

ARTIFICIAL INTELLIGENCE IN WARFARE: EXPLORING CRIMINAL RESPONSIBILITY UNDER INTERNATIONAL HUMANITARIAN LAW

Abstract

The increasing deployment of Artificial Intelligence (AI) in warfare has significant implications for International Humanitarian Law (IHL). As AI systems become more autonomous, attributing criminal responsibility for war crimes committed using these systems becomes increasingly complex. This study examines the challenges posed by AI in warfare to the principles of criminal responsibility under IHL. It analyzes the existing legal frameworks, identifies their limitations, and explores the various scenarios in which AI systems may be used in warfare, including autonomous weapons, cyber warfare, and intelligence gathering. A primary challenge in attributing criminal responsibility for AI-related war crimes is the lack of clear guidelines on human oversight and control. This study argues that adapting the current principles of command responsibility and superior responsibility may be necessary to account for AI system use. The study also examines the potential liability of states and non-state actors for AI-related war crimes, considering the implications of AI system use on the principles of distinction and proportionality, and the potential consequences for civilians and civilian objects. This research proposes a framework for attributing criminal responsibility for AI-related war crimes under IHL, emphasizing the need for further research and clear guidelines to ensure accountability for AI use in warfare.

Keywords: Artificial Intelligence, Warfare, International Humanitarian Law, Criminal Responsibility.

1. Introduction

International humanitarian Law (IHL) aims at regulating the conduct of armed conflict¹ by among other things regulating the means and methods of warfare.² As part of this regulation, IHL also holds individuals criminally liable for their acts or omissions which constitute war crimes. This way, the traditional IHL only assigns criminal liability to humans. This is well captured by the International Military Tribunal, Nuremberg which posited that “war crimes are committed by individuals, not by abstract entities.” This assertion underscores the importance of holding individuals criminally liable for their acts or omissions which constitutes war crimes rather than abstract entities like machines. To attribute criminal liability however, it is necessary to prove the mental element of intent and knowledge.³ The Rome Statute of the International Criminal Court, 1998 affirms this stance thus: “a person shall be criminally responsible and liable for punishment for a crime within the jurisdiction of the court only if the material elements are committed with intent and knowledge.”⁴ Intent in

* **Anita Nwotite, PhD**, Lecturer, Department of Clinical Legal Education, Nnamdi Azikiwe University, Awka, Anambra State, am.nwotite@unizik.edu.ng; +2348039574167

****Chiomu Nwabachili, PhD**, Lecturer, Department of International Law and Jurisprudence, Nnamdi Azikiwe University, Awka, Anambra State; co.nwabachili@unizik.edu.ng +2348063772200

*****Maureen Ugwu**, Lecturer, Department of Clinical Legal Education, Nnamdi Azikiwe University, Awka, Anambra State, mo.ugwu@unizik.edu.ng; +2348036860338

¹ N Melzer, International Humanitarian Law- A Comprehensive Introduction (International Committee of the Red Cross, 2016), 16

² Protocol Additional to the Geneva Conventions 1949 and relating to the Protection of Victims of International Armed Conflict, 1977, (AP I) Articles 35; J Meurant, ‘Inter Arma Caritas: Evolution and Nature of International Humanitarian Law (1987) 24 3, *Journal of Peace Research*, Special Issue on Humanitarian Law of Armed Conflict, 237-248, 237; MK Yadav, “International Humanitarian Law as a Part of International Law with Special Reference to its Implementation in the West and South Asian Religion” (2015) 20 (2), 2279-0845, 2279

³ Rome Statute of the International Criminal Court, 1998, Article 30(2) & (3)

⁴ Article 30

this sense means the deliberate act of the perpetrator aimed at achieving a specific result (intent to harm).⁵ The question is: Does the perpetrator intend to commit the crime or to cause harm to protected persons or objects under IHL. The world is nevertheless witnessing a rapid technological advancement⁶ such that machines are now taking up some human activities,⁷ including in the battlefield. In as much as technology has positively impacted on human life and has in fact become an important aspect of human life, as Harvey rightly points out, “the problem began when computers evolved from “thinking” machines (machines that were programmed to perform defined thought processes/computing) into thinking machines (without quotation marks), or Artificial Intelligence (AI).”⁸

AI ‘is the simulation of human behaviour and cognitive processes on a computer.’⁹ The introduction of AI in the conduct of modern armed conflict has no doubt improved the conduct of warfare, particularly in terms of military efficiency and precision. However, this technological advancement challenges the age long principle of IHL which relies on human intent and knowledge for the attribution of criminal liability for war crimes.¹⁰ Hence, the legal issue raised by the AI entities is whether they are subjected to social control like every other legal entity.¹¹ AI systems do not process intentions as human beings do. The decision-making process of AI systems is based on complex data and algorithms. Thus, as AI systems become autonomous and execute military operations without human intervention, the line between human and machine-decision-making become blurred thereby creating accountability gap. Hence, the question that begs for an answer is: when AI systems commit war crimes, who or what is to be held criminally liable – the developer, the manufacturer, the programmer or the AI itself? Against this background, the paper interrogates criminal liability for war crimes committed by AI systems within the context of IHL. By this investigation the paper seeks to contribute to the existing debate on the future of IHL and criminal liability in the age of autonomous weapons. The paper concedes that AI has indeed revolutionized the conduct of modern armed conflict but then that the present IHL framework and *mens rea* mechanism for criminal liability are inadequate to address the peculiarities of AI systems. The paper recommended among others the adoption of specific legal framework and new *mens rea* mechanisms in order to fill the accountability gaps raised by this emerging technological advancement.

