

CRIMINAL ACCOUNTABILITY FOR AI-GENERATED WAR CRIMES

MICHAEL P. SCHARF* AND ISABELLE BAZZY**

ABSTRACT

The integration of artificial intelligence (AI) into military operations enhances efficiency and precision but raises significant challenges to accountability under International Humanitarian Law (IHL), as illustrated by the use of AI-driven decision support systems (AI-DSS) like Lavender and Gospel during the Gaza conflict. By using the recent Gaza conflict as a case study, this Article explains how the growing use of AI-DSS is creating a dangerous accountability chasm. While some have advocated for a strict liability approach, this Article makes the case for judicial adaptation of the principles of command responsibility and accessory liability through recognition of the concept of “digital subordinates” and “digital co-perpetrators” to close the accountability gap related to AI-generated war crimes.

I.	INTRODUCTION	56
II.	AN ILLUSION OF HUMAN CONTROL	58
	A. <i>Case Study: The Role of AI in the Gaza Conflict</i>	60
	B. <i>Overview of the Technology: Lavender and Gospel</i>	61
III.	FILLING THE GAPS IN HUMAN-CONTROLLED AI IN IHL	63
IV.	COMMAND RESPONSIBILITY AND AI	66
	A. <i>The Invisible Commanders of Algorithmic Warfare</i>	69
	B. <i>Distributed Responsibility Under IHL</i>	70
V.	CO-PERPETRATION AND AI	73
	A. <i>Why not Strict Liability?</i>	76
	B. <i>Accessory Liability for AI-DSS</i>	78
VI.	CONCLUSION	82

* Michael P. Scharf is the Joseph C. Hostetler—BakerHostetler Professor of Law at Case Western Reserve University School of Law, where he served as Dean from 2013-2024. He is also President of the American Branch of the International Law Association and Managing Director of the Public International Law & Policy Group.

** Isabelle Bazy, Case Western Reserve University School of Law JD expected in 2026, is Senior Fellow at the Frederick K. Cox International Law Center.

© 2026, Michael P. Scharf and Isabelle Bazy.

I. INTRODUCTION

Imagine an advanced artificial intelligence (AI) system deployed to analyze satellite imagery and identify military threats in an area of armed conflict. Programmed to prioritize speed and precision, the system flags a facility in a bustling neighborhood as a weapons storage site based on size, heat signatures, and activity patterns. Within minutes, the site is struck by precision missiles. Hours later, humanitarian organizations report that the target was not in fact a weapons cache but rather a shelter housing displaced families. While the thought of robots as autonomous, decision-making entities can be frightening, another insidious risk arises from the ways humans use the most basic AI systems due to the nominal inclusion of a human decisionmaker in the accountability loop. This burgeoning feature of modern warfare raises a critical question: Who is criminally responsible when AI leads humans to make a targeting mistake? Is it the programmers, procurement officers, operators, battlefield commanders, or their superiors?

Existing international humanitarian law (IHL) frameworks struggle to address this type of scenario.¹ While UN expert groups have focused narrowly on outlawing fully autonomous weapon systems, real-world deployments highlight the complexities of decision-support AI tools like the Israel Defense Forces' (IDF) *Gospel* and *Lavender* systems.² Although these tools facilitate military operations and remain under human oversight, concerns about accountability for errors persist.³ This Article examines how AI tools blur the line between human and machine decision-making, creating novel challenges for IHL accountability. While some commentators have advocated for application of strict liability as a solution,⁴ this Article explains why a better approach

1. See, e.g., Michael P. Scharf, *White Paper on the Need to Strengthen International Humanitarian Law to Address the Challenges of 21st Century Warfare*, 57 CASE W. RES. J. INT'L L. 7, 8 (2025).

2. See Tal Mimran et al., *Israel–Hamas 2024 Symposium – Beyond the Headlines: Combat Deployment of Military AI-Based Systems by the IDF*, LIEBER INST. WEST POINT (Feb. 2, 2024), <https://lieber.westpoint.edu/beyond-headlines-combat-deployment-military-ai-based-systems-idf>; Michael N. Schmitt, *Israel–Hamas 2024 Symposium – The Gospel, Lavender, and the Law of Armed Conflict*, LIEBER INST. WEST POINT (June 28, 2024), <https://lieber.westpoint.edu/gospel-lavender-law-armed-conflict>; United Nations Office for Disarmament Affairs, *Lethal Autonomous Weapon Systems*, UNITED NATIONS, <https://disarmament.unoda.org/en/our-work/emerging-challenges/lethal-autonomous-weapon-systems> (last visited Jan. 22, 2026).

3. Mimran et al., *supra* note 2.

4. See Mia Swart, *Constructing “Electronic Liability” for International Crimes: Transcending the Individual in International Criminal Law*, 24 GER. L.J. 589, 597 (2023); Paul R. Williams & Ryan J. Westlake, *A Taste of Armageddon: Legal Considerations for Lethal Autonomous Weapons Systems*, 57 CASE W. RES. J. INT'L L. 187, 187 (2025).

would be to adapt principles of command responsibility and accessory liability to the unique context of AI-generated war crimes.

Parts I and II provide an overview of the operational framework of human-controlled AI systems and explain how these systems' ability to learn and adapt from experience leads to unpredictable outcomes, complicating humans' ability to maintain meaningful control.

Part III analyzes specific AI systems deployed in the recent Gaza conflict to illustrate gaps in the existing IHL framework. It then focuses on how these gaps impact foundational principles like proportionality and distinction, as well as the precautionary obligation to take constant care in military operations.

Part IV advocates for applying the legal doctrine of command responsibility to the use of AI for target selection to ensure compliance with IHL and address criminal liability.⁵ This Part argues commanders can and should be held accountable for AI systems under their control as "digital subordinates," even when these systems exhibit a high degree of autonomy.⁶ Under this approach, not only military commanders, but also AI developers, procurement officers, operators, and an extended chain of decision-makers can be considered commanders and held criminally liable for AI-generated war crimes.

Part V introduces an additional approach to ensuring accountability for AI-generated war crimes. It examines difficulties of establishing *mens rea* in AI-assisted operations and proposes addressing these gaps by adapting the principle of accessory liability for acts of "digital co-perpetrators" as a potential framework, where responsibility would attach to individuals who knowingly rely on AI systems in ways that make unlawful outcomes more likely.

This Article concludes by addressing the accountability challenges of human-controlled AI, arguing that it does not necessitate entirely new rules under IHL or international criminal law. Rather, it provides a

5. See Clive Baldwin, *How Does International Humanitarian Law Apply to Israel and Gaza?*, HUM. RTS. WATCH (Oct. 27, 2023, at 1:45 ET), <https://www.hrw.org/news/2023/10/27/how-does-international-humanitarian-law-apply-israel-and-gaza> (noting that deliberate killings and indiscriminate bombings and airstrikes in Israel and Palestine constitute war crimes under IHL and trigger requirements of criminal prosecution and accountability).

6. *Id.* (introducing the principle that commanders may incur criminal liability for ordering, assisting in, or knowingly failing to prevent or punish war crimes committed by their subordinates); see generally Guido Acquaviva, *Crimes without Humanity? Artificial Intelligence, Meaningful Human Control, and International Criminal Law*, 21 J. INT'L CRIM. JUST. 981 (2023) (noting that the human-centered nature of superior responsibility under Article 28 of the Rome Statute complicates attributing criminal liability for conduct carried out by autonomous weapon systems, potentially leaving a responsibility gap).

roadmap for judicial adaptation of command responsibility and accessory liability approaches to fill existing gaps.

II. AN ILLUSION OF HUMAN CONTROL

AI is increasingly employed in military applications to enhance effectiveness and efficiency in both operational decision-making and battle-field execution.⁷ One prominent tool, AI Decision Support Systems (AI-DSS), utilizes advanced algorithms to display, synthesize, and analyze data.⁸ Unlike fully autonomous weapons, AI-DSS are considered “human-centered” or “human-in-the-loop” systems, wherein humans retain ultimate authority over decisions.⁹ However, in practice, human operators and AI-DSS interact in a complex web of shared agency and distributed responsibility.¹⁰

Technological control is essential in the realm of warfare.¹¹ From fail-safes to defensive systems, control mechanisms ensure mission integrity and mitigate operational risks.¹² Yet, the sufficiency of AI control mechanisms remains uncertain.¹³ Errors often remain invisible until they culminate in collapse, which can have catastrophic consequences.¹⁴

7. See Ruben Stewart & Georgia Hinds, *Algorithms of War: The Use of Artificial Intelligence in Decision-Making in Armed Conflict*, HUMANITARIAN L. & POL’Y BLOG (Oct. 24, 2023), <https://blogs.icrc.org/law-and-policy/2023/10/24/algorithms-of-war-use-of-artificial-intelligence-decision-making-armed-conflict>.

8. AI-DSS make recommendations or predictions to aid human decision-making in warfare. Anna Nadibaidze, *Do AI Decision-Support Systems ‘Support’ Humans in Military Decision-Making on the Use of Force?*, OPINIO JURIS (Nov. 29, 2024), <https://opiniojuris.org/2024/11/29/do-ai-decision-support-systems-support-humans-in-military-decision-making-on-the-use-of-force>.

9. Stewart & Hinds, *supra* note 7; Jovana Davidovic, *What’s Wrong with Wanting a “Human in the Loop”?*, WAR ON THE ROCKS (June 23, 2022), <https://warontherocks.com/2022/06/whats-wrong-with-wanting-a-human-in-the-loop>.

10. Stewart & Hinds, *supra* note 7.

11. See Helga Nowotny, *The Illusion of Control: Living with Digital Others*, 5 GLOB. PERSPS. 2024 1, 2; Cătălin Balamuş et al., *The Impact of Technology on Warfare*, 18 INT’L SCI. CONF. STRATEGIES XXI 171-177 (Dec. 12, 2022).

12. Nowotny, *supra* note 11, at 1.

13. *Id.* at 2.

