AMOGHVARTA

ISSN : 2583-3189

Emerging Ethical and Normative Challenges in Military AI

ORIGINAL ARTICLE

Authors
Brig Eshan Dalal, SM
Research Scholar
Defence & Strategic Studies Department
Jiwaji University
Gwalior, Madhya Pradesh, INDIA

Dr. Girish Sharma
Guide / Professor Military Science
P. G. Department & Research Center
Govt. Science College
Gwalior, Madhya Pradesh, INDIA

Abstract

Advances in robotics and AI are revolutionizing military operations by enhancing speed, precision, and scalability, yet they introduce profound ethical challenges including accountability gaps for lethal autonomous weapon systems (LAWS), algorithmic biases causing civilian harm, over-reliance leading to escalation risks, dual-use dilemmas, and potential AGI threats. Drawing on real-world examples from conflicts in Libya (STM Kargu-2), Gaza (Lavender and Where's Daddy systems), Ukraine (drone swarms in Operation Spiderweb), and Nagorno-Karabakh (AI drones), the chapter highlights issues like misclassification, opacity in decision-making, and erosion of human oversight amid gaps in frameworks such as the Geneva Conventions and CCW.[library] It critiques the lack of binding international laws on AI attribution, proposes extending command responsibility doctrines, and recommends mandatory human-in-the-loop controls, transparent auditing, rigorous ethical testing, updated global governance, and civilian protections to ensure military AI aligns with humanitarian principles urging India to lead ethically under Viksit Bharat @2047.

Key Words

Emerging Ethical, Normative Challenges, Military, Artificial Intelligence.

Introduction

Advances in Robotics and Artificial Intelligence (AI) has made it possible for their integration into military systems, transforming the nature of warfare, decision making and strategic planning. While AI offers significant operational advantages in terms of speed, precision and scalability, it also brings to fore profound ethical and normative challenges. This chapter extrapolates ethical theory into real world and future military contexts, examining accountability, bias, dual use risks and governance challenges. It draws on recent conflicts in Ukraine and Gaza to illustrate these dilemmas and proposes policy recommendations.

Autonomous Accountability: Who is Liable?

The introduction of AI enabled lethal autonomous weapon systems (LAWS) which, once activated can select and engage targets without human oversight fundamentally complicates the attribution of responsibility

for actions undertaken in war. A significant concern that has been raised in academic discourse and global forums revolves around 'whom to be held accountable?' if something goes wrong. Conventional frameworks of accountability in war and conflicts are based on chains of command and the direct involvement of human agents. With increasing involvement of AI systems in threat assessment, target detection, identification and engagement decisions, attribution of human liability becomes a central ethical problem.

The 2020 Libya conflict, in which a Turkish origin loitering munition (*STM Kargu 2*) reportedly operated without direct human oversight, illustrates the ambiguity surrounding responsibility in autonomous engagements.[1] The autonomous weapon system was programmed to attack even in the absence of data connectivity between the operator and the weapon effectively a truly independent "fire, forget and find" weapon system. Sea Hunter, a trans oceanic U.S. Navy Unmanned Surface Vessel (USV) prototype, designed to hunt submarines similarly raises concerns about who would be liable if autonomous navigation leads to a strategic error[2].

The 'Human Factor' in AI

There exists no binding international legal framework delineating responsibility in AI enabled combat. The Geneva Conventions do not directly address autonomous systems and the UN's Convention on Certain Conventional Weapons (CCW) discussions remain inconclusive on the aspect of attributability[3]. While a requirement for "meaningful human control" over the targeting and engagement of all weapons has been proposed at the CCW discussions, there is a view that software and automated systems in general are coded and created by humans. Moreover, the decision to use a particular autonomous weapon and activation of its autonomous function is ultimately authorised and controlled by a human and thus could be seen as a form of human control. While this could be acceptable in non lethal situations, ethicists argue that the decision to use violent force which could result in human casualties, require human in the loop to assess the situation and evaluate the necessity to engage a weapon on a target – in essence, someone who can be apportioned moral and legal responsibility for the consequences of that decision.

