

Briding The Gap: Artificial Intelligence In Addressing The Accountability Gap In International Humanitarian Law

Priya Mondal¹, Dr. Nagesh Sawant², Dr. Ramratan Dhumal³

¹Research Scholar Ajeenkya DY Patil University, Lohegaon, Pune priya.mondal@adypu.edu.in

²PhD. Supervisor Ajeenkya DY Patil University, Lohegaon, Pune nagesh.sawant@adypu.edu.in

³Ph.D. Co-Supervisor Ajeenkya DY Patil University, Lohegaon, Pune ramratan.dhumal@adypu.edu.in

Cite this paper as: Priya Mondal, Nagesh Sawant, Ramratan Dhumal (2024) Briding The Gap: Artificial Intelligence In Addressing The Accountability Gap In International Humanitarian Law. *Frontiers in Health Informatics*, 13 (3), 7933-7945

Abstract :

Despite the negative aspects mentioned, it is important to note that the current scenario is not a result of new technology causing a completely new issue. The advancement of new technologies has made an earlier issue of unintentional civilian injuries, both incidental and inadvertent, more noticeable. Specifically, these cases underscore the absence of a global framework to hold responsible parties accountable for unintentional injury to civilians during armed conflicts.

*This article presents a more positive perspective on artificial intelligence, highlighting two problems for those who hold the belief in the existence of a gap in responsibility. Initially, advocates of accountability gaps must provide more clarification about the specific circumstances in which responsibility gaps arise. Acknowledging a challenging limitation in plausibility regarding the origin of these gaps is hard to deny. It becomes clear that the situations in which responsibility gaps emerge are unclear once we acknowledge this. Furthermore, it is essential to address the reasons behind our worry about accountability gaps, should they arise. This paper addresses the primary issues with responsibility gaps, including the repercussions of such gaps and the breaches of *jus in bello*.*

Keywords : Artificial intelligence , Accountability Gap , International Humanitarian Law, Warfare.

INEFFACEABLE ACCOUNTABILITY GAP

Despite the negative consequences mentioned, this is not a situation where new technology is causing a fresh issue. The emergence of new technology has brought greater attention to the longstanding issue of unintended civilian harms, whether they occur incidentally or accidentally. These instances highlight the absence of a global accountability system for unintentional harm to civilians during times of war. There are certain accountability mechanisms in place, but they fall short when it comes to addressing unintentional civilian harms that are within the bounds of the law. The deliberate targeting of civilians or the commission of grave breaches of international humanitarian law may lead to individuals being held accountable for war crimes. However, the emergence of autonomous weapon systems and AI-enabled systems raises concerns about the potential for harmful actions to be taken without anyone having the necessary criminal intent. States that commit internationally wrongful acts can be held accountable under the law of state responsibility. However, if an act is deemed lawful (which can include collateral damage and accidental harms), it is not considered internationally wrongful, and therefore the law of state responsibility does not apply. The law of armed conflict aims to reduce, rather than completely eliminate, the chances of causing unnecessary harm to civilians. There is still a significant amount of collateral damage allowed. According to international law, there is no legal responsibility for the negative outcomes of lawful actions during armed conflict. That point deserves to be reiterated. According to international law, there is no legal responsibility for the negative outcomes of lawful actions during times of armed conflict. Unfortunately, innocent civilians often bear the brunt of unintended harm. Moreover, the lack of a deterrent in

International Humanitarian Law (IHL) potentially contributes to unintended harm to civilians¹.

The lack of accountability becomes more apparent as technology-enabled behavior underscores the necessity for new accountability systems. Addressing worries about AI's unintentional negative impact and the lack of accountability in international humanitarian law (IHL), some have suggested implementing regulations and policy guidance specifically tailored to technology. For instance, the US Department of Defense (DoD) has recently embraced a collection of ethical guidelines for the application of AI, which highlights that DoD personnel must continue to be “responsible for the development, deployment, and use of AI capabilities”². Others focus instead on revising the legal regimes themselves. Military personnel, advisors, and civilian proponents have presented ethical and strategic justifications for voluntarily compensating civilians who have suffered harm³. Some legal experts are considering ways to enhance the reparations process⁴ while others are suggesting reinterpretations or amendments to current regulations that would increase government accountability for harm to civilians in times of war.

Proposing the establishment of an accountability regime that holds states responsible for both legal and illegal actions during times of war that cause harm to civilians. This would encompass situations that involve the deployment of autonomous weapons or other similar scenarios. A potential structure for such a regime could be an adversarial tribunal, where states are held accountable for the harmful consequences of their actions in armed conflict. Another option could be a no-fault system, where states contribute to a victim's fund that is then distributed to claimants. Alternatively, a hybrid approach could be considered, aiming to combine the strengths of both approaches⁵.

Technological advancements frequently amplify previously overlooked, sporadic, or underemphasised issues, making them more prominent, widespread, or impactful. Although numerous proposals centre around regulating specific weapons technologies to address concerns regarding unintended consequences or accidents, it is important to note that the law has not fallen behind in keeping up with technological advancements. On the other hand, the focus has shifted towards the accountability gap that is inherent in the structure of IHL, thanks to technological advancements. By doing so, the emergence of AI and other new military technologies has brought attention to the importance of establishing accountability mechanisms for any harm caused to civilians.