14. Reports of Turkey’s Kargu-2 drone operating autonomously in Libya in 2020; Russian-made Lancet-3 drones prematurely deploying and disabling yet still engaging targets when communication was lost; a mistaken identification, caused by data overreliance, that resulted in the deaths of ten civilians in Kabul. In each case, the potential for error remained opaque until after harm had occurred. See, e.g., U.N. PANEL OF EXPERTS ON LIBYA, FINAL REPORT, ¶¶ 63-67, U.N. Doc. S/2021/229 (Mar. 8, 2021); Kateryna Stepanenko, *The Battlefield AI Revolution Is Not Here Yet: The Status of Current Russian and Ukrainian AI Drone Efforts*, INST. FOR THE STUDY OF WAR (June 2, 2025), <https://understandingwar.org/backgrounder/battlefield-ai-revolution-not-here-yet-status-current-russian-and-ukrainian-ai-drone>; Eric Schmitt & Helene Cooper, *Pentagon Acknowledges Aug. 29 Drone Strike in Afghanistan Was a Tragic Mistake*,

Integration of human-controlled AI into military strategy introduces additional complexities.¹⁵ AI's ability to process vast amounts of data, recognize complex patterns, and act at superhuman speeds is one of its greatest advantages.¹⁶ However, it also raises fundamental questions concerning whether meaningful human oversight is achievable.¹⁷ Humans may not directly intervene when AI executes an initiative; instead, they may only exert direct control over AI at earlier stages, such as during design or testing.¹⁸ Moreover, as AI systems become increasingly sophisticated, human oversight simultaneously becomes increasingly superficial.¹⁹ For instance, in systems relying on human operators to confirm targets, AI design might lead humans to accept recommendations uncritically.²⁰ This phenomenon, known as "automation bias," erodes capacity for independent judgment.²¹ This illusion of control is further entrenched by a human tendency to anthropomorphize AI.²² AI-DSS, for example, evoke psychological responses such as trust and perceived reliability.²³ This anthropomorphism exacerbates vulnerabilities to manipulation and error.²⁴

N.Y. TIMES (Sep. 17, 2021), <https://www.nytimes.com/2021/09/17/us/politics/pentagon-drone-strike-afghanistan.html>.

15. See Michael Klare & Xiaodon Liang, *Beyond a Human "In the Loop": Strategic Stability and Artificial Intelligence*, ARMS CONTROL ASS'N (Nov. 12, 2024), <https://www.armscontrol.org/issue-briefs/2024-011/beyond-the-loop>.

16. Mimran et al., *supra* note 2.

17. The speed and agility of AI systems often outpace human judgment, making it difficult to ensure safety, accountability, and ethical decision-making. See Jovana Davidovic, *What's Wrong with Wanting a "Human in the Loop"?*, WAR ON THE ROCKS (June 23, 2022), <https://warontherocks.com/2022/06/whats-wrong-with-wanting-a-human-in-the-loop>.

18. *Id.*

19. Developers frequently modify user interfaces and data presentation to align operator behavior with the AI system's desired outcomes, raising concerns about whether the human role in decision-making is meaningful or merely a formality designed to validate AI's decisions. See Cathy Mulligan, *Automated Warfare and the Geneva Convention*, NETZPOLITIK.ORG (Apr. 17, 2024, at 17:09 CT), <https://netzpolitik.org/2024/artificial-intelligence-automated-warfare-and-the-geneva-convention>; Stewart & Hinds, *supra* note 7; Davidovic, *supra* note 17.

20. Nicolas Spatola, *The Efficiency-Accountability Tradeoff in AI Integration: Effects on Human Performance and Over-Reliance*, 2 COMPUTS. IN HUM. BEHAV.: ARTIFICIAL HUMS., Aug.-Dec. 2024, at 1, 1.

21. Automation bias is a cognitive bias that occurs when individuals overly rely on automated systems or inappropriately assign blame to them. See, e.g., Noah Sylvia, *Israel Defense Forces' Use of AI in Gaza: A Case of Misplaced Purpose*, ROYAL UNITED SERVS. INST. (July 4, 2024), <https://www.rusi.org/explore-our-research/publications/commentary/israel-defense-forces-use-ai-gaza-case-misplaced-purpose>.

22. Generative AI systems replicate human language and creativity so seamlessly that they obscure the distinction between technological creations and human-like beings. See James Johnson, *Is There a Human in the Machine? AI and Future Warfare*, WAR ON THE ROCKS (May 30, 2024), <https://warontherocks.com/2024/05/is-there-a-human-in-the-machine-ai-and-future-warfare>.

23. *Id.*

24. See *id.*

Accordingly, approaches to accountability that emphasize humans' capacity to exert control are often counterproductive.²⁵ Insisting on human oversight where it is ineffective may divert attention from exploring alternative solutions, such as robust ethical frameworks, technical safeguards, or even bans on certain types of AI-enabled weaponry.²⁶

A. *Case Study: The Role of AI in the Gaza Conflict*

As one journalist reports, “much of the destruction and associated civilian casualties in Gaza have been linked to the AI platforms.”²⁷ Systems like *Lavender* and *Gospel*, deployed by the IDF, exemplify the evolving relationship between human oversight and machine autonomy.²⁸ While designed with mechanisms for human intervention, *Lavender* and *Gospel* challenge traditional notions of command responsibility.²⁹

The IDF utilizes AI technologies to enhance operational efficiency and strike precision, with divisions like Unit 8200 spearheading AI development in collaboration with military, academic, and private sectors.³⁰ Technology is not neutral; it embodies the goals and interests of its creators.³¹ In high-stakes scenarios like those in Gaza, these technologies are intended to

25. See Spatola, *supra* note 20, at 2.

26. See Matthias Klaus, *Transcending Weapon Systems: The Ethical Challenges of AI in Military Decision-Support Systems*, HUMANITARIAN L. & POL'Y BLOG (Sep. 24, 2024), <https://blogs.icrc.org/law-and-policy/2024/09/24/transcending-weapon-systems-the-ethical-challenges-of-ai-in-military-decision-support-systems>.

27. Davit Khachatryan, *Algorithms of War: AI Ethics Amid Genocide Allegations in Gaza*, EVN REPORT (Aug. 13, 2024), <https://evnreport.com/opinion/algorithms-of-war-ai-ethics-amid-genocide-allegations-in-gaza>.

28. The IDF rank fifth on the 2023 Global AI Index and have integrated AI into military operations. Inbar Dolinko & Liran Antebi, *Embracing the Organized Mess: Defense AI in Israel*, in THE VERY LONG GAME 397, 397-98 (Heiko Borchert et al. eds., 2024); Schmitt, *supra* note 2.

29. See Mimran et al., *supra* note 2; Christopher Elliott, *Expedient or Reckless? Reconciling Opposing Accounts of the IDF's Use of AI in Gaza*, OPINIO JURIS (Apr. 26, 2024), <https://opiniojuris.org/2024/04/26/expedient-or-reckless-reconciling-opposing-accounts-of-the-idfs-use-of-ai-in-gaza>.

30. While it lacks a centralized AI governance structure, Israel's collaboration between government, security services, and the private sector drives military advancements. The State relies on initiatives from entities like the Israel Ministry of Defense and the Israel Innovation Authority. This decentralized model, termed an “organized mess,” has proven advantageous as it enables adaptability to rapid advancements without bureaucratic restrictions. See Dolinko & Antebi, *supra* note 28, at 397-98; see also Mimran et al., *supra* note 2 (detailing the IDF's use of AI systems to identify and recommend targets, including its prior deployment of the “Gospel” during Operation Guardian of the Walls to locate Hamas missile unit leaders and its current use in Operation Iron Swords to generate hundreds of target options within seconds).

31. See *Questions and Answers: Israeli Military's Use of Digital Tools in Gaza*, HUM. RTS. WATCH (Sep. 10, 2024, at 12:01 ET), <https://www.hrw.org/news/2024/09/10/questions-and-answers-israeli-militarys-use-digital-tools-gaza>.

minimize the duration of military operations, resources expended, and lives lost.³² However, the deployment of AI in conflict zones introduces significant risks and ethical controversies.³³ Growing dependence on AI reduces human oversight of critical decisions and shifts life-or-death choices to automated systems lacking judgment and empathy.³⁴

While the structure of AI technologies indicates humans retain oversight at certain critical points, the complexity of these systems often obscures the actual degree of human control.³⁵ The interaction of various AI components can create opaque decision-making processes, where the AI's rationale behind specific recommendations may be difficult for operators to discern, even with specialized training.³⁶ Furthermore, time pressure in conflict scenarios frequently limits operators' ability to thoroughly review or challenge AI-generated recommendations.³⁷

These patterns challenge traditional notions of accomplice liability and command responsibility, as human control mechanisms appear more procedural than substantive.³⁸ Understanding how command responsibility and accountability apply to AI in warfare requires an analysis of the specific AI systems deployed and their distinct functionalities.

B. Overview of the Technology: Lavender and Gospel

The IDF increasingly uses AI-driven systems to bomb “power targets” or political leaders in Gaza.³⁹ In some cases, these efforts to target

32. AI-DSS can process vast amounts of intelligence data rapidly, enabling the identification and targeting of military assets with reduced risk to military personnel and, theoretically, non-combatants. See Stewart & Hinds, *supra* note 7; Khachatrian, *supra* note 27.

33. AI systems place a significant risk to civilian lives as they remain vulnerable to errors in target identification due to biases in data inputs or limitations in contextual understanding. See Sylvia, *supra* note 21.

34. See Mulligan, *supra* note 19.

35. See Alan L. Schuller, *At the Crossroads of Control: The Intersection of Artificial Intelligence in Autonomous Weapon Systems with International Humanitarian Law*, 8 HARV. NAT'L SEC. J. 379, 416 (2017).

36. See Emelie Andersin, *The Use of the 'Lavender' in Gaza and the Law of Targeting: AI-Decision Support Systems and Facial Recognition Technology*, 16 J. INT'L HUMANITARIAN LEGAL STUD. 336, 351-52 (2025).

37. Documentation of deployment practices has revealed instances where system-generated analyses were accepted with minimal scrutiny, either due to the operators' reliance on the perceived infallibility of AI or the sheer complexity of the data presented. See *Questions and Answers: Israeli Military's Use of Digital Tools in Gaza*, *supra* note 31; Elliott, *supra* note 29.

38. See Matthew T. Miller, *Command Responsibility: A Model for Defining Meaningful Human Control*, 11 J. NAT'L SEC. L. & POL'Y 533, 541 (2020); Gabriel Hallevy, *The Criminal Liability of Artificial Intelligence Entities—From Science Fiction to Legal Social Control*, 29 AKRON INTELL. PROP. J. 171, 181-83 (2015).

