**Headline:** The Rise of AI Warfare: How Autonomous Weapons and Cognitive Warfare Are Reshaping Global Military Strategy

**Teaser:** The future of warfare may be decided by AI, but will humans have a role to play?

By Leslie Alan Horvitz

**Author Bio:** Leslie Alan Horvitz is an author and journalist specializing in science and a contributor to the [Observatory.](https://observatory.wiki/Leslie_Alan_Horvitz) His nonfiction books include [*Eureka: Scientific Breakthroughs That Changed the World*, *Understanding Depression*](https://www.google.com/books/edition/Eureka/9j0xJjHWqa8C) with Dr. Raymond DePaulo of Johns Hopkins University, and [*The Essential Book of Weather Lore*](https://www.google.com/books/edition/Essential_Book_of_Weather_Lore/47K3GAAACAAJ?hl=en). His articles have been published by Travel and Leisure, Scholastic, Washington Times, and Insight on the News, among others. He has served on the board of [Art Omi](https://artomi.org/) and is a member of [PEN America](https://pen.org/). Horvitz is based in New York City. You can find him online at [lesliehorvitz.com](https://tinyurl.com/3v8fdh2k).

**Source:** Independent Media Institute

**Credit Line:** *This article was produced for* [*the Observatory*](https://observatory.wiki/) *by the* [*Independent Media Institute*](https://independentmediainstitute.org/)*.*

**Tags:** Science, War

**[Article Body:]**

In the 1983 film [*War Games*](https://en.wikipedia.org/wiki/WarGames), a supercomputer known as WOPR (for War Operation Plan Response) is about to provoke a nuclear war between the United States and the Soviet Union, but because of the ingenuity of a teenager (played by Matthew Broderick), catastrophe is averted. In the first [*Terminator*](https://en.wikipedia.org/wiki/Terminator_%28franchise%29) film, which was released a year later, a supercomputer called “Skynet” decides to exterminate humanity because it’s perceived as a threat to its existence rather than to protect American nuclear weapons.

Although these films offered audiences grim scenarios of intelligent machines running amok, they were also prophetic. Artificial intelligence (AI) is so commonplace that it’s routinely applied during a simple Google search. That it is also being integrated into military strategies is hardly any surprise. It’s just that we have little understanding of the capacity of these high-tech weapons (those that are now ready for use and those in development). Nor are we prepared for systems that have the capacity to transform warfare forever.

Throughout history, it is human intelligence that uses the technology, not the technology itself, which has won or lost wars. That may change in the future when human intelligence is focused instead on creating systems that are more capable on the battlefield than those of the adversary.

**An “Exponential, Insurmountable Surprise”**

Artificial intelligence isn’t a technology that can be easily detected, monitored, or banned, as Amir Husain, the founder and CEO of an AI company, SparkCognition, pointed out in an [essay](https://ndupress.ndu.edu/Media/News/News-Article-View/Article/2846375/ai-is-shaping-the-future-of-war/) for Media News. Integrating AI elements—visual recognition, language analysis, simulation-based prediction, and advanced forms of search—with existing technologies and platforms “can rapidly yield entirely new and unforeseen capabilities.” The result “can create exponential, insurmountable surprise,” Hussain writes.

Advanced technology in warfare is already widespread. The [use of uncrewed aerial vehicles (UAVs)](https://www.chathamhouse.org/sites/default/files/publications/research/2017-01-26-artificial-intelligence-future-warfare-cummings-final.pdf)—commonly known as drones—in military settings has set off warnings about “killer robots.” What happens when drones are no longer controlled by humans and can execute military missions on their own? These drones aren’t limited to the air; they can operate on the ground or underwater as well. The introduction of AI, effectively giving these weapons the capacity for autonomy, isn’t far off.

Moreover, they’re cheap to produce and cheap to purchase. The Russians are buying drones from Iran for use in their war in Ukraine, and the Ukrainians have been putting together a cottage industry constructing drones of their own against the Russians. The relative ease with which a commercial drone can be converted into one with a military application also blurs the line between commercial and military enterprises. At this point, though, humans are still in charge.

