

# AI Has No Mens Rea: The End of War Crimes Prosecution?\*

Mackenzie Miller-Sweeney<sup>†</sup> and Mario Reyes<sup>‡</sup>

## Abstract

*AI-driven autonomous weapons are colliding with an international criminal law system that still assumes a human mind behind every war crime. As militaries move toward humans-out-of-the-loop systems and rely on opaque “black box” architectures, it becomes harder to pin the required mental states—intent, purpose, knowledge, or recklessness—on any individual for atrocities committed in armed conflict. This article traces how existing doctrines in international humanitarian and criminal law depend on mens rea, shows why that framework breaks down once lethal decisions are delegated to machines, and argues that the resulting accountability gap is both real and growing. To close it, this article proposes a strict-liability rule: the first commander who authorizes the use of an AWS that commits a war crime would be criminally responsible, a move designed to keep humans firmly on the hook, preserve remedies for victims, and force states to take legality and testing of autonomous systems far more seriously before they are deployed.*

**Keywords:** autonomous weapons systems, artificial intelligence in warfare, international criminal law, international humanitarian law, war crimes, mens rea, command responsibility, strict liability, accountability gap, black box problem, lethal autonomous weapons, targeting civilians, Rome Statute, International Criminal Court, military ethics

## A Introduction

Artificial Intelligence (AI) is transforming various fields, and warfare is no exception. AI’s integration into autonomous weapons systems (AWS) introduces numerous ethical, legal and practical challenges. One of the most significant issues arises in the realm of international criminal law (ICL): the prosecution of war crimes when AWS are involved. Unlike humans, AWS lack the mental state or intent (“mens rea”) necessary to establish criminal responsibility under traditional legal paradigms.

\* Authors’ note: AI was used to assist with editing this content for grammar and clarity. However, all final content has been reviewed and edited by the authors. No AI was used for research purposes.

† Mackenzie Miller-Sweeney is a Deputy Public Defender for the Colorado State Public Defender currently working in the Colorado Springs office.

‡ Mario Reyes is an associate attorney with Bennett, Boehning, and Clary in Lafayette, Indiana. Both authors are proud 2025 graduates of the Indiana University Robert H. McKinney School of Law.

This article examines the implications of this gap, focusing on the difficulty of prosecuting war crimes committed through AWS and a proposed remedy.

## B Background

### I *AI in Warfare Background*

AI's predecessors in warfare can be traced back to the Cold War. But modern AI is light years ahead of the computer game theory war modelling used back then. Most recently, AI's influence in warfare began to take shape in 2012 when the U.S. Department of Defense issued a policy directive on AWS.<sup>1</sup> This directive created guidelines on how autonomous and semi-autonomous weapons systems would be designed and utilized. Most notably, it stated that these weapons systems would be designed in a manner that allowed decision makers "appropriate levels of human judgment over the use of force".<sup>2</sup>

AWS are generally categorized into three types: humans-in-the-loop, humans-on-the-loop and humans-out-of-the-loop systems.<sup>3</sup> The key distinction between these categories lies in the role of human oversight. Humans-in-the-loop systems require a human operator to make critical decisions about target selection and the use of force. Humans-on-the-loop systems operate autonomously but allow human operators to monitor and override decisions as necessary. Finally, humans-out-of-the-loop systems function with full autonomy, making decisions about target acquisition and lethal force without human intervention.<sup>4</sup> These categories highlight the varying degrees of human involvement and underscore the ethical and legal complexities posed by fully autonomous systems.

The U.S.'s Global War on Terrorism (GWOT) has prominently featured humans-in-the-loop systems, such as the Predator and Reaper drones. These drones are piloted remotely by human operators, often from bases located in the U.S. (or other countries of origin such as the U.K. and Italy), allowing them to maintain control over target selection and force application. Equipped with Air-to-Ground Missile-114 Hellfires, a Predator drone can deliver lethal force while being commanded through line-of-sight or satellite data links.<sup>5</sup> This operational set-up ensures that human operators remain responsible for critical decisions, aligning these systems with the humans-in-the-loop category. However, future iterations of drone technology may shift towards humans-out-of-the-loop systems, where AI could independently acquire targets and deploy lethal force, eliminating human intervention entirely.

