

Scientific Article

DOI: 10.51940/2024.1.189-216

UDC: 341.3:342.7:004.8

341:623:004.8

Joana Gomes Beirão, Jan Wouters***

Towards an International Legal Framework for Lethal Artificial Intelligence Based on Respect for Human Rights: Mission Impossible?

Abstract

This article considers the potential use of autonomous weapons both in and outside armed conflict, including in law enforcement. It analyses the phenomenon from the perspective of human rights law, with a particular focus on the right to life. For over a decade, the international community has debated whether technological advances pertaining to the development of autonomous weapons require the establishment of new rules within the framework of international humanitarian law. In contrast, consideration of such technology from a human rights law perspective has been limited, despite its implications for the right to life and other human rights. In parallel, several international initiatives have emerged in recent years aiming to establish non-binding and binding rules for the development and use of artificial intelligence (AI) based on respect for human rights. This article reviews four such initiatives: the OECD Recommendation on AI, the UNESCO Recommendation on the Ethics of AI, the INTERPOL and UNICRI Toolkit for Responsible AI Innovation in Law Enforcement, and the Council of Europe AI Convention. It examines the extent to which these initiatives address the specific concerns raised by autonomous weapons.

Key words

autonomous weapons, artificial intelligence, human rights, right to life, law enforcement.

* *Junior researcher at the Leuven Centre for Global Governance Studies – Institute for International Law, America Europe Chair on Technology, Innovation and International Regulation, KU Leuven.*

** *Jean Monnet Chair ad personam and Full Professor of International Law and International Organizations, Director, Leuven Centre for Global Governance Studies, Coordinator, America Europe Chair on Technology, Innovation and International Regulation, KU Leuven.*

1. Introduction

Autonomous weapons have been the subject of long discussions and disagreements over whether they can be used in compliance with existing rules of international humanitarian law (IHL) and whether new IHL rules should be created to prohibit or at least regulate them. As stated in 2014 by Christof Heyns, at the time UN Special Rapporteur on extrajudicial, summary or arbitrary executions, “[t]he legal debate about [autonomous weapons] that has emerged during the past few years has largely left human rights out of the picture, and focused primarily on IHL”.¹ A decade later, the statement remains perfectly accurate.

Building on the conclusions of his predecessor,² Heyns recommended in 2013 that the United Nations (UN) Human Rights Council call on States to declare a moratorium on the development, acquisition, deployment, and use of lethal autonomous robots until an international framework could be established to regulate such technology. He also proposed that the UN High Commissioner for Human Rights convene a high-level panel tasked with advancing the establishment of this framework.³ The following year Heyns called on the international community to “adopt a comprehensive and coherent approach to autonomous weapons systems in armed conflict and in law enforcement, one which covers both the international humanitarian law and human rights dimensions”, stressing that “the various international agencies and institutions dealing with disarmament and human rights, such as the Convention on Certain Conventional Weapons and the Human Rights Council, each have a responsibility and a role to play” with regard to autonomous weapons.⁴

Despite such calls, echoed by civil society,⁵ discussions on the potential regulation of autonomous weapons have primarily taken place within the framework of IHL, specifically within the Group of Governmental Experts on Lethal Autonomous Weapons Systems, established under the UN Convention on Certain Conventional Weapons.⁶ In a time where the development⁷ and use⁸ of this technology are well under way, to date such discussions have yielded only modest results. Many questions regarding the appli-

¹ Heyns, 2014b.

² Alston, 2010, § 48.

³ Heyns, 2013, §§ 113–114.

⁴ Heyns, 2014c, § 89.

⁵ Docherty, 2014, p. 4.

⁶ 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, Geneva, 10 October 1980, UNTS 22495.

⁷ Alston, 2010, §§ 27–28.

⁸ Choudhury *et al.*, 2021, §§ 63–64.

cation of existing rules to autonomous weapons remain unanswered, and new rules have not been established due to persistent difficulties in reaching consensus.⁹

In parallel, several initiatives have recently emerged to regulate artificial intelligence (AI), including ensuring that its use respects human rights. As this technology progresses, the international community has considered whether non-binding or binding rules should be established to address the concerns it raises, including its potential impact on the enjoyment of human rights. Noteworthy among these initiatives are the OECD Recommendation on AI,¹⁰ the UNESCO Recommendation on the Ethics of AI,¹¹ the INTERPOL and UNICRI Toolkit for Responsible AI Innovation in Law Enforcement,¹² and the Council of Europe Framework Convention on AI and Human Rights, Democracy and the Rule of Law.¹³ Considering the attention the topic has received, further initiatives to establish an international legal framework for AI may emerge in the future.¹⁴

With these developments in mind, the present article considers the potential use of autonomous weapons in and outside armed conflict, including in law enforcement, analysing the phenomenon from the perspective of human rights law, focusing particularly on the right to life. Subsequently, we reflect on recent initiatives to regulate AI based on respect for human rights, examining to what extent they address the specific concerns autonomous weapons raise.

Before moving forward with our analysis, it is important to note that the concept of autonomous weapons, as “weapons that select and apply force to targets without human

⁹ Reeves, Alcalá & McCarthy, 2021, pp. 102 and 107–110.

¹⁰ OECD, Recommendation of the Council on Artificial Intelligence, OECD/LEGAL/0449, 22 May 2019 (OECD Recommendation on AI).

¹¹ UNESCO, Recommendation on the Ethics of Artificial Intelligence, SHS/BIO/PI/2021/1, 23 November 2021 (UNESCO Recommendation on the Ethics of AI).

¹² UNICRI and INTERPOL, Toolkit for Responsible AI Innovation in Law Enforcement, June 2023 (INTERPOL and UNICRI Toolkit for Responsible AI Innovation in Law Enforcement).

¹³ Council of Europe, Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, 5 September 2024 (Council of Europe AI Convention).

¹⁴ To name but a few examples: ASEAN is developing a guide on AI governance and ethics, although little is currently known about the initiative (Potkin & Wongcha-um, 2023); the United Kingdom announced it will host “the first major global summit on AI safety” which “will bring together key countries, leading tech companies and researchers to agree safety measures to evaluate and monitor the most significant risks from AI” (UK Prime Minister’s Office, 2023); the European Union and the United States are developing a voluntary AI code of conduct (Blenkinsop, 2023); and the UN Secretary-General has supported the proposal to establish an agency, inspired by the International Agency of Atomic Energy, mandated to regulate AI (Guterres, 2023).

intervention”,¹⁵ includes both weapons that incorporate AI¹⁶ as well as weapons which do not use such technology to perform the autonomous selection and application of force. Nevertheless, the focus of this article is on autonomous weapons that incorporate AI, given the increased unpredictability as to how these machines select and apply force.¹⁷

2. Autonomous Weapons and Human Rights Law

While the development and use of autonomous weapons clearly deserves careful consideration from the perspective of IHL, the same is also required from the perspective of human rights law for at least three reasons. Firstly, even if autonomous weapons were an exclusively military technology, human rights law remains applicable during armed conflicts alongside IHL.¹⁸ Although a thorough analysis of the relationship between IHL and human rights law is not possible here,¹⁹ it should be noted that international and regional courts, UN organs, treaty bodies and human rights special procedures have recognised that “both bodies of law apply to situations of armed conflict and provide complementary and mutually reinforcing protection”.²⁰ In this regard, the International Court of Justice has held that:

“[T]he protection offered by human rights conventions does not cease in case of armed conflict, save through the effect of provisions for derogation of the kind to be found in Article 4 of the International Covenant on Civil and Political Rights. As regards the relationship between international humanitarian law and human

¹⁵ Although there is no universally-accepted definition of autonomous weapons, for the purpose of our analysis we consider the definition endorsed by the International Committee of the Red Cross (see: International Committee of the Red Cross, 2021, p. 5). For a comparative analysis of definitions of autonomous weapons, see: Taddeo & Blanchard, 2022.

¹⁶ Conceptualisations of AI differ as there is currently no universally-agreed upon definition of this technology. The OECD defines an AI system as “a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments” (OECD Recommendation on AI, § I). UNESCO defines such systems as “information-processing technologies that integrate models and algorithms that produce a capacity to learn and to perform cognitive tasks leading to outcomes such as prediction and decision-making in material and virtual environments” (UNESCO Recommendation on the Ethics of Artificial Intelligence, § 2). The Council of Europe Committee on Artificial Intelligence defines an AI system as “a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations or decisions that may influence physical or virtual environments” (Article 2 of the Council of Europe AI Convention).

¹⁷ International Committee of the Red Cross, 2021, pp. 6–7.

¹⁸ Brehm, 2017, p. 25; Odon, 2022, pp. 85–89.

¹⁹ See, *inter alia*, Naert (2016).

