



Beyond the Geneva conventions: AI and accountability in modern warfare

Bhavyaa Hardik Mehta

Faculty of Law, GLS University, Gujarat, India

DOI: <https://doi.org/10.66856/ijl.2026.12.3.12279>

Abstract

The growing use of Artificial Intelligence in modern warfare creates new dimensions in military operations. There are multiple benefits to using artificial intelligence in the military, including more effective surveillance, reconnaissance, and improved target identification. However, weapons systems using AI pose a risk to international law and humanitarian principles. A fundamental challenge is raised by the lack of clarity in terms of responsibility in case of an unlawful use of force by autonomous weapons systems. International Humanitarian Law, including the Geneva Conventions and Additional Protocols remains the relevant body of law when it comes to war and conflict, but its applicability is questioned regarding advanced technologies. This article will explore the question of responsibility and accountability of states for autonomous actions committed by military machines. It will also assess the applicability of command responsibility, as well as the potential shortcomings of international humanitarian law in the face of increasingly complex military technology. This article argues that despite the continued applicability of International Humanitarian Law, accountability rules need to be established that would bring clarity and responsibility to autonomous weapons in order to avoid civilian casualties and ensure humanity in warfare.

Keywords: Autonomous weapons systems, modern warfare, international humanitarian law, Geneva Conventions

Introduction

Warfare is never stagnant; the development of new technologies has consistently led to the evolution of the means and methods of fighting. From traditional to cyber and drones, weapons have challenged the laws regulating armed conflicts. Artificial Intelligence (AI) is the latest concept to be incorporated into defence systems. Today, AI is utilised for surveillance, reconnaissance, target acquisition, and weapons systems that could operate with limited human intervention. Although such technology could increase operational efficiency and reduce casualties, the legal and ethical implications of weapons operated by or designed by AI pose several challenges.

This article will consider the laws governing warfare particularly the Geneva Conventions of 1949 and Additional Protocols, and attempt to identify the accountability gap in the laws related to AI driven conflict.

Artificial Intelligence in Contemporary Armed Conflict

Artificial Intelligence is a concept that refers to systems that perform tasks typically requiring human intellect or judgement. It is used to perform a wide range of functions, including, but not limited to intelligence analysis, surveillance, reconnaissance, logistics, cyber-warfare, and weaponisation.

It is the application of AI to weaponry that has caused the most controversy about its legality under International Humanitarian Law (IHL). Although weapons driven by AI and requiring limited human input are not yet fully developed, present-day AI-assisted technologies enable soldiers to take over combat roles traditionally assumed by humans. The rise of AI has sparked debate among researchers about whether machines can think and decide according to humanitarian principles. But International Humanitarian Law particularly the Geneva Conventions and

Additional Protocols still apply to warfare even as technology advances.

The Continued Relevance of International Humanitarian Law

The Geneva Conventions do not become irrelevant when new technologies are introduced. Even though the Conventions were drafted in 1949, the rules govern contemporary armed conflict. All states participating in hostilities must comply with the principles of distinction, proportionality, military necessity, and precautions. However, the application of these principles becomes complicated when weapons are computerised and partially autonomous.

For example, the principle of distinction requires belligerents to differentiate between combatants and civilians. But in situations where cities are used as military objectives and dual-use infrastructure exists it may not be possible for a computer programme to determine friend from foe.

Similarly, the principle of proportionality requires that attacks must not cause foreseeable disproportionate harm to civilians. Such judgement in targeting requires contextual analysis and morality, raising doubts about the ability of an algorithm to determine what an acceptable risk to civilians is.

Finally, the principle of precaution requires parties to take all feasible measures to avoid civilian casualties while the principle of military necessity allows targeting of resources necessary for the enemy's war effort. Computers may assist human beings in warfare by analysing information quickly and efficiently. But errors in calculations or inaccurate information may lead to attacks causing greater civilian harm than anticipated. Thus, although the Geneva Conventions are still relevant their application becomes more challenging with autonomous weapons.

The Accountability Gap

Perhaps the most pressing concern about AI driven conflict is the lack of a legal framework concerning responsibility. Most laws about warfare are based on the assumption that humans commit war crimes and can be prosecuted for them. But with AI-driven weapons, it may become challenging to determine which party bears responsibility when an unlawful attack occurs. For example, if an autonomous drone erroneously targets a hospital killing several civilians and soldiers, who is responsible for authorising such an attack? Could the commander be prosecuted for not reviewing the actions of the autonomous weapon system? Similarly, could the defence lawyers argue that since weapons had human input in their design the responsibility lies with the state that manufactured them? The uncertainty about who bears responsibility for unlawful attacks constitutes what experts have termed an accountability gap. Victims of unlawful attacks may not receive justice if responsibility is not explicitly determined by IHL. Furthermore, states may take advantage of the accountability gap to violate the law with impunity.