2. Conceptual Clarifications

Practice shows that military systems that employ Artificial Intelligence (AI) and machine learning technologies could be used in the conduct of armed conflicts. “AI is the capability of a machine to imitate intelligent behaviour.”¹² It is “the simulation of human behaviour and cognitive processes on a computer.”¹³ In other words, it is the simulation of human intelligence in machines and enabling them to perform functions that basically requires

⁵ Rome Statute (above note 6), Article 30(2)

⁶ Gabriel Harvey, “The Criminal Liability of Artificial Intelligence Entities – from Science Fiction to Legal Social Control,” (2010) vol 1, issue 2, Art. 1, *Akron Intellectual Property Journal*, 172.

⁷ *Ibid*

⁸ *Ibid*.

⁹ *Ibid*

¹⁰ International Military Tribunal, Nuremberg; Rome Statute of the International Criminal Court (above note 6), Article 30; Protocol Additional to the Geneva Conventions, 1949 and relating to the Protection of Victims of International Armed Conflicts, Articles 86 & 87

¹¹ Gabriel Harvey (above note 9), 174

¹² NP Padhy, *Artificial Intelligence and Intelligent Systems 3* (Oxford University Press, 2005) cited in Gabriel Harvey, (above note 7), 172

¹³ Gabriel Harvey, (above note 7), 172

human intelligence. Some notable examples of AI powered weapons are drones, autonomous tanks, smart missiles, robotic soldiers, hypersonic weapons, lethal autonomous weapons and a host of others.

AI has gained increasing prominence as a means of warfare in contemporary armed conflict. The term “means of warfare” is a broad term not just limited to weapons but also includes “all weapon systems, weapon platforms, and associated equipment used to project and deliver force during hostilities.”¹⁴ They refer “to all forms of firearms and firearms system;”¹⁵ or “activities or modes of conduct.”¹⁶

As a branch of public international law that regulates the conduct of warfare,¹⁷ IHL also regulates the means of warfare.¹⁸ Thus, Article 35¹⁹ stipulates that parties are not unlimited in their choice of means and methods of warfare. The violation of IHL in this regard gives rise to individual criminal liability for the perpetrator.

“Criminal liability” in this sense means culpability and accountability of individuals for serious violations of IHL (war crimes). The idea is that a person who committed a crime admits the commission of such an act; and that he deserves to be punished for same so as to serve as deterrence to future violators. IHL has such mechanisms of ensuring accountability and justice for acts and omissions constituting war crimes and other serious violations of IHL.

In IHL, criminal responsibility applies not only to the principal perpetrators (individual criminal responsibility) of war crimes and other violations of International Humanitarian Law;²⁰ but also to superiors (command responsibility) who have ordered the commission of a crime; or who have the duty to act in order to prevent the violation of International Humanitarian Law, but then failed to do so.²¹

Criminal liability is an age long principle of IHL. Hence, the International Criminal Court, Nuremberg noted that crimes are committed by individuals and not by abstract entities. However, IHL traditionally relies on intent, knowledge and control to determine criminal liability. AI’s autonomous and independent decision-making process blurs the line between human and machine decision-making process thereby making it difficult to attribute criminal liability where an AI powered weapon or system commits war crimes or other serious violations of IHL. In essence, the scenario raises such questions as: who is responsible when AI powered weapon commits war crimes, the developer, the commander, the individual deploying it or the AI itself? This accountability gap challenges the traditional principle of criminal liability. The paper explores the complex intersection between AI and criminal liability in IHL and recommends the adoption and development of new framework and mechanisms to address this *lacuna*.

¹⁴ R Liiraja and L Chirop, “Are Enhanced Warfighters Weapons, Means, or Methods of Warfare?”, (2018) 94 *INT’LL STUD.* 161, 178

¹⁵ Khyati Basant, “Means and Methods of Warfare under International Humanitarian Law.” <<https://www.google.com/amp/s/blog/ipleader-in/means-methods-warfare-international-huma...>> accessed 21 May 2023

¹⁶ R Liiraja and L Chirop, (above note 10), 181

¹⁷ N Melzer, 16

¹⁸ AP I, (above note 3), Articles 35 & 36

¹⁹ AP I, (above note 3)

²⁰ It is immaterial that the subordinates acted on the order of a superior.

²¹ AP I, (above note 3), Articles 86(2); & 87(3)

3. IHL and the Regulation of AI-Powered Weapons & Weapon Systems

IHL does not have direct or specific laws regulating AI but then there are general provisions of IHL on the basis of which the regulation of AI can be situated. These provisions basically aim at preventing unnecessary harm to civilians, minimizing harm to civilians, conducting legal review of AI powered weapons, ensuring accountability for AI powered weapons, regulating the development and use of autonomous weapon systems, and maintaining human oversight and control. Below are some of the provisions:

3.1 Protocol Additional to the Geneva Conventions 1949 and relating to the Protection of Victims of International Armed Conflicts, 1977 (hereinafter referred to as AP I)

The AP I embodies some key provisions regulating the use of AI. Article 35 for instance provides:

In any armed conflict, the right of the Parties to the conflict to choose methods or means of warfare is not unlimited.