### II *What Is International Law?*

International law is the law governing relations between states. Historically, it "covers the body of rules that regulate the conduct of sovereign states in relation

1 *Autonomy in Weapon Systems*, Dep't of Def. Directive 3000.09, 3 (21 November 2012), <https://www.esd.whs.mil/portals/54/documents/dd/issuances/dodd/300009p.pdf>.

2 *Ibid.*

3 *Losing Humanity: The Case Against Killer Robots*, Human Rights Watch (19 November 2012), <https://www.hrw.org/report/2012/11/19/losing-humanity/case-against-killer-robots>.

4 *Ibid.*

5 *MQ-1B Predator*, Air Force (September 2015), <https://www.af.mil/About-Us/Fact-Sheets/Display/Article/104469/mq-1b-predator/>.

to each other.”<sup>6</sup> Certain aspects of international law concern the actions of individuals and their respective responsibilities. It defines the states’ legal responsibilities in their conduct with each other, within state’s boundaries and in their treatment of individuals.<sup>7</sup> International law’s domain encompasses a wide range of issues of international concern, such as human rights, disarmament, international crime, refugees, migration, problems of nationality, the treatment of prisoners, the use of force and the conduct of war, among others.<sup>8</sup> International law also regulates the global commons, such as the environment and sustainable development, international waters, outer space, global communications and world trade. International law encompasses many branches, including ICL and international humanitarian law.<sup>9</sup>

Two major courts of international law are the International Court of Justice (ICJ) and International Criminal Court (ICC). The ICJ focuses on disputes between nations rather than individuals or private parties. The ICJ’s jurisdiction is rooted in its status as the official judicial organ of the United Nations.<sup>10</sup> The ICC is a criminal court, tasked with investigating and prosecuting individuals accused of atrocities, including genocide, crimes against humanity, war crimes and aggression.<sup>11</sup> The ICC’s jurisdiction is based on consent, meaning it can prosecute crimes only in nations that have signed and ratified the Rome Statute, the court’s governing treaty. Currently, 124 countries are parties to the Rome Statutes, but notable powers like the United States, China and Russia are not.<sup>12</sup>

In summary, international law serves as the framework governing the relationships and responsibilities of states, as well as certain actions of individuals. Its branches, such as ICL and international humanitarian law, address critical aspects of global governance. The ICJ and ICC exemplify the dual focus of international law, with the ICJ addressing disputes between states and the ICC holding individuals accountable for atrocity crimes. Together, these courts reflect the evolving scope and significance of international law in promoting justice and maintaining order on a global scale.

### III What Is International Criminal Law?

ICL is a set of international rules aimed at prohibiting specific types of conduct, such as war crimes, crimes against humanity, genocide and aggression, while holding individuals who engage in such acts criminally responsible.<sup>13</sup> See Figure 1 for a

6 Legal Concepts, United Nations International, Impartial and Independent Mechanism (IIIM), <https://iiim.un.org/resources/legal-concepts/>.

7 *Fact Sheet 1: Treaty Event 2011 – Towards Universal Participation and Implementation*, United Nations, 1 (2011), [https://treaties.un.org/doc/source/events/2011/press\\_kit/fact\\_sheet\\_1\\_english.pdf](https://treaties.un.org/doc/source/events/2011/press_kit/fact_sheet_1_english.pdf).

8 *Ibid*

9 *Ibid*

10 Franklin Rosenblatt, *The International Criminal Court and International Court of Justice: Similarities and Differences*, Mississippi College Law Review (26 January 2024). ., [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4877734](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4877734).

11 *Ibid*.

12 *Ibid*.