ISSUE OF RESPONSIBILITY

The machines follow logical mechanisms that are determined by the laws of causation, and as a result, they can only be considered causally accountable for their actions, not morally accountable. Based on this approach, it may be more logical to consider the concept of shared responsibility in intricate settings involving technology. Therefore, the following section presents advocates of a distributed perspective on human-machine responsibility. Therefore, suggests a post-phenomenological viewpoint when addressing ambient intelligence and persuasive technologies⁶. In this analysis, Verbeek highlights the fact that our existence is deeply intertwined with the material world, leaving us no choice but to engage with various technologies. As a result, the concept of freedom and intentionality is seen as a combination of human actions and the influence of objects. According to him, Technology significantly influences our ability to exercise freedom of action. Similarly, Technology, despite not having human-like intentions, can still be given intentionality as it has the ability to

¹ Lieblich, E. (2019). The facilitative function of jus in bello. SSRN Electronic Journal. <https://doi.org/10.2139/ssrn.3334526>

² Nguyen, A., Ngo, H. N., Hong, Y., Dang, B., & Nguyen, B.-P. T. (2023). Ethical principles for artificial intelligence in education. *Education and Information Technologies*, 28(4), 4221–4241. <https://doi.org/10.1007/s10639-022-11316-w>

³ Otynshiyeva, A. A., Sairambayeva, Zh. T., Baimagambetova, Z. M., Nyssanbekova, L. B., & Castellino, J. (2022). Human rights institute under english law. *Journal of Actual Problems of Jurisprudence*. <https://doi.org/10.26577/JAPJ.2022.v104.i4.015>

⁴ Autonomous Weapons and the Responsibility Gap in light of the Mens Rea of the War Crime of Attacking Civilians in the ICC Statute.

⁵ Crotoft, R. (2014). The killer robots are here: legal and policy implications. *Cardozo L. Rev.*, 36, 1837.

⁶ Verbeek, P. P. (2011). *Moralizing technology: Understanding and designing the morality of things*. University of Chicago press.

impact and shape our actions and encounters. Therefore, the intricate connection between humans and technology means that freedom cannot be attributed solely to either humans or technologies. However, Verbeek argues that despite the influence of technological mediations, we can still engage in discussions about human responsibility. He believes that these mediations can actually provide opportunities for moral decision making⁷ In addition, the importance of recognising moral agency in intelligent artificial systems that can adapt and respond autonomously to their surroundings. Therefore, an artificial agent can be held morally accountable for a specific action it is causally responsible for. What truly matters is whether an individual's actions can be morally evaluated, based on whether they are capable of performing actions that can be considered good or evil. Through this approach, the authors demonstrate how normative actions can occur without necessarily implying moral responsibility based on intentional states. "Moral accountability alone is not enough to fully establish moral responsibility⁸." The moral responsibility of a person should be taken into account if they are responsible for a particular action or behaviour that can be judged morally. To take moral responsibility for x, the person must show the relevant intentional states. Floridi and Sanders propose that their perspective on agency can contribute to the ethical discourse surrounding accountability in situations involving non-human entities. It is believed that ensuring alignment between human and artificial moral agents will enable the development of a process to censor immoral artificial agents. This process would range from monitoring their actions to ultimately eliminating them from cyberspace, without any possibility of recovery We are no longer fixated on identifying the individual responsible when something bad occurs, as we now recognize that the source of good and evil can sometimes transcend specific individuals or groups. This provides a more nuanced understanding of moral considerations in non-human contexts⁹.

The recommendations posit the potential advantages of extending our consideration beyond the attribution of responsibility and accountability solely to human agency. Likewise, in a martial framework, it furnishes a rationale for responsibility anchored in a relational ontological perspective. In this presentation, the author introduces the idea of a "swarm," relating to the deployment of intelligent swarm systems within the realm of military operations. These systems endeavour to mimic the self-organizing and cognitive dynamics identified in biological entities and wildlife. The ambition is to formulate algorithms that draw inspiration from the aggregate cognitive phenomena manifested by congregations of birds, colonies of ants, swarms of bees, and other self-organizing entities like microbial systems. Building on this concept, a shift in how we approach human-machine interactions. Rather than viewing humans and machines as separate entities, that we recognise their inherent connection and shared activity. This perspective encourages us to consider the broader network and collective movements before focussing on specific connections. Given the absence of a transparent chain of accountability, which acknowledges that this shift in relationships adds complexity to questions of moral responsibility within a military context¹⁰. Nevertheless, it concludes optimistically, expressing optimism for our future efforts in the field of (machine) ethics. Hence, the epistemological principles linked to his shift in relationships could offer a hopeful basis for tackling the challenges of accountability gap.

Regarding Singer's swarm concept, one might argue that its potential moral, social, and epistemic effects could be seen as having a negative impact on the future. The absence of control and predictability is particularly unsettling to contemporary thinking. Nevertheless, it could serve as a source of comfort to contemplate ethical considerations [...] ¹¹may extend into swarms too: Initially unseen, yet possibly influential communities of individuals and concepts that can assist us in gaining deeper insights into what the future holds, in discovering unforeseen paths for transformation, and in revealing various avenues of potential.¹²

⁷ Ibid

⁸ Floridi, L., & Sanders, J. W. (2004). On the morality of artificial agents. *Minds and machines*, 14, 349-379.

⁹ Ibid

¹⁰ Coeckelbergh, M. (2011). Humans, animals, and robots: A phenomenological approach to human-robot relations. *International Journal of Social Robotics*, 3, 197-204.

¹¹ Ibid

¹² Weizmann, N., & Costas Trascasas, M. (2014). Autonomous weapon systems under international law.