A similar problem can be seen in information-gathering systems that have dual uses, including satellites, manned and unmanned aircraft, ground and undersea radars, and sensors, all of which have both commercial and military applications. AI can process vast amounts of data from all these systems and then discern meaningful patterns, identifying changes that humans might never notice. American forces were stymied to some degree in wars in Iraq and Afghanistan because they [could not process large amounts of data.](https://www.chathamhouse.org/sites/default/files/publications/research/2017-01-26-artificial-intelligence-future-warfare-cummings-final.pdf) Even now, remotely piloted UAVs are using AI for autonomous takeoff, landing, and routine flight. All that’s left for human operators to do is concentrate on tactical decisions, such as selecting attack targets and executing attacks.

AI also allows these systems to operate rapidly, determining actions at speeds that are seldom possible if humans are part of the decision-making process. Until now, decision-making speed has been the most important aspect of warfare. If, however, AI systems go head-to-head against humans, AI will invariably come out ahead. However, the possibility that AI systems eliminate the human factor terrifies people who don’t want to see an apocalyptic scenario on celluloid come to pass in reality.

**Automated Versus Autonomous**

A distinction needs to be made between the term “autonomous” and the term “automated.” If we are controlling the drone, then the drone is automated. But if the drone is programmed to act on its own initiative, we would say it is autonomous. But does the autonomous weapon describe the actual weapon—i.e., a missile on a drone—or the drone itself? Take, for example, the [Global Hawk military UAV](https://www.af.mil/About-Us/Fact-Sheets/Display/Article/104516/rq-4-global-hawk/) (drone). It is automated insofar as it is controlled by an operator on the ground, and yet if it loses communication with the ground, the Golden Hawk can land on its own. Does that make it automated or autonomous? Or is it both?

The most important question is whether the system is [safety-critical](https://www.chathamhouse.org/sites/default/files/publications/research/2017-01-26-artificial-intelligence-future-warfare-cummings-final.pdf). Translated, that means whether it has the decision-making capacity to use a weapon against a target without intervention from its human operator. It is possible, for example, for a drone to strike a static military target on its own (such as an enemy military base) but not a human target because of the fear that innocent civilians could be injured or killed as collateral damage. Many countries have already developed drones with real-time imagery capable of acting autonomously in the former instance, but not when it comes to human targets.

Drones aren’t the only weapons that can act autonomously. Military systems are being developed by the U.S., China, and several countries in Europe that can act autonomously in the air, on the ground, in water, and underwater with varying degrees of success.

Several types of autonomous helicopters designed so that a soldier can direct them in the field with a smartphone are in development in the U.S., Europe, and China. Autonomous ground vehicles, such as tanks and transport vehicles, and autonomous underwater vehicles are also in development. In almost all cases, however, the agencies developing these technologies are struggling to make the leap from development to operational implementation.

There are many reasons for the lack of success in bringing these technologies to maturity, including cost and unforeseen technical issues, but equally problematic are organizational and cultural barriers. The U.S. has, for instance, struggled to bring autonomous UAVs to operational status, primarily due to organizational infighting and prioritization in favor of manned aircraft.

**The Future Warrior**

In the battleground of the future, elite soldiers may rely on a [head-up display](https://www.wired.com/story/us-military-hyper-enabled-operator/?utm_source=nl&utm_brand=wired&utm_mailing=WIR_Daily_070824&utm_campaign=aud-dev&utm_medium=email&utm_content=WIR_Daily_070824&bxid=5be9f4663f92a404692f5fba&cndid=26282935&hasha=f9863c1a66b7d467fe7ef67ebff76b37&hashb=f6877553839b3dc1fbd44403a2770a2df97af336&hashc=f373637b64e78ec70a6b19fb56aa496ebb0cbab4c5316dcbd2181f783d8ee6bb&esrc=OIDC_SELECT_ACCOUNT_PAGE&utm_term=WIR_Daily_Active) that feeds them a wealth of information that is collected and routed through supercomputers carried in their backpacks using an AI engine. With AI, the data is instantly analyzed, streamlined, and fed back into the head-up display. This is one of many potential scenarios [presented](https://taskandpurpose.com/news/hyper-enabled-operator-socom/) by U.S. Defense Department officials. The Pentagon has embraced a relatively simple concept: the “[hyper-enabled operator](https://www.defensenews.com/industry/techwatch/2023/05/10/hyper-enabling-special-ops-will-give-transform-missions/).”

The objective of this concept is to give Special Forces “cognitive overmatch” on the battlefield, or “the ability to dominate the situation by making informed decisions faster than the opponent.” In other words, they will be able to make decisions based on the information they are receiving more rapidly than their enemy. The decision-making model for the military is called the “[OODA loop](https://taskandpurpose.com/news/case-against-ooda-loop/)” for “observe, orient, decide, act.” That will come about using computers that register all relevant data and distill them into actionable information through a simple interface like a head-up display.

