



AHMADU BELLO UNIVERSITY LAW JOURNAL

FACULTY OF LAW

AHMADU BELLO UNIVERSITY, ZARIA - NIGERIA.

website: <https://abulj.org.ng> | e-mail: abulj@abu.edu.ng

APPRAISING THE STATUS OF COMBATANTS AND PRISONERS OF WAR STATUS IN ARTIFICIAL INTELLIGENCE TECHNOLOGY IN INTERNATIONAL HUMANITARIAN LAW

ELIZABETH JALO- OKOTIE*

Abstract

This paper examines the evolving role of artificial intelligence (AI) in armed conflict and its implications under International Humanitarian Law (IHL). As militaries increasingly feature AI technologies into warfare, questions arise regarding the applicability of IHL principles—such as distinction, proportionality, and necessity—to autonomous systems. While AI has the potential to enhance operational efficacy and improve targeting accuracy, it also poses significant challenges, including issues of responsibility, ethical considerations, and the risk of unlawful attacks on civilians. The impact of AI on armed conflict necessitates careful scrutiny to ensure compliance with humanitarian standards. Prospective developments in AI warfare underscore the urgent need for robust legal frameworks that adapt to technological advancements while safeguarding human rights. This analysis highlights the necessity for comprehensive legal reviews to address the complications of AI in military operations, ensuring that humanitarian principles remain central to the conduct of armed conflict. Ultimately, ongoing dialogue among legal experts, military leaders, and ethicists is essential to navigate the implications of AI technologies, promoting responsible use in alignment with IHL.

Keywords: Artificial Intelligence, State Liability, Combatants, Prisoners of war and International Humanitarian Law.

1.1 Introduction

Since time immemorial, humanity has shown great disposition to warfare (armed conflict¹) and this continues till the present day. From the beginning of civilization on earth, military forces have been engaged in several wars for one reason or the other. The wars are now mostly narrowed to the territory of one state, with devastating impact on the civilians. The known horrors of armed conflict are indescribable. Loss of lives, properties of all sorts and damages to the natural environment are common occurrences. The impact of these losses are severe on those still living; ranging from psychological, social and physical suffering. This is so, despite the fact that International Humanitarian Law (IHL) ever since its inception to the present time has set limits on the rights of belligerents to the choice of means and method of warfare.² The law also prohibits destruction of property and objects belonging to the civilians,³ this also includes objects belonging to the environment.⁴ The objective of this study is to critically examine the evolving status of combatants and prisoners of war (POWs) within the context of artificial intelligence (AI) technologies under international humanitarian law (IHL). It aims to analyze the legal and ethical implications of using autonomous weapon systems and other AI-driven technologies in armed conflicts, particularly with regard to determining combatant status, lawful targeting, and the treatment and classification of captured persons. The study seeks to assess whether existing IHL frameworks are adequate to address the challenges posed by AI in warfare or

* Lecturer, Dept of Public Law Faculty of law, Taraba State University, Jalingo, Taraba State.

¹ ‘Armed Conflict’ is preferred in this discourse, because it did not give room for argument against the applicability of law of war (Humanitarian law), for even where one of the parties to a conflict denies the existence of a state of war, once troops engaged, it is armed conflict. See Commentary on Geneva Convention 1 p. 32 where it was stated that ‘Armed conflict’ is ‘generic’ and make argument less easy that there is no war.

² Article 22 of Regulation annexed to Hague Convention IV of 1907, Article 35 of 1977 AP I to GC of 1949.

³ Article 23 (1) (g) Hague Convention IV of 1907

⁴ Article 23 (1) (g) Hague Convention IV of 1907

whether new legal instruments are necessary to ensure compliance with humanitarian principles.

1.2 Artificial Intelligence

By copying human intelligence, Artificial Intelligence enables machines and computer systems to learn from, act, and perform tasks like a human. The definition of AI, as proposed within the European Commission's Communication on AI is as follows:

“Artificial intelligence (AI) refers to systems that display intelligent behaviour by analyzing their environment and taking actions – with some degree of autonomy – to achieve specific goals...”⁵

AI technology is now being used in various human endeavors, including the military, with a key role in current and future military applications. AI technologies holds prospect to revolutionize the face of warfare as new means and methods of warfare are increasingly unfolding. But the path to the future of technological advancements is not always clear as AI technologies are rapidly evolving.

Using AI in military operations raises ethical and legal challenges. One of the major challenges is that advances in AI technology are altering the abilities of weapons in modern warfare, lessening or eliminating the need for human oversight and participation. AI can perform most warfare tasks without any human participation; AI could empower fully military weapon systems in algorithmic combat to carry out attacks, or to take decisions autonomously without human involvement. By their nature, such weapons cannot guarantee what will happen when encountering a novel situation jeopardizing non-combatants. If artificial intelligence can perform any warfare task without human

⁵ Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions on Artificial Intelligence for Europe, Brussels, 25.4.2018 COM (2018) 237

participation, humanitarian debates arise as to whether this technology can distinguish between military and civilian targets or not.⁶ Non-combatants are protected by International Humanitarian Law (IHL, one branch of public international law), and therein lies the challenge. Beyond this postulation, the interest of this paper is considering whether AI can be attributed combatant status since it can act in the place of humans and possesses certain attributes and qualities of humans.

