-11, not published at the webpage of the *CCW* and available with the author.

Sweden (2018 d), *(agenda item 6d) Possible options for addressing the humanitarian and international security challenges posed by emerging technologies in the area of lethal autonomous weapons systems in the context of the objectives and purposes of the Convention without prejudging policy outcomes and taking into account past, present and future proposals*, 2018-04-13, not published at the webpage of the *CCW* and available with the author.

General exchanges

Sweden (2014), *Remarks by Sweden at the Expert Meeting on LAWS in the CCW on 13.05.2014*, general remark for the Meeting of Experts (2014-05-13–16), 2014-05-13.

Sweden (2015), *General statement by Sweden at the 2015 CCW Meeting of Experts on Lethal Autonomous Weapons Systems (LAWS)*, general remark for the Meeting of Experts (2015-04-13–17).

Presentations

Heyns, Christof (2014), *Autonomous weapons systems and human rights law*, presentation made at the informal Meetings of Experts (2014-05-13–16).

Briefing papers

Campaign to Stop Killer Robots (2018), *Retaining human control of weapons systems*, briefing note for the Group of Governmental Experts (2018-04-9–13).

Literature

- Bantekas, Ilias & Oette, Lutz, *International human rights law and practice*, 2nd ed., Cambridge, Cambridge University Press, 2016.
- Brehm, Maya, *Defending the Boundary: Constraints and Requirements on the Use of Autonomous Weapon Systems Under International Humanitarian and Human Rights Law*, Geneva Academy Briefing no 9, 2017-05, available at [www.geneva-academy.ch/joomlatools-files/docman-files/Briefing9_interactif.pdf].
- Canellas, Marc C. & Haga, Rachel A., “Toward meaningful human control of autonomous weapons systems through function allocation”, in *2015 IEEE International Symposium on Technology and Society (ISTAS)*, IEEE, 2015, mmxvi, pp. 1–7.
- Cassese, Antonio, *International law*, 2nd ed., Oxford, Oxford University Press, 2005.
- Crootof, Rebecca, “The Varied Law of Autonomous Weapon Systems”, in Williams, Andrew & Scharre, Paul (ed.), *Autonomous systems: issues for defence policymakers*, Norfolk, VA, NATO Communications and Information Agency, 2015, pp. 98–126.
- Crootof, Rebecca, “A Meaningful Floor for “Meaningful Human Control””, in *Temple International & Comparative Law Journal*, vol. 30, 2016, pp. 53–62.
- Crootof, Rebecca, “Autonomous Weapon Systems and the Limits of Analogy”, in *Harvard National Security Journal*, vol. 9, 2018, pp. 51–83.
- Droege, Cordula, “The interplay between international humanitarian law and international human rights law in situations of armed conflict”, in *Israel Law Review*, vol. 40(2), 2007, pp. 310-355.
- Ekelhof, Merel, “Moving Beyond Semantics on Autonomous Weapons: Meaningful Human Control in Operation”, in *Global Policy*, vol. 10(3), 2019, pp. 343–348.
- Gaggioli Gasteyger, Gloria & Kolb, Robert, “A right to life in armed conflicts? : the contribution of the European Court of Human Rights”, in *Israel Yearbook on Human Rights*, vol. 37, 2007, pp. 115–163.
- Grabewarter, Christoph, *European Convention on Human Rights: commentary*, München, C.H. Beck, Hart Publishing, Nomos, Helbing Lichtenhahn Verlag, 2014.
- Hagström, Martin (ed.) (2016 a), *Autonom våldsutövning – hot eller möjligheter: För en strukturerad debatt* [Swedish], FOI Memo 5676, 2016-04-07, available at [www.foi.se/report-summary?reportNo=FOI%20MEMO%205676].
- Hagström, Martin (2016 b), “Characteristics of autonomous weapon systems”, in ICRC, *Autonomous Weapon Systems: Implications of Increasing Autonomy in the Critical Functions of Weapons*, 2016, pp. 23–25, available at [shop.icrc.org/autonomous-weapon-systems.html?__store=default&__ga=2.47996839.1236762131.1576406716-1798777859.1551638227].
- ICRC, *Commentary on the additional protocols of 8 June 1977 to the Geneva Conventions of 12 August 1949*, Geneva, Nijhoff, 1987.
- Park, Ian, *The right to life in armed conflict*, Oxford, Oxford University Press, 2018.
- Roorda, Mark, “NATO’s Targeting Process: Ensuring Human Control Over (and Lawful Use of) ‘Autonomous’ Weapons”, in Williams, Andrew & Scharre, Paul (ed.), *Autonomous systems: issues for defence policymakers*, Norfolk, VA, NATO Communications and Information Agency, 2015, pp. 152–168.
- Schabas, William A., *The European Convention on Human Rights: a commentary*, Oxford, Oxford University Press, 2015.