This display will also offer a “[visual environment translation](https://www.socom.mil/st/PublishingImages/SOCOM%20JATF%20Visual%20Environment%20Translation/SOCOM%20JATF%20Visual%20Environment%20Translation.mp4)” system designed to convert foreign language inputs into clear English in real time. Known as VITA, the system encompasses both a visual environment translation effort and voice-to-voice translation capabilities. The translation engine will allow the operator to “engage in effective conversations where it was previously impossible.”

[VITA](https://www.marktechpost.com/2025/01/05/vita-1-5-a-multimodal-large-language-model-that-integrates-vision-language-and-speech-through-a-carefully-designed-three-stage-training-methodology/), which stands for Versatile Intelligent Translation Assistant, offers users [language](https://www.wired.com/story/us-military-hyper-enabled-operator/) capabilities in Russian, Ukrainian, and Chinese, including Mandarin, a Chinese dialect. Operators could use their smartphones to scan a street in a foreign country, for example, and immediately obtain a translation of street signs in real-time.

**Adversary AI Systems**

Military experts divide [adversarial attacks](https://mwi.westpoint.edu/artificial-intelligence-future-warfare-just-not-way-think/) into [four categories](https://doi.org/10.6028/NIST.IR.8269-draft): evasion, inference, poisoning, and extraction. [These types of attacks](http://papers.nips.cc/paper/7647-adversarial-examples-that-fool-both-computer-vision-and-time-limited-humans.pdf) are [easily accomplished](https://www.sciencedirect.com/science/article/abs/pii/S0031320318302565?via%3Dihub) and often don’t require computing skills. An enemy engaged in evasive attacks could attempt to deceive an AI weapon to avoid detection—hiding a cyberattack, for example, or convincing a sensor that a tank is a school bus. This may require the development of a new type of AI camouflage, such as [strategic tape placement](https://www.sciencedirect.com/science/article/abs/pii/S0031320318302565),that can fool AI.

Inference attacks occur when an adversary acquires information about an AI system that allows evasive techniques. Poisoning attacks target AI systems during training, interfering with access to the datasets used to train military tools—mislabeling images of vehicles to dupe targeting systems, for instance, or manipulating maintenance data designed to classify imminent system failure as a regular operation.

Extraction attacks exploit access to the AI’s interface to learn enough about the AI’s operation to create a parallel model of the system. If AI systems are not secure from unauthorized users, then an adversary’s users could predict decisions made by those systems and use those predictions to their advantage. For instance, they could predict how an AI-controlled unmanned system will respond to specific visual and electromagnetic stimuli and then proceed to alter its route and behavior.

[Deceptive attacks](https://spectrum.ieee.org/adversarial-attacks-and-ai-systems?mkt_tok=eyJpIjoiTlROa1pEZ3dORGhpTlRCaiIsInQiOiJaT0Z3bXNkYnhTTHBLMk5ma01VcDNvem1xaUZEdUFxYUh3clpqbDJ1OHhBdlBmcmlKOUJUVzZCZEZjMFwvK3JobTFMZ1U1N3JlMkMrXC9zaHV0VTRxSmw5ZEgzMmN6YTdYSGRzTUg4R1NGUGJvY2k4QWNIQkxHaFwvV3lvSkFnMFlBayJ9) have become increasingly common, as illustrated by cases involving image classification algorithms that are [deceived](https://spectrum.ieee.org/tag/deepfakes) into perceiving images that aren’t there, confusing the meaning of images, and mistaking a turtle for a rifle, for instance. Similarly, [autonomous vehicles](https://spectrum.ieee.org/transportation/self-driving) could be forced to swerve into the wrong lane or speed through a stop sign.

In 2019, [China](https://warontherocks.com/2022/04/new-tech-new-concepts-chinas-plans-for-ai-and-cognitive-warfare/) announced a new military strategy, [Intelligentized Warfare](https://cyberdefensereview.army.mil/Portals/6/Documents/2024-Fall/Baughman_CDRV9N3-Fall-2024.pdf), which utilizes AI. Officials of the Chinese People’s Liberation Army have stated that their forces can overtake the U.S. military by using AI. One of its intentions is to use this high-tech type of warfare to bring Taiwan under its control without waging conventional warfare. However, [only a few](https://www.defenseone.com/ideas/2021/09/how-chinese-strategists-think-ai-will-power-military-leap-ahead/185409/) of the many Chinese studies on intelligentized warfare have focused on replacing guns with AI. On the other hand, Chinese strategists have made no secret of their intention to [control the enemy’s will directly](http://www.nids.mod.go.jp/publication/chinareport/pdf/china_report_EN_web_2021_A01.pdf).