1.3 Combatants in International Humanitarian Law

International Humanitarian Law (IHL) plays a vital role in regulating the management of armed conflict, providing a legal framework that strives to balance military necessity with humanitarian considerations. Fundamental to this framework is the distinction between combatants and non-combatants. Combatants, as defined by Article 43 of Additional Protocol I to the Geneva Conventions, are members of the armed forces of a party to a conflict, not including medical and religious personnel. This formal identification of combatants is essential, as it legitimizes their involvement in hostilities and distinguishes them from civilians, who are afforded protection under IHL. By establishing clear standards for what constitutes a combatant, IHL aims to promote accountability and order during conflict⁷.

The rights and obligations of combatants under IHL are intricately connected to the concept of combatant immunity. This principle grants lawful combatants protection from prosecution for acts of war, as long as they abide by the rules of warfare. For example, combatants may engage in attacks against enemy forces, but they must observe to IHL's prohibitions against aiming at civilians and other protected persons. If a combatant fails to adhere to these legal standards, they risk

⁶ This is known as the distinction principle of International Humanitarian Law.

⁷International Committee of the Red Cross (ICRC). (2016). *Customary International Humanitarian Law

being classified as an unlawful combatant, losing their protections and possibly facing legal penalties. Article 44 of Additional Protocol I highlights this relationship between compliance and status, emphasizing that a combatant's rights are depending on their adherence to the laws governing armed conflict⁸.

The classification of combatants becomes even more complicated in the context of non-international armed conflicts, where non-state actors often plays important role. In these situations, IHL specifies that members of organized armed groups can be recognized as combatants if they meet specific criteria, such as being part of a command structure and adhering to IHL. Nevertheless, individuals who act independently or outside the confines of an organized group may be labeled as unlawful combatants. This distinction is critical, as it influences the legal protections obtainable to individuals involved in conflicts, highlighting the importance of organized and lawful military conduct⁹.

1.4 Status of Combatants and Prisoner of War

Article 43(1) of the Additional Protocol I operates as the definition for the concept of armed forces. The rule is also acknowledged as a norm of customary international law¹⁰ applicable in international armed conflicts by state practice, including by states not party—or not at the time—to the Additional Protocol I. Article 43(1) augments Article 4 of the Geneva Convention III, which defines the POW, by seeking to identify who are the members of the armed forces¹¹. Additionally, relation between the provisions is that while combatant status can

⁸Geneva Conventions and Additional Protocols. (1977). Additional Protocol I (Art. 43 and 44).

⁹A. Roberts, & R. Guelff, *Documents on the AI of War* (3rd ed.). Oxford University Press. (2000).

¹⁰Henckaerts, Jean-Marie, and Louise Doswald-Beck. *Customary International Humanitarian Law: Volume I: Rules. Libk. I.* Cambridge; New York: Cambridge University Press. 2005.

¹¹Sandoz, Yves, Christophe Swinarski, and Bruno Zimmermann. *Commentary on the Additional Protocols of 8 June 1977 to the Geneva Conventions of 12 August 1949.* 1987 MartinusNijhoff Publishers; Distributors for the United States and Canada, Kluwer Academic Publishers.

only be attained by being a member of the armed forces¹², POW status is a privilege given—but not limited to—the combatants.¹³ Therefore, it is necessary to read these two provisions in conjunction to ascertain the boundaries of the combatant concept.

At one view, Article 43 of the Additional Protocol I defines all armed forces by introducing the following requirements; organisation, command responsibility, and the existence of an internal disciplinary system to enforce respect to the rules of IHL, and does not distinguish regular armed forces from other armed groups and units. Article 4(A2) of the Geneva Convention III, another view, presents six categories that the definition of POW covers and particularly requires irregular armed forces to fulfil the following conditions; command responsibility, having a fixed distinctive sign, carrying arms openly, and acting in accordance with the laws of war to be considered combatant entitled to POW status.

The first main disparity between those two provisions is that while Article 43 of the Additional Protocol I provides a general definition of armed forces and combatant concept, Article 4 of the Geneva Convention III recognizes the issue from an individual perspective by questioning in what form of armed forces a fighter can be entitled to POW status. Second, visibility requirements in Article 4, namely, having a distinctive emblem and carrying arms openly, are absent in Article 43 of the Additional Protocol I as these are criteria more relevant to the entitlement of POW status rather than constituting armed forces. Hence, such issues are regulated under Article 44 of the Additional Protocol I, entitled combatants and prisoners of war. In spite of the relative distinction in the content

¹²However, Article 4 of the Geneva Convention III also introduces categories which are not the members of armed forces but are entitled to have POW status, and those should be treated as POWs. See; Article 4(A)(4), 4(B), Geneva Convention III.

