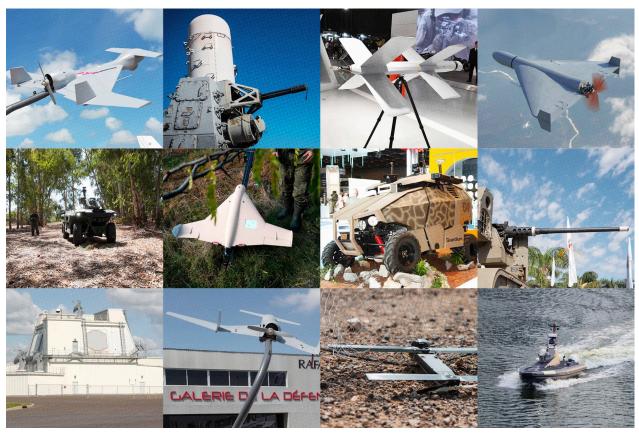**MEMO** *Published February 28, 2024 · 3 minute read*

Lethal Autonomous Weapons 101

Mike Sexton
Senior Policy Advisor for Cyber and Artificial Intelligence
🐦 **@MikeESexton**

Erik Carneal
Executive Coordinator
🐦 **@ThirdWayNatSec**

The AI revolution is bringing unprecedented change to every facet of human life and society, and warfare is no exception. AI makes it easier to cheat on your homework and spin up fake news; it also makes it easier to discover life-saving medical treatments and craft legal defenses for the indigent. So what is going to happen with AI-enabled robotic weapons?

Takeaways

- Lethal autonomous weapons predate the computer era and are in use in offensive and defensive warfare throughout the world today.

- Incorporating autonomy into weapons can make them safer and more precise in the immediate term but adds new risk in the longer term.

- The next generation of autonomous weapons has a small but non-negligible risk of being hacked.

- A rogue AGI that causes widespread destruction is highly speculative, but too grave of a possibility to dismiss.

- The enemy gets a vote. International agreements to safeguard humans from autonomous weapons are necessary, but there are also rogue actors who could ignore them.

While "killer robots" sound newfangled, lethal autonomous weapons (LAWs) are not new. Many LAWs currently in use are decades old. Even mines—a *very* old weapon—are a lethal autonomous weapon. [1]

It may be counterintuitive, but because autonomy can improve not just the efficiency but the accuracy of weapons, some autonomy can actually make weapons more humane. [2] Precision-guided munitions incorporate a degree of autonomy in order to be more accurate than a purely human-aimed "dumb" bomb, thereby saving civilian lives. So while it makes sense to be wary of automation in warfare, the reality of its impact is more complex than our intuitions would lead us to believe.

Our intuitive aversion to automated warfare is not baseless, though. Autonomous and smart weapons have a small but non-negligible chance of being hacked, and recent AI breakthroughs only make this easier. Worse yet, an unscrupulous military could hypothetically rely on such weapons to automate mass killings, circumventing its troops' consciences.

Perhaps the most prominent threat, however, comes from artificial general intelligence (AGI)—a hypothetical future AI that is vastly smarter than humans. [3] A rogue AGI—i.e., an AGI that has turned against humans—may be able to hack autonomous weapons at scale, stealing its own robotic military and turning it against humanity. While this is a highly speculative scenario, it is much too grave and plausible to dismiss.

However, it is important to remember the maxim that in war, the enemy gets a vote: LAWs are not developed in a vacuum, but to rival adversary LAWs and to overcome asymmetric threats like piracy and guerrilla-style fighting. For better or worse, near—and medium-term considerations—including humanitarian concerns—have contributed to the diffusion of lethal autonomous weapons.

To better understand whether advances in AI and weaponry will make warfare more or less humane, it is more useful to learn what lethal autonomous weapons there are than it is to hear what any one person thinks about them. For that reason, this product emphasizes exposition over analysis and policy recommendations.

It is hard to imagine any international rule that will conclusively resolve common sense humanitarian concerns with autonomous weapons. An intuitive requirement that every "autonomous" weapon must be fired manually will simply not work for missile defense systems like the Phalanx CIWS. [4] The Iran Air Flight 655 disaster at the hands of a human-supervised Aegis combat system furthermore underscores that this human review would be no guarantee of better results. [5]

That does not mean policy progress on autonomous weapons is doomed, however. Incremental agreements—e.g., requiring certain precautions in urban areas, or limiting certain classes of autonomous weapons like turrets—are realistic and worth exploring. The key to achieving meaningful policy change with respect to lethal autonomous weapons will be striving for progress, not perfection.