That would include the U.S. president, members of Congress, combatant commanders, and citizens. “Intelligence dominance”—also known as cognitive warfare or “control of the brain”—is seen as the new battleground in intelligentized warfare, putting AI to a very different use than most American and allied discussions have envisioned. According to the Pentagon’s 2022 report on Chinese military developments, the People’s Liberation Army is [being trained and equipped](https://www.thenation.com/article/world/artificial-intelligence-us-military/) to use AI-enabled sensors and computer networks to “rapidly identify key vulnerabilities in the U.S. operational system and then combine joint forces across domains to launch precision strikes against those vulnerabilities.”

Controlling an adversary's mind can affect not just someone’s perceptions of their surroundings but, ultimately, their decisions. For the People’s Liberation Army, cognitive warfare is equal to the other domains of conflict, which are air, land, and sea. In that respect, social media is considered a key battlefield.

Russia has also been developing its own AI capacity. As early as 2014, the Russians inaugurated a National Defense Control Center in Moscow, a centralized command post for assessing and responding to global threats. The center was [designed](https://foxtrotalpha.jalopnik.com/look-inside-putins-massive-new-military-command-and-con-1743399678) to collect information on enemy moves from multiple sources and provide senior officers with guidance on possible responses.

[Russia](https://www.nytimes.com/2023/11/19/technology/russia-ukraine-electronic-warfare-drone-signals.html) has declared that it will eventually develop an AI system capable of running the world. Russians are already using AI in Ukraine to jam wireless signals connecting Ukrainian drones to the satellites they rely on for navigation, causing the machines to lose their way and plummet to Earth. [The Russian Ministry of Defense](https://www.cnas.org/publications/reports/the-role-of-ai-in-russias-confrontation-with-the-west) (MOD) has explored ways in which AI systems can be developed for uncrewed systems for the air, maritime, and ground domains. At the same time, at least in the short term, official policy is predicated on the belief that humans must remain firmly in the loop.

Meanwhile, the Russians are trying to improve UAV capabilities with AI as a mechanism for command, control, and communications. MOD also emphasizes the use of AI for data collection and analysis as a natural evolution from the current “digital” combat technology and systems development.

**“Raven Sentry”: AI in the U.S. War in Afghanistan**

The use of AI on the battlefield by U.S. intelligence, while brief, showed promising results. “Raven Sentry,” an AI tool launched in 2019 by a team of American intelligence officers (known as the “nerd locker”), with help from Silicon Valley expertise, was intended to forecast insurgent attacks. The initial use of AI came at a time when U.S. bases were closing, troop numbers were falling, and intelligence resources were being diverted. Raven Sentry relied on open-source data.

“We noticed an opportunity presented by the increased number of commercial satellites and the availability of news reports on the Internet, the proliferation of social media postings, and messaging apps with massive membership,” [says](https://www.eurasiantimes.com/raven-sentry-how-us-military-weaponized/) Col. Thomas Spahr, chief of staff of the Resolute Support J2 intelligence mission in Kabul, Afghanistan, from July 2019 to July 2020.

The AI tool also drew on historical patterns based on insurgent activities in Afghanistan going back 40 years, which encompassed the Soviet occupation of the country in the 1980s. Environmental factors were also considered. “Historically, insurgents attack on certain days of the year or holidays, for example, or during certain weather and illumination conditions,” Spahr notes. He adds, “The beauty of the AI is that it continues to update that template. The machine would learn as it absorbed more data.” Before its demise in 2021 (with the U.S. withdrawal from Afghanistan), Raven Sentry had demonstrated its feasibility, predicting an insurgent attack with 70 percent accuracy. The AI tool predicted that attacks were more likely to occur when the temperature was above 4 degrees Celsius (or 39.2 degrees Fahrenheit), when lunar illumination was below 30 percent, and when there was no rain. Spahr was satisfied with the results: “We validated that commercially produced, unclassified information can yield predictive intelligence.”