¹³The reference to a person is not introduced as categories of person as it is later clarified by the definition of civilian in Article 50 of the Additional Protocol I.

to those two provisions of IHL, in essence, both describe combatants as those members of a state armed forces or volunteer corps or militias fighting on behalf of a party of a conflict as long as they meet the structural criterion introduced.

Neither the procedural aspect of membership to an armed force enshrined in Article 43 of the Additional Protocol I set a standard of being human nor the scope of armed forces excludes non-human actors explicitly. Merely the POW is defined as “...persons belonging to one of the following categories, who have fallen into the power of the enemy...” in the chapeau of Article 4 of the Geneva Convention III. Although Article 4 describes POWs as persons, it would be difficult to argue that the term “persons” qualifies the categories with combatant status ¹⁴ listed under Article 4. The reason why it would be difficult to infer that the combatant is defined as “persons” only departing from the reference to “persons” in the chapeau of Article 4 will be further exemplified comparatively in the following section.¹⁵

Furthermore, Article 4 also refers to persons in its sub-paragraph A4, for civilians who are authorised to accompany the armed forces. Sub-paragraph A6 uses the term “inhabitants” for the participants in a *levée en masse* to identify who is entitled to POW status.

Seeing that sub-paragraph A4 regulates civilians who are not members of the armed forces; it has no influence on the assessment of whether non-human actors have combatant status.

Nonetheless, in sub-paragraph A6, the term “inhabitants” used—instead of persons refers to the participants in a *levée en masse*, which are exceptionally believed to be combatants in IHL. Nevertheless, it is debatable whether inhabitants imply only human beings. Nor is it defined under international law.

¹⁴(Article 4 (A1), (A2), (A3), (A6))

¹⁵Although Article 4 also refers to persons in its other sub-paragraphs, only those having combatant status are scrutinized here.

Linguistically, the Cambridge Dictionary defines the notion of inhabitant as “a person or animal that lives in a particular place” (Cambridge Dictionary), while the Merriam-Webster Dictionary describes it as “one that occupies a particular place regularly, routinely, or for a period of time” (Merriam-Webster.com Dictionary). Regardless of the ambiguity, even if it is assumed that Article 4(A6) of the Geneva Convention explicitly refers to humans or even implies it, the applicability of the article would be limited to participants in a *levée en masse* since it is a precise regulation and does not define combatant status generally.

In other terms, the combatant status would not be attributable to non-human actors who participate in a *levée en masse* if it was deemed that the provision refers to humans. Still, in any case, it does not involve the other categories of combatant listed in the same Article.

To cap up, it is evident to say that provisions defining combatant status and members of the armed forces do not specifically refer to a human (or a person). Yet, whether the members of the armed forces suggest humans remains ambiguous due to the lack of detailed description. The only possible references to humans may be obtained from the chapeau of Article 4 and Article 4(A6), whose applicability is limited due to the use of the terms “person” and “inhabitants”. Nevertheless, this paper debates that neither the word “person” qualify categories of combatant in the chapeau of Article 4, nor is it easy to come to a conclusion from the use of the term “inhabitants” that combatants are described as a person. Additionally, to what extent does the precise reference to person imply human is questionable and will be dealt with in the next section.

Article 43 of the Additional Protocol I and Article 4 of the Geneva Conventions make evident in what circumstances combatant status is born for the members of armed forces and in what category of armed conflict. Furthermore, while Article 43 of the Additional Protocol I indicate the conditions of forming an armed group, Article 4 of the Geneva Convention III explains under what circumstances

irregular armed forces may become part of an armed forces party to an armed conflict and can be entitled to both combatant and POW status. Article 43(2) nevertheless indicates who will be exempted from the status of combatant despite their membership in the armed forces and introduces rights derived from the combatant status.

Despite the wealth of provisions related to combatant, both Article 43 of the Additional Protocol I and Article 4 of the Geneva Convention III use the word members of the armed forces as the subject of combatant status, yet fail to explain who is eligible to be considered as a member of the armed forces or whether the term members of the armed forces implies human. Thus, it is still unlikely to come to a conclusion whether the combatant status is attributable solely to humans based on both Article 43 of the Additional Protocol I and Article 4 of the Geneva Convention III.

1.5 The Personality of Artificial Intelligence Technology

The postulations presented above entail concluding that it is required to be a human in order to have a combatant status, and both combatant and civilian statuses do not fit the non-human actors in the sense of IHL unless the non-human actors covered by the term person¹⁶.