IMPORTANCE OF BRIDGING ACCOUNTABILITY WITH INTERNATIONAL HUMANITARIAN LAW WITH REGARDS TO AUTONOUS WEAPON SYSTEMS

Under certain circumstances, AWS may struggle to adhere to the regulations of International Humanitarian Law and International Human Rights Law, resulting in potential infringements on crucial rights such as the right to life. We must now identify the responsible party for the breaches involving an Autonomous Weapon System. Autonomous weapons systems that lack 'Meaningful Human Control' can be highly unpredictable in the battlefield or any other context where they are deployed. When there are instances of unintentional law violations, it becomes unclear who should be held legally responsible, resulting in a gap in accountability.¹³ Ensuring accountability in international law is crucial as it directly impacts the ability of victims to seek legal remedies.¹⁴

It is crucial to recognize the seriousness of the concerns brought up by AWS regarding accountability before we examine the details of the condensed arguments. The introduction highlighted that the potential accountability shortfall resulting from AWS will adversely affect victims' ability to seek redress. This field of international law holds significant importance. Without accountability, international law "brutum fulmen – a harmless thunderbolt," which means holds no weight and becomes nothing more than an empty threat. According to Steven Ratner, international law serves the dual purpose of establishing standards for governments, non-state actors, and their agents, as well as outlining the consequences that result from failing to meet those standards. Accountability for failure to abide by international humanitarian law norms, some of which are part of jus cogens¹⁵, is crucial to ensure their effectiveness. Some experts have noted that the lack of accountability for violations could potentially jeopardise the overall maintenance of peace and security¹⁶. Accountability plays a crucial role in international law as it is closely tied to the victim's right to seek redress. When dealing with violations resulting from the use of certain weapons, it's essential to take into account the victims from both civilian and military sectors.

*[Unlawful] use of a weapon will give rise to a right to a remedy or reparation. Such unlawful use of weapons includes: any use of a weapon that has been outlawed in all circumstances, such as biological weapons or, at least for any State Party to the relevant treaty, anti-personnel mines or cluster munitions; the use of indiscriminate weapons or the indiscriminate use of a weapon as a method of warfare in an armed conflict; or the use of force that is disproportionate or excessive during law enforcement. Any wilful or negligent failure to protect victims from harmful weapons, especially explosive weapons delivered from drones, mines, sub-munitions or other victim-activated explosive devices has also been recognised...as unlawful conduct tantamount to a rights violation.*¹⁷

The accountability challenges that are posed by AWS must be taken seriously as they threaten some aspects of

¹³ Akkuş, B. (2022). Autonomous weapon systems under international law. *Güvenlik Bilimleri Dergisi*, 11(2), 333–366. <https://doi.org/10.28956/gbd.1078155>

¹⁴ Burke, M., & PERSI, V. L. (2014). Remedies and reparations. weapons Under international hUMAN rights Law (ed. S. Casey-Maslen), Cambridge University Press Cambridge, 542-590.

¹⁵ Ratner, S. R., Abrams, J. S., & Bischoff, J. L. (2009). Accountability for human rights atrocities in international law: Beyond the Nuremberg legacy (3rd ed). Oxford University Press.

¹⁶ Jones, J. R. W. D., & Powles, S. (2003). International criminal practice: The international criminal tribunal for the former yugoslavia, the international criminal tribunal for rwanda, the international criminal court, the special court for sierra leone, the east timor special panel for serious crimes, war crimes prosecutions in kosovo (3rd ed). Transnational ; Oxford University Press.

¹⁷ Ferstman, C. (2017). International organizations and the fight for accountability: The remedies and reparations gap (First edition). Oxford University Press.

victims' right to remedy. There is an ongoing debate regarding the ethical implications of responsibility gaps in autonomous systems. Some argue that these gaps present a new moral problem, while others hold the opposing view that they do not. Among the authors who discuss the issue of responsibility gaps, there are two main positions. One position highlights that these gaps are not a new phenomenon and can be observed in current practices. The other position argues that while gaps do exist, they should be viewed as unintentional occurrences rather than moral problems.¹⁸ The first position can be attributed to having a practical understanding of the current state of human decision-making in warfare. Patrick Taylor Smith provides a strong defence of this perspective. The argument presented is that while there may be instances where casualties occur due to laws, such outcomes are not unique and happen regularly. The belief that responsibility gaps are unsolvable is based on a mistaken assumption that this situation only occurs when LAWS are used¹⁹. The potential danger of autonomous weapons behaving in ways that commanders did not intend or could not have predicted is a concern. However, this issue is not exclusive to autonomous weapons as responsibility gaps also occur with human warfighters.²⁰ Furthermore, it is important to acknowledge that these accountability gaps are not a recent development, as they are also present in modern warfare. It is noteworthy that critics raise apprehensions regarding accountability gaps specifically for autonomous drones, while similar gaps are accepted for other advanced weaponry. It is important to classify responsibility gaps solely as accidents. The issue of responsibility in human-AI interactions can be compared to our responsibility when using and training non-human animals for specific purposes, like police dogs and racehorses. In this analogy, the focus is on the responsibility for the use of the instrument.

These cases highlight the inherent difficulty of completely eradicating negative outcomes. It is evident that individuals who neglect to take necessary precautions or engage in activities that carry a risk of harm are typically held accountable. However, these cases highlight the possibility of situations where it is justifiable to conclude that no individual can be held accountable due to the fulfilment of all necessary obligations of care. During these circumstances, it is crucial to avoid discussing responsibility gaps in these scenarios and to regard them as accidents instead, as they don't pose a moral dilemma²¹. The adverse outcomes stemming from regulations should be handled and dealt with in a comparable fashion to mishaps involving non-learning systems. The pivotal factor in both scenarios is the level of tolerance.

The minimum level of reliability necessary for a system is determined by the tolerance level, which is a critical factor. For instance, engineers designing a bridge must ensure it is sufficiently strong, and the contractors are responsible for meeting this requirement. When accidents happen within the acceptable tolerance level, such as due to engineers neglecting extreme temperature changes or users exceeding the maximum weight limit of a bridge, accountability needs to be assigned to at least one individual involved, whether it's the engineer, controller, or user. It is possible that no individual is responsible for deaths that fall outside the required tolerance levels, like in the case of a rare rainstorm occurring once every century. In this scenario, it was determined that the bridge wouldn't withstand such an event due to the low probability and high construction cost. Therefore, when all necessary precautions have been taken and an undesirable result still occurs, it should be considered an accident for which no one can be held responsible, especially in the context of laws.²² Therefore, suggesting that the absence of responsibility is not a concern, as no one can be held accountable, results in dismissing the significance of the responsibility gap.