²⁰ Office of the United Nations High Commissioner for Human Rights, 2011, p. 1. See also: Droege, 2007, pp. 320–324.

rights law, there are thus three possible situations: some rights may be exclusively matters of international humanitarian law; others may be exclusively matters of human rights law; yet others may be matters of both these branches of international law”.²¹

For this reason, State conduct, such as the use of autonomous weapons, should be assessed considering both international human rights law (as *lex generalis*) and IHL (as *lex specialis*).²² In essence, the concurrent application of the two regimes means that human rights rules are to be interpreted in light of IHL.²³

Secondly, military technologies regularly find their way outside armed conflict.²⁴ As such, it cannot be ruled out that autonomous weapons may be used in peacetime, including in law enforcement.²⁵ The incorporation of military technologies into law enforcement can already be seen in the increasing use of remote-controlled drones and robots by police (e.g., for bomb disposal,²⁶ surveillance²⁷ and border patrol²⁸).²⁹ Moreover, there is at least one recorded instance in which police used a remote-controlled robot to employ lethal force.³⁰ It is important to note that this unprecedented action, which took place in Texas in 2016, did not result from an official policy change that would allow the use of robots to employ lethal force, but from a “creative” solution reached by police officers facing an extremely dangerous situation. One instance which is perhaps more indicative of the militarisation and depersonalisation of law enforcement is the proposal of the San Francisco Police Department to establish a new policy allowing (remote-controlled) lethal robots to be employed in extreme circumstances which pose an immediate risk to life.³¹ Advocates of such a policy argue that it could save police officers’ lives since they would not have to be physically present in dangerous situations. Such reasoning applies to remote-controlled and autonomous robots alike. However, in conjunction with the (perceived) need to increase the efficiency of law enforcement, it is possible that we will see a push in the future towards the incorporation of autonomous robots into

²¹ International Court of Justice, *Legal Consequences of the Construction of a Wall in the Occupied Palestinian Territory*: Advisory Opinion, 9 July 2004, § 106.

²² International Court of Justice, *Legal Consequences of the Construction of a Wall in the Occupied Palestinian Territory*: Advisory Opinion, 9 July 2004, § 106; Odon, 2022, pp. 85–86.

²³ European Court of Human Rights, *Hassan v. the United Kingdom* [GC], App. No. 29750/09, Judgement, 16 September 2014, §§ 102–104; Odon, 2022, pp. 85–86.

²⁴ Amnesty International, 2015, p. 9.

²⁵ Heyns, 2013, § 84; Heyns, 2014a, § 144; Heyns, 2014c, § 84; Marijan, 2023.

²⁶ Allison, 2016.

²⁷ Singapore Home Team Science and Technology Agency, 2021; Reuters, 2017.

²⁸ The Guardian, 2014; U.S. Department of Homeland Security, 2022.

²⁹ Heyns, 2014c, §§ 77–83; Marijan, 2023.

³⁰ Sinder & Simon, 2016; Fund, 2016.

³¹ Derico & Clayton, 2022; Rodríguez, 2023.

policing since the technology is capable of processing information and responding faster than humans piloting remote-controlled robots.³² For law enforcement, States may be particularly willing to use the so-called less-lethal autonomous weapons as these are generally considered less dangerous and, hence, less controversial. However, such weapons also raise concerns from a human rights perspective, including with regard to the right to life.³³ It should be recalled that the use of less-lethal weapons (such as tasers, rubber bullets and tear gas), whether employed directly by a police officer, remotely-controlled or autonomously, may lead to the death of the targeted person(s) and/or innocent bystanders.³⁴ Since IHL is not applicable outside an armed conflict, any rules which may be created within that field to prohibit or regulate autonomous weapons, including within the context of the Convention on Certain Conventional Weapons³⁵, would not apply to the use of this technology in law enforcement or other domestic settings such as private security. It is thus crucial to carefully assess the potential use of autonomous weapons in domestic settings from the perspective of human rights law.

Thirdly, the use of autonomous weapons may have far-reaching implications for human dignity and human rights.³⁶ Some scholars argue that entrusting the decision to kill a human being to a machine constitutes a grave violation of human dignity, rendering the use of any technology capable of autonomously employing lethal force *a priori* unlawful.³⁷ While the scope of the right to human dignity remains contentious, it is clear that the use of autonomous weapons in and outside armed conflict may impact the right to life and the right to not be subjected to cruel, inhuman, or degrading treatment.³⁸ Moreover, considering the large-scale collection and processing of data required for the functioning of this technology, as well as concerns regarding bias, transparency, and explainability of algorithmic decisions, autonomous weapons may also affect the right to privacy, the right not to be discriminated against, and the right to an effective remedy.³⁹ Given these concerns, the development and use of autonomous weapons deserves careful consideration from the perspective of human rights law.

³² Heyns, 2016, p. 359; Marijan, 2023.

³³ Human Rights Committee, General Comment No. 36 on Article 6: right to life, § 14; Brehm, 2017, pp. 54–55.

³⁴ Heyns, 2014c, § 69; Heyns, 2016, p. 361; Office of the United Nations High Commissioner for Human Rights, 2020, § 1.2.

³⁵ Amnesty International, 2015, pp. 7–8.

³⁶ Heyns, 2014a, § 144.

³⁷ Heyns, 2016, pp. 369–371; Docherty, 2014, pp. 23–24; Brehm, 2017, pp. 63–65.

³⁸ Brehm, 2017, pp. 69–70.

³⁹ *Ibid.*, pp. 56–68; Spagnolo, 2017, pp. 52–56; Spagnolo, 2019, pp. 59–61.

3. Autonomous weapons and the right to life

From a human rights perspective, the most important implication of the use of autonomous weapons both in and outside armed conflict is the potential interference with the right to life.⁴⁰ The Human Rights Committee has recognised the importance of considering this right when it comes to autonomous weapons. Referring to Article 36 of Additional Protocol I to the Geneva Conventions, the Committee held that:

“States parties engaged in the deployment, use, sale or purchase of existing weapons and in the study, development, acquisition or adoption of weapons, and means or methods of warfare, must always consider their impact on the right to life. For example, the development of autonomous weapon systems lacking in human compassion and judgment raises difficult legal and ethical questions concerning the right to life, including questions relating to legal responsibility for their use. The Committee is therefore of the view that such weapon systems should not be developed and put into operation, either in times of war or in times of peace, unless it has been established that their use conforms with article 6 [of the International Covenant on Civil and Political Rights] and other relevant norms of international law.”⁴¹

Importantly, the Human Rights Committee did not categorically state that the use of autonomous weapons in and outside armed conflict is *a priori* incompatible with the right to life. Instead, it noted that, from the perspective of the right to life, autonomous weapons are lawful if and to the extent that they can be used in accordance with the requirements of Article 6 of the International Covenant on Civil and Political Rights (ICCPR). We now turn to those requirements.

As “the supreme right” inherent to every human being whose “effective protection [...] is the prerequisite for the enjoyment of all other human rights”,⁴² the right to life is enshrined in all human rights treaties, as well as in Article 3 of the Universal Declaration of Human Rights and Article I of the American Declaration of the Rights and Duties of Man. Pursuant to Article 6 of the ICCPR and Article 4 of the American Convention on Human Rights (ACHR), “[n]o one shall be arbitrarily deprived of his life”. Both conventions explicitly state that the death penalty, when applied for the most serious crimes, does not constitute an arbitrary deprivation of life. Article 4 of the African Charter on Human and People’s Rights also provides that no one may be arbitrarily deprived of their life, but without explicitly addressing whether the death penalty is to be considered an arbitrary deprivation of life.

In a different formulation, Article 2 of the European Convention on Human Rights (ECHR) stipulates that “[n]o one shall be deprived of his life intentionally” except if sen-

⁴⁰ Heyns, 2013, §§ 36 and 85; Spagnolo, 2019, p. 59.

⁴¹ Human Rights Committee, General Comment No. 36 on Article 6: right to life, § 65.