In this perspective, the issue of whether non-human actors may be considered as a person within the meaning of provisions regulating the combatants and civilians should be examined from the lenses of IHL to conclude whether the nonhuman actors are eligible to be considered a combatant. Even though there is no authoritative definition of a person in the context of IHL, the word is mostly described in the domestic law of states. Technically, when a country allows non-

¹⁶Weller, Chris, "A Robot that Once Said it would "Destroy Humans" Just became a Robot Citizen in Saudi Arabia." *Business Insider Nederland* 2017 < <https://www.businessinsider.nl/sophiarobot-citizenship-in-saudi-arabia-the-first-of-its-kind-2017-10/>>. Accessed July 2024

humans, such as machines and animals, to be included in their definition of a person, this may lead to complicated situations.

For instance, legal citizenship already given to a robot named “Sophia” by Saudi Arabia¹⁷. Sophia was also awarded as the first innovation champion by the United Nations Development Program (UNDP) in Asia and the Pacific¹⁸. Even though giving a machine legal citizenship was a controversial issue and raised contrasting voices worldwide, there may be a possibility that an AI-based robot may also be acknowledged as a person in the future. Accordingly, such a situation may trigger a debate whether those non-human actors become a member of the armed forces, combatants or civilians in times of armed conflict. At this point, this paper emphasizes that it is necessary to elaborate on what “person” means and how it is characterized under IHL.

Despite the lack of explicit human attribution in the definitions of combatant and civilian, as discussed expounded above, both definitions refer to a person and the term person is used to imply a human according to the universal spirit of the Geneva Conventions and, in particular, regulations preserving civilians. The cases of expressions of Geneva Conventions related to the term person and words used in combination with the person are; sick and aged person, personal right to have correspondence with the family, respect to the honor, family rights, religious beliefs, and practices, and manners and customs of persons, treatment of persons according to their state of health, age, sex.¹⁹ In this context, albeit an AI-based machine may become a citizen of a country legally, it never becomes a person within the context of IHL taking into account human-based insinuations while referring to the notion of “person” such as family, religious beliefs, and gender.

¹⁷Weller *Note 17 Supra*

¹⁸Xia, Zhang. 2017. “Sophia the AI Robot is First Non-human Recipient of UN Technology Innovation Competition Gong.” <<https://www.yicai-global.com/news/sophia-the-ai-robot-is-first-non-human-recipient-of-un-technology-innovation-competition-gong>>. Accessed July 24th 2024

¹⁹Article 27(3), Geneva Convention IV

As such, neither animal can be included in the scope of the term person²⁰ despite greater relevance with the aforementioned characterizations compared to machines. Those are only a handful of examples that demonstrate that when the Geneva Convention IV uses the term “person”, it is actually referring to the human being.

Furthermore, considering the notions such as man, woman, child, and special protections provided to them, especially women and children, such categorizations would fundamentally be invalid for robots both from combatant and civilian stand points. To demonstrate, regulations such as; the separation of dormitories according to the sex of the inmates,²¹ same-sex supervision for women and their imprisonment in separate quarters,²² are not in nature adaptable to non-human actors under the Convention relating to the treatment of prisoners of war. Furthermore, the special protection provided to women to protect them from rape, forced prostitution and indecent assault,²³ protection of pregnant women and mothers of a child below seven years,²⁴ protection of children under fifteen years old by establishing (if needed) hospitals and safety zones in occupied areas,²⁵ measures related to child welfare,²⁶ evacuations of children from besieged or encircled areas,²⁷ protection of children in the occupied territory,²⁸ including the provision of education to children by the occupying power²⁹ are the few instances

²⁰Animals have also families and gender differences like humans. However, the protections as a whole provided for civilians hamper including animals to the scope of the person concept while referring to correspondence with the family, religious convictions, and age standards.

²¹Article 29(2), Geneva Convention III

²²Article 108(2), Geneva Convention III

²³Article 27(2), Geneva Convention IV

²⁴Articles 38(5) and 132(2), Geneva Convention IV

²⁵Article 14(1), Geneva Convention IV

²⁶Article 24, Geneva Convention IV

²⁷Article 17, Geneva Convention IV

²⁸Article 50, Geneva Convention IV

²⁹Articles 50(1) and 94(2), Geneva Convention IV

that are not written in nature to cover non-human actors under the Convention related to the safeguard of civilian persons.

From another viewpoint, the prohibition of starvation of civilians regulated under Article 54 of the Additional Protocol I can only be appropriated for living species for obvious reasons and would not include AI-driven machines since they do not starve.³⁰ Additionally, considering that suffering is a feeling peculiar to living species, Article 35(2) of the Additional Protocol I, prohibiting attacks that cause unnecessary pain, would not have an effect on non-living beings either, at least in the lenses of IHL. In this context, it is possible to conclude at this point that the examples presented above show that the term “person” is used to refer to “human” under IHL. As a result; it would not be possible to discuss military AI technologies or any other non-human subjects within the context of combatant and civilian notions.

