Laws for LAWS: The Governance of Killer Robots

by Barron Omega

Edited by
Lara Pesce Ares
Sanaaz Nourkhaladj
Emma Fulweiler

Photo by
Gustavo Denuncio

Abstract:

Barron Omega explores the proliferation of lethal autonomous weapons systems (LAWS) and their evolving role in modern warfare, narrowing in on the Russia-Ukraine conflict as a case example. Through evaluating technological developments and battlefield applications, this article demonstrates how autonomous drones create favorable cost-benefit ratios while posing significant ethical challenges. This article recommends that these systems should only be used defensively against armored machine targets.

In *Slaughterbots*, UC Berkeley Professor Stuart Russell and the Future of Life Institute (FLI) stunned millions with their warning of a dystopian future laced with lethal autonomous weapons. FLI's short film depicted tiny quad-copter drones armed with explosives designed for surgical military strikes that were repurposed to identify and engage civilians for mass assassinations.[1] While we have not seen such targeted executions, lethal autonomous weapons systems (LAWS) are instrumental in waging modern conflict and are influencing its evolution. Currently, LAWS occupy a tiny fraction of even the most advanced militaries. However, the combatants in the ongoing Russia-Ukraine conflict are demonstrating how that paradigm is changing and emphasize the perceived necessity of such weapons.[2] Specifically, they illustrate how these weapons excel at effectively conducting crippling strikes and imposing costly trade-offs against an adversary.

Proliferate development and use of LAWS risk both increased false positives *and* false negatives, needlessly endangering noncombatants and friendly forces alike.[3] Yet, we face so-called "unstoppable incentives"—geopolitics, favorable economic tradeoffs, and diffusive technological norms—that demand nations to pursue such technologies to remain competitive.[4] Given these incentives and the lack of international agreement, critics from academia, civil society, and government must acknowledge that LAWS *will* have a role in future conflict; further protestation is futile.[5] As a major purveyor of autonomous weapons for the United States, the Department of Defense's (DoD) Defense Innovation Unit (DIU) has foundational governance capabilities. I argue that governance over LAWS' *use* and *design*—specifically through constraining their deployment conditions and payload size—will help mitigate the risk of errant targeting.

Below, I briefly outline what lethal autonomous weapons are, the driving forces behind their steadily increasing adoption in modern militaries, and discuss the importance and implications of LAWS on the battlefield. I then recommend some minimal governance standards and conclude with comments on the DIU's Replicator Initiative.

The next generation of weapons platforms

Defining LAWS has perplexed governments since the mid-2010s and hindered international control efforts.[6] Per the DoD's Directive 3000.09, the United States characterizes it as "weapon systems that, once activated, are able to select and engage their targets without further intervention from a human operator."[7] The lack of consensus has not prevented the keenest of states from developing their own LAWS—the United States, United Kingdom, Australia, Turkey, Israel, South Korea, China, and Russia have all reportedly either deployed or are actively investing in such weapons.[8]

These weapons are categorically different from their predecessors because of their independence. Like many of their semiautonomous counterparts, LAWS use an array of sensors—optical, infrared, electronic signal, etc.—to accurately navigate and identify objects in their environment. Yet, they differ in their capability of "mission autonomy"—the ability to operate or "loiter" in an engagement zone, determine whether an entity is a viable target, and when necessary, carry out a strike without human control or approval.[9]

Many of these advances are due to the confluence of massive data availability, advances in machine learning, and unprecedented processing speed at low costs. This speaks to the omni-use nature of technology as the same revolutions exciting civilians about ChatGPT and Waymo and Tesla's self-driving cars are powering military technologies as well.[9] However, unlike their civilian counterparts, these systems have no shortage of data to train on, as myriad intelligence, surveillance, and reconnaissance (ISR) systems and simulated data provide a cornucopia of supplementary training data.[10] Moreover, these weapons currently do not require the same degree of careful navigation as the aforementioned Tesla—there is no need to yield to stop signs, traffic, or pedestrians on the battlefield—making them cheaper and easier to train.[10]

Various NGOs and less-militarized states have repeatedly attempted to ban these weapons through international law.[10] Their advocacy efforts—usually in the form of media campaigns, UN discussion forums, and article publications—have emphasized the risks associated with LAWS. They have highlighted weapons proliferation, inadvertent escalation, dehumanization, and erroneous targeting as fundamental dangers to LAWS development. Such efforts have fallen on deaf ears, though, as the UN's Group of Governmental Experts meeting in August achieved no meaningful outcomes after a small minority—chiefly the United States and Russia—blocked consensus efforts.[10] International law has proven ineffective in governing yet another emerging technology.