Scharre, Paul, “The Opportunity and Challenge of Autonomous Systems”, in Williams, Andrew & Scharre, Paul (ed.), *Autonomous systems: issues for defence policymakers*, Norfolk, VA, NATO Communications and Information Agency, 2015, pp. 3–26.

Scharre, Paul, *Army of none: autonomous weapons and the future of war*, New York, W.W. Norton & Company, 2018.

Shaw, Malcolm Nathan, *International law*, 8th ed., Cambridge, Cambridge University Press, 2017.

Williams, Andrew, “Defining Autonomy in Systems: Challenges and Solutions”, in Williams, Andrew & Scharre, Paul (ed.), *Autonomous systems: issues for defence policymakers*, Norfolk, VA, NATO Communications and Information Agency, 2015, pp. 27–62.

Articles published online and blogposts

Hambling, David (2019 a), “The US Army is developing AI missiles that find their own targets”, in *NewScientist*, 2019-08-14, available at [www.newscientist.com/article/2212982-the-us-army-is-developing-ai-missiles-that-find-their-own-targets/].

Hambling, David (2019 b), “Autonomous killer drones set to be used by Turkey in Syria”, in *NewScientist*, 2019-09-20, available at [www.newscientist.com/article/2217171-autonomous-killer-drones-set-to-be-used-by-turkey-in-syria/].

Kania, Elsa, “China’s Strategic Ambiguity and Shifting Approach to Lethal Autonomous Weapons Systems”, in *Lawfare*, 2018-04-17, available at [www.lawfareblog.com/chinas-strategic-ambiguity-and-shifting-approach-lethal-autonomous-weapons-systems/].

Li, Qiang & Xie, Dan, “Legal regulation of AI weapons under international humanitarian law: A Chinese perspective”, in *Humanitarian Law & Policy*, 2019-05-02, available at [blogs.icrc.org/law-and-policy/2019/05/02/ai-weapon-ihl-legal-regulation-chinese-perspective/].

Lövin, Isabella & Linde, Ann, “Sverige ska ta en ledande roll i arbetet för ett förbud mot dödliga autonoma vapensystem”, in *Sydsvenskan*, 2019-11-03, available at [www.sydsvenskan.se/2019-11-03/sverige-ska-ta-en-ledande-roll-i-arbetet-for-ett-forbud-mot-dodliga-autonoma-vapensystem/].

Sandoz, Yves, “The International Committee of the Red Cross as guardian of international humanitarian law”, 1998-12-31, available at [www.icrc.org/en/doc/resources/documents/misc/about-the-icrc-311298.htm].

Sparrow, Rob, “Ethics as a source of law: The Martens clause and autonomous weapons”, in *Humanitarian Law & Policy*, 2017-11-14, available at [blogs.icrc.org/law-and-policy/2017/11/14/ethics-source-law-martens-clause-autonomous-weapons/].

Tucker, Patrick, “China Is Exporting Killer Robots to the Mideast”, in *DefenseOne*, 2019-11-05, available at [www.defenseone.com/technology/2019/11/secdef-china-exporting-killer-robots-mideast/161100/].

Institutional reports, policy papers, etc.

Amnesty, *Autonomous Weapons Systems: Five Key Human Rights Issues for Consideration*, 2015, available at [www.amnesty.org/en/documents/act30/1401/2015/en/].

Article 36, *Key areas for debate on autonomous weapons systems*, briefing paper, 2014-05, available at [www.article36.org/wp-content/uploads/2014/05/A36-CCW-May-2014.pdf].

Article 36 (2016 a), *Key elements of meaningful human control*, background paper, 2016-04, available at [www.article36.org/wp-content/uploads/2016/04/MHC-2016-FINAL.pdf].

Article 36 (2016 b), *Meaningful Human Control, Artificial Intelligence and Autonomous Weapons*, briefing paper, 2016-04, available at [www.article36.org/wp-content/uploads/2016/04/MHC-AI-and-AWS-FINAL.pdf].

Article 36 (2019 a), *Critical Commentary on the “Guiding Principles”*, policy note, 2019-11, available at [www.article36.org/wp-content/uploads/2019/11/Commentary-on-the-guiding-principles.pdf].