39. “Power targets” are a set of civilian infrastructure that exerts “civil pressure.” See Yuval Abraham, ‘A Mass Assassination Factory’: Inside Israel's Calculated Bombing of Gaza, +972 MAGAZINE

specific Hamas leaders have resulted in hundreds of civilian deaths.⁴⁰ For example, *Gospel*, an AI system also known as “Habsora,” is routinely utilized by Unit 8200, Israel’s signals intelligence division,⁴¹ to enable large-scale strikes on residences.⁴² The tool leverages Palestinians’ personal data to predict threats and identify targets.⁴³ *Gospel* employs machine learning algorithms capable of analyzing data, recognizing patterns, and making inferences without explicit programming.⁴⁴ Despite its increasing role in IDF military operations, the system lacks transparency and falls outside the purview of international agreements.⁴⁵

This is particularly concerning given that reports indicate *Gospel* often prioritizes damage over accuracy.⁴⁶ Since the October 7, 2023 attacks, over 60,000 Palestinians, many of them civilians, have been killed, reflecting a shift toward speed and operational efficiency over precision.⁴⁷ Roughly 25,000 of those deaths resulted from strikes on targets identified by the IDF’s AI systems.⁴⁸

(Nov. 30, 2023), <https://www.972mag.com/mass-assassination-factory-israel-calculated-bombing-gaza>.

40. *Id.*

41. Joseph Gedeon & Maggie Miller, *Israel Under Pressure to Justify Its Use of AI in Gaza*, POLITICO (Mar. 3, 2024), <https://www.politico.com/news/2024/03/03/israel-ai-warfare-gaza-00144491>.

42. *Gospel*’s strikes are guided by a collateral damage scale that pre-approves civilian casualty estimates before strikes. Abraham, *supra* note 39. While it has drastically increased target generation, from tens per year to hundreds daily, it has also contributed to a significant rise in civilian casualties, including strikes on residences with no known Hamas presence. *Id.*

43. *Questions and Answers: Israeli Military’s Use of Digital Tools in Gaza*, *supra* note 31.

44. While specifics remain classified, *Gospel* processes various aggregated surveillance sources. These include imagery (satellite/drone), communications, behavioral patterns, and infrastructure use. The data is anonymized in reporting but clearly involves identifying structures linked to militant activity. Geoff Brumfiel, *Israel Is Using an AI System to Find Targets in Gaza. Experts Say It’s Just the Start*, WISC. PUB. RADIO (Dec. 14, 2023), <https://www.wpr.org/news/israel-using-ai-system-find-targets-gaza-experts-say-its-just-start>.

45. For example, Israel has not signed a United States-backed treaty on responsible AI use. Gedeon & Miller, *supra* note 41; see *Use of automated targeting system in Gaza*, STOP KILLER ROBOTS (Dec. 1, 2023), <https://www.stopkillerrobots.org/news/use-of-automated-targeting-system-in-gaza>.

46. Abraham, *supra* note 39.

47. See, e.g., Emma Farge & Nidal Al-Mughrabi, *Gaza Death Toll: How Many Palestinians Has Israel’s Offensive Killed?*, REUTERS (July 29, 2025), <https://www.reuters.com/world/middle-east/how-many-palestinians-has-israels-gaza-offensive-killed-2025-01-15>.

48. In the first 27 days, more than 12,000 targets were completed. As a force multiplier, AI eliminates resource constraints, enabling the targeting of “junior militants” who would ordinarily be spared, given that their deaths would have little impact on military objectives. See Brianna Rosen, *Unhuman Killings: AI and Civilian Harm in Gaza*, JUST SEC. (Dec. 15, 2023), https://www.justsecurity.org/90676/unhuman-killings-ai-and-civilian-harm-in-gaza/?utm_source=chatgpt.com; Nidal Al-Mughrabi & Emma Farge, *How many Palestinians has Israel’s Gaza offensive killed?*, REUTERS (July 29, 2025, at 1:00 PM

Lavender, another AI system developed by Unit 8200, pinpoints individuals linked to Hamas and Palestinian Islamic Jihad, two militant groups with the stated objective of seeking to establish an independent Palestinian state through armed resistance against Israel.⁴⁹ Reports from the Israeli Palestinian publication *+972 Magazine* and the Hebrew-language outlet *Local Call* indicate *Lavender* initially flagged as many as 37,000 Palestinian men as potential militants.⁵⁰ Due to AI's inability to incorporate complex contextual cues into its decision making,⁵¹ *Lavender* has an error margin of up to ten percent, leading to the targeting of thousands of individuals with minimal or no Hamas affiliation.⁵² Enhanced criminal liability is necessary to catalyze positive change.

III. FILLING THE GAPS IN HUMAN-CONTROLLED AI IN IHL

Although digital tools used in military operations are not classified as weapons under the Geneva Conventions, they are still subject to IHL constraints.⁵³ As highlighted by the International Committee of the Red Cross, “any new technology of warfare must be used, and must be capable of being used, in compliance with existing rules of international humanitarian law.”⁵⁴ The Geneva Conventions and their Additional Protocols form the foundation of IHL, governing the conduct of armed conflict.⁵⁵ IHL aims to balance the necessity of

ET), <https://www.reuters.com/world/middle-east/how-many-palestinians-has-israels-gaza-offensive-killed-2025-01-15>.

49. Lavender analyzes personal and behavioral metadata like WhatsApp group ties to militants, frequent phone or address changes, device history, social links, names in militant files, and broader surveillance patterns. The system then scores Palestinians on a 1–100 scale for similarity to Hamas operatives, designating those above a shifting threshold as targets. Gaby Del Valle, *Report: Israel Used AI to Identify Bombing Targets in Gaza*, VERGE (Apr. 4, 2024, at 3:38 PM ET), <https://www.theverge.com/2024/4/4/24120352/israel-lavender-artificial-intelligence-gaza-ai>; Rania Aby Shamala, *Israeli army is using artificial intelligence to generate kill lists in Gaza: Report, Anadolu Agency* (Apr. 25, 2024), <https://www.aa.com.tr/en/artificial-intelligence/israeli-army-is-using-artificial-intelligence-to-generate-kill-lists-in-gaza-report/3183446>.

50. Del Valle, *supra* note 49.

51. Sylvia, *supra* note 21; Peter J. Denning & John Arquilla, *The Context Problem in Artificial Intelligence*, 65 COMM'N OF THE ACM 18, 19 (2022).

52. See, e.g., Sylvia, *supra* note 21; Abraham, *supra* note 39.

53. Laurent Gisel et al., *Twenty Years On: International Humanitarian Law and the Protection of Civilians Against the Effects of Cyber Operations During Armed Conflicts*, 102 INT'L REV. RED CROSS 287, 303 (2020).

54. *Artificial Intelligence and Machine Learning in Armed Conflict: A Human-Centered Approach*, 102 INT'L REV. RED CROSS 463, 464 (2020).

55. See, e.g., Geneva Convention for the Amelioration of the Condition of the Wounded and Sick in Armed Forces in the Field art. 1, Aug. 12, 1949, 75 U.N.T.S. 31; Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International

weakening adversaries with the imperative to reduce unnecessary suffering.⁵⁶

IHL imposes four key obligations on military operations: military necessity, distinction, proportionality, and the duty to take precautions, each of which raises specific concerns in the context of AI-driven warfare.⁵⁷ This Article solely addresses distinction and proportionality, as these obligations raise particular concerns for the military use of AI.

The distinction principle mandates that attacks target only military objectives and attempt to avoid harm to civilians; accordingly, it prohibits the use of inherently indiscriminate weapons.⁵⁸ AI has the potential to align with the principle of distinction under IHL, as its ability to deliver highly targeted strikes could reduce collateral damage.⁵⁹ Traditionally, humans have made these nuanced judgments on the battlefield, considering factors like threatening actions, surrender, or incapacitation.⁶⁰ However, AI systems often rely on data and algorithms, which may not capture battleground realities, increasing wrongful targeting risk.⁶¹ For example, *Gospel*, once described as a “factory” for producing lethal targets, has caused catastrophic damage, raising concerns that it functions indiscriminately rather than enhancing precision.⁶² While machines excel in observation and recognition, such as mapping and weapons detection, their ability to make context-sensitive judgments remains deficient.⁶³ Experts predict AI may become autonomous enough to achieve this level of sophistication to make judgments by 2040.⁶⁴

The proportionality principle ensures collateral damage is not excessive relative to the anticipated military advantage.⁶⁵ AI systems can

Armed Conflicts (Protocol I) pmbL., June 8, 1977, 1125 U.N.T.S. 3 [hereinafter Additional Protocol I].

56. See *Fundamental Principles of IHL*, INT’L COMM. OF THE RED CROSS 60 https://casebook.icrc.org/a_to_z/glossary/fundamental-principles-ihl (last visited Jan. 13, 2025).

57. *Id.*; see U.S. DEP’T OF DEF., DEPARTMENT OF DEFENSE LAW OF WAR MANUAL, 50-66, 188-97 (2023); Sylvia, *supra* note 21.

58. Additional Protocol I, *supra* note 55, art. 57; *Rule 7. The Principle of Distinctions between Civilian Objectives and Military Objectives*, INT’L COMM. RED CROSS, <https://ihl-databases.icrc.org/en/customary-ihl/v1/rule7> (last visited July 9, 2025).

59. See Elliot Winter, *The Compatibility of Autonomous Weapons with the Principles of International Humanitarian Law*, 27 J. CONFLICT & SEC. L. 1, 13-14 (2022).

60. *Id.*

61. See Klaus, *supra* note 26.

62. Harry Davies et al., *The Gospel: how Israel uses AI to select bombing targets in Gaza*, GUARDIAN (Dec. 1, 2023, at 5:03 AM ET), <https://www.theguardian.com/world/2023/dec/01/the-gospel-how-israel-uses-ai-to-select-bombing-targets>.

63. Winter, *supra* note 59, at 13-14.

64. *Id.* at 15.

65. Additional Protocol I, *supra* note 55, arts. 51(5)(b), 57(2)(a)(iii).

calculate risks and advantages using data-driven models, but an AI system might suggest executing an attack where collateral damage outweighs military gain because it cannot fully comprehend qualitative factors to the same extent as a human commander.⁶⁶ AI lacks the necessary reference points to interpret patterns using cultural context and judgment.⁶⁷ However, just because commanders can understand proportionality does not mean they will always act accordingly. In Gaza, some on-duty unit commanders were described as “more trigger happy than others.”⁶⁸ One former IDF officer, who worked on strike planning until 2021, recalled instances where there was “doubt about a target,” yet “we killed what I thought was a disproportionate amount of civilians.”⁶⁹ Such circumstances suggest that growing deference to AI-DSS may erode commanders’ independent proportionality judgments.