In the absence of a human in the loop till the final target engagement phase, it could otherwise be argued that the accountability lies distributed between a multitude of human actors which include: the developer – who has designed and manufactured the system, the operator – who has activated the system, the commander – who has issued the orders to use the system, or the state – who has procured the system to be put in use. However, it has been counter argued that, taking into account the machine's self learning ability, humans may lack sufficient control or awareness of AI behaviour and hence cannot be held responsible for the consequences.

'Command Responsibility' in Autonomous Systems

Notwithstanding the above argument, one solution that frequently emerges to address responsibility gaps in AI weapon systems is the concept of 'command responsibility', wherein military leaders overseeing or ordering deployment of AI enabled autonomous weapons may be held responsible because they perform a supervisory role over the (subordinate) AI system. The doctrine of command responsibility is part of customary international law, as codified in Article 28 of the Rome Statute of the International Criminal Court (ICC) and Article 86–87 Additional Protocol I to the Geneva Conventions.[4] The doctrine holds military leaders criminally responsible for crimes committed by forces under their effective authority and control, constituting a form of liability for omissions related to the acts of subordinates.

Although the viability of applying the doctrine of command responsibility to scenarios involving 'autonomous and artificial subordinates' has come under lot of debate, ethically, the application may seem plausible, as commanders are generally considered responsible for adverse outcomes caused by human subordinates, despite these subordinates often having significant discretionary power. The adaptation of the concept of command responsibility, to address new challenges involving artificial entities, therefore may make sense from an ethical and legal standpoint.

BIAS, Civilian Harm and Algorithmic Misjudgement

AI models, especially those reliant on machine learning are susceptible to biases embedded in their training data and algorithmic designs. In the military context, biased or incomplete datasets and flawed algorithms can produce catastrophic humanitarian consequences. Such algorithmic biases present deeply ethical concerns: if a facial recognition system is more likely to flag certain ethnic groups as threats due to flawed data or deliberately induced biases, the resulting harm could be systematic and grievous. Algorithmic bias threatens core principles of proportionality under International Humanitarian Law, especially when opaque systems automate life or death decisions without human oversight.[5]

During the Israel Gaza conflict, AI systems were widely used by Israeli Defence Forces (IDF) to profile individuals as targets based on behavioural patterns such as frequent SIM changes, resulting in unverified strikes[6]. The Israeli Defence Forces, also reportedly used an AI Decision Support System (AI DSS) in Gaza, known as the 'Lavender', which uses Facial Recognition Techniques (FRT) to identify targets based on their facial features. Another AI system called 'Where's Daddy?' was used to track suspected militants and signalled to the Israeli military when they entered their home, marking it for bombing even when family members were present. These AI based systems are reportedly operated by the IDF's elite intelligence Unit 8200.[7] Similar reports have also emerged in Russia – Ukraine conflict, where drone based visual recognition systems misclassified civilians due to flawed imagery inputs[8].

Civilian casualties have long been a harsh reality of warfare, but AI driven operations add a layer of opacity and scale. Automated targeting systems are touted for their precision, but errors such as misclassifying civilian infrastructures such as a school or hospital as a target can have catastrophic outcomes when they occur at machine speed. Therefore, mandating meaningful human control and oversight – guided by situational awareness and ethical reasoning is essential for responsible deployment.

The complexity of modern battlefields, characterised by blurred enemy lines, extension of the battlespace into urban centres, amorphous actors and information overload, places AI systems at risk of making incorrect decisions when faced with untrained scenarios. Worse, adversaries may resort to AI manipulation through deception (a tactic known as "data poisoning"), causing it to misclassify targets with catastrophic consequences. These vulnerabilities highlight the urgent need for AI enabled military systems to undergo rigorous tests in simulated combat scenarios, include robust error reporting mechanisms and operate under persistent human oversight.

Over Reliance on AI and Decision Making Risks

AI systems can compress decision timelines and reduce the fog of war, but over reliance on these systems risks strategic misjudgements and may unintentionally escalate the conflict. In a 2024 wargame simulation model designed by Stanford University to evaluate the escalation risk of Large Language Model (LLM) based AI agents, it was found that all models show forms of escalation and difficult to predict escalation patterns that lead to greater conflict and, in some cases, the use of nuclear weapons. It also emerged that the only model that did not undergo reinforcement learning with human feedback a safety technique to align models to human instructions, presented the most escalatory and unpredictable decisions[9].