1. It is prohibited to employ weapons, projectiles and material and methods of warfare of a nature to cause, or to leave the victim with, superfluous injury or unnecessary suffering.

2. It is prohibited to employ methods or means of warfare which are intended, or may be expected, to cause widespread, long-term and severe damage to the environment.

On the other hand, Article 36²² provides that:

In the study, development, acquisition or adoption of a new weapon, means or method of warfare, a High Contracting Party is under an obligation to determine whether its employment would in some or all circumstances, be prohibited by this Protocol or by any other rule of international law applicable to the High Contracting Party.

This provision provides a framework for ensuring that future weapons comply with the dictates and principles of IHL. This framework among other things requires States to assess new means of warfare, determine if their use would be prohibited under IHL or international law; and to ensure that the use of these weapons is not inconsistent with the principles of IHL. This way, accountability and transparency is promoted in the development and employment of weapons, including AI powered weapons.

More so, Article 48²³ requires Parties to the conflict to distinguish at all times between the civilian population and combatants on the one hand; and between civilian objects and military objectives on the other hand; and accordingly to direct attacks only against military objectives. This provision is the fountain of the principle of distinction and in fact the cornerstone of IHL that aims at restricting attacks to legitimate targets and preventing indiscriminate attacks.

Further to that, Article 51²⁴ provides general protection to civilians and civilian population against military operations and against terror; and provides exceptions under which civilians may be targeted; while Article 52²⁵ protects civilian objects and infrastructure, ensures that

²² AP I, (above note 3)

²³ AP I, (above note 3)

²⁴ *Ibid*

²⁵ *Ibid*

only military objectives are targeted. Article 52²⁶ further defines what constitutes military objectives for the purposes of military operations; and presumes civilian use where objects have dual use.

Again, Article 53²⁷ protects cultural objects and places of worship against attacks during armed conflict. Article 54²⁸ protects objects indispensable to the survival of the civilian population²⁹ while Article 55 protects the natural environment against attacks which have severe and long-lasting effects.

Furthermore, Article 57³⁰ requires parties to verify that their target is a military objective before launching an offensive; to choose means and methods of warfare that will avoid civilian harm, to avoid disproportionate attack; and to provide advance notice before an attack is launched. The aim of Article 57³¹ is to minimize harm to civilians and civilian objects. This way, Article 57³² protects civilians and civilian infrastructure, promote careful planning and execution of attacks and to ensure a balance between military necessity and concern for humanity.

Article 58³³ requires Parties to the conflict to take precautions to protect civilians from the dangers arising from armed conflicts.

On the other hand, Article 80³⁴ requires High Contracting Parties to suppress violations of this Protocol;³⁵ to prevent and repress grave breaches of this Protocol;³⁶ to take measures to bring perpetrators to justice;³⁷ and to assist the International Fact-Finding Commission in its work of investigating violations of IHL.³⁸

Article 85³⁹ defines what constitutes serious violations⁴⁰ of the Protocol.

On the whole, the above provisions are meant to prevent unnecessary harm to civilians and to protect civilian objects and infrastructure.

3.2 Protocol III Additional to the CCW 1980

This Protocol restricts the use of incendiary weapons including AI powered weapons.

3.3 The Marten clause:

This is a basic principle of International Humanitarian Law that ensure that the aim of IHL is achieved even where there are no explicit provisions in that regard or there are not rules at all. This clause allows for the adaptation to new situations and technologies AI powered

²⁶*Ibid*

²⁷*Ibid*

²⁸*Ibid*

²⁹ Such as food, water and medical supplies

³⁰ AP I, (above note 3)

³¹*Ibid.*

³²*Ibid.*

³³*Ibid*

³⁴*Ibid.*

³⁵*Ibid* Article 80(1)(a)

³⁶*Ibid*, Article 80(1)(b)

³⁷*Ibid*, Article 80(1)(c)

³⁸*Ibid* , Article 80(1)(d)

³⁹ AP I, (above note 3)

⁴⁰ Making civilians the object of attack; launching indiscriminate attacks against civilians and civilian objects; and launching of attack against civilian object with the knowledge that it will cause disproportionate harm

weapons. The clause is a mechanism that aims at ensuring that emerging technologies comply with the principles of IHL such as the principle of distinction, proportionality, humanity, among others and accountability in the development and employment of new technologies including AI in the conduct of warfare. The clause provides:

In cases not covered by this Protocol, or by other international agreements, civilians and combatants remain under the protection and authority of the principles of international law derived from established custom, from the principles of humanity, and from the dictates of public conscience.⁴¹

The Marten clause encapsulates some elements necessary for regulating AI powered weapons such as the principle of humanity preventing unnecessary suffering; the dictates of public conscience setting the standard of public morality; and established and recognized customary international law to fill in the *lacuna*.