1.6 Liability of the use of Artificial Intelligence Technology under International Humanitarian Law

One of the most pressing subjects for the implementation and integration of AI is the uncertainties regarding obligation and accountability. Since robots cannot be held legally responsible, the perceived “accountability gap” regarding who is responsible for robot violations is a critical component for AI acquisition and procurement. There are many investors in the lifecycle from AI development to implementation; acquisition officials will be forced to consider liability within various frameworks—state liability, individual liability, commander responsibility (a subset of individual responsibility), and corporate responsibility. Each vantage point plays a role in machine development and performance.

³⁰Article 14, Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of Non-International Armed Conflicts (Protocol II), of 8 June 1977; Article 8(2)(b)(xxv), 1998 International Criminal Court (ICC) Statute.

1.6.1 State Liability

The law of state responsibility is a crucial institution in international law and defines when a state has violated an international obligation, the consequences of the violation, and the appropriate measures to be taken to implement the consequences of the breach. The rules on state responsibility are articulated in the International Law Commission (ILC) articles on “Responsibility of States for Internationally Wrongful Acts.”³¹

The ILC articles stipulates the principles governing the liability of states in instances where the state has committed an internationally unlawful act. Article 1 of the ILC articles outlines how “every internationally wrongful act of a state entails the international liability of that state, and thus gives rise to the new international legal relations additional to those which existed before the act took place.”³² An internationally unlawful act occurs when two conditions are met: (1) conduct consisting of an action or omission is attributable to the state under international law; and (2) when that action or omission constitutes a breach of an international obligation.³³

The first element of an internationally unlawful act, attribution to the state, only applies to conduct that is “attributed to the State...that is of its organs of government, or of others who have acted under the direction, instigation or control of those organs, i.e. as agents of the State.”³⁴

³¹Vincent Boulanin, NettaGoussac, and Laura Bruun, “Autonomous Weapon Systems and International Humanitarian Law: Identifying Limits and the Required Type and Degree of Human–Machine Interaction. Stockholm International Peace Research Institute 2021.

³²International Law Commission, Draft Articles on Responsibility of State for Internationally Wrongful Acts, with commentaries (2001)

³³International Law Commission, Responsibility of States for Internationally Wrongful Acts, A/56/49, December 2001, Article 2. [Hereinafter ILC Articles (2001)].

³⁴ILC Articles, with commentaries (2001), Chapter 2, para. 2.

An “organ of the state” refers to any person or body that carries that status in accordance with the internal law of that state.³⁵ AI does not in itself constitute an organ of the state that is competent of acting on behalf of the state. The ILC has clarified, “the ‘act of the State’ must involve some act or omission by a human being or group.”³⁶ The more suitable organ of the state is the commander who takes the decision to deploy AI and decides the circumstances to use AI.

The second element of an internationally unlawful act, a breach of the state’s international legal responsibilities, occurs when a state’s actual conduct breaches a primary rule of international law that is binding on it.³⁷ Because the law on state liability is intentionally written to be broad and applicable in many instances, the origin of the obligation matters not; for the purpose of this report, however, the responsibilities to consider are the obligations within IHL. State responsibility is a distinct framework from other frameworks discussed below and does not have the difference between “civil” or “criminal” liability found in internal legal systems.

States that have committed internationally wrongful acts have two obligations: They must cease the conduct that is in violation of international legal obligations if it is continuing, and the state must make full compensation for the injury caused by the internationally unlawful act.³⁸

Based on these requisite circumstances of the law of state liability, there are three ways a state will assume responsibility for wrongful behavior by AI.³⁹ First, when a state agent deploys AI and the robot violates IHL rules; recently, Israel deployed a drone where it killed the Hesbollah leader inside the underground bunker, in

³⁵ILC Articles (2001), Article 4.

³⁶ILC Articles, with commentaries (2001), Article 2(5).

³⁷ILC Articles, with commentaries, Chapter 3, para. 3.

³⁸ILC Articles (2001), Articles 30 and 31, respectively.

³⁹Thompson Chengeta, “Accountability Gap: Autonomous Weapon Systems and Modes of Responsibility in International Law,” *Denver Journal of International Law and Policy* 45, no. 1 (2016), particularly pp. 47–49.

Beirut. Palestine is claiming it killed civilians alongside Hassan Nasr Allah. This is the most straightforward instance of the action being attributable to a state organ and constituting a violation of an international legal obligation. Second, state liability could be incurred if the “authorization, acquiescence, complicity or acknowledgement from state agents, a non-state actor deploys AI which violates protected rights.”⁴⁰ In this instance, non-state actors could be groups or contractors working under or together with state organs in which the state acknowledges liability for unlawful outcomes of non-state AI.