CONSEQUENCES OF NOT BRIDGING THE ACCOUNTABILITY GAP

¹⁸ Mazzeschi, R. P. (2003). Reparation claims by individuals for state breaches of humanitarian law and human rights: an overview. *J. Int'l Crim. Just.*, 1, 339.

¹⁹ Oimann, A. K. (2023). The responsibility gap and LAWS: A critical mapping of the debate. *Philosophy & Technology*, 36(1), 3.

²⁰ Six, P. (2002). Ethics, regulation and the new artificial intelligence, part II: autonomy and liability. *Information, Communication and Society*, 4(3), 406-434.

²¹ Gerdes, A. (2018). Lethal autonomous weapon systems and responsibility gaps. *Philosophy Study*, 8(5), 231-239.

²² Wemmers, J. A. (2009). Victim reparation and the International Criminal Court. *International Review of Victimology*, 16(2), 123-126.

Concerns about autonomous weapons systems and their potential to create gaps in responsibility are significant. It is important to be able to identify those accountable for war crimes, as the lack of ability to do so would make prosecuting them meaningless. This, in turn, could have disastrous consequences for the future conduct of wars. The deployment of autonomous systems in the military sector raises concerns about war crimes and military autonomous systems. However, his apprehension goes beyond this and also encompasses non-military autonomous systems. The potential consequences of non-military autonomous systems are widely discussed in the artificial intelligence and accountability literature.²³ The significance of ensuring that individuals are held responsible lies in its ability to motivate developers to prioritize the safety and dependability of these devices. The presence of gaps in accountability is a source of worry. When people engaged in the design, regulation, and operation of a system can escape accountability for their misconduct, it diminishes their incentive to prevent such behaviour.²⁴

Concerns arise in the absence of accountability, which can lead to a decrease in motivation to minimize harm and damage. When there are no proper incentives to reduce harm and damage, individuals responsible for developing and operating autonomous systems may fail to exercise due care or, in some instances, intentionally cause harm. This is because they have the ability to do so without facing consequences. Their reckless or deliberately malicious actions will result in catastrophic outcomes for the broader community. The concern revolves around the need for individuals to act responsibly and avoid causing harm, or rather the absence of consequences for acting recklessly and with malicious intent. Responsibility gaps undermine the incentive structure that benefits society. It would be beneficial to provide two additional clarifications in order to gain a better understanding of this concern. First, it is worth considering that the lack of responsibility may remove any deterrents to acting without proper care. However, it is important to clearly define the nature of these deterrents to categorise three types of disincentives that may be considered insufficient in the presence of responsibility gaps²⁵:

- (1) assigning blame,
- (2) establishing moral liability for damages, and
- (3) imposing punishment.

Additionally, it is crucial to acknowledge that we are discussing the potential disincentives individuals may rightfully encounter, rather than the ones they are currently experiencing. The argument is that in cases where there is a gap in responsibility, it would be unfair to hold individuals accountable, demand compensation for damages, or impose punishment. They would lack these disincentives because it is important to refrain from engaging in unjust actions. Therefore, we have three distinct variations of this concern²⁶:

- (1) In cases where there are gaps in responsibility, it would be unjust to assign blame to individuals, as it is unfair to criticize those who are not accountable. Consequently, individuals would lack this incentive to act responsibly.
- (2) The cost of damages shouldn't fall on individuals if there are gaps in responsibility. It's unfair to expect those who aren't responsible to bear the consequences. This would remove the incentive for individuals to act carefully.
- (3) Individuals who do not have any responsibility should not be punished because it would be unfair to do so. Punishing people who are not responsible is unjust, and it could deter individuals from behaving responsibly. The idea that individuals can act recklessly or intentionally cause harm without facing consequences due to potential gaps in responsibility is unacceptable. The lack of a negative outcome can be attributed to the discovery

²³ Taddeo, M., & Blanchard, A. (2022). Accepting moral responsibility for the actions of autonomous weapons systems—a moral gambit. *Philosophy & Technology*, 35(3), 78.

²⁴ Santoni de Sio, F., & Mecacci, G. (2021). Four responsibility gaps with artificial intelligence: Why they matter and how to address them. *Philosophy & Technology*, 34(4), 1057-1084.

²⁵ Ibid

²⁶ Nucci, E. D., & Sio, F. S. D. (2016). Drones and responsibility: Legal, philosophical, and sociotechnical perspectives on remotely controlled weapons (E. Di Nucci & F. S. De Sio, Eds.; 1st ed.). Routledge.
<https://doi.org/10.4324/9781315578187>

that the occurrence of gaps in responsibility is influenced by the plausibility constraint mentioned earlier. If any responsibility gaps occur, they do not occur when individuals act without exercising proper caution. The concept of blame. Pointing fingers at individuals may deter them from participating in actions that are detrimental to society.²⁷ The absence of responsibility gaps will eliminate the motivation to avoid participating in such actions. Nonetheless, the plausibility constraint mentioned earlier indicates that responsibility gaps do not arise from situations of negligence or ill will. As a result, this worry is unfounded. Indeed, if the absence of accountability allowed individuals to escape blame for their negligent, reckless, or malicious actions, it is likely that people would be more tempted to engage in such behaviour, as they would not face rightful blame. However, it is highly unlikely to assume that the behaviour of an autonomous system can be justified solely based on its high degree of autonomy. If someone's actions, whether they were careless, thoughtless, or intentionally harmful, resulted in harm caused by an autonomous system, it is evident that the individual responsible, whether it be the manufacturer, programmer, operator, or anyone else involved, should be held accountable. It is therefore fair to hold this individual accountable and discourage such behaviour in this manner. Gaps in responsibility may occur in specific circumstances, provided that their appearance is bound by the earlier mentioned plausibility requirement. They do not justify negligent or malicious conduct.