The risk of misjudgement is accentuated especially in time critical situations when algorithmic recommendations by AI DSS can lead to automation bias, where humans defer to the machine's automated recommendations while ignoring empirical data or intelligence that may suggest the opposite. In Ukraine, while ISR systems have improved battlefield awareness, there is growing concern over automation bias where operators increasingly tend to rely too heavily on algorithmic outputs[10].

Delegating critical decisions to opaque algorithms can erode the capacity for human judgment in morally challenging situations and also provide an alibi for non complicity in potentially unethical actions. As AI enabled

warfare becomes the norm, there is a risk that violence becomes normalised and morally distanced, making it easier for commanders and operators to disassociate from the moral and ethical consequences of their actions.

It is therefore, crucial not only to consider the technical robustness of AI systems but also to maintain and strengthen the ethical frameworks that govern their use. Policymakers and military leaders should proceed with utmost caution when confronted with proposals to use LLMs and LLM based agents in military decision making. Democratic nations abiding by humanitarian values and international law must lead the way to ensure that over reliance on AI does not come at the expense of rigorous oversight, meaningful accountability and the preservation of humanitarian values even if this means sacrificing some degree of efficiency.

Dual Use Dilemmas and Governance Problems

Many AI technologies, including large language models, have dual use potential. GPT based systems have demonstrated utility in cyber operations and strategic planning, raising concerns about repurposing civilian technologies for military ends and vice versa[11]. Systems relying on facial recognition, gait analysis and predictive behavioural algorithms, while developed for military or national security purposes, are increasingly finding applications in the civilian domain. The transgression of AI systems from battlefield to civilian society raises urgent concerns about dual use technologies: tools originally designed for defence or intelligence that are later adopted for peacetime law enforcement, public administration, or commercial surveillance.

No global mechanism exists to monitor dual use transitions. Current international law and arms control treaties were largely designed with traditional weapons in mind. The rapid development and adaptability of AI outpaced these frameworks, leaving significant gaps in oversight and accountability. Moreover, there is no global consensus on which AI capabilities should be restricted, how these restrictions should be enforced, or what constitutes legitimate military versus illicit use in emerging domains such as cyber and autonomous systems. In the absence of strict legal constraints, AI enabled battlefield surveillance tools could be repurposed to monitor political opponents, media personnel, or minority communities. Numerous global examples, from Ethiopia to Myanmar, illustrate how defence AI systems, can become instruments of authoritarian governance. There is also a profound psychological toll on civilians living under persistent AI powered surveillance.

The rapid commercialisation and easy access of open source AI technologies further raises the risk that advanced systems, originally developed for benign purposes could be repurposed for military use by state and non state actors alike. In some cases, state and commercial entities may unintentionally contribute to the proliferation of AI enabled military capabilities among unauthorised actors, undermining efforts to contain escalation and maintain ethical standards.

Addressing these challenges requires novel governance mechanisms tailored to the specific characteristics of AI. These include robust life cycle auditing, transparency requirements, reporting mechanisms and international agreements on the development and use of military AI. Such frameworks must be adaptive, capable of evolving as technology does and involve input from a wide range of stakeholders, including ethical watchers, developers, practitioners and civil society.

AGI Risks, Control Mechanisms and 'Kill Switch' Ethics

Current military AI systems are narrow in scope, performing designated tasks with precision. As military organizations aggressively invest in artificial intelligence, the possibility of developing Artificial General Intelligence (AGI) AI systems that rival or surpass human intelligence presents a new spectrum of ethical and security risks. Unlike current AI systems, AGI systems may exceed human reasoning and develop independent operational strategies. If such a system were deployed militarily, the potential for unintended escalation, mission drift, or catastrophic decision making would rise sharply. Should such an AGI system take actions beyond human intent, its speed, unpredictability and autonomy may preclude any efforts to regain control, posing existential risks in conflicts involving states possessing weapons of mass destruction.