Within these first two conditions, states are liable for the unlawful conduct and have a duty to provide compensations to individuals or other entities for an internationally wrongful act committed by a robot.⁴¹ The third instance requires a deeper evaluation. A state can incur liability where a private technology firm (or any private entity) contributes to the production of AI and that system, due to potentially low development standards, breaches international legal obligations. In this case, both Articles 5 and 11 of the ILC Report designate State responsibility as follows:

Article 5: The conduct of a person or entity which is not an organ of the State...but which is empowered by the law of that State to exercise elements of that governmental authority shall be considered an act of the State under international law, provided the person or entity is acting in that capacity in the particular instance.⁴²

This article expands attribution beyond formal State organs (such as ministries or the military) to cover non-State actors or semi-private bodies that exercise delegated public powers. Examples include: private security firms carrying out policing functions, corporations running prisons, or private contractors

⁴⁰ILC Articles (2001), 40–42.

⁴¹ILC Articles (2001), Article 31.

⁴²ILC Articles (2001), Article 5.

administering customs or border control on behalf of the State. The key condition is that the entity must be empowered by domestic law to perform functions of governmental authority, and the wrongful act must occur while exercising that authority. If so, their conduct is attributed to the State under international law, and the State bears responsibility for any internationally wrongful act committed in that context.

1.6.2 Individual Criminal Responsibility

Acquisition officials must furthermore consider the risks for individual criminal liability for violations committed by AI as a framework with the potential to establish some kind of legal liability for their armed forces. Establishing individual criminal responsibility for war crimes requires a high degree of *mens rea* (discussed below). For acquisition officials, this framework is complicated to establish in the AI context, and the risk for individual criminal liability is much less than state responsibility.⁴³

Individual criminal liability is a basic tenet of international criminal law (ICL) and overall requires the presence of *mens rea* before an individual incurs criminal liability, as discussed below. A variety of important individuals can fall within this legal framework, and analyses of individual criminal liability generally refer to individual military operators and commanders. While command responsibility will be addressed separately below, this section specifically addresses the risks for individual military operators being held responsible for war crimes occasioned from AI violating IHL.⁴⁴

⁴³Carrie McDougall, “Autonomous Weapon Systems and Accountability: Putting the Cart before the Horse,” *Melbourne Journal of International Law* 20 (2019);

⁴⁴UN General Assembly, *Rome Statute of the International Criminal Court (last amended 2010)*, July 17, 1998, ISBN No. 92-9227-227-6 [hereafter *Rome Statute*].

The risk of AI performance resulting in unlawful outcomes is of particular concern, as these unlawful consequences can come in numerous and unpredictable ways. For example, one scenario that garners noteworthy concern and attention is the risks in intentionally targeting the civilian populace, a crime in both international armed conflicts (IAC) and non-international armed conflicts (NIAC). Under Article 8 of the Rome Statute of the International Criminal Court (the Rome Statute), intentionally attacking civilians is a war crime. To establish this crime, five elements must be satisfied: (1) the perpetrator directed an attack, (2) the object of the attack was a civilian population or individual civilians not taking direct part in hostilities, (3) the perpetrator intended the civilian population or individuals not taking direct part in hostilities to be the object of the attack, (4) the conduct took place in the context of and was associated with an international armed conflict/armed conflict not of an international character, (5) the perpetrator was aware of factual circumstances that established the existence of an armed conflict. Article 30 of the Rome Statute requires that, unless otherwise stated, the elements of war crimes must be “committed with intent and knowledge.”⁴⁵ Article 30 (2)(b) further provides that if a person has intentions relating to a consequence, then “that person means to cause that consequence or is aware that it will occur in the ordinary course of events.”⁴⁶

It should be noted that International Criminal Law has not generally acknowledged recklessness or negligence as sufficient to establish criminal liability for most international crimes, aside from a few exceptions.⁴⁷ There are a little number of crimes that represent exceptions to Article 30 and provide a lower

⁴⁵*Rome Statute art. 30*

⁴⁶*Rome Statute art. 30(2)(b).*

⁴⁷McDougall, “Autonomous Weapon Systems,” Rebecca Crotoft, “War Torts: Accountability for Autonomous Weapons,” *University of Pennsylvania Law Review* 164, no. 6 (2016).

mental elements threshold, one of which is individual criminal liability within command responsibility, which is detailed below.⁴⁸

Some scholars have identified certain scenarios in which individuals could still (relatively) clearly be held responsible for war crimes committed by AI, two of which are worth emphasizing here.⁴⁹ First a clear case where any person could incur criminal liability is intentionally programming targeting parameters to violate international targeting obligations.⁵⁰ This scenario would likely, and most clearly, satisfy the mental elements and necessary intention to be considered a war crime.

Second, a case in which a commander or senior official was to authorize employment of AI they knew, or due to the circumstances should have known, to be unpredictable in complex conditions and could be virtually certain of unlawful conduct.⁵¹ This scenario would necessarily require the commander or individual giving authorization to be aware of the circumstances and possibility of legal violations by the AI.