The same principle can be applied to (2) and (3). The idea behind this concept is that it would not be just to hold manufacturers, programmers, or operators financially accountable for any harm they cause due to their negligence. Such a requirement could potentially discourage them from taking certain actions.²⁸ The concern about the potential consequences of responsibility gaps in prosecuting war criminals, which could have laid a significant impact on the conduct of future wars. It seems that the reasoning behind this line of thought is that if no one is held responsible, leaders of independent weapons may not have a strong enough deterrent to stop them from committing war crimes. This is due to the belief that it would be unjust to punish or hold them accountable for their actions. Consequently, there is an increased chance of future war crimes taking place.²⁹

The assertion that it is unfair to hold individuals accountable when they are not responsible is essentially a tautology, as responsibility is inherently linked to being deserving of blame. In contrast, (2) and (3) raise more controversial arguments regarding the fairness of making individuals who are not at fault bear the cost of damages or subjecting them to punishment. This raises a further objection for (2) and (3). When autonomous systems are involved, it is improbable to assume that careless or malicious behaviour would not be deemed blameworthy. Furthermore, there is room for debate regarding the notion that it is unjust to hold individuals accountable for damages or impose penalties on them, even if they are not directly responsible for the negative consequences they have caused. The autonomy of autonomous systems does not automatically imply that it would be unfair to hold careless or malicious individuals accountable for the damages they have caused, regardless of whether it absolves them of blame.³⁰ Similarly, it could be argued that taking action to discourage careless or malicious behaviour, even if it is not deserving of blame, is justifiable. Nonetheless, the primary flaw of the argument from consequences is that it unrealistically assumes that gaps in responsibility would justify and thus promote negligent, reckless, or malicious behaviour.

Responsibility gaps might exist, but they do not occur when individuals behave negligently or with malicious intent. Those involved in the development, production, and operation of autonomous systems may rightfully face criticism, legal responsibility, and consequences for their actions. This will discourage socially harmful behaviour.

UNBRIDGED ACCOUNTABILITY IN INTERNATIONAL HUMANITARIAN LAW AND WAR

The primary issue revolves around the tension between the concepts of fairness in times of war and the absence of responsibility for the harm inflicted by these armaments. It is crucial to recognise that in the context of a just

²⁷ Ibid

²⁸ Roff, H. M. (2013). Killing in war: Responsibility, liability, and lethal autonomous robots. In *Routledge Handbook of Ethics and War* (pp. 352-364). Routledge.

²⁹ Ibid

³⁰ ibid

war, there must be accountability for the loss of enemy lives. It is crucial that accountability is established for the loss of civilian lives. The use of autonomous weapons systems raises ethical concerns due to the lack of accountability for their actions, which goes against the principles of *jus in bello*. Many consider opposition to autonomous weapons systems to be a major reflection of ethical concerns about artificial intelligence and accountability. However, it also falls short when subjected to careful examination. In the previous section, instead of denying the possibility of responsibility gaps, these gaps may not be as problematic as they have been portrayed. Considering the possibility of responsibility gaps is essential, and it's important to note that there might be cases where a military general or commander of an autonomous weapons system is not held responsible for the actions of the system because of its autonomy. However, more clarity is necessary to define the specific situations in which this occurs. Just war theory doesn't find these lapses in accountability to be worrisome. The first and second principles of *jus in bello* are especially significant from this standpoint.

1. The importance of distinguishing between military objectives and civilians cannot be overstated. It is crucial for belligerents to intentionally target only military objectives, while ensuring the safety of civilians.

2. When considering the consequences of military actions, it is essential to weigh the foreseen but unintended harms against the military advantage gained. Proportionality is key in maintaining a just and balanced approach.

3. Necessity: It is important to prioritise the use of the least harmful means possible.

If there are gaps in responsibility, it can lead to a breach of the mentioned principles or create challenges in following them.

The assignment and/or distribution of accountability is essential for the effective application of *jus in bello* principles. One example is the principle of discrimination, which emphasises the importance of distinguishing between legitimate and illegitimate targets during combat. This principle operates under the assumption that we are able to identify and hold accountable those responsible for any attacks that may potentially violate it. However, it is unclear why the issue of responsibility gaps would conflict with the aforementioned principles of *jus in bello*. Autonomous weapons systems should be programmed and deployed in a way that either complies with the stated principles or leads to breaches of these conditions. In this particular situation, if civilians are not deliberately targeted and any resulting harm is unintentional but reasonable, while also ensuring that the least harmful methods available are utilized, it can be determined that the three criteria of *jus in bello* have been met, thus making the war just in this aspect.³¹ The principles of *jus in bello* can be followed even in situations where no one can be held accountable. If civilians are targeted, the potential unintended consequences may be disproportionate, and using excessively harmful means clearly violates the principles of *jus in bello*. Consequently, the war may be deemed unjust. However, the lack of someone who can be held responsible does not factor into the violation of these principles. If it were indeed possible to assign rightful responsibility for the actions of weapons, it would still result in a violation of principles. The principles are considered violated due to the civilian casualties, the unintended but disproportionate harm caused, or the use of means that were not the least harmful option available. Compliance with the three most commonly applied principles of *jus in bello* is not synonymous with responsibility. Utilizing autonomous weapons systems without resolving responsibility gaps would be viewed as a violation of *jus in bello*. Therefore, the crucial issue to ponder is the rationale behind endorsing this additional principle of *jus in bello*.