The Čelebići Case ^[1] and the evolution of the command responsibility

In *Prosecutor v. Delalić et al.* (The Čelebići Case), the ICTY expanded on the principle of command responsibility, finding that a commander can be held responsible if they have knowledge or grounds for belief that their subordinate has committed is committing or is about to commit a violation of International Humanitarian Law and did not take all necessary and feasible measures to prevent such violations or to punish their perpetrators. The tribunal held that for command responsibility to arise the commander must have had knowledge or grounds for belief that his subordinates had committed violations. The commander's knowledge or belief does not need to be based on specific information about the unlawful acts; it is enough for the commander to have received generalized information that could raise reasonable suspicion about the subordinates' violations. Therefore, the commander is obliged to investigate the facts and circumstances in order to determine whether there are grounds for suspicion of violations of International Humanitarian Law.

Although the Čelebići ruling preceded the development of artificial intelligence. The principle of command responsibility can be applied to AI enabled military operations. In situations where military commanders use AI technology that they know or should have known to be limited in its capabilities, command responsibility for unlawful targeting decisions would apply. Therefore, commanders cannot rely on the limitations of artificial intelligence to avoid responsibility for potential violations of International Humanitarian Law.

Why Existing Laws May Not Be Enough

The Geneva Conventions of 1949 ^[2] represent the core principles of IHL. But despite the humanitarian principles enshrined in the treaties. They do not directly address several challenges associated with AI driven warfare such as liability for software errors the legality of fully autonomous weapons or the responsibility of states for attacks carried out by computers. Moreover, the laws do not explicitly state how humanity could utilise new technology while still

adhering to the principles of distinction, proportionality and military necessity. Contemporary laws only provide general principles in making it difficult to apply the existing treaties to emerging technologies. Instead of ignoring the inadequacy of the current laws, States should work towards supplementing IHL with additional laws explicitly addressing the challenges posed by AI driven warfare to humanitarian principles.

The Battlefield has changed but has the law?

With the increasing incorporation of Artificial Intelligence in the military sphere this issue is becoming increasingly topical from the perspective of modern international law. The latest cases of armed conflict like the Russia Ukraine War and the conflict in Gaza show that the application of technologies enhanced by Artificial Intelligence becomes more widespread in order to gather intelligence data and identify targets.

Although such technological advances can increase the effectiveness of military actions they give rise to several issues related to the safety of civilians and accountability for military actions. The Geneva Conventions constitute the fundamental source of law regulating the matters of armed conflict, nevertheless they have been written at the time when there were no autonomous systems, *let alone* those which could be enhanced by AI. Therefore, the legal sources do not provide any solutions to the issue of determining responsibility for an attack conducted with the help of AI technology.

The Way Forward

The best way to ensure compliance with International Humanitarian Law while benefiting from the advantages offered by AI is to ensure that human involvement in warfare is always present. Human decisions can be reviewed and prosecuted if they violate International Humanitarian Law ensuring that accountability remains with state actors. Additionally, states must adhere to the weapons review process mandated by Additional Protocol I Article 36 ^[3] to ensure that AI enabled weapons are reliable, predictable, and consistent with humanitarian principles. Finally, the international community needs to engage in multilateral negotiations to establish binding treaties that govern autonomous weapon systems.

Such negotiations could take place under the aegis of the UN Convention on Certain conventional weapons which has already begun addressing some challenges posed by autonomous weapons.

Conclusion

Artificial Intelligence is a concept that has introduced unprecedented challenges to the laws governing warfare. The principles of International Humanitarian Law continue to apply to autonomous weapons but the application of existing laws has become complicated by the nature of combat roles performed by computers. Even more worrying is the lack of a legal framework clarifying which entity bears responsibility for unlawful attacks carried out by AI driven weapons. The future of armed conflict will undoubtedly include AI but ensuring that technology does not undermine humanitarian principles requires States to

clarify accountability gaps while keeping the principles of International Humanitarian Law intact.

References

1. Geneva Convention related to the protection of civilians in the time of war (Fourth Geneva Convention), 1949
2. Rome Statue of the International Criminal Court, 1998
3. Protocol Additional to Geneva Convention (12th August, 1949)
4. International Committee on Red Cross (ICRC), reports on AI-enabled warfare.