13 United Nations International, *supra* note 6.

**Figure 1** *Four categories of core international crimes within the jurisdiction of the International Criminal Court (ICC): genocide, crimes against humanity, war crimes, and the crime of aggression ( <https://iasgoogle.com/news/international-criminal-court-icc-structure-and-function>)*

| <div>  <h2>4 Categories of crimes under ICC</h2> </div>                                  |                                                                                                                                                                                                                                                                                                                  |                                                                                                                                                                                                                                                                                                     |                                                                                                                                                             |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  <b>Genocide</b>                                                                           |  <b>Crimes against humanity</b>                                                                                                                                                                                                 |  <b>War crimes</b>                                                                                                                                                                                                 |  <b>Crime of aggression</b>                                                |
| <ul style="list-style-type: none"> <li>Specific intent to <b>destroy a national, ethnic, racial or religious group</b> by killing its members or by other means.</li> </ul> | <ul style="list-style-type: none"> <li>Serious violations committed through <b>large-scale attack against civilian population</b></li> <li><b>5 forms listed in the Rome Statute</b> includes offences like murder, rape, imprisonment, enforced disappearances, enslavement, torture, apartheid etc.</li> </ul> | <ul style="list-style-type: none"> <li><b>Grave breaches of the Geneva conventions</b> in the context of armed conflict</li> <li>Includes use of <b>child soldiers; killing or torture of civilians/prisoners of war</b>; intentional attacks against hospitals, historic monuments etc.</li> </ul> | <ul style="list-style-type: none"> <li>Use of armed force by a State <b>against the sovereignty, integrity or independence</b> of another State.</li> </ul> |

description of each of the four categories of international crimes under the ICC. These rules empower or require states to prosecute and punish those responsible for these crimes. Additionally, ICL governs the proceedings of international criminal courts and tribunals tasked with prosecuting and trying individuals accused of these offences.<sup>14</sup>

ICL provides the framework for international criminal courts and tribunals to prosecute individuals accused of these serious offences. Accountability extends beyond those who directly commit these crimes. Individuals in positions of authority can also be held criminally responsible under the doctrine of command or superior responsibility.<sup>15</sup> This principle underscores the duty of military and civilian leaders to prevent crimes by subordinates or to hold them accountable when crimes occur.<sup>16</sup> The scope of ICL reinforces the international community's collective resolve to combat impunity for atrocities and uphold the rule of law.

#### IV What Is International Humanitarian Law?

Intertwined with ICL, International Humanitarian Law (IHL) is a body of regulations designed to minimize the impact of armed conflict for humanitarian purposes.<sup>17</sup> These laws aim to safeguard individuals who are not engaged in combat or who

<sup>14</sup> *Ibid.*

<sup>15</sup> *International Criminal Law*, ICRC Casebook Glossary, International Committee of the Red Cross (n.d.), [https://casebook.icrc.org/a\\_to\\_z/glossary/international-criminal-law](https://casebook.icrc.org/a_to_z/glossary/international-criminal-law).

<sup>16</sup> *Ibid.*

<sup>17</sup> *What Is International Humanitarian Law?*, Int'l Comm. of the Red Cross (July 2004), [https://www.icrc.org/sites/default/files/external/doc/en/assets/files/other/what\\_is\\_ihl.pdf](https://www.icrc.org/sites/default/files/external/doc/en/assets/files/other/what_is_ihl.pdf).

ceased to participate, and they also set limits on the methods and weapons used in warfare. IHL is often referred to as the law of war or the law of armed conflict and forms part of international law, which oversees interactions between nations. This legal framework is established through treaties and conventions between nations, customary practices deemed legally binding and overarching legal principles.<sup>18</sup>

The origins of IHL can be traced back to the principles and customs of ancient civilizations and religious teachings, which have long imposed rules on warfare.<sup>19</sup> The formal codifications of IHL at a global level started in the nineteenth century.<sup>20</sup> Since then, nations have established a set of practical regulations informed by the harsh realities of modern conflict. These rules carefully weigh the need for humanitarian protection against the military needs of states. As the global community expanded, more nations participated in shaping these laws, leading to the creation of a universally recognized legal framework today.

The core of IHL is found in the four Geneva Conventions of 1949, which have been accepted by almost every nation worldwide.<sup>21</sup> These Conventions were later expanded through two Additional Protocols in 1977, which focus on safeguarding victims of armed conflicts.<sup>22</sup> Other treaties have since been adopted to ban specific weapons and combat strategies, as well as to provide protections for particular groups of people and types of property.<sup>23</sup> Today, many elements of IHL are recognized as customary law, meaning they are considered binding rules that apply universally to all nations.<sup>24</sup>