⁴² *Ibid.*, § 2.

tenced to death by a court or if the use of force is absolutely necessary to defend a person from unlawful violence, to effect a lawful arrest, to prevent the escape of a detainee or to quell a riot or insurrection. The ECHR allows States to derogate from the right to life but only with respect to deaths resulting from lawful acts of war.⁴³ Even though the ICCPR⁴⁴ and ACHR⁴⁵ allow no derogations from the right to life, deaths resulting from lawful acts of war are not considered arbitrary deprivations of life and thus do not contravene the right to life under these treaties.⁴⁶ The same applies to the situations in which the use of force is absolutely necessary to defend a person from unlawful violence, effect a lawful arrest, prevent the escape of a detainee or quell a riot or insurrection.⁴⁷

Importantly, the rules governing the use of lethal force under human rights law are more stringent than those under IHL. Human rights law only tolerates the use of lethal force in exceptional circumstances in accordance with the principles of legality, necessity and proportionality. Firstly, any use of lethal force must have a sufficient legal basis; it must be authorised and sufficiently regulated by law.⁴⁸ Secondly, the use of lethal force must be strictly necessary to protect life or prevent serious injury from an imminent threat. In adhering to the principle of necessity, any alternatives to the use of lethal force must have been exhausted, unless they are not possible or adequate to protect the interest in question.⁴⁹ Thirdly, the amount of force employed must be proportional to the interest protected. Thus, the principle of proportionality requires that the amount of force employed does not exceed what is strictly necessary to respond to the threat.⁵⁰

As recognised in the preamble of the UN Basic Principles on the Use of Force and Firearms by Law Enforcement Officials, “law enforcement officials have a vital role in the protection of the right to life, liberty and security of the person”.⁵¹ For this reason,

“[t]he use of potentially lethal force for law enforcement purposes is an extreme measure that should be resorted to only when strictly necessary in order to protect life or prevent serious injury from an imminent threat”.⁵²

⁴³ ECHR, Article 15(2).

⁴⁴ ICCPR, Article 4(2).

⁴⁵ ACHR, Article 27(2).

⁴⁶ Brehm, 2017, pp. 24–25.

⁴⁷ Human Rights Committee, General Comment No. 36 on Article 6: right to life, § 10.

⁴⁸ *Ibid.*, § 11.

⁴⁹ *Ibid.*, § 12.

⁵⁰ *Ibid.*, § 12.

⁵¹ Basic Principles on the Use of Force and Firearms by Law Enforcement Officials, adopted by the Eighth United Nations Congress on the Prevention of Crime and the Treatment of Offenders, Havana, Cuba, 27 August to 7 September 1990.

⁵² Human Rights Committee, General Comment No. 36 on Article 6: right to life, § 12. See also: Article 3 of the Code of Conduct for Law Enforcement Officials, adopted by General Assembly resolution 34/169 of 17 December 1979.

The Basic Principles further specify that firearms shall only be used

“in self-defence or defence of others against the imminent threat of death or serious injury, to prevent the perpetration of a particularly serious crime involving grave threat to life, to arrest a person presenting such a danger and resisting their authority, or to prevent his or her escape, and only when less extreme means are insufficient to achieve these objectives”.⁵³

Moreover, law enforcement officials

“shall, as far as possible, apply non-violent means before resorting to the use of force and firearms. They may use force and firearms only if other means remain ineffective or without any promise of achieving the intended result”.⁵⁴

When law enforcement officials do use force, they must exercise restraint, act in proportion to the seriousness of the offence and the legitimate objective to be achieved, minimise damage and injury, and ensure that medical assistance is rendered at the earliest possible moment.⁵⁵

In addition to the prohibition of unlawfully interfering with the right to life, States have positive obligations pertaining to the right to life. States have the duty to protect the right to life, including by establishing an appropriate legal framework that ensures the full enjoyment of this right, protects it from foreseeable threats, establishes with sufficient precision the grounds on which lethal force may be used and puts in place procedures to prevent, investigate and prosecute potential cases of unlawful deprivation of life.⁵⁶ With regard to law enforcement, States must put in place “all necessary measures to prevent arbitrary deprivation of life by their law enforcement officials, including soldiers charged with law enforcement missions”.⁵⁷ Such measures include adopting

“appropriate legislation controlling the use of lethal force by law enforcement officials, procedures designed to ensure that law enforcement actions are adequately planned in a manner consistent with the need to minimize the risk they pose to human life, mandatory reporting, review and investigation of lethal incidents and other life-threatening incidents, and supplying forces responsible for crowd control with effective, less-lethal means and adequate protective equipment in order to obviate their need to resort to lethal force”.⁵⁸

⁵³ Basic Principles on the Use of Force and Firearms by Law Enforcement Officials, § 9.

⁵⁴ *Ibid.*, § 4.

⁵⁵ *Ibid.*, § 5. On less-lethal weapons, see: Office of the United Nations High Commissioner for Human Rights, 2020, § 2.1-2.11.

⁵⁶ Human Rights Committee, General Comment No. 36 on Article 6: right to life, § 18–20.

⁵⁷ *Ibid.*, § 13.

⁵⁸ *Ibid.*, § 13; Basic Principles on the Use of Force and Firearms by Law Enforcement Officials, §§ 1–3, 6–7, 11, 22–26; European Court of Human Rights, 2022, § 91–96. On less-lethal weapons, see: Office of the United Nations High Commissioner for Human Rights, 2020, § 3.1–4.8.2.

Considering the negative and positive obligations of the State with regard to the right to life, the development and use of autonomous weapons must be carefully assessed. A State intending to use autonomous weapons must ensure that any use of potentially lethal force therein complies with the principles of legality, necessity and proportionality. However, it remains unclear whether autonomous weapons are capable of complying with these principles as they require contextual value judgements, which machines may not be able to make reliably.⁵⁹ To determine whether recourse to lethal force is necessary, autonomous weapons would need to assess in a limited time if a person poses an imminent threat, including by ascertaining that person's intent to kill or seriously injure another person, which may be particularly difficult for a machine to accurately assess.⁶⁰ Similarly, the balancing exercise required to comply with the proportionality principle may be challenging for autonomous weapons to perform, since it requires an assessment, which has to be performed in a limited time, of the amount of force strictly needed to respond to the threat in question.⁶¹

Moreover, under human rights law, any use of force requires an individual assessment of the circumstances that justify recourse to force. Since autonomous weapons are programmed to some extent beforehand, the requirement to individuate the use of force may not be met.⁶² For this reason, and considering the doubts as to whether autonomous weapons can reliably make the value judgements necessary to assess the necessity and proportionality of using lethal force, some scholars argue that autonomous weapons which employ lethal force without meaningful human control contravene the right to life.⁶³ Accordingly, in order to comply with the right to life, the use of autonomous weapons would need to comprise human agents which "remain constantly and actively (personally) engaged in every individual application of force", essentially ruling out the use of fully autonomous weapons.⁶⁴

Given the grave consequences of an erroneous assessment by an autonomous weapon—namely an unlawful deprivation of life—States must exercise particular caution with this technology. Arguably, the aforementioned doubts regarding the ability of fully autonomous weapons to comply with the principles of necessity and proportionality provide sufficient reason for States to refrain from using such technology, at least while such doubts persist. Ensuring meaningful human control over the technology may contribute to ensuring compliance with the prohibition of unlawful interference with the

⁵⁹ Heyns, 2014c, § 85; Spagnolo, 2017, p. 48.

⁶⁰ Heyns, 2016, pp. 364–366.

⁶¹ *Ibid.*

⁶² Brehm, 2017, pp. 45–48; Heyns, 2016, pp. 370–371.

⁶³ Kiai & Heyns, 2016, § 67(f); Heyns, 2016, pp. 374–376; Brehm, 2017, p. 48.

⁶⁴ Brehm, 2017, p. 48; Asaro, 2012, p. 708.

right to life.⁶⁵ However, many questions remain regarding the technical and operational requirements necessary to effectively operationalise the concept of meaningful human control.⁶⁶ Thus, States must still exercise caution when assessing whether to use partially autonomous weapons. Ultimately, a death will be unlawful if it does not strictly comply with the principles of legality, necessity, and proportionality, regardless of whether force was directly employed by a person, by remotely-controlled technology, by a fully autonomous weapon or by a partially autonomous weapon. The use of autonomous weapons, regardless of their level of autonomy, does not excuse States from complying with the prohibition of unlawfully interfering with the right to life.

Furthermore, States intending to use autonomous weapons must also respect their positive obligations to ensure the right to life, including by establishing an appropriate legal framework regulating the use of autonomous weapons, ensuring they are designed to minimise the risk to human life, and adequately training the persons responsible for exercising control and oversight over the technology.⁶⁷ A notable risk of using autonomous weapons is that humans who engage with them may overly rely on the machine's assessments that the use of force is legal, necessary and proportional, thereby limiting their role to an automatic approval of the machine's decisions.⁶⁸ States must provide adequate training to avoid such risk, as well as ensure that there are sufficient human resources to effectively exercise control over the weapons. Additionally, States must investigate and prosecute potential cases of unlawful deprivation of life resulting from the use of autonomous weapons. However, it may be challenging for a State to fulfil such duties where the decision to use lethal force was made by an autonomous system without meaningful human control, as there will only be an indirect link between the actions of the persons involved (e.g., the public body which approved the use of autonomous weapons in a certain context, its developers, etc) and the decision to kill.⁶⁹ Even when human control and oversight are present, there is a risk that persons involved in the use of the system may claim that unlawful deprivations of life were caused by technical errors. The

⁶⁵ Although discussions on the concept of meaningful human control have mostly concerned its role in ensuring compliance with international humanitarian law, many of the considerations therein can be applied to human rights law. Since a parallel can be drawn between the difficulty in ensuring that autonomous weapons comply with the principles of distinction and proportionality under international humanitarian law and the difficulty in ensuring that they comply with the principles of necessity and proportionality under human rights law, the concept of meaningful human control may be useful for both bodies of law.