The concept that this rule is supported by the same elements that form the ethical foundation of the three previously mentioned essential principles of justice in war, such as demonstrating consideration for the humanity of our opponents. The principles of justice in war are rooted in our respect for our enemies, as stated in their work ³². Engaging in actions that result in the death of our adversaries, particularly innocent civilians, without holding anyone accountable for it would go against the fundamental principles of *jus in bello*. Such actions would demonstrate a lack of respect for our enemies and undermine the core values that underpin *jus in bello*. To quote sparrow:

It is a minimal expression of respect due to our enemy

³¹ Schwarz, E. (2021). Autonomous weapons systems, artificial intelligence, and the problem of meaningful human control. *Philosophical Journal of Conflict and Violence*, 5(1), 53–72. <https://doi.org/10.22618/TP.PJCV.20215.1.139004>

³² Sparrow, R. (2007). Killer robots. *Journal of applied philosophy*, 24(1), 62–77.

*— if war is going to be governed by morality at all—
that someone should accept responsibility, or be capable of being held responsible, for the decision to take
their life. If we fail in this, we treat our enemy like
vermin, as though they may be exterminated without
moral regard at all. The least we owe our enemies is
allowing that their lives are of sufficient worth that
someone should accept responsibility for their deaths³³.*

If this is the explanation for why autonomous weapons systems breach the principles of jus in bello, they would breach them even if the three traditional principles of jus in bello are adhered to. Nonetheless, the case for embracing responsibility is not particularly compelling.

First, the importance of the message conveyed by autonomous weapons seems to be minimal, providing only a weak argument against their use at best. When it comes to the difficulty of calculating the weight of items, deciding whether to show respect by rejecting autonomous weapons or to save the life of any individual, irrespective of their position, should be an obvious decision. The importance of displaying respect is outweighed by other important factors in warfare. Furthermore, demonstrating respect for opponents can be done in multiple ways. One possible approach is to exercise caution when it comes to taking someone's life, unless there is clear and justifiable accountability. However, there are alternative methods to demonstrate respect. As an expert in law and policy research, it is worth considering potential actions that could be taken to pay tribute to the victims of autonomous weapons. One possibility is the construction of monuments by the government to honour those who have lost their lives. Additionally, commanders of autonomous weapons could periodically observe a minute of silence as a gesture of respect for the deceased.

Alternatively, there's an argument proposing a distinct jus in bello principle which asserts that individuals should be answerable for war-time fatalities, but this too is not persuasive. It's important for me to clarify that I have not supported the deployment of autonomous weaponry. There might be additional concerns regarding 'killer robots'. However, the issue lies in their failure to adhere to jus in bello principles, resulting in responsibility gaps. While the previous discussion has primarily focused on military robots, its importance extends beyond the realm of the military. One could argue that the concept of disrespect in the absence of accountability also applies to non-military autonomous systems, such as self-driving cars or work robots. This idea would thus provide a more moral justification for objecting to the development and use of autonomous systems, regardless of the industry they are used in. There are sound reasons why Sparrow's disrespect argument does not withstand close examination. These reasons also pertain to non-military contexts, effectively addressing this broader opposition to autonomous systems that is not limited to any specific field.

RECCOMENDATION AND SUGGESTIONS:

The progress of self-governing technology presents a notable danger of diminishing the traditional human involvement in using force. Completely autonomous weapons would select and engage targets without substantial human supervision. The lack of human oversight in these systems, also known as lethal autonomous weapons systems or "killer robots," raises various legal and ethical issues³⁴. Starting in 2014, the Convention on Conventional Weapons (CCW) has held eight extensive meetings to address the issue of lethal autonomous weapons systems. Extensive analysis by researchers has been conducted on the significant challenges posed by these systems, and they have recognized the importance of preserving human control over the use of force.³⁵ The advancement towards a suitable international solution, however, has been sluggish. If governments do not transition promptly from theoretical discussions to formal agreements, the advancement of technology will surpass the progress of international diplomacy.

³³ Ibid

³⁴ Waluyo, B. D., & Fahd, R. M. (2020). Telaah konstruktivis partisipasi cina dalam the campaign to stop killer robots. Indonesian Journal of International Relations, 4(2), 92–113. <https://doi.org/10.32787/ijir.v4i2.123>

³⁵ CCW Meeting of High Contracting Parties, "Final Report," CCW/MSP/2019/9, December 13, 2019, <https://undocs.org/CCW/MSP/2019/9> (accessed May 21, 2020), para. 31.

An argument in law suggesting that fully autonomous weapons go beyond the acceptable threshold and should be banned through a new global agreement. The challenges presented by fully autonomous weapons, particularly in the context of international humanitarian law. Subsequently, it elucidates the reasons why a legally enforceable document is the most effective approach to tackle such difficulties. Ultimately, it suggests essential components of a fresh agreement aimed at maintaining significant human authority in the utilisation of force and prohibiting weaponry systems that function autonomously. The use of completely self-governing weaponry likewise poses a significant deficiency in responsibility.

Individuals must be held criminally accountable for war crimes and serious violations of the Geneva Conventions in accordance with international humanitarian law. If military leaders or operators deploy a completely autonomous weapon with the intention of committing a crime, they may be held accountable and found guilty. However, it could pose legal challenges and might be considered unfair to hold an operator responsible for the unpredictable behaviour of an autonomous robot. Ultimately, fully independent weapons violate the Martens Clause, a provision found in multiple international treaties on humanitarian law. This clause states that in the absence of formal legislation, individuals are still protected by the principles of humanity and the dictates of public conscience. Autonomous weapons, lacking the ability to show compassion or uphold human dignity, would undercut the core principles of humanity. A strong and widespread opposition to the use of completely autonomous weapons exists among various groups such as church leaders, scientists, technology professionals, civil society organizations, and the general public. This opposition suggests that this new technology goes against the moral principles and values held by the public³⁶.