IHL is applicable exclusively during situations of armed conflict and does not extend to internal disturbances or isolated incidents of violence. It comes into effect only once hostilities have commenced and applies impartially to all parties, regardless of who initiated the conflict.<sup>25</sup> IHL differentiates between international and non-international conflicts. International armed conflicts involve at least two nations, whereas non-international conflicts occur within a single country, involving either state forces against armed rebel groups or clashes between such groups.<sup>26</sup> Internal conflicts are regulated by a narrower set of rules, found in Article 3 common to the Geneva Conventions and Additional Protocol II.<sup>27</sup> While extensive rules cover international conflicts, including those in the four Geneva Conventions and Additional Protocol I, it is important to distinguish IHL

18 *Ibid.*

19 *Ibid.*

20 *Ibid.*

21 *Ibid.*

22 *Ibid.*

23 These agreements include the 1954 Convention for the Protection of Cultural Property in the Event of Armed Conflict, plus its two protocols; the 1972 Biological Weapons Convention; the 1980 Conventional Weapons Convention and its five protocols; the 1993 Chemical Weapons Convention; the 1997 Ottawa Convention on anti-personnel mines; the 2000 Optional Protocol to the Convention on the Rights of the Child on the involvement of children in armed conflict.

24 *What Is International Humanitarian Law?*, *supra* note 17.

25 *Ibid.*

26 *Ibid.*

27 *Ibid.*

from human rights law.<sup>28</sup> Although they overlap, these two legal frameworks have separate origins and are from different treaties. For instance, human rights law, unlike IHL, applies in peacetime, and many of its provisions may be suspended during an armed conflict.

Two main areas are covered by IHL: protecting those who are not, or are no longer, participating in combat and limiting the methods and weapons used in warfare. So, what does 'protection' mean? IHL safeguards individuals who are not involved in fighting, like civilians and medical or religious military staff. It also covers those who have stopped taking part in combat, such as wounded soldiers, shipwrecked individuals, sick fighters and prisoners of war.<sup>29</sup> These groups have the right to have their lives and their physical and mental well-being respected. They are guaranteed legal protections and must be treated humanely at all times, without discrimination. More specifically, it is prohibited to kill or injure an enemy who has surrendered or can no longer fight.<sup>30</sup> The wounded and sick must be gathered and cared for by the controlling party. Medical staff, supplies, hospitals and ambulances must be safeguarded. IHL also includes specific rules about the treatment of prisoners of war and the handling of civilians under enemy control, ensuring they receive food, shelter and medical care and can communicate with their families.<sup>31</sup> Additionally, IHL uses clear symbols to mark protected people, places and objects.<sup>32</sup> IHL prohibits any weapons or tactics that fail to distinguish between combatants and non-combatants, including civilians, in order to safeguard civilian lives, property and infrastructure; cause excessive injury or unnecessary suffering; or inflict significant or lasting harm on the environment.<sup>33</sup> As a result, IHL has outlawed several weapons, such as exploding bullets, chemical and biological weapons, blinding laser weapons and anti-personnel mines.<sup>34</sup>

Unfortunately, violations of IHL are widespread, and civilians increasingly bear the brunt of war.<sup>35</sup> Ensuring respect for IHL requires proactive measures, though achieving full compliance can be challenging. States have a responsibility to educate both their armed forces and the public about IHL. They should take steps to prevent violations and impose penalties when breaches occur, even if enforcement is not always consistent. Specifically, governments need to enact legislation to punish serious violations of the Geneva Conventions and Additional Protocols, recognized as war crimes.<sup>36</sup> At the international level, progress has been made, albeit imperfectly. Tribunals were established to address atrocities committed during the conflicts in the former Yugoslavia and Rwanda.<sup>37</sup> Furthermore, the International Criminal Court, formed under the 1998 Rome Statute, was created to prosecute

28 *Ibid.*

29 *Ibid.*

30 *Ibid.*

31 *Ibid.*

32 Including the red cross, red crescent and symbols for cultural sites and civil defence facilities.

33 *What Is International Humanitarian Law?*, *supra* note 17.

34 *Ibid.*

35 *Ibid.*

36 *Ibid.*

37 *Ibid.*

war crimes, among other offences.<sup>38</sup> While compliance with IHL is undoubtedly a complex and ongoing effort, individuals, governments and organizations can all contribute meaningfully to its enforcement and promotion.<sup>39</sup> Despite the difficulties, striving for adherence to IHL remains a vital endeavour.