The future of AGI remains uncertain, but strategic planning and international cooperation could help minimize risks and guide the development of these technologies in a safe direction. The concept of a "kill switch" a manual override to disable rogue systems becomes imperative. It will be necessary to have failsafe architectures that require multi factor, multi level human authorization for all critical operations, enmeshed with hardwired protocols enabling human operators to immediately disable or quarantine AGI systems, if risk thresholds are breached.

However, these measures are fraught with both technical and ethical complexities. Overly stringent kill switch protocols might render systems vulnerable to exploitation by adversaries. Defining the exact point at which a "kill switch" is ethically justified relies on transparent criteria, moral judgment and international consensus.[12] Ethical questions also arise around timing, legitimacy and proportionality in activating these controls.[13] Overusing such tools could diminish the strategic advantage of autonomous systems, while underusing them poses risk of unintended harm. Designing kill switches, therefore involves navigating a trade off between human dignity and operational necessity a balance that must be grounded in robust legal frameworks and ethical reasoning.

Ethical Reflections on Recent Conflict Practices

Israel Hamas: AI Enabled Targeting and Civilian Harm: In the recent Isreal Hamas conflict, AI powered systems such as "Lavender" "Gospel and "Where's Daddy" have been employed to generate target lists, automate target acquisition and threat prioritization. Using AI DSS, Israeli forces generated real time target lists and assessments, based on massive surveillance data aiming to increase operational precision.[14] While it has been contended that this enhances discrimination between combatants and civilians, independent observers and humanitarian organizations have raised concerns stating that the use of AI has intensified civilian casualties. This is plausible, as rapid, algorithm driven strikes diminish opportunities for human review and proportionality assessments, risking accidental or disproportionate harm to civilian populations.

Human rights groups have documented cases where alleged algorithmic misjudgements have led to significant civilian casualties and infrastructure loss, highlighting the ethical imperative for strict oversight of AI augmented targeting.[15] UN Secretary General António Guterres specifically condemned the delegation of life and death decisions to algorithms, highlighting the ethical and legal dangers of AI use in densely populated areas that resulted in large scale civilian casualties.[16]

The absence of transparent auditing and external review mechanisms hampers public trust and post conflict accountability. These events underscore the risk that, even where AI is justified as enhancing operational efficiency, its deployment can dramatically erode humanitarian protections in war, particularly when meaningful human oversight is sacrificed.

Russia Ukraine: Autonomous Weapons, Escalation and Oversight Gaps

The Russia Ukraine conflict has featured extensive use of AI enabled technologies, from drone swarms to automated surveillance and targeting platforms. Both sides have deployed advanced machine learning algorithms for real time battlefield intelligence and autonomous reconnaissance, which have offered apparent tactical advantages but also raised serious ethical concerns amidst reports of AI based systems that identify and prioritize targets without sufficient human vetting, sometimes resulting in strikes against misclassified or ambiguous objects.

While Ukraine publicly emphasizes "human in the loop" oversight, in practice the speed and complexity of battlefield decisions have often led to AI systems taking over targeting functions, sparking concerns about transparency and compliance with humanitarian law. Ukrainian officials and researchers admit that efficiency sometimes outweighs ethical considerations and regulatory frameworks often lag behind wartime innovation.[17]

During 'Operation Spiderweb', Ukranian drones upon losing signal, switched to performing a mission using artificial intelligence reportedly destroying or damaging 41 Russian aircraft on ground. Such capabilities

while showcasing the potential of AI for high precision, high impact operations, also highlight escalation risks, particularly if AI misjudges ambiguous targets leading to catastrophic consequences and in a worst case scenario – a nuclear response if the victim's tolerance threshold is breached.

Ukraine's experiences highlight both the tactical advantages and the urgent regulatory and moral dilemmas attending the military use of AI. International observers have called for external monitoring of such technologies and the establishment of norms to ensure that their use does not eclipse humanitarian and ethical restraints.