Article 36 (2019 b), *Target profiles*, discussion paper, 2019-08, available at [www.article36.org/wp-content/uploads/2019/08/Target-profiles.pdf].

Article 36 (2019 c), *Autonomy in weapons systems: mapping a structure for regulation through specific policy questions*, policy note, 2019-11, available at [www.article36.org/wp-content/uploads/2019/11/regulation-structure.pdf].

Boulanin, Vincent & Verbruggen, Maaïke, *Mapping the Development of Autonomy in Weapon Systems*, report published by Stockholm International Peace Research Institute (SIPRI), 2017-11, available at [www.sipri.org/publications/2017/other-publications/mapping-development-autonomy-weapon-systems].

Campaign to Stop Killer Robots (2019 b), *Key Elements of a Treaty on Fully Autonomous Weapons*, 2019-11, not published online and available with the author.

Human Rights Watch, *Mind the Gap: The Lack of Accountability for Killer Robots*, 2015, available at [www.hrw.org/report/2015/04/09/mind-gap/lack-accountability-killer-robots].

ICRC (2019 d), *Autonomy, artificial intelligence and robotics: Technical aspects of human control*, 2019-08, available at [www.icrc.org/en/document/autonomy-artificial-intelligence-and-robotics-technical-aspects-human-control].

iPRAW (2018 a), *Focus on the Human-Machine Relation in LAWS*, 2018-03, available at [www.ipraw.org/human-machine-relation/].

iPRAW (2018 b), *Focus on Ethical Implications for a Regulation of LAWS*, 2018-08, available at [www.ipraw.org/ethical-implications/].

iPRAW (2018 c), *Concluding Report: Recommendations to the GGE*, 2018-12, available at [www.ipraw.org/recommendations/].

Rantakokko, Jouni, *Tekniköversikt autonoma och obemannade system - Del 1: Historik* [Swedish], FOI report FOI-R--4680--SE, 2018-12, copy with author or available to order at [www.foi.se/report-summary?reportNo=FOI-R--4680--SE].

UNIDIR, *The Weaponization of Increasingly Autonomous Technologies: how Meaningful Human Control might move the discussion forward*, 2014, available at [www.unidir.org/publication/weaponization-increasingly-autonomous-technologies-considering-how-meaningful-human].

United States Defense Innovation Board, *AI Principles: Recommendations on the Ethical Use of Artificial Intelligence by the Department of Defense*, 2019, available at [media.defense.gov/2019/Oct/31/2002204458/-1/-1/0/dib_ai_principles_primary_document.pdf].

United States Department of Defense, *Department of Defense directive 3000.09: Autonomy in Weapon Systems*, 2012-11-21.

United States Department of Defense, *Department of Defense Law of War Manual*, 2016.

Websites

Article 36, *About*, available at [www.article36.org/about/].

BAE Systems, *Bofors 155mm BONUS Munition*, available at [www.baesystems.com/en/product/155-bonus].

Campaign to Stop Killer Robots, *About*, available at [www.stopkillerrobots.org/about/].

Campaign to Stop Killer Robots, *Learn*, available at [www.stopkillerrobots.org/learn/].

UNOG, *Background on Lethal Autonomous Weapons Systems in the CCW*, available at [[www.unog.ch/80256EE600585943/\(httpPages\)/8FA3C2562A60FF81C1257CE600393DF6?OpenDocument](http://www.unog.ch/80256EE600585943/(httpPages)/8FA3C2562A60FF81C1257CE600393DF6?OpenDocument)].

Other sources

Stop Autonomous Weapons, YouTube video "*Slaughterbots*", 2017-11-12, available at [www.youtube.com/watch?v=9CO6M2HsoIA].

FOI, Swedish Defence Research Agency, is a mainly assignment-funded agency under the Ministry of Defence. The core activities are research, method and technology development, as well as studies conducted in the interests of Swedish defence and the safety and security of society. The organisation employs approximately 1000 personnel of whom about 800 are scientists. This makes FOI Sweden's largest research institute. FOI gives its customers access to leading-edge expertise in a large number of fields such as security policy studies, defence and security related analyses, the assessment of various types of threat, systems for control and management of crises, protection against and management of hazardous substances, IT security and the potential offered by new sensors.



FOI
Defence Research Agency
SE-164 90 Stockholm

Phone: +46 8 555 030 00
Fax: +46 8 555 031 00

www.foi.se