For example, in late 2023, after months of being unable to locate Ibrahim Biari, a senior Hamas commander believed to be hiding within Gaza’s tunnel network, Israel deployed an experimental AI-powered audio analysis tool to locate him.⁷⁰ By intercepting Biari’s calls, the AI system generated an estimated location.⁷¹ “A strike on the area was greenlit despite warnings within the military that several apartment buildings would need to be destroyed to ensure Biari was killed.”⁷² While the strike successfully killed Biari, it also resulted in the deaths of more than 125 civilians.⁷³ The commanders’ reliance on the AI’s output in authorizing a strike with foreseeably disproportionate civilian costs illustrates how AI complicates accountability.⁷⁴ Furthermore, such disregard for anticipated civilian harm underscores the need for clearer criminal accountability.

66. See Klaus, *supra* note 26.

67. Davies et al., *supra* note 62.

68. *Id.*; Abraham, *supra* note 39.

69. Davies et al., *supra* note 62.

70. Ibrahim Biari was a high listed senior Hamas leader who Israeli forces believed to have played a central role in planning the October 7th attacks. Sheera Frenkel & Natan Odenheimer, *Israel’s A.I. Experiments in Gaza War Raise Ethical Concerns*, N.Y. TIMES (Apr. 25, 2025), <https://www.nytimes.com/2025/04/25/technology/israel-gaza-ai.html>.

71. Avi Hasson, CEO of Startup Nation Central, stated that Israeli military AI innovation has been significantly advanced by reservists from tech companies like Meta, Google, and Microsoft. According to Hasson, these individuals provided “know-how and access to key technologies” previously unavailable to the military. *Id.*

72. TOI Staff, *Israel using AI to pinpoint Hamas leaders, find hostages in Gaza tunnels — report*, TIMES OF ISRAEL (Apr. 26, 2025, at 10:57 AM ET), <https://www.timesofisrael.com/israel-using-ai-to-pinpoint-hamas-leaders-find-hostages-in-gaza-tunnels-report>.

73. Frenkel & Odenheimer, *supra* note 70.

74. *See id.*

Incidents like the Biari strike illustrate the high stakes of deploying AI-driven systems without adequate legal safeguards or accountability mechanisms. Under IHL, tools like *Gospel* and *Lavender* require legality reviews conducted by legal advisors to ensure they comply with distinction and proportionality principles.⁷⁵ A legality review assesses new weapons, means, or methods of warfare to ensure they comply with international law before deployment.⁷⁶ These formal assessments establish a documented process that helps reduce the risk of unlawful harm and enhances accountability for military decision-makers.⁷⁷ While Israel is a party to the original Geneva Conventions of 1949, it has not ratified the First Additional Protocol of 1977.⁷⁸ Consequently, Israel is not legally bound by Article 36.⁷⁹

In the absence of a customary international law obligation requiring Israel to conduct legal reviews of new wartime technologies, there are limited effective constraints on the development and use of such systems. At the same time, broader principles of international law and international human rights law support the need for robust assessments of emerging wartime technologies.⁸⁰ Moreover, broader principles of international law and human rights also call for robust assessments of emerging technologies used in warfare.⁸¹

IV. COMMAND RESPONSIBILITY AND AI

Commanders hold a unique position within the IHL framework, with legal obligations extending beyond duties of ordinary combatants. Their responsibilities are not defined by direct control over every individual action, such as firing a weapon, but rather by their authority over the broader military operation or attack.⁸² Therefore, the concept of “meaningful human control” should require that commanders

75. Legality reviews are prescribed under Article 36 of the First Additional Protocol to the Geneva Conventions. Additional Protocol I, *supra* note 55, art. 36.

76. See Mimran et al., *supra* note 2.

77. See *id.*

78. See Khachatryan, *supra* note 27.

79. *Id.*

80. U.N. GAOR, Sixth Committee, 75th Sess., 12th plen. mtg. at 45, U.N. Doc. A/C.6/75/SR.12 (Nov. 4, 2020); Khachatryan, *supra* note 27.

81. See Yahli Shereshevsky, *International humanitarian law-making and new military technologies*, 104 INT’L REV. RED CROSS 2132, 2139-41 (Nov. 2022), <https://international-review.icrc.org/articles/international-humanitarian-law-making-and-new-military-technologies-920>.

82. See United States, *Implementing International Humanitarian Law in the Use of Autonomy in Weapon Systems* ¶ 4 (Grp. of Governmental Experts Working Paper, CCW/GGE.1/2019/WP.5 2019), [https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-_Group_of_Governmental_Experts_\(2019\)/CCW_GGE.1_2019_WP.5.pdf](https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-_Group_of_Governmental_Experts_(2019)/CCW_GGE.1_2019_WP.5.pdf).

maintain enhanced oversight over AI-DSS, ensuring they exert necessary control to prevent IHL violations.⁸³ This ensures commanders are accountable for actions within their scope of responsibility, even when using AI systems.

To fully understand how command responsibility applies to AI-driven warfare, three primary challenges must be addressed: AI unpredictability and opacity, the necessity of shared responsibility across multiple actors, and the need for proactive measures to preemptively prevent violations.⁸⁴

The command responsibility doctrine traditionally encompasses three primary duties: (1) proper training of subordinates, (2) intervening to prevent violations, (3) and disciplining subordinates after violations occur.⁸⁵ With respect to AI, ensuring proper “training” requires designing, developing, and deploying these systems in compliance with IHL.⁸⁶ This includes programming AI systems to follow distinction and proportionality principles, rigorously testing such systems in controlled environments, and updating AI algorithms to address errors or biases.⁸⁷

While IHL typically requires a purposeful or knowing standard of intent (*mens rea*) for the conviction of war crimes,⁸⁸ it imposes criminal liability on commanders under a negligence standard—namely, when commanders “should have known”—for unlawful acts of subordinates.⁸⁹ Applying this standard to commanders overseeing the use of AI systems would incentivize more responsible deployment and establish a basis for criminal liability when problems arise.

Unlike human subordinates, whose decision-making can be scrutinized and guided through training and accountability measures, AI systems may learn and adapt in ways not fully transparent, raising

83. See U.N. Secretary-General, *Lethal Autonomous Weapons Systems: Report of the Secretary-General*, U.N. Doc. A/79/88 (July 1, 2024), <https://docs.un.org/en/A/79/88>.

84. See generally *id.* (examining the implications of lethal autonomous weapons systems for international humanitarian law, including responsibility, accountability, and compliance challenges).

85. Jamie Allan Williamson, *Some Considerations on Command Responsibility and Criminal Liability*, 90 INT’L REV. OF THE RED CROSS 303, 303 (2008).

86. Clémence Duranleau-Hendrickx, *Artificial Intelligence Use for Military Purposes*, in *HOW DOES LAW PROTECT IN WAR?* (3rd ed. 2014), <https://casebook.icrc.org/case-study/artificial-intelligence-use-for-military-purposes> (last visited Mar. 25, 2026).

87. *Id.*

88. INT’L CRIM. CT., *ELEMENTS OF CRIMES 1* (2011), <https://www.icc-cpi.int/sites/default/files/Publications/Elements-of-Crimes.pdf>; Rome Statute of the International Criminal Court, art. 30, July 17, 1998, 2187 U.N.T.S. 90 [hereinafter *Rome Statute*].

89. *Rule 153. Command Responsibility for Failure to Prevent, Repress or Report War Crimes*, INT’L COMM. RED CROSS, <https://ihl-databases.icrc.org/en/customary-ihl/v1/rule153> (last visited July 9, 2025).

concerns about maintaining compliance over time.⁹⁰ Furthermore obscurity of AI decision-making processes, often referred to as the “black box” problem, complicates the ability of commanders to anticipate or address potential violations.⁹¹ Commanders must implement mechanisms to monitor AI operations in real time and establish safeguards to override or halt these operations when necessary.⁹² However, the speed and unpredictability of AI actions can make timely intervention difficult.⁹³ The “commanders should have known about violations” standard is particularly challenging to meet when AI systems operate in opaque ways.⁹⁴ A commander’s ability to intervene effectively hinges on the explainability of AI systems, which remains a significant technological and legal hurdle.⁹⁵

The duty to discipline subordinates after violations occur requires a reframing when the subordinate in question is an AI system.⁹⁶ Investigations would need to rely on analyzing “black box” data to trace the AI’s decision-making process, as these records provide an objective account of the AI’s internal operations.⁹⁷

In this context, it is crucial to conceptualize AI and AI-DSS as analogous to subordinates under the command responsibility doctrine. By framing AI systems as “digital subordinates,” commanders must approach their obligations to these systems similarly to their responsibilities toward human personnel.⁹⁸ Importantly, AI-DSS should be viewed as extensions of the commander’s authority, with their design and functionality reflecting intent and oversight of human leadership. This means while AI should be viewed as a subordinate in terms of training and monitoring, the discipline prong of commander responsibility should not be placed

90. See *Questions and Answers: Israeli Military’s Use of Digital Tools in Gaza*, *supra* note 31.

91. The black-box problem in AI refers to the opacity of AI systems, where their inputs and decision-making processes remain unclear to users. See Scott Sullivan, *Targeting in the Black Box: The Need to Reprioritize AI Explainability*, LIEBER INST. W. POINT (Aug. 28, 2024), <https://lieber.westpoint.edu/targeting-black-box-need-reprioritize-ai-explainability/>. This issue arises due to intentional secrecy, technological illiteracy, and a fundamental cognitive mismatch between AI operations and human understanding. *Id.*

92. See Miller, *supra* note 38, at 536.

93. See *Questions and Answers: Israeli Military’s Use of Digital Tools in Gaza*, *supra* note 31.

94. See Rome Statute, *supra* note 88, art. 28.

95. See Davidovic, *supra* note 17.

96. See Alex Sarch & Ryan Abbott, *Punishing Artificial Intelligence: Legal Fiction or Science Fiction*, 53 UNIV. CAL. DAVIS L. REV. 323, 345 (Nov. 2019).

97. See *Questions and Answers: Israeli Military’s Use of Digital Tools in Gaza*, *supra* note 31.

98. This approach is appropriate because it recognizes that AI does not merely function as a weapon, but as a decision-shaping actor within the chain of command. See INT’L COMM. RED CROSS, *supra* note 89.

onto AI, but rather its human counterpart. This is because AI-DSS are not sentient and are not making decisions autonomously. Therefore, prosecutors can and should identify accountable or blameworthy human beings.

Finally, commanders not only have a duty to act once they discover a potential problem, they also have a duty to seek out information reasonably available to them.⁹⁹ This duty prevents commanders from unreasonably relying on assurances from their subordinates.¹⁰⁰ In the context of AI-DSS, this duty underscores the need for commanders to critically evaluate systems acting as subordinates and not blindly accept their recommendations.