## V War Crimes

As mentioned, war crimes are violations of IHL.<sup>40</sup> The perpetrators are prosecuted under ICL in the ICCs. The Rome Statute of the International Criminal Court defines war crimes as “serious violations of the laws and customs applicable in international armed conflict” and “serious violations of the law and customs applicable in international character.”<sup>41</sup> Unlike the other categories of crimes in the International Criminal Courts, war crimes always take place in the context of an armed conflict, despite intentions.<sup>42</sup> A common misconception among the public is that any violation of the laws of armed conflict automatically qualifies as a war crime. This reasoning is flawed because it disregards the *mens rea* requirement, or mental requirement.<sup>43</sup> A breach of the laws of armed conflict, absent the necessary *mens rea*, does not inherently amount to a war crime. Neither the Rome Statute nor the Elements of Crimes define the term “armed conflict.”<sup>44</sup> In the *Lubanga Dyllo* judgment, the ICC adopted the definition established by the International Criminal Tribunal for the former Yugoslavia (ICTY) in *The Prosecutor v. Tadić* (1995) case:

An armed conflict exists whenever there is a resort to armed force between States or protracted violence between governmental authorities and organized armed groups or, between such groups within a State.<sup>45</sup>

Once an armed conflict is established, IHL comes into effect.

Article 8(2)(a) of the Rome Statute incorporates the grave breaches provisions outlined in the Geneva Conventions of 12 August 1949.<sup>46</sup> These provisions encompass crimes such as wilful killing, torture or inhumane treatment, intentionally inflicting severe suffering, extensive destruction or appropriation of property not justified by military necessity and carried out unlawfully, and forcibly conscripting prisoners of war or other protected persons into the armed forces of a hostile power.<sup>47</sup>

While Articles 8(2)(a) and 8(2)(b) address war crimes committed during international armed conflicts, Article 8(2)(c) extends to violations of the laws and

38 *Ibid.*

39 *Ibid.*

40 United Nations International, *supra* note 6.

41 Geert-Jan Alexander Knoops, *Mens Rea at the International Criminal Court*, 16 (2016) (ebook), <https://brill.com/display/title/32537?language=en&srsId=AfmBOopde7bqu2vixxT5890QH-D9aq1pWEnd48sQ-dJlGFuUW-4Zmq3a>.

42 *Ibid.*

43 *Ibid.*

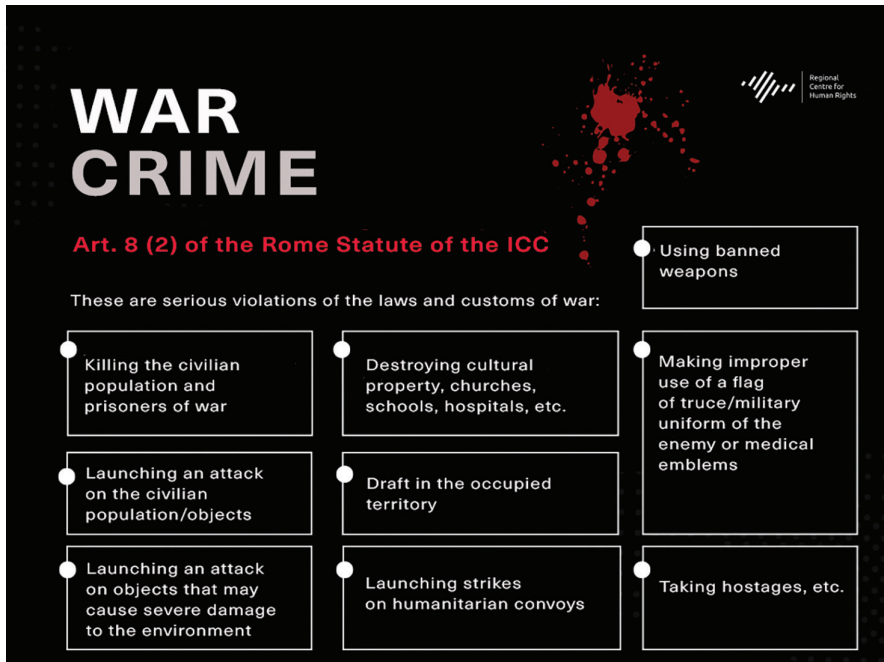
44 *Ibid.*, at 67.