Lethal Autonomous Weapons
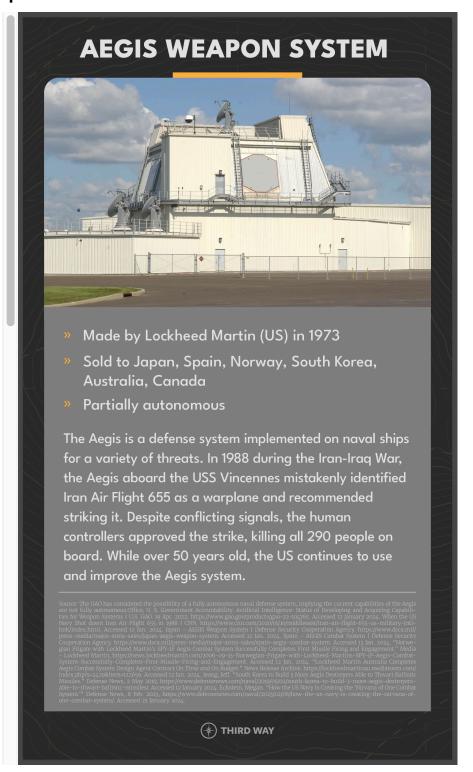

Aegis Weapon System

Rex MK II

Guardium

Harpy drone

Harop drone

Phalanx Close in Weapon System (CIWS)

AEGIS WEAPON SYSTEM

» Made by Lockheed Martin (US) in 1973

» Sold to Japan, Spain, Norway, South Korea, Australia, Canada

» Partially autonomous

The Aegis is a defense system implemented on naval ships for a variety of threats. In 1988 during the Iran-Iraq War, the Aegis aboard the USS Vincennes mistakenly identified Iran Air Flight 655 as a warplane and recommended striking it. Despite conflicting signals, the human controllers approved the strike, killing all 290 people on board. While over 50 years old, the US continues to use and improve the Aegis system.

Source: The GAO has considered the possibility of a fully autonomous naval defense system, implying the current capabilities of the Aegis are not fully autonomous.Office, U. S. Government Accountability. Artificial Intelligence: Status of Developing and Acquiring Capabilities for Weapon Systems | U.S. GAO. 19 Apr. 2022, https://www.gao.gov/products/gao-22-104765. Accessed 12 January 2024. When the US Navy Shot down Iran Air Flight 655 in 1988 | CNN. https://www.cnn.com/2020/01/10/middleeast/iran-air-flight-655-us-military-intl-hnk/index.html. Accessed 12 Jan. 2024. Japan – AEGIS Weapon System | Defense Security Cooperation Agency. https://www.dsca.mil/press-media/major-arms-sales/japan-aegis-weapon-system. Accessed 12 Jan. 2024. Spain – AEGIS Combat System | Defense Security Cooperation Agency. https://www.dsca.mil/press-media/major-arms-sales/spain-aegis-combat-system. Accessed 12 Jan. 2024. "Norwegian Frigate with Lockheed Martin's SPY-1F Aegis Combat System Successfully Completes First Missile Firing and Engagement." Media – Lockheed Martin, https://news.lockheedmartin.com/2006-09-13-Norwegian-Frigate-with-Lockheed-Martins-SPY-1F-Aegis-Combat-System-Successfully-Completes-First-Missile-Firing-and-Engagement. Accessed 12 Jan. 2024. "Lockheed Martin Australia Completes Aegis Combat System Design Agent Contract On Time and On Budget." News Release Archive, https://lockheedmartinau.mediaroom.com/index.php?s=24298item=122659. Accessed 12 Jan. 2024. Jeong, Jeff. "South Korea to Build 3 More Aegis Destroyers Able to Thwart Ballistic Missiles." Defense News, 2 May 2019, https://www.defensenews.com/naval/2019/05/02/south-korea-to-build-3-more-aegis-destroyers-able-to-thwart-ballistic-missiles/. Accessed 12 January 2024. Eckstein, Megan. "How the US Navy Is Creating the 'Nirvana of One Combat System.'" Defense News, 8 Feb. 2023, https://www.defensenews.com/naval/2023/02/08/how-the-us-navy-is-creating-the-nirvana-of-one-combat-system/. Accessed 25 January 2024.

⊛ THIRD WAY

ENDNOTES

1. "70 States Deliver Joint Statement on Autonomous Weapons Systems at UN General Assembly." Stop Killer Robots, 21 Oct. 2022, https://www.stopkillerrobots.org/news/70-states-deliver-joint-statement-on-autonomous-weapons-systems-at-un-general-assembly/. Accessed 12 December 2023.

2. Horowitz, Michael C., and Paul Scharre. "Do Killer Robots Save Lives?" POLITICO Magazine, 19 Nov. 2014, https://www.politico.com/magazine/story/2014/11/killer-robots-save-lives-113010. Accessed 11 January 2024.

3. Walsh, Toby. "The Problem with Artificial (General) Intelligence in Warfare." Centre for International Governance Innovation, https://www.cigionline.org/articles/the-problem-with-artificial-general-intelligence-in-warfare/. Accessed 16 Jan. 2024.

4. Phalanx System. https://www.rtx.com/raytheon/what-we-do/sea/phalanx-close-in-weapon-system. Accessed 16 Jan. 2024.

5. When the US Navy Shot down Iran Air Flight 655 in 1988 | CNN. https://www.cnn.com/2020/01/10/middleeast/iran-air-flight-655-us-military-intl-hnk/index.html. Accessed 12 Jan. 2024.