3.4 Other Initiatives and Guidelines

Aside these provisions drawing directly from the Geneva regimes regulating the conduct of armed conflicts, there are also various initiatives and guidelines proposed or adopted to that effect. Below are some of them:

- (a) International Committee of the Red Cross Guidelines on the Use of Artificial Intelligence in Warfare: the Guidelines provides practical guidelines on the use of AI powered weapons in IHL while focusing on human responsibility, accountability and transparency.
- (b) The UN Secretary-General's Report on Autonomous Weapons which proposed a ban on complete autonomous weapons.
- (c) European's Parliament's Resolution on Autonomous Weapons which also calls for a total ban on complete autonomous weapons.
- (d) NATO's AI Strategy which emphasizes responsible use of AI in defense in accordance with the principles of IHL.
- (e) UN Group of Government Experts (GGE) on Lethal Autonomous Weapons (LAWS) which developed guidelines for the development and use of lethal autonomous weapons.
- (f) DoD Directive 3000.09 (US) which establishes guidelines for the development and use of autonomous systems in warfare.

The aim of these guidelines and initiatives is to ensure responsible use of AI and the use of AI powered weapons comply with the basic tenets of IHL.

4. Basis for Adjudging the Applicability or Otherwise of Means of Warfare in IHL

As a body of law which among others regulate the means of warfare, IHL adjudges the applicability or otherwise of the means of warfare on the bases of the following principles of IHL:

4.1 Principle of Distinction

This principle is the cornerstone of IHL founded on Article 48.⁴² The principle requires that parties to the conflict distinguish between civilians and combatants on the one hand and

⁴¹ Hague Convention II on the Laws and Customs of War on Land, 1899, Preamble; Geneva Conventions 1949, common Articles 1,2 &3; AP I, (above note 9) Article 1; Protocol Additional to the Geneva Conventions 1949 and relating to the Protection of Victims of Non-International Armed Conflicts, 1977 (AP II), Article 1;

⁴² V Chetail, 'The Contribution of the International Court of Justice to International Humanitarian Law' (2003 85 (850) *IRRC* 235; Nuclear Weapon Case, 78-79

civilian objects and military objectives on the other hand and to direct attacks only against combatants and military objectives.⁴³ The principles also abhors indiscriminate attacks and the use of means of warfare that cannot discriminate between civilians and combatants and between civilian objects and military objectives.⁴⁴

4.2 Principle of Proportionality

This has a close relationship with the principle of military necessity⁴⁵ and requires that the expected harm to civilians and civilian objects is not excessive in relation to the military advantage sort to be achieved. In other words, the principles prohibits attacks that are 'expected to cause incidental loss of civilian life, injury to civilians, damage to civilian objects, or a combination thereof, which would be excessive in relation to the concrete and direct military advantage anticipated.'⁴⁶ Hence, even where a target is a military objective, the attack must nevertheless not be disproportionate to the military advantage anticipated. The aim of this principle is to maintain a balance between military necessity and humanity.⁴⁷ The implication is that IHL envisages collateral damage in the conduct of warfare but then it frowns at excessive harm when placed side by side with the military advantage sought to be achieved.

4.3 Principle of Humanity (Unnecessary Suffering)

The principle of humanity is also known as the principle of unnecessary suffering. The principle requires parties to employ means of warfare that does not cause unnecessary harm to combatants.

4.4 Military Necessity

The only object which parties must seek to achieve in any armed conflict is to weaken the military strength of the enemy⁴⁸ and nothing more. This is exemplified in the principle of military necessity which seeks to limit the use of means of warfare to achieving legitimate military objective or advantage. This principle as noted earlier is tied to the principle of proportionality.⁴⁹

4.5 Principle of Precaution

This principle requires parties to the conflict to assess, analyze, evaluate and decide on the use of a particular means of warfare before launching an attack, so as to minimize harm to civilians.⁵⁰

On the issue of assessment, parties are required to determine the effects and potential harm of the means of warfare they intend to adopt. The means of warfare must also be weighed to ensure their compliance with the basic principles of IHL,⁵¹ and their necessity and proportionality must also be put into consideration.⁵² The above decision on the choice of the appropriate means of warfare to be employed by parties to the conflict must be taken in

⁴³ AP I, (above note 3), Article 48

⁴⁴ Anita Nwotite, 'Protecting Civilians in the Russia-Ukraine War: The Obligation to respect the Basic Principles of International Humanitarian Law' (2022) 2 (2) *UCC Law Journal*, 128

⁴⁵ Brown, cited in Anita Nwotite (above note 40), 129

⁴⁶ AP I, (above note 3), Article 57(2)(b)

⁴⁷ Anita Nwotite (above note 45), 129

⁴⁸ St. Petersburg Declaration, 1868, Preamble

⁴⁹ Brown, cited in Anita Nwotite (above note 45), 129

⁵⁰ AP I, (above note 3), Article 57

⁵¹ Principles of distinction, proportionality, humanity and necessity.

⁵² AP I, (above note 3), Article 57

consultation with legal advisers, military commanders and other experts. On the basis of the above considerations, parties shall make informed decision on the applicable means of warfare.

4.6 Article 35 of the AP I

Article 35⁵³ is to the effect that in any armed conflict, the right of the Parties to the conflict to choose methods or means of warfare is not unlimited. As such it is prohibited to employ weapons, weapon systems, or method of warfare that is expected or that has the nature of inflicting unnecessary or disproportionate harm or damage to protected persons and objects or the natural environment.