45 *Ibid.*

46 *Ibid.*, at 66.

47 *Ibid.*

**Figure 2** *Examples of war crimes under Article 8(2) of the Rome Statute of the International Criminal Court, including attacks on civilians, protected objects, the environment, and humanitarian operations (<https://voxukraine.org/en/they-will-be-held-accountable-for-everything-what-war-crimes-against-ukraine-will-russia-be-punished-for>)*



customs of war during internal armed conflicts.<sup>48</sup> Many war crimes listed in the Rome Statute require a more specific mens rea than the general intent level defined in Article 30. Examples include “willfully” causing severe suffering or serious injury, “intentionally” targeting civilian structures and “intentionally” using starvation as a method of warfare.<sup>49</sup> These provisions emphasize that both the act itself and its consequences must be intentional. This approach has stayed consistent through international case law. In the interlocutory appeal in *The Prosecutor v. Tadić* (1995), the Appeals Chamber of the International Criminal Tribunal for the former Yugoslavia stated, “[T]he violation of the rule [of international humanitarian law] must entail, under customary or conventional law, the individual criminal responsibility of the person breaching the rule.”<sup>50</sup> Further, in cases involving violations of Additional Protocol I, the ICTY assessed whether these infractions

<sup>48</sup> *Ibid.*

<sup>49</sup> *Ibid.*, at 66-67.

<sup>50</sup> *Customary International Humanitarian Law: Rule 156, International Humanitarian Law (IHL) Databases*, Int’l Comm. of the Red Cross (n.d.), <https://ihl-databases.icrc.org/en/customary-ihl/v1/rule156>.

entailed individual criminal responsibility under customary international law, even when not classified as grave breaches under the protocol.<sup>51</sup> This approach does not preclude a state from designating other breaches of IHL as war crimes under its national laws. However, such classifications remain domestic in scope, without creating international obligations to prosecute or establishing universal jurisdiction.<sup>52</sup>

Figure 2 highlights the breadth of actions classified as war crimes under Article 8(2) of the Rome Statute of the International Criminal Court. These examples underscore the severe nature of war crimes and reinforce the necessity for specific mens rea requirements to establish criminal responsibility. Further, Table 1 is a table that connects the Rome Statute’s emphasis on mens rea requirements to the example war crimes in Figure 2. This demonstrates how specific levels of intent shape criminal liability in the international courts.

**Table 1** *Illustrative mens rea standards and corresponding examples of war crimes under international criminal law*

| Mens Rea           | Definition                                                                                                                 | Example War Crimes                                                                                      |
|--------------------|----------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------|
| Willfully          | Acts committed deliberately, with knowledge of the consequences and intent to achieve the prohibited result. <sup>53</sup> | Willfully causing severe suffering or serious injury (torture or inhumane treatment).                   |
| Intentionally      | Conscious and deliberate intent to perform an act or bring about a specific outcome. <sup>54</sup>                         | Intentionally targeting civilian structures (destroying cultural property, churches, hospitals).        |
| Knowingly          | Awareness that a particular action will result in specific consequences, even if not actively desired. <sup>55</sup>       | Using banned weapons with knowledge of their destructive and indiscriminate impact.                     |
| Negligently (rare) | Failure to act with reasonable care, resulting in harm, though not necessarily deliberate. <sup>56</sup>                   | Launching attacks on objects causing severe environmental damage (failure to consider consequences).    |
| Purposefully       | Acting with the aim or purpose of achieving a specific unlawful outcome. <sup>57</sup>                                     | Launching strikes on humanitarian convoys to disrupt air or inflict harm (specific intent to obstruct). |
| Recklessly         | Consciously disregarding a substantial and unjustifiable risk that a harmful consequence will occur. <sup>58</sup>         | Launching attacks on civilian populations (knowing risk of harm but proceeding anyway).                 |