⁶⁶ Boutin & Woodcock, 2022, p. 2.

⁶⁷ Spagnolo, 2019, p. 67.

⁶⁸ For an explanation of the phenomenon of over reliance on algorithmic decision, known as automation bias, see: Jones-Jang & Park, 2022, p. 2.

⁶⁹ Heyns, 2016, p. 373; Spagnolo, 2017, pp. 50–51.

opacity of the technology may render such claims difficult or impossible to assess.⁷⁰ If a State intends to use autonomous weapons while respecting its obligations under human rights law, it must ensure that responsibility for unlawful deaths is not evaded.

Overall, many questions remain regarding how States can ensure that they respect the right to life when using autonomous weapons. Considering the fundamental nature of this right, it is critical that the international community discusses the concerns autonomous weapons raise. The next section analyses the extent to which recent initiatives to regulate AI address these concerns.

4. Regulating AI but not lethal AI?

As “[t]he use of force against the human person, including the use of deadly or potentially deadly force by agents of the State, is a central human rights concern”,⁷¹ it would be expected that any initiative to regulate AI based on respect for human rights would carefully examine the concerns autonomous weapons raise with regard to the right to life. From this lens, this section reflects on the OECD Recommendation on AI, the UNESCO Recommendation on the Ethics of AI, the INTERPOL and UNICRI Toolkit for Responsible AI Innovation in Law Enforcement, and the Council of Europe AI Convention.

4.1. OECD Recommendation on AI

In May 2019, the OECD Council adopted a Recommendation on AI, which “aims to foster innovation and trust in AI by promoting the responsible stewardship of trustworthy AI while ensuring respect for human rights and democratic values”.⁷²

Although devoid of binding force, the Recommendation is “an important political and moral commitment at the intergovernmental level” recognising not only that AI may pose harm to human rights and democratic values but also that these concerns need to be addressed at both intergovernmental and national levels.⁷³ The Recommendation was endorsed by all 36 OECD Members, as well as Argentina, Brazil, Columbia, Costa Rica, Peru, and Romania, and formed the basis of the G20 AI Principles adopted by G20 Leaders that same year.

The Recommendation sets forth five complementary principles for responsible stewardship of trustworthy AI: inclusive growth, sustainable development and well-being; human-centred values and fairness; transparency and explainability; robustness, security,

⁷⁰ Bo, Bruun & Boulanin, 2022, pp. 46–49.

⁷¹ Heyns, 2014c, § 65.

⁷² OECD Recommendation on AI, p. 3.

⁷³ Yeung, 2020, p. 28.

and safety; and accountability. Furthermore, it provides five recommendations regarding the development of national policies and international cooperation, namely investing in AI research and development, fostering a digital ecosystem for AI, shaping an enabling policy environment for AI, building human capacity and preparing for labour market transformation, and promoting international cooperation for trustworthy AI.

Of particular relevance to the subject of our analysis is the set of five principles for responsible stewardship of trustworthy AI, which sets forth that:

“a) AI actors should respect the rule of law, human rights and democratic values, throughout the AI system lifecycle. These include freedom, dignity and autonomy, privacy and data protection, non-discrimination and equality, diversity, fairness, social justice, and internationally recognised labour rights.

b) To this end, AI actors should implement mechanisms and safeguards, such as capacity for human determination, that are appropriate to the context and consistent with the state of art” (§1.2).

AI actors, i.e., any actors who play an active role in the lifecycle of an AI system, should further “commit to transparency and responsible disclosure regarding AI systems [...] to enable those adversely affected by an AI system to challenge its outcome” and ensure that AI systems are

“robust, secure and safe throughout their entire lifecycle so that, in conditions of normal use, foreseeable use or misuse, or other adverse conditions, they function appropriately and do not pose unreasonable safety risk” (§ 1.3–1.4).

Finally,

“AI actors should be accountable for the proper functioning of AI systems and for the respect of the [...] principles [set forth in the recommendation], based on their roles, the context, and consistent with the state of art” (§ 1.5).

Interestingly, the right to life is not mentioned anywhere in the document, nor are the specific concerns autonomous weapons raise reflected in its text. In line with the declaration that the use of AI should respect human rights, the document does recommend that “mechanisms and safeguards” are implemented, and that safety and accountability is ensured. While these principles are relevant for the development and use of autonomous weapons, they are likely insufficient to ensure that the right to life is respected. Consider, for example, the recommendation to implement “capacity for human determination”. Designing and using an autonomous weapon that allows human intervention if a malfunction is detected but does not require prior human approval for the use of lethal force may not meet human rights law requirements for the use of lethal force, as detailed in section 3. Overall, the Recommendation does not significantly contribute to clarifying how States can ensure they respect the right to life when developing and using autonomous weapons.

4.2. *UNESCO Recommendation on the Ethics of AI*

In November 2021, the General Conference of UNESCO adopted a Recommendation on the Ethics of AI,

“a standard-setting instrument developed through a global approach, based on international law, focusing on human dignity and human rights, as well as gender equality, social and economic justice and development, physical and mental well-being, diversity, interconnectedness, inclusiveness, and environmental and ecosystem protection”.

The Recommendation addresses ethical issues concerning AI to the extent that they are within UNESCO’s mandate, focusing particularly on its central domains, namely education, science, culture, communication and information (§ 1–3).⁷⁴

The document sets forth a set of values and principles, operationalised in eleven policy areas: ethical impact assessment; ethical governance and stewardship; data policy; development and international cooperation; environment and ecosystems; gender; culture; education and research; communication and information; economy and labour; and health and social well-being. For the subject of our analysis, the first value set forth in Recommendation is of particular relevance, as it stresses that “[h]uman rights and fundamental freedoms must be respected, protected and promoted throughout the life cycle of AI systems”, and

“[n]o human being or human community should be harmed or subordinated, whether physically, economically, socially, politically, culturally or mentally during any phase of the life cycle of AI systems”.

The need to respect human dignity is emphasised: “persons should never be objectified, nor should their dignity be otherwise undermined” when interacting with an AI system (§ 13–16).

Among the principles set out in the Recommendation, three should be emphasised. Pursuant to the principle of proportionality and “do no harm”, the choice to use an AI system should be appropriate and proportional to the aim pursued and should not infringe on human rights. For this reason,

“[i]n scenarios where decisions are understood to have an impact that is irreversible or difficult to reverse or may involve life and death decisions, final human determination should apply” (§ 26).

The principle of human oversight and determination requires States to

“ensure that it is always possible to attribute ethical and legal responsibility for any stage of the life cycle of AI systems, as well as in cases of remedy related to AI systems, to physical persons or to existing legal entities” (§ 35).

⁷⁴ Law enforcement is specifically mentioned in the Recommendation, which classifies it as a “human rights-sensitive use case” (UNESCO Recommendation on the Ethics of Artificial Intelligence, § 62).

While humans may decide to delegate certain decisions to AI systems, “an AI system can never replace ultimate human responsibility and accountability” and “[a]s a rule, life and death decisions should not be ceded to AI systems” (§ 36). Finally, pursuant to the principle of responsibility and accountability,

“ethical responsibility and liability for the decisions and actions based in any way on an AI system should always ultimately be attributable to AI actors corresponding to their role in the life cycle of the AI system” (§ 42).

Despite its soft law nature, the Recommendation deserves praise for explicitly considering the possibility of life and death decisions being delegated to AI systems and cautioning against it. While autonomous weapons are not specifically mentioned,⁷⁵ the values and principles of the Recommendation point to the need to maintain human control, oversight and responsibility over this technology, whether used for law enforcement, defence or other purposes. In particular, the requirement that final human determination should apply to life-and-death decisions excludes the use of fully autonomous weapons.

4.3. *INTERPOL and UNICRI Toolkit for Responsible AI Innovation in Law Enforcement*

In June 2023, INTERPOL and UNICRI released a Toolkit for Responsible AI Innovation in Law Enforcement. The foundation of the toolkit is a set of soft law principles

“designed to guide law enforcement agencies across the world in integrating AI systems into their work in ways that align with good policing practices and AI ethics, and respect human rights”.⁷⁶

Based on five core principles, the document argues that

“responsible AI innovation in law enforcement consists of developing, procuring, and using AI systems in a way that is lawful, minimizes harm, respects human autonomy, is fair, and is supported by good governance”.⁷⁷

The document reiterates that, as with any action carried out by law enforcement, the use of AI by police must respect human rights.⁷⁸ For this reason,

⁷⁵ When referring to decisions which may “have an impact that is irreversible or difficult to reverse or may involve life and death decisions”, the Recommendation only mentions social scoring and mass surveillance, stating that AI systems should not be used for such purposes (UNESCO Recommendation on the Ethics of Artificial Intelligence, § 26).