4.7 Article 36 of the AP I

In addition to the above principles, IHL also requires High Contracting Parties to comply with the provisions of Article 36⁵⁴ to the effect that in the study, development, acquisition or adoption of a new weapon, means or method of warfare, they are obliged to determine whether its employment would in some or all circumstances, be prohibited by this Protocol or by any other rule of international law applicable to the High Contracting Party. The above rules together constitute the basis for adjudging the applicability or otherwise of the employment of any means of warfare. These, together constitute the yardstick for assessing the legality or otherwise of the employment of a particular means of warfare. Suffice it to say that these standards also applies to AI-related weapons as they must comply with basic principles of IHL.⁵⁵ More so, as an emerging area in IHL, States are required to ensure that in their development of any weapons or weapon system that they comply with the relevant provisions of IHL and be bound by other provisions of international treaties and Conventions to which they are Parties.⁵⁶

5. Criminal Liability in IHL

IHL sets out the standard of conduct required of parties during the conduct of warfare⁵⁷ so as to prevent impunity. The violation of these rules may constitute war crimes⁵⁸ or other serious violations of IHL depending on the context of the violations. War crime are grave breaches of the law and customs of war,⁵⁹ and the perpetrators are criminally held liable for their acts or omissions constituting war crimes.

Criminal liability arises when individuals or groups violate the rules of IHL giving rise to such crimes as war crimes, crimes against humanity and genocide.⁶⁰ These are also referred to as grave breaches of the Geneva Conventions⁶¹ or the Additional Protocol I.⁶² In other words, criminal liability refers to accountability of individuals⁶³ or bodies for acts or omissions

⁵³ AP I, (above note 3)

⁵⁴ *Ibid.*

⁵⁵ *Ibid.* Article 35, 36, 48, 52, 57

⁵⁶ Weizmann, Nathalie and Costas Trascasas, Milena, "Autonomous Weapon Systems under International Law," Academy Briefing, Issue 8, 2014, 13 and 27

⁵⁷ AP I, (above note 3), Articles 35, 36, 48, 52, 57

⁵⁸ Rome Statute, (above note 4), Article 8(2)(a)(iii)(iv)(vi) & 8(2)(b)(xxv)

⁵⁹ Which includes wilful killing, torture, inhuman treatment, destruction of property, willfully causing great suffering, attacks on civilians and civilian objects; and cultural property and places of worship

⁶⁰ Rome Statute, (above note 6), Article 8(2)(a)(iii)(iv)(vi) & 8(2)(b)(xxv)

⁶¹ Common Article 3; Geneva Convention I, Article 50; Geneva Convention II, Article 51; Geneva Convention III, Article 130; and Geneva Convention VI, Article 147

⁶² Articles 11 and 85

⁶³ AP I, (above note 3) Articles 86 & 87; Statute of the International Criminal Tribunal for the former Yugoslavia, Article 7(1) & (3)

constituting grave breaches (war crimes) and other serious violations of IHL. Criminal liability is hinged on the fact that individuals bear personal responsibility for their actions. Liability may be as a result of direct⁶⁴; or indirect commission;⁶⁵ aiding and abetting;⁶⁶ and joint criminal enterprise (participation in joint criminal outfit or sharing a common purpose).

The jurisprudence of criminal liability in IHL is well enunciated in the decisions of the various international criminal tribunals⁶⁷ and courts;⁶⁸ and is reinforced by treaty provisions.⁶⁹

Criminal liability exists at 2 levels, to wit, individual criminal liability,⁷⁰ and command responsibility.⁷¹

On the issue of individual responsibility, individuals are held personally criminally liable for their acts or omissions constituting grave breaches or serious violation of IHL, whether committed directly by them or through other persons by instigation, aiding and abetting or ordering.⁷²

On the other hand, commanders and other superiors are also held criminally liable for acts or omissions of their subordinates or persons under their effective control that constitutes grave breaches or serious violation of IHL.⁷³ Command responsibility may be direct or indirect.⁷⁴ For direct command responsibility, it suffices that the military commander or superior orders the act or omission constituting the violation of IHL.⁷⁵ Indirect command responsibility arises from the military commander's or superior's failure to prevent, punish, or report the violation of IHL to the competent authority.⁷⁶ On the whole, the aim of criminal liability is to ensure accountability, prevent impunity, deter future offenders and ensure justice for the victims of IHL violations.

⁶⁴ Committing a crime directly

⁶⁵ Committing a crime through another person(s)

⁶⁶ Supporting the commission of a crime

⁶⁷ International Military Tribunal, Nuremberg, International Criminal Tribunal for the former Yugoslavia, Tokyo Tribunal, among others

⁶⁸ International Criminal Court

⁶⁹ AP I, (above note 3), Articles 86 & 87; Statute of the International Criminal Tribunal for the former Yugoslavia, Article 7(1) & (3)

⁷⁰ AP I, (above note 3), Article 86 & Statute of the International Criminal Tribunal for the former Yugoslavia, Article 7(1)

⁷¹ AP I, (above note 3), Article 87; Statute of the International Criminal Tribunal for the former Yugoslavia, Article 7(3)