Since AI lacks the capacity for agency or intent, accountability requires identifying and addressing human contributors to the AI system's unlawful actions—such as commanders, operators, procurement officers, and developers responsible for the AI system's design, deployment, and oversight.¹⁰¹

A. *The Invisible Commanders of Algorithmic Warfare*

The traditional view of command responsibility focuses on military superiors who issue direct orders in the field.¹⁰² However, in the age of machine-influenced targeting, particularly with AI-DSS like *Lavender* and *Gospel*, this model is increasingly outdated.¹⁰³ The chain of command no longer runs solely through physical ranks but is distributed across a diffuse network of actors who influence how these systems operate.¹⁰⁴ In this context, IHL must reconsider commanders, operators, and even developers as actors who each exercise a form of authority over digital subordinates.

Some of these individuals may not be present on the battlefield but still shape outcomes as directly as any field officer, often with greater influence over target selection, interpretation, and validation, or

99. Rome Statute, *supra* note 88, art. 28; Médecins Sans Frontières, *The Practical Guide to Humanitarian Law: Duty of Commanders*, <https://guide-humanitarian-law.org/content/article/3/duty-of-commanders/> (last visited Jan. 11, 2025).

100. *See* Médecins Sans Frontières, *supra* note 99.

101. Thomas C. King et al., *Artificial Intelligence Crime: An Interdisciplinary Analysis of Foreseeable Threats and Solutions*, 26 SCI. & ENG'G ETHICS 89, 95 (2020); Brumfiel, *supra* note 44.

102. Command responsibility attaches to military superiors who exercise “effective control” over subordinates with the material ability to prevent or punish the commission of international crimes. *See* Rome Statute, *supra* note 88, art. 28; Evan Wallach & Maxine Marcus, *Command Responsibility*, in INTERNATIONAL CRIMINAL LAW 1, 1 (3rd ed., 2008).

103. *See* ANNA NADIBAIDZE ET AL., AI IN MILITARY DECISION SUPPORT SYSTEMS 6-7 (2024).

104. *Id.*

elimination of AI recommendation.¹⁰⁵ Accordingly, each actor functions as de facto commander, exercising “command” over different layers of AI system deployment.

For example, if programmers fail to ensure AI systems comply with distinction and proportionality principles, the resulting errors may violate international law.¹⁰⁶ Biases embedded in training datasets, particularly where civilian and combatant identifiers are culturally or contextually nuanced, can result in disproportionate and indiscriminate targeting.¹⁰⁷ Such systemic design flaws are not neutral mistakes as they reflect human judgment, prioritization, and sometimes recklessness.¹⁰⁸ Therefore, responsibility should trace back not only to those who launch strikes but to those whose coding decisions made that strike possible.¹⁰⁹ The dispersed nature of AI decision-making thus does not shield individual actors; instead, it reinforces the importance of extending responsibility across the chain of command, including to those previously seen as merely technical.¹¹⁰

B. *Distributed Responsibility Under IHL*

While distributing responsibility seems copacetic in theory, additional layers of complexity arise for commanders in AI-assisted warfare.¹¹¹ This diffusion of control raises urgent questions regarding legal responsibility when civilian harm results.

The principle of proportionality becomes especially difficult to assess where decision-making is no longer the sole purview of a single battlefield commander. Traditionally, commanders must weigh the anticipated

105. These individuals can range from developers to training and evaluation officers who design, implement, and test the technical aspects of these AI-DSS. *See id.* at 20.

106. *See* Additional Protocol I, *supra* note 55, art. 48.

107. *See* Andersin, *supra* note 36, at 349; Ingvild Bode, *Falling Under the Radar: The Problem of Algorithmic Bias and Military Applications of AI*, HUMANITARIAN L. & POL’Y BLOG (Mar. 14, 2024), <https://blogs.icrc.org/law-and-policy/2024/03/14/falling-under-the-radar-the-problem-of-algorithmic-bias-and-military-applications-of-ai>.

108. *See id.* at 344.

109. For example, a developer who programs a machine learning model using biased training data, such as videos and images predominantly of people of color based on the prejudiced belief that they are more likely to be terrorists, can directly influence the system’s targeting behavior and, ultimately, the legality and ethics of its outcomes. King et al., *supra* note 101, at 108-10.

110. *See* Isaac Taylor, *Collective Responsibility and Artificial Intelligence*, 37 PHILOS. & TECH. 27, 6 (2024).

111. *See generally* Ann-Katrien Oimann & Adriana Salatino, *Command Responsibility in Military AI Contexts: Balancing Theory and Practicality*, 5 AI AND ETHICS 1757 (2024) (examining the theoretical foundations of command responsibility and the potential risks and advantages of extending the doctrine to military AI systems).

military advantage of an action against the expected civilian harm and take precautions to minimize collateral damage.¹¹² A commander exercises significant discretion over the use of force when planning an attack by implementing precautions to mitigate risks and ensure compliance with proportionality standards.¹¹³ While AI systems can assist in calculating risks, human commanders remain essential for making context-sensitive judgments.¹¹⁴ This limitation directly critiques operational reliance on AI systems like *Lavender*, which utilize predefined parameters to recommend targets.¹¹⁵ By pre-filtering targets, such systems shift meaningful discretion away from battlefield commanders and upstream such decisions to designers and programmers who set the algorithm's logic, thresholds, and training data.¹¹⁶ As a result, those individuals effectively participate in the targeting process, even if they are geographically or temporally removed from combat.¹¹⁷ Automating proportionality assessments without human oversight heightens the risk of unlawful outcomes, as algorithmic logic cannot fully capture the nuances of each strike.¹¹⁸

With respect to these nuances, Sarah Harrison, a former U.S. Department of Defense lawyer and current analyst at Crisis Group, highlighted inherent challenges in determining proportionality in warfare.¹¹⁹ She emphasized that proportionality cannot be reduced to a fixed number of “acceptable” collateral deaths for specific categories of targets.¹²⁰ “While there may be certain occasions where 15 collateral civilian deaths could be proportionate, there are other times where it

112. See Miller, *supra* note 38, at 537.

113. See United States, *supra* note 82, at 4-5.

114. See *International Humanitarian Law and the Challenges of Contemporary Armed Conflicts*, 106 INT'L REV. RED CROSS 1357, 1422-25 (2024).

115. Both *Lavender* and *Gospel* do not generate collateral damaged estimates. Integrating these estimates into AIDSS would likely shift the point of potential failure upstream to those who designed or validated the AI's collateral damage estimate algorithms. See Schmitt, *supra* note 2; Bethan McKernan & Harry Davies, ‘The machine did it coldly’: Israel used AI to identify 37,000 Hamas targets, *GUARDIAN* (Apr. 3, 2024, at 9:53 ET), <https://www.theguardian.com/world/2024/apr/03/israel-gaza-ai-database-hamas-airstrikes>; Abraham, *supra* note 30.

116. See Andersin, *supra* note 36, at 349-50.

117. MARTA BO ET AL., RETAINING HUMAN RESPONSIBILITY IN THE DEVELOPMENT AND USE OF AUTONOMOUS WEAPON SYSTEMS: ON ACCOUNTABILITY FOR VIOLATIONS OF INTERNATIONAL HUMANITARIAN LAW INVOLVING AWS 34, 38 (2022); Julia Gawlas, *Use of AI DSS in Military Operations: An Assessment under International Humanitarian Law*, ASA INT'L L. (May 25, 2025), <https://www.asainternationallaw.nl/post/use-of-ai-dss-in-military-operations-an-assessment-under-internationalhumanitarian-law>.

118. See *Questions and Answers: Israeli Military's Use of Digital Tools in Gaza*, *supra* note 31.

119. See McKernan & Davies, *supra* note 115.

120. *Id.*

definitely wouldn't be," Harrison noted.¹²¹ "You can't just set a tolerable number for a category of targets and say that it'll be lawfully proportionate in each case."¹²²

Automation bias further complicates the application of command responsibility to AI-driven warfare.¹²³ Commanders may over-rely on AI systems while underestimating limitations, undermining meaningful human oversight and eliminating ethical judgment and contextual awareness from decision-making.¹²⁴ This automation bias is often embedded by system designers, demonstrating that automation bias is not always a cognitive failure by the end-user but a systemic condition shaped by those who design AI systems in the first place.¹²⁵

In Gaza, IDF commanders, at least superficially, retain the ultimate decision-making authority over offensive operations.¹²⁶ However, in the aftermath of the October 7 attacks, commanders demanded speedier target identification, prioritizing operational outcomes over meticulous oversight.¹²⁷ In furtherance of these aims, AI-driven tools like *Lavender* became indispensable in sustaining a continuous pipeline of potential targets.¹²⁸ However, the algorithm's data inputs, processing methods, and decision-making criteria remain opaque, even to those involved.¹²⁹ When *Lavender's* accuracy in correctly identifying targets reached ninety percent, the tool became widely used as a target recommendation system, despite the concerning ten percent error rate that remained.¹³⁰

The ongoing conflict in Gaza demonstrates that relying on human intervention to overcome machine decision-making is misaligned with

121. An international law expert at the U.S. State Department stated they had "never remotely heard of a one to 15 ratio being deemed acceptable, especially for lower-level combatants." *Id.*

122. *Id.*

123. See Mimran et al., *supra* note 2.

124. See Andersin, *supra* note 36, at 348, 351-52, 361.

125. See *id.* at 350 (demonstrating that algorithmic bias is often pre-embedded in AI systems at the training phase, where skewed or non-representative data reflecting human assumptions and prejudices leads to systematically higher error rates for marginalized groups and increases the risk of misidentification).

126. Khachatryan, *supra* note 27.

127. Regardless of the legal or moral justifications offered for Israel's bombing strategy, some of its own intelligence officers have reportedly begun to question the approach set by their commanders. One officer described the post-October 7 atmosphere within the IDF as "painful and vindictive." See McKernan & Davies, *supra* note 115.

128. *Id.*

129. *Id.*; Andersin, *supra* note 36, at 351.

130. "[T]he numbers changed all the time, because it depends on where you set the bar of what a Hamas operative is." McKernan & Davies, *supra* note 115.

the realities of warfare.¹³¹ Commanders may officially retain ultimate decision-making authority but are heavily dependent on AI systems to deliver actionable intelligence at exceptional speed.¹³² One intelligence officer described being pressured to “[b]ring us more targets” and told to “fuck up Hamas, no matter what the cost,” adding “[w]hatever you can, you bomb.”¹³³ This kind of directive pressure undermines meaningful human control, which is the cornerstone of lawful AI use in warfare.¹³⁴

Ultimately, applying command responsibility to AI subordinates highlights significant challenges, including the unpredictability and opacity of AI systems, the need for shared responsibility across diverse actors, and the necessity of proactive measures to prevent violations.