1.6.3 Command Liability

A subset of individual criminal liability is the responsibility of the commander. It is fundamental to consider the framework of command liability, as some states, notably including the United States, have expressed their commitment to placing importance on the relationship between commanders and AI. The United States of America has argued that, rather than simply prioritizing human control over robots, the key issue is instead “ensuring machines help effectuate the intention

⁴⁸Ibid

⁴⁹Ibid

⁵⁰Ibid

⁵¹Dan Saxon, *Drones and Responsibility: Legal, Philosophical and Socio-Technical Perspectives on Remotely Controlled Weapons*, eds. Ezio Di Nucci and Filippo Santoni De Sio (2016).

of commanders and the operators of the system.”⁵² From this view, AI would not require human supervision, but rather reflect commander intentions and broader objectives; as with subordinate human troops. This makes the commander–robot relationship all the more important to consider.

Command liability, or superior responsibility for high-ranking civilians, is not meant to punish commanders for directly partaking in criminal behavior that can be shown to have planned, ordered, committed, or aided and abetted crimes undertaken by others.⁵³ Rather, the principle under international law refers to a form of responsibility for the commission of crimes taken by subordinates; or the failure to prevent or punish crimes. There are three elements necessary to establish command responsibility as follows:

- i. The existence of a superior–subordinate relationship between the defendant–superior and the perpetrators of the underlying offense.
- ii. The superior (commander) knew or had reason to know that a subordinate was about to commit such acts or had done so.
- iii. The superior failed to take necessary and reasonable measures to prevent such acts or to punish perpetrators thereof.⁵⁴

The first element obliges a superior–subordinate relationship, whether it be *de jure*, meaning that the commander’s authority comes as a position with the purpose of commanding or leading subordinates (e.g., being appointed, elected, or otherwise assigned to an authoritative position); or a *de facto* relationship, in which a commander exercises authority based on inter-personal relationships or

⁵²Karl Chang, U.S. Mission to International Organization in Geneva, Consideration of the Human Element in the Use of Lethal Force, Address Before the Convention on Certain Conventional Weapons Group of Governmental Experts on Emerging Technologies in the Area of AI (March 26, 2019).

⁵³Guénaél Mettraux, *The Law of Command Responsibility* (Oxford: Oxford University Press, 2009).

⁵⁴*Ibid*

other factual or personal factors.⁵⁵ There are two requirements within this relationship. The first is a chain of command or hierarchical relationship, whether direct or indirect. This requirement does not demand a direct command over subordinates, but rather the commander “by virtue of his position, must be senior in some sort of formal or informal hierarchy to the perpetrator.”⁵⁶ The second requirement is effective control, meaning that the commander must have “the material skill to avoid offences or punish the principal offenders.”⁵⁷ Effective control is the minimal threshold necessary to establish commander responsibility for subordinate offences.

In the context of AI, we can comprehend AI to be a “subordinate” of the commander, as the robot will be programmed with the commander’s intentions and objectives. While a machine does not undergo a training program and process with the commander as with human troops, the machine must reflect the strategic objectives from the commander and conduct the mission in compliance with international law and rules of engagement just as with humans. The effective control requirement is more intricate to establish for three reasons. Recall that effective control principle requires the power to take necessary steps to avoid and punish crimes committed by subordinates. Firstly, it is not likely that a commander, even observing in real-time, would have the speed necessary to direct machine behavior. Secondly, the commander would likely require a certain level of technical proficiency. It would be necessary to study the computer code and make adjustments as necessary for safety. This could also be done with contracted software operator or engineers, but doing so can slow the process. Thirdly, there must be a high level of machine predictability for effective commander control.

This is less of an issue for hand-coded systems, but machine learning systems can formulate their own decisions in response to changing environmental conditions.

⁵⁵Supra note 53

⁵⁶Halilovic, ICTY Appeal Judgement, par. 59. 16.10.2007.

⁵⁷ Supra Note 53

It is unlikely to establish effective control over a machine learning system, which may perpetuate a responsibility loophole.⁵⁸ The second element is a mental requirement, which the Rome Statute Article 28(a)(i) defines as, “a military commander or a person either knew or, due to the circumstances at the time, should have known that the forces were committing or about to commit such crimes.”⁵⁹

This “should have known” standard places obligation for commanders or superiors to be informed of the threats associated with an operation.

The issues for AI are related to the first element in that commanders may not necessarily have insight into machine learning systems and how AI will choose to adapt to the environment; especially as commanders are unlikely to program the machine themselves.⁶⁰ Similar to the ICC “should have known” standard, the International Criminal Tribunal for the Former Yugoslavia (ICTY) held that commanders can be responsible if they had information to put them “on notice of the risk” that is “sufficiently alarming to justify further inquiry.”⁶¹ It is unclear what will constitute a notice of threat for AI, but the language could be a useful framework for incorporating a standard of threat for commanders to evaluate appropriateness of use for a particular system.