⁷² AP I, (above note 3), Article 86; Statute of the International Criminal Tribunal for the former Yugoslavia, Article 7(1)

⁷³ AP I, (above note 3), Article 87

⁷⁴ *Celebici Camp, Prosecutor v Delalic (Zejnil) and Ors*, Trial Judgment, Case No.IT-96-21-T, ICL 95 (ICTY1998), 16th November 1998, Trial Chamber II (ICTY), para. 333

⁷⁵ Rome Statute of the International Criminal Court, (above note 4), Article 25(3)(b)

⁷⁶ AP I, (above note 3), Article 87; Statute of the International Criminal Tribunal for the former Yugoslavia, Article 7(3); Rome Statute of the International Criminal Court, (above note 6), Article 28; Levine II, James D, "The Doctrine of Command Responsibility and its Application to Superior Civilian Leadership: Does the International Criminal Court have the correct Standard?" (2007) vol 193 *Mil. L. Rev.*, 53

5.1 Requirements for Imposing Criminal Liability under the Rome Statute of the International Criminal Court, 1998

To determine criminal liability generally, the maxim *actus reus non facit reum nisi mens sit rea* holds sway.⁷⁷ In this sense, attributing criminal liability requires establishing the physical element, that is the criminal conduct (*actus reus*); and the mental element, that is knowledge or intent in relation to the criminal conduct (*mens rea*) of the crime.⁷⁸ This standard also apply *mutatis mutandi* under the Rome Statute of the International Court.⁷⁹

The Rome Statute of the International Criminal Court established the International Criminal Court⁸⁰ and vests it with the jurisdiction over, among others, war crimes.⁸¹ For the purposes of determining criminal liability for war crimes, for instance, the Rome Statute⁸² provides:

- (1) Intent and knowledge: A person shall be criminally responsible and liable for punishment for a crime within the jurisdiction of the court only if the material elements are committed with intent and knowledge.
- (2) Intent: For the purposes of this Article, a person has intent where:
In relation to a consequence, that person means to cause that consequence or is aware that it will occur in the ordinary course of events.
- (3) Knowledge: For the purposes of this Article “knowledge” means awareness of a circumstance or consequence when:
He or she is aware of the substantial likelihood that the circumstance exists or that the consequence will occur.
Exceptions: this Article does not affect the normal operation of the rule of international law that require other forms of mental element.

The *actus reus* and the *mens rea* must exist conjunctively for there to be a criminal, otherwise, no crime is committed and no criminal liability can be imposed.⁸³

5.1.1 *Actus Reus*

The *actus reus* is the external element of the crime otherwise called the criminal conduct or the physical element of a crime.⁸⁴ This is expressed by acts or omissions.⁸⁵ However, sometimes, other external elements may be required in addition to the physical conduct or criminal conduct. Such additional elements may include the specific results of that conduct and the circumstances surrounding the conduct.⁸⁶

5.1.2 *Mens rea*

This is the mental element or internal element of the crime.⁸⁷ This exists at different levels.⁸⁸ While knowledge represent the highest level of *mens rea*⁸⁹ (and is sometimes accompanied by the requirement of intent);⁹⁰ negligence or strict liability represent the lower levels.⁹¹

⁷⁷ Rome Statute (above note 4), Article 30

⁷⁸ *Ibid*, Article 30(1); Gabriel Harvey, (above note 7).

⁷⁹ Article 30

⁸⁰ Article 1

⁸¹ Rome Statute (above note 4), Article 8

⁸² Article 30

⁸³ Gabriel Harvey, (above note 7), 177

⁸⁴ *Ibid*.

⁸⁵ Walter Harrison Hatcher, *The Physical Element of Crime*, (1934) 39 Dick. L. REV., 95; Michael Moore, *Act and Crime: The Philosophy of Action and its Implications for Criminal Law* (1993), 156-157

⁸⁶ Gabriel Harvey, (above note 7), 177

⁸⁷ *Ibid*.

⁸⁸ Joshua Dressler, *Cases and Materials on Criminal Law* (4th edn, West Academic Publishing, 2007), 980-981 cited in Gabriel Harvey, (above note 7)

⁸⁹ *Ibid*

5.1.3 Intent

This is the deliberate action of the accused aimed at achieving a specific result.⁹² The question here is whether the accused person intended to commit the crime or to cause harm to protected persons or objects under IHL (intent to harm).

5.1.4 Knowledge

This is the accused awareness of the result or the circumstances of his action or omissions constituting a war crime.⁹³ The question is whether the accused knew or should have known that his acts or omissions could cause harm to persons or objects protected under IHL (knowledge of harm or consequences of the actions). Therefore, the implication of Article 30⁹⁴ is that, the mental (*mens rea*) element of knowledge and intent must exist in order to impose criminal liability.⁹⁵ By so doing, it has been opined that the International Criminal Court sets the level of *mens rea* required to ground a war crime high,⁹⁶ while leaving out the other aspect of *mens rea*.⁹⁷ This does not however, affect the general rules of international law that requires that other forms of mental elements such as negligence and recklessness, be proved in addition to intent and knowledge.⁹⁸ In other words, *actus reus* and *mens rea* are the only two requirements needed to impose criminal liability under the Rome Statute.⁹⁹ These requirements must exist conjunctively, otherwise no crime is committed and by implication, no criminal liability can be imposed.¹⁰⁰ These standards apply irrespective of whether the entity in question is a human beings, corporations or AI.¹⁰¹