V. CO-PERPETRATION AND AI

IHL is enforced via criminal prosecution at national and international levels.¹³⁵ Under the doctrine of command responsibility, military leaders can be criminally liable for failing to prevent or address IHL violations if they knew, or should have known, about such crimes.¹³⁶ Importantly, criminal liability is not confined to direct perpetrators and their commanders. Other individuals who contribute to, or enable, violations may also bear responsibility under international law.¹³⁷ The prosecution of potential violations in AI-assisted military operations faces unprecedented evidentiary and procedural challenges.¹³⁸ Liability under international criminal law is based on individual criminal responsibility, which is attributed only to humans; thus, assigning fault when AI-DSS cause harm is particularly

131. See, e.g., Mimran et al., *supra* note 2.

132. See McKernan & Davies, *supra* note 115.

133. See *id.*

134. *Id.*

135. Prosecutions before international tribunals or foreign courts using universal jurisdiction are the common methods of enforcing IHL. However, international prosecutions are typically reserved for the “most serious crimes of concern to the international community.” See Rome Statute, *supra* note 88, pmb., arts. 17, 25(2).

136. See *id.* art. 28.

137. See *Individual Criminal Responsibility*, INT’L COMM. OF THE RED CROSS, [hereinafter *Individual Criminal Responsibility*], https://casebook.icrc.org/a_to_z/glossary/individual-criminal-responsibility (last visited Jan. 13, 2025). As the late United Nations Special Rapporteur on extrajudicial killings, Christof Heyns, noted, “without the promise of accountability, deterrence and prevention are reduced, resulting in lower protection of civilians and potential victims of war crimes.” Christof Heyns (Special Rapporteur on extrajudicial, summary or arbitrary executions), *Rep. of the Special Rapporteur on Extrajudicial, Summary or Arbitrary Executions*, Christof Heyns, ¶ 75, U.N. Doc. A/HRC/23/47 (Apr. 9, 2013).

138. See, e.g., Andersin, *supra* note 36, at 370.

difficult.¹³⁹ This section examines criminal prosecution challenges in the application of *mens rea* (guilty mind)¹⁴⁰ in cases involving AI-DSS, causation and its evidentiary difficulties, limitations of strict liability, and accessory liability as a potential alternative framework.

Traditional criminal liability requires both a physical action (*actus reus*) and a *mens rea*, with the latter presenting difficulties when applied to AI-DSS.¹⁴¹ One of the central features of *mens rea* is the requirement for intent or knowledge.¹⁴² However, when military personnel rely on AI-DSS, they may not directly intend the specific outcomes of the system's decisions.¹⁴³ Yet, under criminal law, liability often attaches if the operator intended the underlying action regardless of whether they foresaw the precise resultant harm.¹⁴⁴ For example, if an AI system used in military targeting misidentifies civilians as combatants, the operator still intended to strike the designated target but accidentally relied on faulty AI-generated intelligence. This suggests that the issue is not one of intent to cause harm but rather a question of whether operator reliance on faulty AI-generated intelligence constitutes a reasonable mistake of fact.

Additionally, the level of interaction between human operators and AI-DSS can vary significantly depending on system autonomy.¹⁴⁵ In many cases the human operator may provide oversight or make final decisions, but the system is performing most decision-making autonomously or with minimal human input.¹⁴⁶ In other cases, automation bias may come into effect causing human operators to over-rely on AI-

139. *Individual Criminal Responsibility*, *supra* note 137.

140. *See* Rome Statute, *supra* note 88, art. 30 (explaining that (1) to establish a crime under the Rome Statute, the requisite *mens rea* must be proven; and (2) the general standard of *mens rea*—apart from intent—is *knowledge*, defined as “awareness that a circumstance exists or that a consequence will occur in the ordinary course of events.”).

141. *See id.* arts. 35, 30.

142. The requirement of intent and knowledge for components of a crime are more demanding to prove than lesser mental states. Consequently, lower thresholds like recklessness or negligence are typically considered insufficient to establish liability for many international crimes. *See* Christine Carpenter, *Whose [Crime] is it Anyway? Adapting the Crime of Aggression to Grapple with AI and the Future of International Crimes*, 23.1 J. INT'L CRIM. JUST. 69, 73-74 (2025); Rome Statute, *supra* note 88, art. 30.

143. When an AI-DSS make a recommendation, decision, or takes action, it often does so based on its programming, data inputs, and algorithms. *See* Johnson, *supra* note 22; *Questions and Answers: Israeli Military's Use of Digital Tools in Gaza*, *supra* note 31.

144. *See* King et al., *supra* note 101, at 95.

145. *See* *Questions and Answers: Israeli Military's Use of Digital Tools in Gaza*, *supra* note 31; DR. MAXIMILIAN PORETSCHKIN ET AL., GUIDELINE FOR DESIGNING TRUSTWORTHY ARTIFICIAL INTELLIGENCE 52-52 (2021).

146. *See* McKernan & Davies, *supra* note 115.

DSS due to belief that the system is more reliable or objective than human judgment.¹⁴⁷ Moreover, the nature of machine learning-based systems complicates accountability, as complex algorithms can obscure the reasoning behind specific outputs, making it even harder to pinpoint human responsibility.¹⁴⁸

Even if an operator does not have intent or knowledge regarding a harmful act by the AI, the concept of recklessness or negligence might still apply.¹⁴⁹ However, recklessness and negligence require proving that the operator acted in a manner demonstrating a disregard for potential consequences.¹⁵⁰ This can be difficult to prove in the context of AI-DSS.¹⁵¹ AI-DSS are often used because they are designed to optimize decisions and reduce human error.¹⁵² If the operator followed the system's recommendation with reasonable reliance on its accuracy, liability may depend on whether such reliance was justified given the known limitations of the AI.¹⁵³

Given the opacity of machine decision-making, establishing causation in AI-assisted warfare presents unique hurdles.¹⁵⁴ Prosecutors typically rely on direct links between an accused's actions and the resulting harm.¹⁵⁵ However, AI-DSS complicate this analysis, as harm often results from an interplay of multiple AI systems and human operators.¹⁵⁶ Traditional legal standards for evidence collection and preservation may also prove inadequate, as its decision-making processes are often not fully explainable, even to their designers.¹⁵⁷ If a system has evolved since the time of an alleged violation, forensic analysis may

147. E.g., *Questions and Answers: Israeli Military's Use of Digital Tools in Gaza*, *supra* note 31.

148. See Swart, *supra* note 4, at 600.

149. *AI and Criminal Law: Who Is Responsible When a Machine Commits a Crime?*, L. & MORE ATTY'S (Oct. 14, 2025), <https://lawandmore.eu/blog/ai-and-criminal-law-who-is-responsible-when-a-machine-commits-a-crime>.

150. See Model Penal Code § 2.02 (A.L.I., Proposed Official Draft 1962).

151. See *AI and Criminal Law: Who Is Responsible When a Machine Commits a Crime?*, *supra* note 149.

152. Andersin, *supra* note 36, at 348; see Wen Zhou & Anna Rosalie Greipl, *Artificial intelligence in military decision-making: supporting humans, not replacing them*, HUMANITARIAN L. & POL'Y BLOG (Aug. 29, 2024), <https://blogs.icrc.org/law-and-policy/2024/08/29/artificial-intelligence-in-military-decision-making-supporting-humans-not-replacing-them>.

153. Gawlas, *supra* note 117; Andersin, *supra* note 36, at 363.

154. Andersin, *supra* note 36, at 355.

155. See A.B.A. CRIM. JUST. SECTION, CRIM. JUST. STANDARDS FOR THE PROSECUTION FUNCTION § 3-4.4 (4th ed. 2017), https://www.americanbar.org/groups/criminal_justice/resources/standards/prosecution-function.

156. See Chantal Grut, *The Challenge of Autonomous Lethal Robotics to International Humanitarian Law*, 18 J. CONFLICT & SEC. L. 5, 19-20 (2013).

157. See Zhou & Greipl, *supra* note 152.

struggle to reconstruct its precise functioning at the critical moment.¹⁵⁸ This lack of a fixed decision-making framework creates additional barriers to effective legal scrutiny, as prosecutors must rely on technical expertise to interpret the AI's role in the chain of causation.¹⁵⁹

A. *Why not Strict Liability?*

In light of difficulties in establishing *mens rea* in cases involving AI-DSS, strict liability has been proposed as a potential solution.¹⁶⁰ Strict liability encompasses crimes that do not require proof of intent, knowledge, recklessness, or negligence concerning one or more elements of the *actus reus*.¹⁶¹ However, this approach faces significant resistance in international criminal law, particularly for serious crimes where the moral element has traditionally been considered essential.¹⁶²

Some scholars have attempted to circumvent this resistance by reinterpreting existing doctrines to support stricter forms of liability.¹⁶³ For example, Paul Williams and Ryan Westlake rely on precedent from the International Criminal Tribunal for the Former Yugoslavia (ICTY) to advocate for the imposition of strict liability for contributions to international crimes.¹⁶⁴ Their recent article, *A Taste of Armageddon*, invokes the concept of joint criminal enterprise III (JCE III) liability to argue that the mere participation in a collective system, such as an AI-assisted military infrastructure, should be sufficient to establish criminal responsibility regardless of individual intent.¹⁶⁵

158. Jason Brooks, *Deploy Now, Explain Never? Why AI Needs Forensic Parity*, MEDIUM (June 24, 2025), <https://medium.com/%40muckypaws/deploy-now-explain-never-why-ai-needs-forensic-parity-294676beb8f8>.

159. Individual criminal responsibility and legal personality are not obstacles to prosecuting human operators of AI systems, as existing frameworks already recognize their liability. See *Individual Criminal Responsibility*, *supra* note 137; Andersin, *supra* note 36, at 355.

160. See Swart, *supra* note 4, at 597; Williams & Westlake, *supra* note 4, at 187.

161. *Strict Liability*, LEGAL INFO. INST., https://www.law.cornell.edu/wex/strict_liability (last visited Jan. 14, 2026).

162. For example, the Rome Statute rejects strict liability because it requires a specific state of mind or degree of fault for criminal responsibility. See Swart, *supra* note 4, at 599; Rome Statute, *supra* note 88, art. 30.