⁷⁶ UNICRI and INTERPOL Toolkit for Responsible AI Innovation in Law Enforcement: Principles for Responsible AI Innovation, p. 3.

⁷⁷ *Ibid.*, p. 6.

⁷⁸ *Ibid.*, p. 8.

“law enforcement agencies should ensure legitimacy, necessity, and proportionality whenever they engage with AI systems in ways that could have an impact on human rights”.⁷⁹

Moreover, AI systems should “not pose a threat to the physical or mental well-being of individuals, their property or the environment”.⁸⁰ AI systems must be “safe, meaning that they include sufficient safeguards to prevent unacceptable harm and minimize unintentional and unexpected harm”.⁸¹ Furthermore, the document stresses the importance of respecting human autonomy, which

“requires that any decisions that impact humans are ultimately taken by humans, especially in a high-stakes context such as law enforcement”.⁸²

Thus, “[e]nsuring human control and oversight of an AI system is [...] essential to upholding human autonomy” and entails “protecting the independence and dignity of every individual or group that interacts with or is affected by the use of an AI system”.⁸³ The need to uphold human control and oversight of AI systems in the law enforcement context is further stressed,

“considering that the work of law enforcement agencies is at the very core of the functioning of society, justice and political systems, and therefore has a significant influence on individuals and their rights”.⁸⁴

The document cautions against the use of “AI systems with a high degree of autonomy—meaning, those which are able to make decisions about the “real world” and act on them without human supervision and intervention”, stating that they “are generally not recommended, as their decisions can have a direct impact on people’s lives”.⁸⁵ Guidance is provided on how law enforcement agencies should ensure human control and oversight: they should

“verify that the AI systems they currently use or intend to use are built with the functionalities needed to ensure that humans remain in charge during use, as well as to certify that the necessary organizational structures are in place to ensure that humans have the last word regarding certain decisions”.⁸⁶

Interestingly, the need to ensure human control and oversight over AI systems is explicitly related to accountability for decisions taken with the assistance of such systems. Essentially, it is argued that the personnel interacting with AI systems will be ultimately

⁷⁹ *Ibid.*, p. 9.

⁸⁰ *Ibid.*, p. 14.

⁸¹ *Ibid.*

⁸² *Ibid.*, p. 20.

⁸³ *Ibid.*

⁸⁴ *Ibid.*, p. 21.

⁸⁵ *Ibid.*

⁸⁶ *Ibid.*, p. 20.

responsible for any decisions taken therein, and as such, they should ensure they maintain control and oversight over the systems.⁸⁷ Moreover, law enforcement should

“ensure that, when AI-supported decisions have an unjust negative impact, those affected are able to formally seek redress through adequate and accessible processes”.⁸⁸

Mechanisms need to be

“put in place to enable stakeholders to clearly determine who is responsible for the decisions made with the support of the AI system, and the consequences of those decisions”.⁸⁹

Unlike the UNESCO Recommendation on the Ethics of AI, the principles put forth by INTERPOL and UNICRI do not explicitly consider the possibility of AI systems being used to make life-and-death decisions. Indeed, the potential use of autonomous weapons in law enforcement was not explicitly considered; neither were the concerns that such use raises. Nevertheless, it can be argued that the emphasis placed on ensuring human control and oversight over AI systems when they make decisions with significant impacts implies that human control, oversight and accountability should be maintained over autonomous weapons used in law enforcement.

4.4. *Council of Europe AI Convention*

Since 2019, the Council of Europe has been exploring the possibility of establishing a legal framework on the development, design and application of AI systems based on human rights, democracy and rule of law standards. Building upon its predecessor's work,⁹⁰ the Committee on Artificial Intelligence (CAI) was tasked with establishing an international negotiation process and elaborating such a framework until November 2023.⁹¹ The CAI brings together representatives of the 46 Member States of the Council of Europe and observer states (Argentina, Australia, Canada, Costa Rica, the Holy See, Israel, Japan, Mexico, Peru, the United States of America and Uruguay), as well as representatives of other Council of Europe bodies, international organisations (including the European Union, the Organisation for Security and Co-operation in Europe, the

⁸⁷ *Ibid.*, p. 21.

⁸⁸ *Ibid.*, p. 32.

⁸⁹ *Ibid.*, p. 34.

⁹⁰ The Ad Hoc Committee on Artificial Intelligence was mandated from 2019 to 2021 to “examine the feasibility and potential elements on the basis of broad multi-stakeholder consultations, of a legal framework for the development, design and application of artificial intelligence, based on Council of Europe's standards on human rights, democracy and the rule of law” (Decision CM/Del/Dec(2019)1353/1.5-app adopted at the 1353rd meeting of the Ministers' Deputies, 11 September 2019).

⁹¹ Council of Europe Committee of Ministers, Decision CM(2021)131-addfinal.

Organisation for Economic Co-operation and Development, and the United Nations Educational, Scientific and Cultural Organisation), the private sector, civil society, and research and academic institutions. The work of the CAI culminated in the landmark adoption of the Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law by the Committee of Ministers of the Council of Europe on 17 May 2024. The Convention is open for signature by the Member States of the Council of Europe, the non-member States which participated in its elaboration and the European Union. It will enter into force three months after five signatories, including at least three Member States of the Council of Europe, express their consent to be bound by the Convention.⁹²

Although a full analysis of the Convention is not the aim of this article, it is necessary to provide a few contextual notes regarding the object and purpose of this treaty. According to its Explanatory Report, the Convention does not set out to regulate all AI systems, focusing instead on those systems which have the potential to interfere with human rights, democracy and the rule of law.⁹³ As such, its provisions “aim to ensure that activities within the lifecycle of artificial intelligence systems are fully consistent with human rights, democracy and the rule of law”.⁹⁴ Importantly, the Convention does not intend to create new human rights obligations but rather “to facilitate the effective implementation of the applicable human rights obligations of each Party in the context of the new challenges raised by artificial intelligence”.⁹⁵ To achieve this, the Convention sets forth legally binding obligations that Parties must give effect to through appropriate legislative, administrative or other measures.⁹⁶ The drafters of the Convention intended for Parties to

“enjoy a certain margin of flexibility as to how exactly to give effect to the provisions of the [...] Convention, in view of the underlying diversity of legal systems, traditions and practices among the Parties and the extremely wide variety of contexts of use of artificial intelligence systems in both public and private sectors”.⁹⁷

However, in giving effect to the Convention, Parties must take into account and tailor measures according to the level of risk posed by AI systems in different contexts of use.⁹⁸ Of particular relevance among the obligations that Parties must give effect to are: the

⁹² Council of Europe AI Convention, Article 30.

⁹³ Council of Europe, Explanatory Report to the Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, 5 September 2024, § 12.

⁹⁴ Council of Europe AI Convention, Article 1(1).

⁹⁵ Council of Europe, Explanatory Report to the Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, 5 September 2024, § 13; Article 21 Council of Europe AI Convention.

⁹⁶ Council of Europe AI Convention, Article 1(2).

⁹⁷ Council of Europe, Explanatory Report to the Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, 5 September 2024, §16.

⁹⁸ Council of Europe AI Convention, Articles 1(2) and 16.

protection of human rights,⁹⁹ respect for human dignity and autonomy,¹⁰⁰ transparency and oversight,¹⁰¹ accountability and remedies,¹⁰² equality and non-discrimination,¹⁰³ privacy and personal data protection,¹⁰⁴ and reliability.¹⁰⁵ To ensure effective implementation, the Convention foresees a follow-up mechanism and international co-operation.¹⁰⁶

Specifically with regard to autonomous weapons and the right to life two considerations should be highlighted. First, the scope of the AI Convention excludes “activities within the lifecycle of artificial intelligence systems related to the protection of [...] national security interests” and “matters relating to national defence”.¹⁰⁷ Thus, the Convention will apply to the design, development and application of autonomous weapons in law enforcement and other domestic settings, but not in national security or defence matters. Arguably, this limitation is a missed opportunity to positively influence the development and use of autonomous weapons in the defence field by clarifying (some of) the requirements such conduct must comply with to respect human rights, especially the right to life. While it is true that IHL is the specialised legal framework to be applied in the conduct of hostilities and that there are ongoing discussions to establish specific rules on autonomous weapons within that area, the difficulty in achieving consensus within the UN Group of Governmental Experts on Lethal Autonomous Weapons Systems suggests that such rules may not emerge in the foreseeable future.¹⁰⁸ In this context, a human rights treaty such as the Council of Europe’s AI Convention could contribute to filling some of the legal gaps pertaining to the development and use of autonomous weapons in armed conflict. As discussed in section 2, human rights law, including the ECHR, continues to apply in armed conflict, with its rules being interpreted in light of IHL.¹⁰⁹ If, for example, the Council of Europe’s AI Convention were to include a provision requiring States to ensure meaningful human control over autonomous weapons used in armed conflict, such provision would have to be interpreted in light of IHL, including the principles of military necessity, distinction, proportionality and precaution. Since there is an ongoing unsettled debate on whether States using au-

⁹⁹ *Ibid.*, Article 4.