Fully independent weapons pose several extra dangers that go beyond concerns about following international humanitarian law. For many people, delegating life-and-death choices to robots would cross a moral line. The deployment of entirely automated weaponry, even in law enforcement scenarios, would weaken the basic rights to life, justice, and human dignity.

EXPLORING THE COMPONENTS FOR A JUST TREATY

Since 2014, discussions have been ongoing among the parties to the CCW about the issues presented by fully autonomous weapons and the adequacy of international humanitarian law. A new treaty proposes a broad scope of implementation. The treaty should apply to any weapon system that carries out target selection and engagement using sensor processing instead of depending on human guidance.³⁷ A thorough evaluation is conducted on all systems within the category to prevent any problematic systems from evading regulation, regardless of their current or future status. The prohibitions and restrictions below are forward-thinking and specifically aimed at fully autonomous weapons.

HUMAN CENTRIC CONTROL

The importance of meaningful human control cannot be overstated in the context of the new treaty. The ethical, legal, and accountability issues that arise from fully autonomous weapons can be largely attributed to the absence of this control. Understanding the potential dangers involved, the majority of states have acknowledged the importance of human involvement in the utilisation of force. Although various states and experts have employed different language, there is a preference for the term "meaningful human control." Control is a more powerful term compared to alternatives like "intervention" and "judgement". It has a broader scope that includes both of these concepts, and it is also widely recognised in the field of international law. Ensuring that control

³⁶ Marauhn, T. (2014, May). An analysis of the potential impact of lethal autonomous weapons systems on responsibility and accountability for violations of international law. In CCW Expert Meeting on Lethal Autonomous System

³⁷ Sauer, F. (2020). Stepping back from the brink: Why multilateral regulation of autonomy in weapons systems is difficult, yet imperative and feasible. *International Review of the Red Cross*, 102(913), 235–259.
<https://doi.org/10.1017/S1816383120000466>

reaches a significant level is crucial for it to be considered "meaningful."³⁸

Different groups, such as governments, international bodies, non-governmental organizations, and individual experts, have recognized several elements of substantial human authority.³⁹ Such as follows:

1. The elements of decision-making equip people with the essential knowledge and abilities to assess whether the application of force conforms with legal and moral guidelines. It is important to note that a human operator needs specific capabilities. These encompass a deep understanding of the operational environment, a comprehensive grasp of the system's functioning and its target identification criteria, and sufficient time for thorough deliberation.
2. The elements of technology play a crucial role in a weapon system by increasing the amount of control humans have. Important considerations in technological components involve predictability and reliability, the system's ability to communicate information to a human operator, and the opportunity for human intervention following activation.
3. The operational parts impose constraints on the timing, place, and objectives of a weapon system. Constraints may include the duration between a person's lawful assessment and the use of force by a system, the period of time the system functions, and the specifics and scope of the geographical area in which it operates.⁴⁰

None of these elements alone are enough, but together they enhance the significance of control, often working together. It is important to note that the list provided may not include all possible technologies. Additional research into both current and upcoming technologies may uncover additional options.

In any case, a new legally binding agreement should encompass components that set forth criteria for substantial human oversight.

ESSENTIAL OBLIGATIONS AND PROHIBITIONS

A comprehensive structure for a legally enforceable agreement on fully autonomous weapons should consist of a fundamental requirement, as well as explicit bans and essential affirmative duties.

LEGAL OBLIGATION

It is vital for the treaty to include a broad mandate for countries to maintain substantial human oversight in the use of force. This requirement establishes a core principle that serves as a guiding framework for interpreting the remaining parts of the treaty. The expansive nature of this structure helps prevent potential loopholes that could arise in the more specific mandates. The focus on behaviour—specifically the "use of force"—rather than particular technology ensures that the treaty's demands remain applicable in the future. It acknowledges the impossibility of foreseeing all potential problematic technologies. The reference to the use of force can be applied to both scenarios of armed conflict and law enforcement operations. Within this scope, there are mandates that necessitate the retention of meaningful human control over all other systems involved in choosing targets and engaging. While these systems may not be explicitly banned under the treaty, there is a worry that they could be utilized without meaningful human control. The mandates in place pertain to all systems that select and engage targets using sensor processing. These demands establish criteria to ensure that human control over these systems is significant⁴¹.

³⁸ Docherty, B. L. (with Human Rights Watch (Organization) & Harvard Law School). (2016). Making the case: The dangers of killer robots and the need for a preemptive ban. Human Rights Watch ; IHRC, International Human Rights Clinic.

³⁹ Marks, P. (2011). Reaching critical point. *New Scientist*, 212(2841), 25. [https://doi.org/10.1016/S0262-4079\(11\)62960-6](https://doi.org/10.1016/S0262-4079(11)62960-6)

⁴⁰ Docherty, B. (2019). Key elements of a treaty on fully autonomous weapons. Campaign to Stop Killer Robots.

⁴¹ Gunawan, Y., Aulawi, M. H., & Ramadhan, A. R. (2019). Command responsibility of autonomous weapons systems under international humanitarian law. *Jurnal Cita Hukum*, 7(3), 351–368. <https://doi.org/10.15408/jch.v7i3.11725>

The elements that play a part in establishing substantial human authority, as previously stated, can help in recognizing the necessary conditions to guarantee that systems are used only under such authority.