163. See, e.g., Swart, *supra* note 4, at 597.

164. Williams & Westlake, *supra* note 4, at 215-16.

165. JCE is a mode of liability in international criminal law that holds individuals criminally responsible for crimes committed by a group when they intentionally participate in a common plan or purpose that results in the commission of crimes. While JCE I and II require that the accused intend the crimes committed, JCE III is the most expansive form, attaching liability for crimes that were merely the natural and foreseeable consequences of the joint enterprise. See *Prosecutor v. Tadić*, IT-94-1-A, Appeals Chamber Judgment, ¶¶ 185–229 (Int'l Crim. Trib. for the

There are two problems with that approach. First unlike strict liability, JCE III does not do away with *mens rea* altogether. Rather, as conceived by the Appeals Chamber of the ICTY in the 1999 *Tadić* decision,¹⁶⁶ the “extended form of JCE” requires that the crimes committed by the co-perpetrator be “reasonably foreseeable”—in other words, a negligence standard.¹⁶⁷

Second, although JCE III has been applied by the ICTY, the International Criminal Tribunal for Rwanda,¹⁶⁸ and the Special Court for Sierra Leone,¹⁶⁹ commentators have been highly critical of the expansive mode of liability (calling it “guilt by association”) and its weak foundations prior to the *Tadić* decision.¹⁷⁰ The Rome Statute of the International Criminal Court (ICC) does not include a reference to the extended form of JCE, and the ICC has not applied JCE III to any of its cases.¹⁷¹

It is noteworthy that the Extraordinary Chambers in the Courts of Cambodia (ECCC) denied the prosecution’s argument for application of JCE III in *Prosecutor v. Nuon Chea et al.*, holding that JCE III was not recognized under customary international law or Cambodian law as of 1975.¹⁷² The Pre-Trial, Trial, and Supreme Court Chambers of the ECCC consistently held that while JCE I and II were accepted forms of liability at the relevant time, JCE III was not, emphasizing that

Former Yugoslavia July 15, 1999); Williams & Westlake, *supra* note 4, at 215-16; Kai Ambos, *Joint Criminal Enterprise and Command Responsibility*, 5 J. OF INT’L CRIM. JUST. 159, 160 (2007).

166. *Tadić*, IT-94-1-A at ¶ 220.

167. *Joint criminal enterprise*, UNITED NATIONS INTERNATIONAL RESIDUAL MECHANISM FOR CRIMINAL TRIBUNALS, <https://cld.irmct.org/notions/show/488/joint-criminal-enterprise#> (last visited Jan. 14, 2026); *Tadić*, IT-94-1-A at ¶ 220 (reflecting a standard of *dolus eventualis*, also called “advertent recklessness” in some legal systems).

168. *Prosecutor v. Ntakirutimana*, ICTR-96-10-A & ICTR-96-17-A, Judgment, ¶ 468 (Int’l Crim. Trib. for Rwanda Dec. 13, 2004); Lachezar Yanev, *Joint Criminal Enterprise*, in *MODES OF LIABILITY IN INTERNATIONAL CRIMINAL LAW* 135, 162 (Jérôme de Hemptinne, Robert Roth, & Elies van Sliedregt 2019).

169. Michael P. Scharf, *The Once and Future Doctrine of Joint Criminal Enterprise*, in *THE LEGACY OF AD HOC TRIBUNALS IN INTERNATIONAL CRIMINAL LAW: ASSESSING THE ICTY’S AND THE ICTR’S MOST SIGNIFICANT LEGAL ACCOMPLISHMENTS* 161, 162 (Milena Sterio and Michael Scharf eds., 2019).

170. Jens David Ohlin, *Three Conceptual Problems with the Doctrine of Joint Criminal Enterprise*, 5 J. OF INT’L CRIM. JUST. 69, 70, 74-76 (2007); Miren Odriozola-Gurrutxaga, *The Doctrine of Joint Criminal Enterprise at the Ad Hoc Tribunals and Its Applicability in the Rome Statute of the ICC*, *ELEC. REV. IAPL* (2013); Susie Hulley & Tara Young, *Joint Enterprise in England and Wales: Why Problems Persist Despite Legal Change*, 37 *CURRENT ISSUES CRIM. JUST.* 134, 134 (2025).

171. See Rome Statute, *supra* note 88, art. 25(3) (a); Univ. of Exeter, *International Criminal Court should adopt ‘joint criminal enterprise’ criminal liability mode, team argues*, *PHYS. ORG* (Aug. 22, 2024), <https://phys.org/news/2024-08-international-criminal-court-joint-enterprise.html>.

172. *Prosecutor v. Nuon Chea et al.*, 002/19-09-2007/ECCC/TC, Decision on Appeal Against Closing Orders, ¶ 3707 (Extraordinary Chambers in the Courts of Cambodia Nov. 16, 2018).

international criminal responsibility requires a more individualized and culpability-based assessment.¹⁷³ This decision reflects growing judicial resistance to doctrines that overly dilute *mens rea* requirements and impose liability as a principal without sufficient evidentiary linkage between the accused's mental state and the criminal outcome.

B. *Accessory Liability for AIDSS*

At the same time, the ICC's 2014 *Katanga* case illustrates a more balanced approach—one that preserves the importance of individual culpability while allowing for flexibility in applying *mens rea* through knowledge-based and contribution-focused theories of responsibility.¹⁷⁴ Initially charged as a principal perpetrator for his role in an attack on the village of Bogoro, Congolese rebel Germain Katanga was ultimately convicted as an accessory, which carries a lower penalty.¹⁷⁵ The majority reasoned that, although Katanga did not directly carry out the crimes, he knowingly made significant contributions to a group committing those crimes and did so with awareness of the likelihood such crimes would occur.¹⁷⁶ This shift from principal to accessory liability demonstrates the ICC was willing to impose liability based not on direct intent to commit the underlying crimes, but on conscious participation in a system or structure of violence with foreseeable outcomes.¹⁷⁷

Under Article 25(3)(d) of the Rome Statute, individuals may be held criminally responsible as accessories if they contribute to the commission of crimes by a group acting with a common purpose, even without direct participation in the underlying act.¹⁷⁸ This framework lowers the evidentiary threshold by focusing on an individual's contribution and awareness that their actions would assist the group's criminal conduct.¹⁷⁹

173. Case of NUON Chea et al., 002/19-09-2007-ECCC/SC, Case 003 Defense Submission in Intervention or Amicus Curiae Brief on JCE III Applicability at 1-2, 4, 15, 20, 24 (2015).

174. Prosecutor v. Katanga, ICC-01/04-01/07, Judgment Pursuant to Article 74 of the Statute, ¶¶ 1670, 1682 (Mar. 7, 2014).

175. *Id.* § XII Disposition; *Case Information Sheet: Situation in the Democratic Republic of Congo, The Prosecutor v. Germain Katanga*, INT'L CRIM. COURT (July 2021), <https://www.icc-cpi.int/sites/default/files/CaseInformationSheets/KatangaEng.pdf>; *Accessory*, CRIM. L. (2012), <https://courses.lumenlearning.com/suny-criminallaw/chapter/7-3-accessory>.

176. *Katanga*, ICC-01/04-01/07 at ¶¶ 1467-73.

177. *See id.* ¶¶ 1478-79.

178. *See* Rome Statute, *supra* note 88, art. 25(3)(d).

179. *See id.*; KAI AMBOS, TREATIES ON INTERNATIONAL CRIMINAL LAW: VOLUME I: FOUNDATION AND GENERAL PART 285-90 (2d ed. 2021); Yanev, *supra* note 168, at 133.

While international criminal law traditionally emphasizes *dolus directus*, AI-DSS complicate that model.¹⁸⁰ Since commanders or developers contribute to systems they know carry a substantial risk of unlawful harm, and they proceed anyway, their conduct aligns more closely with *dolus eventualis*.¹⁸¹ Applied to AI-assisted warfare, this standard allows for accountability across a wide range of actors, including programmers, procurement officers, operators, and commanders at all levels, by examining their functional roles in enabling or facilitating harm, rather than requiring direct intent or control over each outcome.

A commander or operator who knowingly deploys an AI system with known or reasonably foreseeable risks could be liable as an accessory to war crimes resulting from the actions of what this Article terms a “digital co-perpetrator.”¹⁸² They may not intend the specific outcome, but their knowing contribution to high-risk operations could satisfy the *mens rea* threshold under more flexible interpretations akin to *Katanga*.¹⁸³

Additionally, an accessory liability approach eliminates incentives for those responsible for AI-DSS to remain ignorant regarding the functional details of the systems.¹⁸⁴ These systems, often powered by algorithms comprising millions of lines of code, are typically developed by

180. *Dolus directus* refers to the form of intent where the accused acts with the clear objective or desire to bring about a specific criminal consequence. See Jindan-Karena Mann, *Autonomous Weapons Systems and the Liability Gap, Part One: Introduction to Autonomous Weapons Systems and International Criminal Liability*, RETHINKING SLIC* (July 15, 2019), <https://www.rethinkingslic.org/blog/criminal-law/51-autonomous-weapons-systems-and-the-liability-gap-part-one-introduction-to-autonomous-weapon-systems-and-international-criminal-liability>; AL-HAQ, AL-HAQ LEGAL BRIEF I: SPECIAL INTENT (DOLUS SPECIALIS)_REQUIRED TO CLASSIFY ACTS AS GENOCIDE I; Prosecutor v. Bemba et al., ICC-01/05-01/13-1989-Red, Trial Chamber VII, (Oct. 19, 2016).

181. *Dolus eventualis* is recognized as an independent basis for criminal responsibility and, as such, supports the inclusion of lower forms of intent within the legal standard of Article 30 of the Rome Statute, demonstrating that international law already tolerates reduced *mens rea* thresholds in certain doctrines of accessory liability. Mohamed Elewa Badar, *Dolus Eventualis and the Rome Statute Without It?*, 12 NEW CRIM. L. REV. 433, 435, 452-55 (2009); see Mann, *supra* note 180.

182. An operator or programmer of AI-DSS may bear responsibility as accessories if their contributions facilitate the commission of war crimes. See Marco Sassóli, *Autonomous Weapons and International Humanitarian Law: Advantages, Open Technical Questions and Legal Issues to Be Clarified*, 90 INT'L L. STUD. 308, 325 (2014).

183. Prosecutor v. Katanga, ICC-01/04-01/07, Judgment Pursuant to Article 74 of the Statute, ¶¶ 1670, 1682 (Mar. 7, 2014).