¹⁰⁰ *Ibid.*, Article 7.

¹⁰¹ *Ibid.*, Article 8.

¹⁰² *Ibid.*, Articles 9, 14 and 15.

¹⁰³ *Ibid.*, Article 10.

¹⁰⁴ *Ibid.*, Article 11.

¹⁰⁵ *Ibid.*, Article 12.

¹⁰⁶ *Ibid.*, Articles 1(3), 23–26.

¹⁰⁷ *Ibid.*, Articles 3(2) and 3(4).

¹⁰⁸ Reeves, Alcala & McCarthy, 2021, pp. 101–118.

¹⁰⁹ European Court of Human Rights, *Hassan v. the United Kingdom* [GC], App. No. 29750/09, Judgement, 16 September 2014, § 102–104.

onomous weapons can comply with such principles if meaningful human control is not ensured,¹¹⁰ the aforementioned provision would hold particular weight. In this regard, one should, of course, be aware that, pursuant to its Statute, “[m]atters relating to national defence do not fall within the scope of the Council of Europe”.¹¹¹ However, that does not exclude that treaties developed within the Council of Europe apply to matters related to defence. Notably, this is the case of the ECHR, which remains applicable in times of war, although the High Contracting Parties may derogate from some of their obligations “to the extent strictly required by the exigencies of the situation”.¹¹² Indeed, the European Court of Human Rights has extensive case law on the application of the ECHR to State conduct in armed conflict.¹¹³

Secondly, even though the Convention applies to the design, development and application of autonomous weapons in law enforcement and other domestic settings, its text does not explicitly address the grave issues raised by this technology with regard to the right to life.

It is undeniable that the Convention obliges Parties to “adopt or maintain measures to ensure that the activities within the lifecycle of artificial intelligence systems are consistent with obligations to protect human rights”, which obviously includes the right to life.¹¹⁴ Moreover, Parties are obliged to tailor measures to the degree of risk posed by AI systems in different contexts of use, considering in particular the “severity and probability of the occurrence of adverse impacts on human rights [...]”.¹¹⁵ Thus, any Party intending to use autonomous weapons in law enforcement would need to consider the extremely severe risks of unlawful deprivation of life discussed in section 3 of this article. However, Parties to the Convention enjoy a margin of appreciation of these risks. As long as Parties apply the general risk and impact management framework foreseen in

¹¹⁰ See, for example: Amoroso & Tamburrini, 2020, pp. 188–189.

¹¹¹ Statute of the Council of Europe, European Treaty Series, No. 1, 5 May 1949, Article 1(d).

¹¹² ECHR, Article 15.

¹¹³ For a collection of caselaw of the European Court of Human Rights on the application of the ECHR to armed conflicts, see: European Court of Human Rights, 2023. One case of particular relevance to discussions on autonomous weapons is *Streletz, Kessler and Krenz v. Germany* [GC], App. nos. 34044/96, 35532/97 and 44801/98, Judgement, 22 March 2001. The Court considered the use of anti-personnel mines and automatic-fire systems by the German Democratic Republic (GDR) for border control, and held that this practice breached “the obligation to respect human rights and the other international obligations of the GDR, which, on 8 November 1974, had ratified the International Covenant on Civil and Political Rights, expressly recognising the right to life and to the freedom of movement” (§ 73). To reach this conclusion, the Court considered, among other elements, the “automatic and indiscriminate effect” of anti-personnel mines and automatic-fire systems (§ 73).

¹¹⁴ Council of Europe AI Convention, Articles 1(2) and 4.

¹¹⁵ *Ibid.*, Article 1(2) and 16.

the treaty,¹¹⁶ they may reach different decisions on whether the perceived benefits of this technology outweigh its risks, as well as on the conditions for its use. The Convention itself does not explicitly ban the use of autonomous weapons for law enforcement purposes or set forth limits for such use (such as the requirement of meaningful human control). These choices are left to the discretion of the Parties. Thus, a priori, it cannot be said that the obligations set forth by the Convention preclude Parties from using fully or partially autonomous weapons for law enforcement.

Interestingly, the Convention foresees the possibility of Parties imposing bans or moratoriums on certain uses of AI systems which, for example, pose an unacceptable risk to human rights.¹¹⁷ However, it is up to the discretion of each Party to determine what is an unacceptable risk to human rights that would warrant the imposition of a ban or moratorium. Thus, Parties to the Convention may reach different decisions on whether a ban or moratorium on the use of autonomous weapons is necessary.

A similar logic applies to the obligation foreseen in the treaty to ensure that effective procedural safeguards are available where an AI significantly impacts upon the enjoyment of human rights,¹¹⁸ as would be the case of the use of autonomous weapons in law enforcement. According to the Explanatory Report,

“[w]here an artificial intelligence system substantially informs or takes decisions impacting on human rights, effective procedural guarantees should, for instance, include human oversight, including *ex ante* or *ex post* review of the decision by humans”.¹¹⁹

However, what procedural safeguards are required for such impactful AI systems is left to the discretion of Parties. Ultimately, Parties to the Convention may reach different decisions on whether *ex ante* human review of the decision to use force is required.

On the one hand, the open-ended risk-based approach that underlines the Convention makes it suitable to be applied to a broad range of AI systems across the public and private sectors, including systems which have not yet been developed.¹²⁰ On the other

¹¹⁶ *Ibid.*, Article 16.

¹¹⁷ *Ibid.*, Article 16(4).

¹¹⁸ *Ibid.*, Article 15.

¹¹⁹ Council of Europe, Explanatory Report to the Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, 5 September 2024, § 103.

¹²⁰ As recounted in the Explanatory Report, the provisions of the Convention “are purposefully drafted at a high level of generality, with the intention that they should be overarching requirements that can be applied flexibly in a variety of rapidly changing contexts” (§ 49). According to the same document, the open-ended risk-based approach underlining the Convention “is based on the assumption that the Parties are best placed to make relevant regulatory choices, taking into account their specific legal, political, economic, social, cultural, and technological contexts, and that they should accordingly enjoy a certain flexibility when it comes to the actual governance and regulation which accompany the processes” (§ 106).

hand, the fundamental nature of the right to life and the grave risks posed by autonomous weapons arguably call for a red line to be drawn. The Convention thus missed an opportunity to unequivocally establish a ban or a moratorium on the use of autonomous weapons for law enforcement, or to set forth requirements for that use (such as the requirement of meaningful human control).

While it is clear that States must respect the right to life if they intend to use autonomous weapons, the issue at hand is *whether* and *how* they can ensure that such standards are met when using a technology that may not be able to reliably make the value judgments necessary to assess the necessity and proportionality of the use of lethal force. Arguably, the fundamental nature of the right to life calls for an unequivocal statement that decisions to kill should not be delegated to machines; hence, States must not employ autonomous weapons for law enforcement or at a minimum ensure meaningful human control over them. Overall, although the Convention should be praised for being the first AI human rights treaty to ever be adopted, it does not significantly contribute to clarifying whether and how States can ensure that they respect the right to life if they intend to develop or use autonomous weapons for law enforcement and other domestic purposes.

This omission may be explained by the assumption that autonomous weapons will only be used in armed conflict, resulting in a tendency to only consider the right-to-life concerns this technology raises with regard to its military use. To illustrate these observations, we briefly discuss two resolutions of the Parliamentary Assembly of the Council of Europe.

The first resolution, adopted in October 2020, concerns the role of AI in police and criminal justice systems. It notes that AI applications for use by the police and criminal justice systems have been developed and introduced in many countries, and

“include facial recognition, predictive policing, the identification of potential victims of crime, risk assessment in decision making on remand, sentencing and parole, and identification of ‘cold cases’ that could now be solved using modern forensic technology”.¹²¹

The Assembly expressed concerns over the use of such applications, namely in light of lack of transparency, unfairness, responsibility gaps, unsafety and disregard for privacy,¹²² and called on Member States to mitigate the risks of such applications seriously impacting human rights.¹²³ The resolution does not consider the potential use of autonomous weapons by police and the concerns it raises with regard to the right to life.