For now, the LAWS seeing the greatest experimentation, deployment, and discussion in Ukraine are in the form of unmanned aerial vehicles (UAVs), and I will be focusing on such UAVs here. Drone makers, like the Ukrainian-based Saker UAV, have fixated on constructing small, relatively cheap products akin to commercially available quad-copters.[11] These are often retrofitted with anti-personnel or anti-vehicle munitions and are used as delivery systems, either dropping their payloads atop a target or diving into them. Notably, most began as first-person view (FPV) drones, piloted by hobbyists and gamers before—and after—their military conversion.[12]

The autonomous moment

Why the sudden drive for autonomous drones now? In short, it is a product of converging forces: the relentless competition of asymmetric warfare and the burgeoning surplus of data and machine learning capabilities.[13] Apart from the Russia-Ukraine conflict, we have not seen the widespread use of autonomous drones, and their use has been confined to open battlefield environments devoid of noncombatants.

Regarding asymmetric competition, one widely reported impetus behind the shift is the cat-and-mouse game of drones and of jammers. As stated, the Ukrainians relied on cheap, human-piloted drones to safely drop munitions on Russian tanks and other assets. Russia responded by placing electronic jammers around such valuable assets to disrupt or disable operator control over piloted drones, essentially rendering the drones useless.[14] In turn, both Ukrainian software developers and frontline troops have increasingly adjusted the weapons' software toward fully autonomous capabilities that enable the drone to identify and engage a target despite losing operator control.[15]

I must stress the less-discussed driver of the omni-use nature of unprecedented compute capabilities and the availability of data on which to train these UAVs. Similar to other groundbreaking AI systems celebrated in the civilian world, the preponderance of data generated or collected through ISR capabilities has been used to train anything from autonomous fighter aircraft to drones.[16] Indeed, many Silicon Valley defense technology startups have come to Ukraine to augur the country's expansive software engineering capabilities and have re-tailored tools like image classification software to recognize combatant forces like tanks instead of everyday objects like fruit.[17]

A more timeless driver of this dawning era of autonomous weapons is the marked change to the cost calculus of conflict, both in terms of human lives and equipment. For the former, these drones are the latest in infantry-capable "standoff" weapons, allowing them to strike targets without fear of reprisal. For the latter, LAWS are significantly cheaper than their conventional weapon alternatives and *drastically* cheaper than their intended machinery targets.[18] This creates a favorable force exchange between them and their target, enabling greater low-risk, high-reward exchanges.

For example, the Ukrainians have reported using up to ten quad-copter drones—valued at roughly $1,000 each—to destroy Russian T-90, T-80, and T-72 battle tanks, or disable them for follow-up artillery.[19] Excluding the artillery, this equates to a value exchange of roughly $10,000 for 2.5 to 4 million USD with minimal risk to the operator.[20] One popular video showcases the perspective from such a UAV that identifies a tank in the distance as it loses the camera feed to static as it comes into closer proximity; the video then cuts to a reconnaissance drone much further away capturing a sizable explosion on the tank, prefaced by faint imagery of a small white drone flying into it.[20]

Indeed, these drone strikes have accounted for roughly two-thirds of Ukraine's successful strikes against Russian tanks, making them an extremely cost-effective countermeasure.[21] Compare this against the shoulder-mounted U.S. Javelin, arguably the safest and most capable "smart" anti-tank infantry weapon. Firing one rocket costs approximately $200,000 and has an effective range of only 2.5 to 4 kilometers, putting the operator in closer proximity than the drone's 10-kilometer range.[21]

Governance recommendations

These latest developments have been a Rubicon crossing as techniques to enable autonomy will invariably proliferate among other weapons systems. Since major powers make international law ineffective and militaries are increasingly looking to adopt commercially and economically viable solutions, I contend that the primary means to resolution lies in guiding the development and deployment of LAWS. Indeed, Dr. Russell has argued that developers could decouple the firing control mechanism from the onboard computer, such that said action always remains under human control.