184. Traditional fault-based liability would likely encourage developers or operators to avoid delving into the nuances of their systems to shield themselves from liability. Accessory liability compels countries to proactively invest in safety measures, transparent design practices, and thorough post-deployment monitoring to minimize risks effectively. See King et al., *supra* note 101, at 108-10 (providing background on fault-based liability to demonstrate its relevance in supporting accessory liability as a viable alternative framework for AI-DSS).

teams with only partial knowledge of the overall design.¹⁸⁵ This makes it nearly impossible to predict every possible outcome, as interactions between code components can lead to unforeseen effects.¹⁸⁶ Accessory liability addresses this challenge by prioritizing knowing contributions to the AI-DSS over the intricate and often inscrutable pathways through which the AI-DSS ultimately produces a targeting decision.¹⁸⁷

Katanga's conviction thus illustrates not only a flexible application of *mens rea*, but also a broad and functional understanding of causation within international criminal law.¹⁸⁸ This causal logic has compelling parallels in the context of AI-DSS used in military operations.¹⁸⁹ In many ways, Katanga's role as a facilitator of violence mirrors the role AI-DSS might play in enabling, directing, or materially contributing to the commission of international crimes.¹⁹⁰ Katanga did not pull the trigger; but his provision of weapons and support was indispensable to the group's capacity to carry out atrocities.¹⁹¹ Likewise, AI-DSS do not kill or destroy, but may critically enable targeting, classification, or prioritization decisions resulting in unlawful attacks.¹⁹²

Ultimately, the *Katanga* judgment offers a principled starting point for evolving legal responsibility in the face of technological transformation.¹⁹³ By emphasizing contribution and causation over mental state, *Katanga* serves as a bridge toward holding human actors liable as accessories for international crimes facilitated by AI-DSS.¹⁹⁴

Furthermore, while fault-based liability aligns most clearly with the retributive aims of international criminal law,¹⁹⁵ the evolving application of accessory

185. Andersin, *supra* note 36, at 343-44.

186. *Id.* at 333.

187. See Swart, *supra* note 4, at 600 (demonstrating that strict liability's focus on causation risks unjust outcomes in the AI context).

188. See generally *Katanga*, ICC-01/04-01/07 (demonstrating that international criminal responsibility may rest on indirect contribution coupled with sufficient awareness of the criminal outcome).

189. See Ernest J Weinrib, *Causation and Wrongdoing*, 63 CHICAGO-KENT L. REV. 407, 416 (1987).

190. As a commander and logistical coordinator in Aweba, Katanga exercised operational control over the movement of arms and fighters and facilitated the coordination of the attack on Bogoro. AI-DSS function similarly by shaping operational choices for others to follow. See *Katanga*, ICC-01/04-01/07 at ¶ 1339; Brumfiel, *supra* note 44.

191. See *Katanga*, ICC-01/04-01/07 at ¶ 1635.

192. See Marwa Fatafta & Daniel Leufer, *Artificial Genocidal Intelligence: how Israel Is automating human rights abuses and war crimes*, ACCESS NOW (May 9, 2024), <https://www.accessnow.org/publication/artificial-genocidal-intelligence-israel-gaza>.

193. See generally *Katanga*, ICC-01/04-01/07.

194. See *id.* ¶ 1691.

195. Swart, *supra* note 4, at 599.

liability suggests moral blameworthiness can be applied more flexibly in contexts involving complex, distributed decision-making.¹⁹⁶ Although retribution traditionally requires punishment proportionate to the defendant's intent and direct participation, international criminal law also serves broader purposes: deterrence, accountability, promotion of peace and security, the creation of a historical record, truth-telling, and protection of victims.¹⁹⁷ These broader objectives suggest the legal framework can potentially evolve to address contemporary issues, including those posed by AI.

In the Gaza conflict, the integration of *Lavender* and *Gospel* raises critical questions about criminal liability and the implementation of accessory liability. Since *Lavender* merely provides recommendations that must be reviewed and approved by human analysts, it seems logical to hold those analysts accountable for errors.¹⁹⁸ However, the practical implementation of *Lavender* suggests a deeper challenge. In the conflict's early stages, military protocol reportedly allowed officers to accept *Lavender*'s "kill lists" with minimal scrutiny of underlying intelligence.¹⁹⁹ IDF analysts often spent only twenty seconds per target at the identification stage, focusing primarily on whether the marked individual was male rather than examining the basis for the recommendation.²⁰⁰ This practice left IDF intelligence personnel effectively serving as "rubber stamps," rather than engaging in meaningful oversight.²⁰¹

196. Federica I. Paddeu, *Shared Non-responsibility in International Law? Defences and the Responsibility of Co-perpetrators and Accessories in the Guiding Principles*, 31 EUR.J. OF INT'L L. 1264, 1264-65 (2020).

197. See Mark Klamberg, *What are the Objectives of International Criminal Procedure?*, 79 NORDIC J. INT'L L. 279, 279 (2010); Swart, *supra* note 4, at 599; Alec Walen, *Retributive Justice*, STAN. ENCYCLOPEDIA OF PHIL. (July 13, 2020), <https://plato.stanford.edu/archives/spr2025/entries/justice-retributive>.

198. Atay Kozlovski, *When Algorithms Decide Who is a Target: IDF's Use of AI in Gaza*, TECH POL'Y PRESS (May 13, 2024), <https://www.techpolicy.press/when-algorithms-decide-who-is-a-target-idfs-use-of-ai-in-gaza/>.

199. This challenge is often described as the "problem of many hands," which refers to the difficulty of assigning accountability when numerous actors are involved in designing, regulating, deploying, and maintaining different AI systems. Since responsibility is so widely dispersed "no single actor will have contributed enough to the action to be blamed." *Id.*

200. *Lavender* does not objectively identify who is a Hamas operative; rather, it constructs that definition through its algorithm, creating an illusion of scientific precision based on patterns in the data it is given. *Id.*; see Kadambini Sharma, *2 New Recruits Help Israel Track Hamas Operatives in Gaza: They Are Both AI*, NDTV (Apr. 28, 2024, at 10:33 AM IST), <https://www.ndtv.com/world-news/explained-what-are-https://www.ndtv.com/world-news/explained-what-are-lavender-and-gospel-ai-israel-is-using-to-bomb-gaza-5540086>.

201. Kozlovski, *supra* note 198.

Solely holding analysts accountable, however, will not address failures of military protocol.²⁰² Accessory liability is the appropriate answer to these complexities. Responsibility must extend to system creators and commanders who authorize and perpetuate deficient review procedures.²⁰³ Under the principle of accessory liability, criminal responsibility may extend to those who knew or should have known their contributions enabled a system that bypassed safeguards required under international humanitarian law.²⁰⁴

By focusing on humans as accessories and treating AI as “digital co-perpetrators,” this approach ensures liability does not dissipate across technical and institutional layers but instead reflects the collective structure facilitating harm. This framework acknowledges that modern conflict no longer hinges on single trigger-pullers and responsibility must follow influence, especially when human review has become a mere formality.²⁰⁵

VI. CONCLUSION

Accountability for battlefield errors in target selection has always been challenging.²⁰⁶ This problem is compounded by the lack of transparency in AI systems, increasing the difficulty of tracing and understanding decision-making processes behind military actions.²⁰⁷ While AI promises efficiency, it upends traditional notions of responsibility, especially as AI-DSS influence decision-making processes.²⁰⁸ In Gaza, Israel is using tools like *Lavender* and *Gospel* to gain a strategic advantage over its adversary.²⁰⁹ These systems were initially designed to assist decision-making; but they evolved into validation tools accelerating military actions at the cost of civilian lives.²¹⁰

202. *See id.*

203. Zoe Porter et al., *Unravelling Responsibility for AI*, J. OF RESPONSIBLE TECH. 4-9 (forthcoming 2025), <https://arxiv.org/pdf/2308.02608>.

204. *See* ELIES VAN SLIEDRECT, *THE CRIMINAL RESPONSIBILITY OF INDIVIDUALS FOR VIOLATIONS OF INTERNATIONAL HUMANITARIAN LAW* 157-60 (2003).

205. *See* Sylvia, *supra* note 21.

206. Mimran et al., *supra* note 2.

207. *See generally* Stewart & Hinds, *supra* note 7 (highlighting that AI-DSS outputs can be difficult to verify and may unduly influence human decision-makers, thereby increasing the risk of unpredictable, unaccountable harm in armed conflict).

208. Klaus, *supra* note 26.

209. *See* Heather Ashby, *From Gaza to Ukraine: AI Is Transforming War*, INKSTICK MEDIA (Mar. 6, 2024), <https://inkstickmedia.com/from-gaza-to-ukraine-ai-is-transforming-war/#:~:text=The%20battle%20between%20Russia%20and,in%20intelligence%20gathering%20and%20surveillance;Schmitt, supra note 2>.

210. *See* Ashby, *supra* note 209; Kozlovski, *supra* note 198; Elliott, *supra* note 29.

Some experts have proposed radical changes such as creating a strict liability regime for those involved with AI-DSS.²¹¹ Those changes would do away with fault for crimes of moral turpitude.²¹² Moreover, they would require negotiating a new IHL treaty and amending the ICC Statute.²¹³ But in the current divisive geo-political climate, countries are increasingly reluctant to adopt binding IHL instruments perceived as constraining national security interests.²¹⁴ In addition, IHL experts are concerned that such an effort “could open up existing Geneva Conventions and related International Humanitarian Law treaties to revision that would water down or weaken their provisions.”²¹⁵

Instead, international and domestic legal systems could effectively address the problems identified in this Article by adapting command responsibility and accessory liability concepts to the unique context of AI-DSS by treating such systems as digital subordinates and digital co-perpetrators. This would encourage proactive supervision and training of AI-DSS, and impose a higher standard of diligence on programmers, procurement officers, operators, and commanders at all levels. This does not require the adoption of new legal instruments but rather flexibility in the application of existing principles to new technological contexts. To address the accountability gap, international legal approaches must evolve alongside technological advancements.

This proposed approach would essentially treat AI-DSS as unindicted co-perpetrators. Doing so would not require recognizing that AI has legal personality, though such a conclusion may eventually arise by analogy to the evolution of corporate rights and criminal liability.²¹⁶

211. See Swart, *supra* note 4, at 597-98; Williams & Westlake, *supra* note 4, at 187.

212. Swart, *supra* note 4, at 598.

213. See Cordula Droegge & Eirini Giorgou, *How International Humanitarian Law Develops*, 104 INT'L REV. RED CROSS 1798, 1809 (Nov. 2022); Rome Statute, *supra* note 88, art. 121.

214. Scharf, *supra* note 1, at 8.

215. *Id.* at 8.

216. See generally Tanner W. Mathison, *Recognizing Right: The Status of Artificial Intelligence*, 19 J. BUS. & TECH. L. 105 (2023), <https://digitalcommons.law.umaryland.edu/jbtl/vol19/iss1/4> (arguing that extending a tailored form of legal personality to AI could address accountability gaps by enabling AI systems to bear limited rights and duties within existing legal frameworks).