Armenia Azerbaijan: AI Enabled Drones and Asymmetric Warfare

The 2020 Nagorno Karabakh war between Armenia and Azerbaijan signalled the entrance of AI assisted warfare in smaller scale regional disputes. Reports indicate both sides used autonomous drones and facial recognition tools, in conjunction with cyber operations demonstrating the profound impact AI powered drones could have on regional conflicts. Azerbaijan deployed Israeli made Harop loitering munitions and Turkish Bayraktar TB2 drones, leveraging AI for target identification and precision strikes. These drones conducted autonomous attacks against high value Armenian military assets, often with little or no human intervention at the moment of engagement.[18]

The rapid deployment of AI systems without robust oversight mechanisms exemplifies how technological diffusion can outpace ethical and legal controls. The operational opacity of such systems have made it difficult to assess the true extent of civilian harm and exposes a governance gap on how emerging AI military technologies are monitored and regulated in active conflict zones.

These recent conflicts underscore the urgent need for transparent governance, robust auditing and independent oversight to ensure that advances in military AI are aligned with international humanitarian law and ethical norms. Risks of misclassification, indiscriminate targeting, rapid escalation raise serious ethical questions surrounding proportionality, target discretion and human oversight.

Conclusion and Recommendations

Recent trends in warfare demonstrate that military AI is already reshaping the ethical and strategic landscape of modern warfare, often with profound, sometimes unanticipated, humanitarian consequences. While these technologies can increase operational efficiency and reduce direct risk to military personnel, they also introduce significant new dangers: loss of meaningful human oversight, heightened risk of civilian harm, escalation of violence and persistent gaps in accountability and transparency.

The responsible integration of AI into military operations demands ongoing dialogue among technologists, ethicists, policymakers and civil society. Key recommendations that need to be addressed to harness the benefits of military AI while minimizing its risks to global security and human dignity, are summarised below:

Maintain Meaningful Human Control: AI systems should never operate without persistent, explainable human oversight, especially in decisions involving lethal force. The principle of "human in the loop" must be rigorously enforced to prevent automation bias and maintain moral authority. Mandating meaningful human control and oversight – guided by situational awareness and ethical reasoning is essential for responsible deployment.

(a) **Enhance Transparency and Accountability:** Militaries deploying AI must establish robust, auditable mechanisms for tracking algorithmic decisions and their humanitarian impacts. Independent external review and transparent reporting are essential to hold all actors developers, commanders and operators accountable for AI driven actions. The adaptation of the concept of command responsibility, to address new challenges involving artificial entities, appears logical from an ethical and legal standpoint.

(b) **Strengthen International Governance:** Global norms and legally binding agreements are urgently needed to regulate the development, export and use of military AI. Existing frameworks such as the Geneva Conventions and Arms Trade Treaty should be updated to explicitly address the unique

challenges posed by autonomous and AI enabled weapon systems. Such frameworks must be adaptive, capable of evolving as technology does and involve input from a wide range of stakeholders, including ethical watchers, developers, practitioners and civil society.

(c) **Invest in Ethical Auditing and Testing:** AI systems used in military contexts must undergo rigorous tests in simulated combat scenarios and pre deployment ethical audits to identify and mitigate biases, errors and potential for misuse. It must be ensured that advances in military AI are aligned with international humanitarian law and ethical norms.

(d) **Protect Civilian Populations:** Military planners must prioritize the protection of civilians in the design, deployment and evaluation of AI systems, ensuring that technological advances do not come at the cost of fundamental humanitarian principles. Overarching principles of humanitarian values and international law must prevail to ensure that advances in military AI are aligned with international humanitarian law and ethical norms, even if this means sacrificing some degree of efficiency.

As India progressively transforms its strategic posture under Viksit Bharat @2047, it must embed ethical principles in Defence planning. This chapter underscores that while rapid advancements in autonomous warfighting systems poses new challenges, as a responsible democracy upholding ethical and moral values, India must lead the way to ensure that military AI remains aligned with principles of justice, accountability and human dignity.