The second resolution, adopted in January 2023, concerns the emergence of lethal autonomous weapons and their necessary apprehension through European human rights law. This resolution considers the risks associated with the development and use of lethal

¹²¹ Parliamentary Assembly of the Council of Europe, Resolution 2342 (2020) Justice by algorithm – The role of artificial intelligence in policing and criminal justice systems, § 6.

¹²² *Ibid.*, § 7.

¹²³ *Ibid.*, § 9.

autonomous weapons in armed conflict and the need for such systems to comply with IHL and human rights law, especially the right to life. In order to meet the requirement that the right to life be protected by law, the Assembly stressed that the States “must introduce a legal framework defining the limited circumstances in which the use of these weapons is authorised”.¹²⁴ The Assembly further maintained that

“[f]rom the viewpoint of international humanitarian law and human rights law, regulation of the development and above all of the use of [lethal autonomous weapon systems] is therefore indispensable”

and that

“[r]espect for the rules of international humanitarian and human rights law can only be guaranteed by maintaining human control [...] over lethal weapons systems at all stages of their life cycle”.¹²⁵

For this reason, the Assembly supported the adoption of non-binding and binding instruments by the UN Group of Governmental Experts on Lethal Autonomous Weapons Systems and invited its Member States to consider initiating such work at the Council of Europe if a consensus does not emerge within a reasonable period of time in that forum.¹²⁶ This resolution does not consider the potential use of lethal autonomous weapons in law enforcement and other domestic contexts nor the concerns such use raises regarding the right to life.

Given its fundamental importance, protecting the right to life should be an absolute priority when establishing a legal framework for AI based on human rights. Arguably, this includes regulating the potential use of autonomous weapons in and outside armed conflict and carefully considering their serious implications for the right to life.

Although not specifically reflected in the text of the Council of Europe AI Convention, right-to-life considerations with regard to autonomous weapons can—and should—be taken into account by the Parties when implementing the risk-based approach foreseen in the treaty. Given the reporting obligation foreseen in the Convention,¹²⁷ it will be interesting to see if Parties adopt and report on any measures in this regard. Moreover, once the Conference of Parties is convened, it will be interesting to see if right-to-life considerations feature in discussions regarding the interpretation and application of the Convention or possibly regarding the supplementation of the Convention.¹²⁸

¹²⁴ Parliamentary Assembly of the Council of Europe, Resolution 2485 (2023) Emergence of lethal autonomous weapons systems (LAWS) and their necessary apprehension through European human rights law, § 6.4.

¹²⁵ *Ibid.*, § 7.

¹²⁶ *Ibid.*, § 14–18.

¹²⁷ Council of Europe AI Convention, Article 24.

¹²⁸ *Ibid.*, Article 23.

5. Conclusion

The increasing use of AI across most, if not all, domains of human life raises legal and societal concerns that should be addressed proactively. This article does not in any way contest the need to ensure that the use of AI across sectors respects human rights, such as the right to privacy and the right to not be discriminated against. What is argued in this article is rather that the specific concerns raised by the possibility of machines autonomously making the decision to kill, deserves the same careful consideration, if not more. Although errors are inevitable when using any technology, caution must be especially acute when such errors may lead to death.

Considering the implications of autonomous weapons for the right to life, this article analysed the different extents to which four recent initiatives to regulate AI considered the potential delegation of decisions on the use of lethal force to AI. While all initiatives stressed the importance of respecting human rights, none explicitly referred to the right to life or to the development and use of autonomous weapons. Only one initiative, the UNESCO Recommendation on the Ethics of Artificial Intelligence, explicitly considered and cautioned against the possibility of AI systems being used to make life-and-death decisions. Arguably, the fundamental nature of the right to life requires that initiatives to regulate AI carefully consider such possibility and unequivocally state that decisions to kill should not be delegated to machines. As “the supreme right” inherent to every human being whose “effective protection [...] is the prerequisite for the enjoyment of all other human rights”,¹²⁹ it is crucial that the development and use of AI in and outside armed conflict is fully aligned with the negative and positive obligations of States in relation to the right to life.

Discussions on the creation of an international legal framework for AI based on respect for human rights will likely continue and intensify in the future, as technology progresses. If the difficulty in achieving consensus in the UN Group of Governmental Experts on Lethal Autonomous Weapons Systems is indicative, agreeing upon a framework which specifically addresses the concerns raised by autonomous weapons may prove challenging. Nevertheless, as precisely this technology entails the most serious consequences, the hope should be expressed that, regardless of the forum at hand, right-to-life considerations feature more prominently in future discussions to regulate AI.

¹²⁹ Human Rights Committee, General Comment No. 36 on Article 6: right to life, § 2.

References

- Allison, P.R. (2016) *What does a bomb disposal robot actually do?*, BBC, <www.bbc.com/future/article/20160714-what-does-a-bomb-disposal-robot-actually-do> (accessed 15 June 2023).
- Alston, P. (2010) *Interim report of the Special Rapporteur on extrajudicial, summary or arbitrary executions*, UN Doc. No. A/65/321.
- Amnesty International (2015) *Autonomous weapons systems: five key human rights issues for consideration*, <www.amnesty.nl/content/uploads/2015/04/autonomous_weapons_systems.pdf?x41799> (accessed 31 January 2024).
- Amoroso, D., & Tamburrini, G. (2020) 'Autonomous Weapons Systems and Meaningful Human Control: Ethical and Legal Issues', *Current Robotics Reports* 1, pp. 187–194.
- Asaro, P. (2012) 'On banning autonomous weapon systems: human rights, automation, and the dehumanization of lethal decision-making', *International Review of the Red Cross* 94(886), pp. 687–709.
- Blenkinsop, P. (2023) *EU tech chief sees draft voluntary AI code within weeks*, Reuters, <www.reuters.com/technology/eu-tech-chief-calls-voluntary-ai-code-conduct-within-months-2023-05-31/> (accessed 15 June 2023).
- Bo, M., Bruun, L., & Boulanin, V. (2022) *Retaining Human Responsibility in the Development and Use of Autonomous Weapons Systems: On Accountability for Violations of International Humanitarian Law Involving AWS*, Stockholm International Peace Research Institute, October 2022.
- Boutin, B., & Woodcock, T. (2022) *Aspects of Realizing (Meaningful) Human Control: A Legal Perspective*, Research paper series, Asser Institute Center for International and European Law.
- Brehm, M. (2017) *Defending the boundary: constraints and requirements on the use of autonomous weapon systems under international humanitarian and human rights law*, Geneva Academy Briefing N° 9.
- Choudhury, M.R. et al. (2021) *Letter dated 8 March 2021 from the Panel of Experts on Libya established pursuant to resolution 1973 (2011) addressed to the President of the Security Council*, UN Doc. No. S/2021/229.
- Derico, B., & Clayton, J. (2022) *San Francisco to allow police 'killer robots'*, BBC, <www.bbc.com/news/technology-63816454> (accessed 13 June 2023).
- Docherty, B. (2014) *Shaking the Foundations: The Human Rights Implications of Killer Robots*, Human Rights Watch.

- Droege, C. (2007) 'The interplay between international humanitarian law and international human rights law in situations of armed conflict', *Israel Law Review* 40(2), pp. 310–355.
- European Court of Human Rights (2022) *Guide on Article 2 of the European Convention on Human Rights: Right to Life*, <www.echr.coe.int/documents/d/echr/Guide_Art_2_ENG> (accessed 23 June 2023).
- European Court of Human Rights (2023) *Factsheet – Armed Conflicts*, <www.echr.coe.int/documents/fs_armed_conflicts_eng.pdf> (accessed 23 June 2023).
- Fung, B. (2016) *Meet the Remotec Andros Mark V-A1, the robot that killed the Dallas shooter*, The Washington Post, <www.washingtonpost.com/news/the-switch/wp/2016/07/11/meet-the-remotec-andros-mark-v-a1-the-robot-that-killed-the-dallas-shooter/> (accessed 15 June 2023).
- Guterres, A. (2023) *Press Conference: Secretary-General Urges Broad Engagement from All Stakeholders towards United Nations Code of Conduct for Information Integrity on Digital Platforms*, UN Doc. No. SG/SM/21832.
- Heyns, C. (2013), *Report of the Special Rapporteur on extrajudicial, summary or arbitrary executions*, Christof Heyns, UN Doc. No. A/HRC/23/47.
- Heyns, C. (2014a) *Report of the Special Rapporteur on extrajudicial, summary or arbitrary executions*, Christof Heyns, UN Doc. No. A/HRC/26/36.
- Heyns, C. (2014b) *Presentation made at the informal expert meeting organized by the state parties to the Convention on Certain Conventional Weapons 13 – 16 May 2014, Geneva, Switzerland by Christof Heyns, Professor of human rights law, University of Pretoria* United Nations Special Rapporteur on extrajudicial, summary or arbitrary execution, <www.ohchr.org/en/statements/2014/07/presentation-made-informal-expert-meeting-organized-state-parties-convention> (accessed 16 June 2023).
- Heyns, C. (2014c) *Report of the Special Rapporteur on extrajudicial, summary or arbitrary executions*, UN Doc. No. A/69/265.
- Heyns, C. (2016) 'Human Rights and the use of Autonomous Weapons Systems (AWS) During Domestic Law Enforcement', *Human Rights Quarterly* 38(2), pp. 350–378.
- International Committee of the Red Cross (2021) *ICRC position and background paper on autonomous weapon systems*, <www.icrc.org/en/download/file/166330/icrc_position_on_aws_and_background_paper.pdf> (accessed 15 June 2023).
- Jones-Jang, S.M., & Park, Y.J. (2022) 'How do people react to AI failure? Automation bias, algorithmic aversion, and perceived controllability', *Journal of Computer-Mediated Communication* 28(1), pp. 1–8.