This follows the human in/on the loop (HITL/HOTL) paradigm that has dominated human-computer interactions in the military.[22] While this has been a principled approach to governance, it may not apply to this new breed of weaponry, as one of the biggest draws of AI-enabled weapons is their response speed. Weapons that rely on human permission will always function at the speed of human processing. As these drones are beginning to be deployed in swarms of dozens or hundreds, human control can become a hindrance.

I assert another practical option would be to limit the size of the payload or munitions used to be no larger than necessary to incapacitate armored targets. Platforms like the American Switchblade 300 or the Israeli Hero-30 use payloads between 1.5 to 2 kilograms—excellent for anti-personnel uses but ineffective against anything beyond "soft-skinned" commercial vehicles.[23] Their bigger brothers, the Switchblade 600 and the Hero-90 or -120, are geared toward larger armored targets, and while they could still be used against personnel, their use may be considered wasteful.[24]

I argue that the ideal use case for these weapons is against armored targets. The key characteristics of these LAWS—namely their compact, cheap nature and relatively limited sensor networks—indicate a circumscribed use case: debilitating strikes against uniquely identifiable targets, predominantly in non-urban environments. For all intents and purposes, these weapons are essentially diminutive smart(er) missiles that maximize their value by hitting targets conventional forces could not safely attack.

The U.S. military should therefore institute them as loitering sentries and force multipliers, not autonomous hunter-killers. Since LAWS rely so heavily on onboard sensors and processors, developers should design them for targets that produce the greatest probability of unique signals and reduce opportunities for signal misinterpretation to maximize their value. Relative to differentiating combatants from non-combatants, it is easier to correctly identify a valid threat in these two contexts given the distinct signals and signatures at play. Assuming no data poisoning, training data on incoming missiles or attack aircraft look unlike virtually anything else and should have high external validity. Similarly, focusing weapons development on armored vehicles promotes greater use of multi-spectral targeting systems, from Electronic Intelligence (ELINT) to Infrared (IR), to most accurately identify a valid target. By limiting the use cases and deployment environments, operators limit the potential for costly false positives and false negatives.

As such, I argue these weapons should primarily be characterized and deployed as defensive assets, not offensive ones. They excel at keeping warfighters *safer* specifically because they distance them from major threats. As such, the DoD should be guiding developers to lean into that niche, designing and marketing LAWS as force *protection* weapons deployed to keep soldiers safe from larger threats. Why risk a squad of soldiers to destroy a tank if you can use a disposable autonomous weapon instead?

Some early successes of the Replicator Initiative

The United States of course noticed the success in Ukraine and launched the DoD's Replicator Initiative in 2023 as a promise to "deliver all-domain attritable autonomous systems (ADA2) to warfighters at a scale of multiple thousands, across multiple warfighting domains, by August 2025."[25] Ostensibly, their vision for future warfare scenarios includes cheap, squad-deployable drones as a critical tool for the warfighter, a major departure from the United State's history of exquisite—i.e. expensive and cutting edge—weaponry.

Unfortunately, there is limited publicly available information on the specifications for Replicator apart from the guiding ethos and the announced partner companies. This is likely because they cannot disclose further details for fear of adversarial exploitation. Still, their public statements on experimenting, fielding, and adopting the *Switchblade 600*, Anduril's *Ghost X* and *Altius 600* systems, and Performance Drone Works *C-100 UAS* does indicate a certain target archetype.[26] Each of the drones mentioned carries payloads ranging from 10 to 20 pounds, well suited for commercial or armored vehicles, as well as fortified infantry positions.[27,28] This is significant as it signals the DoD either aligns with my argued ideal use case or it saliently recognizes the risk of errant targeting. We may not know how the U.S. military will use these weapons until the next major conflict nor whether these particular drones will still be relevant as countermeasures adapt. Nonetheless, the transformative nature of these weapons mandates the utmost scrutiny upon their development and deployment.