References

1. United Nations Security Council. Report of the Panel of Experts on Libya Pursuant to Resolution 1973. March 2021. https://undocs.org/S/2021/229, Accessed on 08/10/2025.
2. Myers, Joe. Does this self-driving warship take us a step closer to robotic warfare?. World Economic Forum, April 2016. https://www.weforum.org/stories/2016/04/does-this-self-driving-warship-take-us-a-step-closer-to-robotic-warfare/, Accessed on 06/10/2025.
3. Asaro, Peter. "Autonomous Weapons and the Ethics of Artificial Intelligence." International Review of the Red Cross 94, no. 886 (2019). https://peterasaro.org/writing/Asaro%20Oxford%20AI%20Ethics%20AWS.pdf https://international-review.icrc.org/articles/autonomous-weapons-and-ethics-artificial-intelligence-986, Accessed on 20/09/2025.
4. Oimann, Ann-Katrien and Salatino, Adriana (2024) Command Responsibility in Military AI Contexts: Balancing Theory and Practicality," *AI and Ethics*, 5:25, no. 2 (July 22, 2024): 1757–67, https://doi.org/10.1007/S43681-024-00512-8.
5. International Civil Aviation Organization (ICAO). Artificial Intelligence and the Future of Aviation Safety. ICAO Publications, 2022. https://www.icao.int/safety/AI-aviation-safety.pdf, Accessed on 18/09/2025.
6. Gross, Yuval. "Targeted by Algorithm: Israel's AI in Gaza." +972 Magazine, November 30, 2023. https://www.972mag.com/israel-gaza-targets-ai-algorithm-war, Accessed on 26/09/2025.
7. Andersin, Emelie (2025) The Use of the Lavender in Gaza and the Law of Targeting: Ai-Decision Support Systems and Facial Recognition Technology, *Journal of International Humanitarian Legal Studies*, 1, no. aop (May 23, 2025): 1–35, https://doi.org/10.1163/18781527-BJA10119.
8. Liam, Wilson (2023) AI Misclassification in Ukraine's Tactical Operations. *Defence Review Quarterly* 45, no. 3, 112–124. https://Defencereviewquarterly.org/ai-ukraine-misclassification, Accessed on 30/09/2025.
9. Russell, Perrault; et al. (2022) AI and Escalation in Conflict Simulations: A Stanford HAI Wargame." Stanford Human-Centreed AI Institute, July 2022. https://hai.stanford.edu/research/ai-conflict-escalation-study, Accessed on 06/09/2025.

10. International Civil Aviation Organization (ICAO). Artificial Intelligence and the Future of Aviation Safety. ICAO Publications, 2022. https://www.icao.int/safety/AI-aviation-safety.pdf, Accessed on 06/10/2025.

11. Miles, Brundage; et al. (2020) Toward Trustworthy AI Development: Mechanisms for Supporting Verifiable Claims. OpenAI Research Paper, July 2020. https://openai.com/research/toward-trustworthy-ai-development, Accessed on 02/10/2025.

12. Nick, Bostrom (2014) Superintelligence: Paths, Dangers, Strategies. Oxford University Press, Oxford.

13. Heyns, Christof (2018) Autonomous Weapons in Armed Conflict and the Right to a Dignified Life: An African Perspective, *South African Journal on Human Rights* 33, no. 1 (September 25, 2018): 46–71, https://doi.org/10.1080/02587203.2017.1303903.

14. Gaza: Israel's AI Human Laboratory – The Cairo Review of Global Affairs, https://www.thecairoreview.com/essays/gaza-israels-ai-human-laboratory/, Accessed on 22/07/2025.

15. Algorithmic Targeting: The Role of Artificial Intelligence in Israeli Strikes in Gaza and Its Ethical Implications - Groupe de Recherche et d'information Sur La Paix et La Sécurité, https://www.grip.org/algorithmic-targeting-the-role-of-artificial-intelligence-in-israeli-strikes-in-gaza-and-its-ethical-implications/, Accessed on 22/07/2025.

16. UN Chief Voices Concern over Reports of Israel Using AI to Identify Targets in Gaza, https://ddnews.gov.in/en/un-chief-voices-concern-over-reports-of-israel-using-ai-to-identify-targets-in-gaza, Accessed on 22/07/2025.

17. How AI Is Eroding the Norms of War | AI Frontiers, https://ai-frontiers.org/articles/how-ai-is-eroding-the-norms-of-war, Accessed on 22/07/2025.

18. AI: The New Paradigm of War | Annenberg, https://www.asc.upenn.edu/research/Centres/milton-wolf-seminar-media-and-diplomacy-8, Accessed on 22/07/2025.

——==00==——