The third element requires commanders to take necessary and reasonable actions to prevent or punish subordinates committing offenses, such as through pre-deployment testing or other safety measures. But if a commander, in addition to

⁵⁸Human Rights Watch, “Mind the Gap: The Lack of Accountability for Killer Robots,” (2015); Rebecca Crooto, “War Torts: Accountability for Autonomous Weapons,” *University of Pennsylvania Law Review* (2016); Peter Asaro, “On Banning Autonomous Weapon Systems: Human Rights, Automation, and the Dehumanization of Lethal Decision-making,” *International Review of the Red Cross*, 687 (2012).

⁵⁹Mettraux, *Command Responsibility*, 193–226. *Supra Note 53*

⁶⁰Daniel Hammond, “Autonomous Weapons and the Problem of State Accountability,” *Chicago Journal of International Law* 15, (2) (2015).

⁶¹Prosecutor v. Strugar, Case No. IT-01-42-A, Appeals Chamber Judgement paras. 297–89; quoted in Crooto, “War Torts.”

other elements, conceals unlawful outcomes and keeps the AI in circulation despite threats of errors or compromised safety and security standards, they could breach this requirement.

1.7 Conclusion

This study has examined the complex interplay between artificial intelligence (AI) technologies and the legal status of combatants and prisoners of war (POWs) under international humanitarian law (IHL). It explored how the deployment of AI, particularly autonomous weapons systems (AWS), challenges traditional definitions of combatants, the principles of distinction and proportionality, and the mechanisms for capturing and treating POWs. As warfare evolves from human-controlled operations to machine-influenced or autonomous actions, fundamental questions arise about accountability, legal status, and the protection of individuals under the Geneva Conventions.

The study concludes that existing IHL frameworks—while robust—are increasingly strained by the rapid development of AI in armed conflict. The traditional legal categorizations of combatants and POWs were drafted in an era devoid of machine decision-making. As such, the classification and treatment of individuals involved in AI-enabled operations may fall into legal gray areas, risking violations of humanitarian principles, particularly where accountability is diffused or human control is minimized.

This research presents key findings on the study in the context of the growing deployment of artificial technology in armed conflicts as follows:

1. **Human-Centric Nature of IHL:** Existing IHL instruments such as the Geneva Conventions and their Additional Protocols are drafted with human actors in mind. The deployment of Artificial Intelligence (AI) in armed conflict threatens to erode the centrality of human dignity and

humanity in warfare. A new legal protocol is necessary to reaffirm that the protection of persons—not machines—remains the cornerstone of IHL.

2. One of the most urgent imperatives in regulating Artificial Intelligence (AI) in armed conflict is the establishment of meaningful human control (MHC) as a binding legal requirement. Current International Humanitarian Law (IHL) assumes human judgment at all stages of military decision-making, particularly in the application of the principles of distinction, proportionality, and precaution. However, autonomous weapon systems and AI-driven decision platforms risk delegating life-and-death decisions to machines devoid of moral reasoning, empathy, or contextual judgment. This creates a significant normative and operational gap.
3. The integration of Artificial Intelligence (AI) into armed conflict underscores the urgent need for specialized military training and capacity building to ensure compliance with International Humanitarian Law (IHL). Traditional military training frameworks, while effective in imparting knowledge of the law of armed conflict, are insufficient to address the complexities introduced by autonomous weapon systems, algorithmic decision-making, and AI-driven surveillance technologies. Without enhanced training and institutional capacity building, there is a heightened risk of IHL violations, accountability gaps, and misuse of AI in warfare.
4. The global nature of Artificial Intelligence (AI) development and its deployment in armed conflict necessitates international collaboration and

regulation to ensure uniform compliance with International Humanitarian Law (IHL). Existing treaties, including the Geneva Conventions, were not designed to regulate transnational technological innovation. Without a coordinated international framework, there is a risk of divergent state practices, regulatory loopholes, and an unchecked arms race in AI-enabled warfare.

In light of these findings, the study recommends the following:

1. **Development of New Legal Protocols:** States and international legal bodies should initiate discussions towards the formulation of new IHL protocols or amendments that directly address the use of AI in warfare, including specific provisions on combatant status and POW protection.
2. **Ensuring Meaningful Human Control:** All AI systems used in armed conflict must be subject to meaningful human oversight to ensure compliance with IHL, particularly regarding the lawful use of force and humane treatment of captured individuals.
3. **Training and Capacity Building:** Military and legal personnel must be trained on the implications of AI use in armed conflict to ensure informed decision-making and adherence to legal obligations.
4. **International Collaboration and Regulation:** Global cooperation is essential to regulate the deployment of autonomous systems and develop a universally accepted framework to protect combatants and POWs in the context of AI-driven warfare.

By integrating legal innovation with technological advancement, the international community can uphold the fundamental principles of humanity and justice in modern armed conflict.