Concluding thoughts

So are the FLI's concerns of an impending drone apocalypse justified? While the answer is unknown because the data is limited, I would argue no, they are not. Developers and governments alike appear to be factoring in use cases to these weapons' construction and for now, they still exceed non-state actor accessibility. That said, lethal autonomous weapons, while excellent cost-saving tools, further dehumanize not just combatants but the conduct of war itself. At best, soldiers and civilians alike simply become objects to watch on a screen and at worst, numbers on a casualty report. We *should* bear some psychological and sociological cost for waging war beyond the spent resources and lost lives. War is instructive, not only for militaries, but for the wider populace, and waging it forces us to reconcile whether state interests are truly worth our blood, sweat, and treasure. If we deny that question, we risk dismissing the true cost of conflict and may make it even more likely in our ignorance. At its core, war is a human institution. It must remain so.

Endnotes

[1] Future of Life Institute. "Slaughterbots." November 13, 2017.

[2] Fontes, Robin, and Jorrit Kamminga. "Ukraine a Living Lab for AI Warfare." National Defense Magazine, March 24, 2024.

[3] Khan, Azmat. "Hidden Pentagon Records Reveal Patterns of Failure in Deadly Airstrikes." The New York Times, December 18, 2021.

[4] Suleyman, Mustafa. The Coming Wave. New York: Crown, 2023.

[5] Varella, Laura. "CCW Report." September 5, 2024.

[6] Adam, David. "Lethal AI Weapons Are Here: How Can We Control Them?" Nature News, April 23, 2024.

[7] United States Department of Defense. "DoD Directive 3000.09, Autonomy in Weapon Systems." January 25, 2023.

[8] Luca, Laura M., and Robert F. Trager. "Killer Robots Are Here—and We Need to Regulate Them." Foreign Policy, May 11, 2022.

[9] Bode, Ingvild. "Loitering Munitions and Unpredictability: Autonomy in Weapon Systems and Challenges to Human Control." AutoNorms, June 7, 2023.

[10] Hunder, Max. "Ukraine Collects Vast War Data Trove to Train AI Models." Reuters, December 20, 2024.

[11] Mozur, Paul, and Adam Satariano. "A.I. Begins Ushering in an Age of Killer Robots." The New York Times, July 2, 2024.

[12] Marson, James. "The Nerdy Gamers Who Became Ukraine's Deadliest..." The Wall Street Journal, November 3, 2024.

[13] Bondar, Kateryna. "Understanding the Military AI Ecosystem of Ukraine." CSIS, November 12, 2024.

[14] "AI Will Transform the Character of Warfare." The Economist, June 20, 2024.

[15] Hambling, David. "Ukraine Rolls out Target-Seeking Terminator Drones." Forbes, March 21, 2024.

[16] Lipton, Eric. "A.I. Brings the Robot Wingman to Aerial Combat." The New York Times, August 27, 2023.

[17] Allen, T.S. "Command, Control, and Autonomy in the Russo-Ukrainian War." CCW Emerging Threats & Technology Group, April 27, 2023.

[18] Rennolds, Nathan. "Ukrainian Forces Strike Russian Airbase with at Least 70 Drones, Targeting Su-34 Jets Used to Drop Glide Bombs." Business Insider, June 15, 2024.

[19] Axe, David. "Sixty-Year-Old T-62s Are about to Become the Russian Army's Main Tanks." Forbes, July 10, 2024.

20. Schmidt, Eric, and Will Roper. "Ukraine Shows How Drones Are Changing Warfare." Time, September 28, 2023.

21. Detsch, Jack. "Ukraine's Cheap Drones Are Decimating Russia's Tanks." Foreign Policy, April 9, 2024.

22. Congressional Research Service. "U.S. Policy on Lethal Autonomous Weapon Systems." February 1, 2024.

23. AeroVironment, Inc. "Switchblade 300 Loitering Munition Systems: Tactical Missile Systems: Suicide Drone: Kamikaze Drone."

24. AeroVironment, Inc. "Switchblade 600 Loitering Munition Systems: Kamikaze Drones: Suicide Drone: Tactical Missile Systems."

25. Department of Defense. The Replicator Initiative.

26. U.S. Department of Defense. "Deputy Secretary of Defense Kathleen Hicks Announces Additional Replicator All-Domain Attritable Autonomous Capabilities." November 13, 2024.

27. Anduril. "Anduril Unveils 'ghost-X' Autonomous Drone for Greater Mission Flexibility in Challenging Environments." December 9, 2023.

28. "The Battlefield Is Undergoing a Major Revolution and Small Robotics Are Leading the Way." PDW.