6. AI and Criminal Liability

Attributing criminal liability, as earlier noted requires among others, the proof of *actus reus* and *mens rea*- that is, the physical criminal conduct and the deliberate act(s) of the actor aimed at causing harm.¹⁰² By the clear provisions of Article 30(2) “a person has intent where, in relation to a consequence, that person means to cause that consequence or is aware that it will occur in the ordinary course of events.” It is pertinent to note that intent is a conscious human action and not an action of abstract entities like machines. Thus, the traditional IHL framework and liability approach focuses basically on human agency and does not envisage criminal liability for machines. To that effect, individuals are held personally criminally liable for their acts or omissions constituting war crimes or other serious violations of IHL.¹⁰³ On the other hand, military commanders and other superiors are held criminally responsible for the acts of their subordinates or persons under their control.¹⁰⁴ This is the standard. The

⁹⁰ L.L.I. J. Edwards, “The Criminal Degrees of Knowledge” (1954) 17 *MOD. L. REV.*, 294 cited in Gabriel Harvey, (above note 7), 178

⁹¹ Jeremy Horder, “Strict Liability, Statutory Construction and the Spirit of Liberty”, (2002) 118 *Law Q. REV.*, 458

⁹² Rome Statute (above note 4), Article 30(2).

⁹³ *Ibid*, Article 30(3).

⁹⁴ Rome Statute, (above note 4).

⁹⁵ *Ibid*, Article 30(2) &(3).

⁹⁶ Such as negligence and recklessness.

⁹⁷ Lasva Valley, Prosecutor v Blaskic (Tihomir), Appeal Judgment, Case No. IT-95-14-A, ICL 34 (ICTY 2004), 29th July 2004, Appeals Chamber (ICTY), para. 42.

⁹⁸ Rome Statute (above note 4), Article 30, Article 30.

⁹⁹ *Ibid*

¹⁰⁰ Joshua Dressler, cited in Gabriel Harvey, (above note 7), 177.

¹⁰¹ *Ibid*.

¹⁰² Rome Statute (above note 4), Article 30(2).

¹⁰³ AP I, (above note 3), Article 86

¹⁰⁴ *Ibid*, Article 87

question that begs for an answer is: whether AI system have the capacity for intent as to make them criminally liable for their acts or omission constituting war crimes?

AI systems are a species of autonomous weapon systems that are able to select and engage targets without human intervention.¹⁰⁵ Although, some have argued that autonomous weapon systems (including AI) will not be able to do what their human programmers do not want them to do,¹⁰⁶ it is nevertheless a common knowledge that autonomy is an important characteristic of AI systems.¹⁰⁷ In other words, AI systems have an independent decision-making process like humans do. They function by combining large amounts of data with complex algorithms to enable them learn, reason, and perform tasks that are typically required of human intelligence. This notwithstanding, AI systems do not process consciousness, emotions or intentions as human beings would do. AI's decisions are based on algorithms, data and programming rather than human-like intentions. This is rightly captured in the following words: "An autonomous weapon system, no matter how smart they are or the level of legal compliance attained, is an object."¹⁰⁸ More so, the International Military Tribunal captures the situation aptly thus: "war crimes are committed by individuals not by abstract entities." Over and above all, the elements of *mens rea* (intent and knowledge) provided under the Rome Statute¹⁰⁹ for grounding criminal liability, are human attributes. Suffice it to say that AI systems cannot be assigned human character or be imposed with a legal obligation due only to humans.¹¹⁰ The decisions of AI systems are determined by programming and input data which are not to be equated with human intent. Without human intent it becomes difficult or almost impossible to hold AI systems criminally accountable for war crimes as the lack of requisite *mens rea* challenges the existence of the crime.¹¹¹ Against this background, Harvey, in his perpetrator-via-Another Liability Model posited that "the AI entity is considered an innocent agent."¹¹² As such a machine is a machine and not a human entity.¹¹³ However, some quarters are of the view that smart machines should be assigned legal personality so that they could be held liable for the violations of IHL.¹¹⁴ It has been argued further that combatant and autonomous weapon systems which also includes AI systems cannot be equated since they are not in the same category.¹¹⁵ Although AI entities have some capabilities particularly of autonomy and independent decision-making process, these though important are not sufficient basis for imposing criminal liability since they are likened to the capabilities of a mentally limited person such as a child,¹¹⁶ mentally

¹⁰⁵ Human Rights Watch cited in Xvavier J Ramirez, "Requirement of Mens Rea for War Crimes in the Light of the Development of Autonomous Weapons Systems, (2021) vol 21 *Anu.Mex.Der.inter*

¹⁰⁶ Implying that they are not completely autonomous and as such remain predictable

¹⁰⁷ Pilling Michael, "Issues regarding the Future Application of Autonomous Systems to command and Control (C2)" DSTO-TR-3112, DSTO Defence Science and Technology Organisation, 2015, 1 and 2

¹⁰⁸ Xvavier J Ramirez, "Requirement of Mens Rea for War Crimes in the Light of the Development of Autonomous Weapons Systems, (2021) vol 21 *Anu.Mex.Der.inter*

¹⁰⁹ Article 30

¹¹⁰ Autonomous weapon system which also includes AI systems.