51 *Ibid.*  
52 *Ibid.*  
53 International Criminal Court, *Rome Statute of the International Criminal Court*, United Nations (2011), <https://www.icc-cpi.int/resource-library/documents/rs-eng.pdf>.  
54 *Ibid.*  
55 *Ibid.*  
56 *Ibid.*  
57 *Ibid.*  
58 *Ibid.*

## C AWS in Warfare

### I *The Arms Race*

AI in warfare presents a catch-22 of sorts. Not utilizing the technology for the sake of avoiding ethical dilemmas would be prudent. But nations that are slow to adapt technologies to their armies will be outpaced by the ones that show little restraint to the ethical dilemmas. For instance, the U.S. Department of Defense has stated, “Failure to adopt AI will result in legacy systems irrelevant to the defense of our people.”<sup>59</sup> If the U.S. views failure of AI into weapons systems as a national defence issue, it is doing so because other nations are moving frantically towards integration.

### II *Current Usage*

The ongoing conflict in Gaza demonstrates how AI is reshaping modern warfare. Israel has employed advanced AI systems for targeting such as ‘Gospel’, and ‘Lavender’.<sup>60</sup> These systems, while currently limited to target acquisition, have already sparked significant ethical debates. For instance, Lavender uses an algorithm to generate a database of people it identifies as potential militant fighters associated with Hamas or PIJ (Palestinian Islamic Jihad).<sup>61</sup> Ethical concerns arise from the system’s reliance on algorithms to make such classifications, potentially leading to errors in biases in identifying targets.

Reports also indicate that the Israeli Defense Force (IDF) has pre-authorized collateral damage thresholds, allowing up to twenty civilian deaths in a bombing operation aimed at eliminating a single, low-level, militant target.<sup>62</sup> Coupled with the IDF’s apparent treatment of the AI outputs as equivalent to “a human decision”,<sup>63</sup> these practices illustrate how fully AWS may soon become a reality. This highlights the urgent need for robust legal frameworks to address the ethical and accountability challenges posed by AI-driven warfare.

### III *Difficulty in Prosecuting War Crimes with Autonomous Weapons Systems*

The prosecution of war crimes such as torture, the intentional targeting of civilian structures, and launching attacks on civilian populations relies heavily on proving the mental state, or mens rea, of the perpetrator. ICL, as codified in instruments like the Rome Statute of the International Criminal Court previously discussed, requires that such crimes be committed intentionally, purposefully or, at the very

59 *Summary of the 2018 Dep’t of Def. Artificial Intelligence Strategy: Harnessing AI to Advance Our Security and Prosperity*, Dep’t of Def. (12 February 2019), <https://media.defense.gov/2019/feb/12/2002088963/-1/-1/1/summary-of-dod-ai-strategy.pdf>.

60 Michael N. Schmitt, *Israel-Hamas 2024 Symposium-The Gospel, Lavender, and the Law of Armed Conflict*, Lieber Institute West Point (28 June 2024), <https://lieber.westpoint.edu/gospel-lavender-law-armed-conflict/>.

61 Bethany McKernan & Harry Davies, ‘The machine did it coldly’: Israel Used AI to Identify 37,000 Hamas Targets, *The Guardian* (3 April 2024), <https://www.theguardian.com/world/2024/apr/03/israel-gaza-ai-database-hamas-airstrikes>.

62 *Ibid.*

63 David Wallace-Wells, *What War by A.I. Actually Looks Like*, *The New York Times* (10 April 2024), <https://www.nytimes.com/2024/04/10/opinion/war-ai-israel-gaza-ukraine.html>.

least, recklessly.<sup>64</sup> AWS complicate this framework because they lack a human mental state, operating solely based on pre-programmed algorithms and machine learning processes.