PROHIBITIONS

This involves restrictions on weapons systems that have inherent ethical and legal concerns due to their ability to choose and attack targets. A potential approach for the treaty could involve the inclusion of provisions that explicitly forbid the creation, manufacturing, and deployment of systems that are fundamentally deemed unacceptable. The clear and explicit prohibitions make it easier to monitor, ensure compliance, and enforce the rules. Their inflexible disposition heightens the stigma, which can in turn affect entities that are not directly involved and non-governmental entities. The suggested agreement encompasses two specific categories of prohibitions. Above all, the limitations cover systems that independently choose and attack targets without substantial human intervention⁴². These systems might operate using machine learning, which would make them too complex for people to understand and control. Additionally, it is crucial to think about the possibility of extending restrictions to cover other systems that have similar challenges. These systems depend on different kinds of data, such as weight, heat, or sound, to represent individuals, regardless of whether they are considered combatants. Causing harm or injury to individuals based on such information would devalue human life and remove the humanity from violence.

CONCLUSION

If this analysis is accurate, it suggests a reason for cautious optimism regarding artificial intelligence and the issue of responsibility gaps. It is important to note that my previous statements do not definitively dismiss the possibility of concerning responsibility gaps. Nevertheless, according to my analysis, it seems that there is minimal evidence to back the idea of problematic responsibility gaps. It's important to note, however, that this article does not conclusively settle the issue. After a thorough examination of the pessimism surrounding responsibility gaps and the identification of two main challenges, it is now the responsibility of those worried about responsibility gaps to demonstrate that these challenges can be overcome. Those who are doubtful about responsibility gaps would have to first pinpoint the specific scenarios in which these gaps arise, while also considering the plausibility of their occurrence. Furthermore, they would need to explain why these types of responsibility gaps are a cause for concern.

REFERENCES

- Akkuş, B. (2022). Autonomous weapon systems under international law. *Güvenlik Bilimleri Dergisi*, 11(2), 333–366. <https://doi.org/10.28956/gbd.1078155>
- CCW/MSP/2019/9. (n.d.). Retrieved August 6, 2024, from <https://undocs.org/Home/Mobile?FinalSymbol=CCW%2FMSP%2F2019%2F9&Language=E&DeviceType=Desktop&LangRequested=False>
- Dinstein, Y. (2018a). Autonomous weapons and international humanitarian law. In W. Heintschel Von Heinegg, R. Frau, & T. Singer (Eds.), *Dehumanization of Warfare* (pp. 15–20). Springer International Publishing. https://doi.org/10.1007/978-3-319-67266-3_2
- Dinstein, Y. (2018b). Autonomous weapons and international humanitarian law. In W. Heintschel Von Heinegg, R. Frau, & T. Singer (Eds.), *Dehumanization of Warfare* (pp. 15–20). Springer International Publishing. https://doi.org/10.1007/978-3-319-67266-3_2

⁴² Riebe, T., Schmid, S., & Reuter, C. (2020). Meaningful human control of lethal autonomous weapon systems: The ccw-debate and its implications for vsd. *IEEE Technology and Society Magazine*, 39(4), 36–51. <https://doi.org/10.1109/MTS.2020.3031846>

- Docherty, B. L. (with Human Rights Watch (Organization) & Harvard Law School). (2016). Making the case: The dangers of killer robots and the need for a preemptive ban. Human Rights Watch ; IHRC, International Human Rights Clinic.
- Ferstman, C. (2017). International organizations and the fight for accountability: The remedies and reparations gap (First edition). Oxford University Press.
- Gunawan, Y., Aulawi, M. H., & Ramadhan, A. R. (2019). Command responsibility of autonomous weapons systems under international humanitarian law. *Jurnal Cita Hukum*, 7(3), 351–368. <https://doi.org/10.15408/jch.v7i3.11725>
- Jones, J. R. W. D., & Powles, S. (2003). International criminal practice: The international criminal tribunal for the former yugoslavia, the international criminal tribunal for rwanda, the international criminal court, the special court for sierra leone, the east timor special panel for serious crimes, war crimes prosecutions in kosovo (3rd ed). Transnational ; Oxford University Press.
- Marks, P. (2011). Reaching critical point. *New Scientist*, 212(2841), 25. [https://doi.org/10.1016/S0262-4079\(11\)62960-6](https://doi.org/10.1016/S0262-4079(11)62960-6)
- News in brief. (2004). *Nature Medicine*, 10(1), 9–9. <https://doi.org/10.1038/nm0104-9>
- Ratner, S. R., Abrams, J. S., & Bischoff, J. L. (2009). Accountability for human rights atrocities in international law: Beyond the Nuremberg legacy (3rd ed). Oxford University Press.
- Riebe, T., Schmid, S., & Reuter, C. (2020). Meaningful human control of lethal autonomous weapon systems: The ccw-debate and its implications for vsd. *IEEE Technology and Society Magazine*, 39(4), 36–51. <https://doi.org/10.1109/MTS.2020.3031846>
- Sauer, F. (2020). Stepping back from the brink: Why multilateral regulation of autonomy in weapons systems is difficult, yet imperative and feasible. *International Review of the Red Cross*, 102(913), 235–259. <https://doi.org/10.1017/S1816383120000466>
- Schulzke, M. (2013). Autonomous weapons and distributed responsibility. *Philosophy & Technology*, 26(2), 203–219. <https://doi.org/10.1007/s13347-012-0089-0>
- Sparrow, R. (2007). Killer robots. *Journal of applied philosophy*, 24(1), 62-77.
- Waluyo, B. D., & Fahd, R. M. (2020). Telaah konstruktivis partisipasi cina dalam the campaign to stop killer robots. *Indonesian Journal of International Relations*, 4(2), 92–113. <https://doi.org/10.32787/ijir.v4i2.123>
- Zarychta, A. (2019). Twarzą do tego co umyka. Przypadek fotografii „ukrytych matek” w projekcie artystycznym Lindy Fregni Nagler. *Widok. Teorie i Praktyki Kultury Wizualnej*, 24. <https://doi.org/10.36854/widok/2019.24.1947>