¹¹¹ Xvavier J Ramirez, (above note 109).

¹¹² Gabriel Harvey, (above note 7) 179; Lawrence B. Solum, "Legal Personhood for Artificial Intelligences" (1992) N.C.L REV., 1231.

¹¹³ *Ibid.*

¹¹⁴ Calverley, David, "Imagining a Non-Biological Machine as a Legal Person, (2008) vol. 22, *AI & Soc.*, 523; Krishnan, Armin, *Killer Robots, Legality and Ethnicity of Autonomous Weapons* (New York: Routledge, 2009), 105.

¹¹⁵ M Sassoli, "Autonomous Weapons and International Humanitarian Law: Advantages, Open Technical Issues to be Clarified," (2014) *Int'l L. Stud.*, vol. 90, 323.

¹¹⁶ *Maxey v United States*, 30 App. D.C. 63, 80 (App. D.C. 1907).

incompetent person¹¹⁷ or of a person lacking a criminal mind.¹¹⁸ Hence, there is no basis for assigning criminal liability to AI systems.

7. Challenges to Assigning Criminal Liability to AI Systems

Notwithstanding the novelties AI has introduced in the conduct of modern warfare, it also created some new challenges for IHL in terms of accountability for war crimes. Below are some of the challenges:

7.1 Lack of Human Intent

As noted earlier, establishing the mental element of a crime is important to attributing criminal liability. However, AI systems are not driven by human intent but by the combination of complex data and algorithms.¹¹⁹ With the absence of the mental element here, it becomes difficult or almost impossible to attribute liability to AI systems.

7.2 Autonomous Decision-Making (Lack of Direct Human Intervention)

AI system's complex and algorithms decision-making process also poses a challenge to establishing the mental element of the crime for the purposes of determining criminal liability. They lack direct human intervention. They have autonomous and independent decision-making process thereby making difficult or impossible to link their decisions with any human actor.

1.3 Difficulty in Establishing Causation

To attribute criminal liability, it is imperative to link an act or omission with the harm or damage caused. However, the complex decision-making process, autonomy, technical complexities, actions of many actors and systems, lack of transparency, among others makes it difficult to establish causal link between the actions and resultant harm or damage in AI-related crimes.

7.4 Lack of the Prerequisite Expertise

Courts and investigators/prosecutors lack the prerequisite expertise for dealing, investigating and prosecuting AI related crimes.

7.5 Inadequate Legal Framework/Mechanisms for AI-Related Cases in IHL

There are inadequate legal frame work and mechanisms for accountability for AI-related matters. The present IHL frame work and accountability mechanisms are inadequate to address the peculiarities of AI-powered weapons.

8. Conclusion and Recommendations

The paper examined criminal liability in AI within the context of IHL. The paper noted that AI has no doubt improved the conduct of armed conflict particularly in terms of precision and efficiency of military operations. These notwithstanding, the paper observed that AI creates accountability gap thereby making it difficult or almost impossible to establish the *mens rea* necessary to attribute criminal liability in the event of war crimes or other serious violation of IHL. To address this challenge, the paper recommended as follows:

9.1 Development of New and Specific Legal Regime

Addressing the challenge posed by AI especially with respect to criminal liability, requires the development and adoption of new and specific legal framework that will address the

¹¹⁷ *Johnson v State* 142 Ala. 70, 71 (1904).

¹¹⁸ *United States v Bryan*, 483 F. 2d 88, 92 (3d Cir. 1973).

¹¹⁹ Xavier J Ramirez, (above note 109).

peculiarities of AI-related weapons as a means of warfare as the present IHL regime is inadequate in this regard.

9.2 Establishment of AI-Specific Standard of *Mens Rea*

There is also the need to establish AI-specific standard of *mens rea* required to prove AI-related violations of IHL. This is by for instance attributing criminal liability for the conduct by placing the mens rea on the human actors¹²⁰ surrounding the operation of the AI. This will help address the characteristic challenge of lack of human intent and motivations of AI-related weapons.

9.3 Adoption of Specific Evidentiary Rules

As an evolving means of warfare, there is the need to adopt a specific evidentiary rules for AI-related cases. This will help address the gap in proving and attributing criminal liability in AI-related cases.

9.4 Focusing on Human oversight and Control

Focusing on human oversight and control in the development and deployment of AI is also necessary to address the causal link for the purposes of attributing criminal liability.

9.5 Enhancing the relevant Technical Expertise

Part of the challenge of AI is lack of the relevant technical expertise required to investigate and prosecute AI-related cases. However, this could be addressed by enhancing technical expertise in investigations and prosecutions.

9.6 Facilitating International Cooperation and Encouraging Responsible use of AI in IHL

The international community must also engage in global dialogue to ensuring responsible development and use of AI in IHL through the provisions of guidelines and initiatives in that regard.

9.7 Continuous Monitoring and Evaluation

To address, the accountability gap in AI powered weapons, it is also important to ensure continuous monitoring and evaluation of the employment of AI in IHL, adoption of laws and practices and the modifications of existing laws in that regard.

¹²⁰ Such as programmers and manufacturers