- Kiai, M., & Heyns, C. (2016) *Joint report of the Special Rapporteur on the rights to freedom of peaceful assembly and of association and the Special Rapporteur on extrajudicial, summary or arbitrary executions on the proper management of assemblies*, UN Doc. No. A/HRC/31/66.
- Marijan, B. (2023) *Allowing Killer Robots for Law Enforcement Would Be a Historic Mistake*, Centre for International Governance Innovation, <www.cigionline.org/articles/allowing-killer-robots-for-law-enforcement-would-be-a-historic-mistake/> (accessed 31 January 2024).
- Naert, F. (2016) 'Human rights and (armed) conflict', in J. Wouters, Ph. De Man and N. Verlinden (eds.), *Armed Conflicts and the Law*, Oxford – Antwerp, Intersentia, 2016, pp. 187–218.
- Odon, D.I. (2022) *Armed conflict and human rights law: protecting civilians and international humanitarian law*. London: Routledge.
- Office of the United Nations High Commissioner for Human Rights (2011) *International Legal Protection of Human Rights in Armed Conflict*, UN Doc. No. HR/PUB/11/01.
- Office of the United Nations High Commissioner for Human Rights (2020) *UN Guidance on Less-Lethal Weapons in Law Enforcement*, UN Doc. HR/PUB/20/1.
- Potkin, L., & Wongcha-um, P. (2023) *Exclusive: Southeast Asia to set 'guardrails' on AI with new governance code*, Reuters, <www.reuters.com/technology/southeast-asia-set-guardrails-ai-with-new-governance-code-sources-2023-06-16/> (accessed 15 June 2023).
- Reeves, S.R., Alcalá, R.T., & McCarthy, A. (2021). 'Challenges in regulating lethal autonomous weapons under international law', *Southwestern Journal of International Law* 27(1), pp. 101–118.
- Reuters (2017) *Robocop joins Dubai police to fight real life crime*, <www.reuters.com/article/us-emirates-robocop-idUSKBN18S4K8> (accessed 15 June 2023).
- Rodríguez, G. (2023) *SFPD may resubmit proposal for 'killer robots' after policy was blocked, reigniting debate*, ABC News, <www.abc7news.com/killer-robots-san-francisco-sfpd-deadly-force/12761983/> (accessed 13 June 2023).
- Sinder, S., & Simon, M. (2016) *How robot, explosives took out Dallas sniper in unprecedented way*, CNN, <www.edition.cnn.com/2016/07/12/us/dallas-police-robot-c4-explosives/index.html> (accessed 15 June 2023).
- Singapore Home Team Science and Technology Agency (2021) *HTX Ground Robot on Trial at Toa Payoh Central to Support Public Officers in Enhancing Public Health and Safety*, <www.htx.gov.sg/news/media-release-htx-ground-robot-on-trial>

- at-toa-payoh-central-to-support-public-officers-in-enhancing-public-health-and-safety> (accessed 15 June 2023).
- Spagnolo, A. (2017) 'Human rights implications of autonomous weapon systems in domestic law enforcement: sci-fi reflections on a lo-fi reality', *QIL Zoom-in* 43, pp. 33–58.
- Spagnolo, A. (2019) 'What Do Human Rights Really Say About the Use of Autonomous Weapons Systems for Law Enforcement Purposes?', in Carpanelli, E., & Lazzarini, N. (eds.) *Use and Misuse of New Technologies*, Springer, pp. 55–72.
- Taddeo, M., & Blanchard, A. (2022) 'A Comparative Analysis of the Definitions of Autonomous Weapons Systems', *Science and Engineering Ethics* 28(37).
- The Guardian (2014) *Half of US-Mexico border now patrolled only by drone*, <www.theguardian.com/world/2014/nov/13/half-us-mexico-border-patrolled-drone> (accessed 15 June 2023).
- UK Prime Minister's Office (2023) *Press release: UK to host first global summit on Artificial Intelligence*, <www.gov.uk/government/news/uk-to-host-first-global-summit-on-artificial-intelligence> (accessed 15 June 2023).
- U.S. Department of Homeland Security (2022) *Feature Article: Robot Dogs Take Another Step Towards Deployment at the Border*, <www.dhs.gov/science-and-technology/news/2022/02/01/feature-article-robot-dogs-take-another-step-towards-deployment> (accessed 15 June 2023).
- Yeung, K. (2020) 'Introductory Note to the Recommendation of the Council on Artificial Intelligence (OECD)', *International Legal Materials* 59(1), pp. 27–34.

Znanstveni članek

DOI: 10.51940/2024.1.185-212

UDK: 341.3:342.7:004.8

341:623:004.8

Joana Gomes Beirão, Jan Wouters

Na poti do mednarodnega pravnega okvira za smrtonosno umetno inteligenco, ki temelji na spoštovanju človekovih pravic: misija nemogoče?

Avtorja obravnavata morebitno uporabo avtonomnega orožja tako v oboroženih spopadih kot tudi zunaj njih, vključno z odkrivanjem kaznivih dejanj in kazenskim pregonom (na primer v policijskih postopkih). Ta pojav proučujeta z vidika prava človekovih pravic, s posebnim poudarkom na pravici do življenja. Mednarodna skupnost že več kot desetletje razpravlja o tem, ali tehnološki napredek na področju razvoja avtonomnega orožja zahteva oblikovanje novih pravil v okviru mednarodnega humanitarnega prava. Na drugi strani pa je bila obravnava tovrstne tehnologije z vidika prava človekovih pravic doslej omejena, čeprav ima pomembne posledice za pravico do življenja in druge človekove pravice. Vzporedno s temi razpravami se je v zadnjih letih pojavilo več mednarodnih pobud, ki si prizadevajo oblikovati nezavezujoča in zavezujoča pravila za razvoj in uporabo umetne inteligence na podlagi spoštovanja človekovih pravic. Ta članek preučuje štiri take pobude: Priporočilo Organizacije za gospodarsko sodelovanje in razvoj (OECD) o umetni inteligenci, Priporočilo Unesca o etiki umetne inteligence, orodje Interpolu in UNICRI za odgovorne inovacije umetne inteligence pri odkrivanju kaznivih dejanj in kazenskem pregonu ter Konvencijo Sveta Evrope o umetni inteligenci. Avtorja analizira, koliko te pobude zajemajo konkretne pomisleke, ki jih sproža avtonomno orožje.

Ključne besede

avtonomno orožje, umetna inteligenca, človekove pravice, pravica do življenja, odkrivanje kaznivih dejanj in kazenski pregon.

Scientific Article

DOI: 10.51940/2024.1.185-212

UDC: 341.3:342.7:004.8

341:623:004.8

Joana Gomes Beirão, Jan Wouters

Towards an International Legal Framework for Lethal Artificial Intelligence Based on Respect for Human Rights: Mission Impossible?

This article considers the potential use of autonomous weapons both in and outside armed conflict, including in law enforcement. It analyses the phenomenon from the perspective of human rights law, with a particular focus on the right to life. For over a decade, the international community has debated whether technological advances pertaining to the development of autonomous weapons require the establishment of new rules within the framework of international humanitarian law. In contrast, consideration of such technology from a human rights law perspective has been limited, despite its implications for the right to life and other human rights. In parallel, several international initiatives have emerged in recent years aiming to establish non-binding and binding rules for the development and use of artificial intelligence (AI) based on respect for human rights. This article reviews four such initiatives: the OECD Recommendation on AI, the UNESCO Recommendation on the Ethics of AI, the INTERPOL and UNICRI Toolkit for Responsible AI Innovation in Law Enforcement, and the Council of Europe AI Convention. It examines the extent to which these initiatives address the specific concerns raised by autonomous weapons.

Key words

autonomous weapons, artificial intelligence, human rights, right to life, law enforcement.