For example, the crime of intentionally targeting civilian structures under Article 8 of the Rome Statute requires proof that the act was deliberate. If an AWS erroneously targets a hospital due to a sensor malfunction or data misclassification, attributing intent becomes nearly impossible. The AWS itself, as a machine, cannot form intent, and its operators or programmers may argue that they acted with due diligence during its design and deployment. This creates a significant accountability gap.<sup>65</sup>

Similarly, the reckless launching of attacks on civilian populations poses challenges when AWS are involved. Recklessness in legal terms implies a conscious disregard of a substantial and unjustifiable risk. However, AWS lack consciousness and cannot assess or disregard risks in the way humans can. The opacity of AI decision-making processes, often referred to as the 'black box problem', makes it difficult to determine whether the actions of an AWS were foreseeable or preventable.<sup>66</sup>

The issue is further exacerbated in cases of torture, where intent plays a critical role. Torture, as defined under international law, involves the infliction of severe pain or suffering for purposes such as obtaining information or punishment. An AWS tasked with interrogation might use excessive force due to programming flaws or biased data, but such an outcome would lack the purposeful intent required to constitute torture. Without human involvement in the decision-making process, the prosecution of such crimes faces significant legal and evidentiary hurdles.

The absence of mens rea in AWS challenges the foundational principles of ICL. As AI systems continue to evolve, legal frameworks must adapt to address these gaps. One potential solution involves creating a presumption of responsibility for commanders deploying AWS unless they can demonstrate rigorous compliance with IHL standards. Additionally, the development of international treaties specific to AWS could help establish clearer guidelines for accountability and use.

#### *IV Proposing a Standard of Strict Liability for Commanders*

To address the accountability gap created by the absence of mens rea in AWS, this article proposes a standard of strict liability for the first commander in the chain of command who authorizes the use of an AWS that commits a war crime. Strict liability removes the need to prove intent, focusing instead on the outcome of the actions taken by the AWS and the commander's responsibility to prevent violations of IHL.

Under this standard, the commander who authorizes the deployment of an AWS would bear full legal responsibility for any war crimes committed by the system. This approach ensures accountability while emphasizing the importance of human oversight in the use of lethal autonomous technologies. By holding

64 Rome Statute of the International Criminal Court Art. 8, July 17, 1998, 2187 U.N.T.S. 90.

65 Gov't Accountability Off., *Artificial Intelligence: Status of Developing and Acquiring Capabilities for Weapon Systems* (2022), <https://www.gao.gov/assets/gao-22-104765.pdf>.

66 Gurman Narula & Sharad Khemka, *Black Box Problem in Artificial Intelligence: Transparency and Accountability Challenges*, National Law Institute University Bhopal Cell for Law & Technology (2021).

commanders strictly liable, military leadership would be compelled to thoroughly assess AWS for compliance with ethical and operational standards before deploying them. It would also incentivize rigorous testing, auditing and training protocols, ensuring that AWS are as compliant with IHL as possible before deployment.

Strict liability addresses several core issues. First, it ensures that victims of war crimes committed by AWS are not left without legal recourse due to the absence of human intent. Second, it provides a straightforward legal standard that eliminates the need to untangle the complex and often opaque decision-making processes of autonomous systems. The black box problem underscores the lack of transparency in AI-driven systems, where the computations leading to a decision remain concealed and difficult to trace.<sup>67</sup> Finally, it reinforces the principle that humans, not machines, bear ultimate responsibility for adherence to IHL.

While some may argue that strict liability could discourage commanders from utilizing AWS for fear of legal repercussions, this concern does not outweigh the imperative of maintaining accountability for war crimes. In fact, the implementation of strict liability would likely lead to more cautious and deliberate deployment of AWS, aligning military operations more closely with international legal and ethical standards. This standard does not aim to stifle technological advancement but to ensure that the integration of AI in warfare does not come at the expense of accountability or the principles of justice.

By focusing on the responsibility of commanders, strict liability provides a clear and enforceable framework to address the legal and ethical challenges posed by AWS. It underscores the absolute necessity of maintaining human accountability, even in an era increasingly dominated by autonomous technologies.

## D Conclusion

The rapid development and deployment of AWS present profound challenges to IHL and the principles of accountability. The lack of a human mental state in AWS disrupts traditional frameworks for prosecuting war crimes, particularly those requiring intent or recklessness. This article has explored how the black box problem, technological opacity and gaps in legal standards further complicate the issue, leaving victims without meaningful recourse. By implementing a strict liability standard for commanders who authorize the use of AWS, the legal system can help close this accountability gap. Such a framework ensures that those in positions of command bear responsibility for preventing violations of international law, while reinforcing the essential role of human oversight in the rapidly evolving landscape of AI-driven warfare.

<sup>67</sup> *Ibid.*