**Ukraine as Testing Ground for AI**

Ever since the Russian invasion, launched in 2022, [Ukraine](https://lieber.westpoint.edu/continuing-autonomous-arms-race/) has become a testing ground for AI in warfare. Outgunned and outmanned, Ukrainian forces have resorted to improvisation, jerry-rigging off-the-shelf devices to transform them into lethal autonomous weapons. The Russian invaders, too, have employed AI, conducting cyberattacks and GPS-jamming systems.

Ukraine’s Saker Scout quadcopters “can [find, identify, and attack](https://www.forbes.com/sites/davidhambling/2023/10/17/ukraines-ai-drones-seek-and-attack-russian-forces-without-human-oversight/?sh=3abfce3966da) 64 types of Russian ‘military objects’ on their own.” These drones are [designed to operate autonomously](https://www.defenseone.com/ideas/2024/04/ai-revolution-already-here/395722/), and unlike other drones that Ukrainian forces have deployed, Russia cannot jam them.

By using code found online and hobbyist computers like [Raspberry Pi](https://www.raspberrypi.com/), easily obtained from hardware stores, Ukrainians are able to construct innovative killer robots. Apart from drones, which can be operated with a smartphone, Ukrainians have built a gun turret with autonomous targeting operated with the same controller used by a PlayStation or a tablet. The gun, called Wolly because it bears a resemblance to the Pixar robot WALL-E, can auto-lock on a target up to 1,000 meters (3,280 feet) away and shift between preprogrammed positions to quickly cover a broad area.

The manufacturer is also developing a gun capable of hitting moving targets. It can automatically identify targets as they come over the horizon. The gun targets and aims automatically; all that’s left for the operator to do is press the button and shoot. Many Ukrainian drones, which look like those you can find at Walmart, are called First Person View (FPV) drones. Capable of flying 100 miles per hour, FPV drones have four propellers and a mounted camera that uses wireless to send footage of their flights back to operators. With a bomb on board, an FPV can be converted into a weapon that can take out a tank. They’re cheap, too; one manufacturer, [Vyriy](https://www.seattletimes.com/nation-world/ai-begins-ushering-in-an-age-of-killer-robots/), charges $400 each, a small price to pay to disable a tank worth millions of dollars. Vyriy derives its name from a mythical land in Slavic folktales.

If one kamikaze drone is good, dozens of them are better insofar as the greater their number, the greater the chance there is of several reaching their targets. In nature, a swarm of ants behaves as a single living organism, whether the task is collecting food or building a nest. Analogously, a [swarm of autonomous drones](https://www.pbs.org/newshour/show/how-militaries-are-using-artificial-intelligence-on-and-off-the-battlefield) could act as a single organism—no humans necessary—carrying out a mission regardless of how many are disabled or crash to the ground or whether communication from the ground is disrupted or terminated.

Although humans are still in the “loop,” these weapons could equally be made entirely autonomous. In other words, they could decide which targets to strike without human intervention.

It isn’t as if Ukraine has adopted [AI weaponry](https://www.csis.org/analysis/ukraines-future-vision-and-current-capabilities-waging-ai-enabled-autonomous-warfare) without any tech experience. In the words of New York Times reporter Paul Mozer, “Ukraine has been a bit of a back office for the global technology industry for a long time.” The country already had a substantial pool of coders and skilled experts who, under emergency conditions, were able to make the transition from civilian uses (such as a dating app) to military purposes. As Mozer reported: “What they’re doing is they’re taking basic code that is around, combining it with some new data from the war, and making it into something entirely different, which is a weapon.”

The reality is, “there’s a lot of cool, exciting stuff happening in the big defense primes,” says [P.W. Singer](https://www.pwsinger.com/), an author who writes about war and tech. “There’s a lot of cool, exciting stuff happening in the big-tech Silicon Valley companies. There’s a lot of cool, exciting stuff happening in small startups.”

One of those smaller startups is [Anduril](https://www.anduril.com/). After selling the popular virtual reality headset Oculus to Facebook (now Meta), Palmer Luckey, an entrepreneur in his early thirties, went on to found an AI weapons company that is supplying drones to Ukraine. “Ukraine is a very challenging environment to learn in,” he says. “I’ve heard various estimates from the Ukrainians themselves that any given drone typically has a lifespan of about four weeks. The question is, “Can you respond and adapt?” Anduril, named after a sword in *The* *Lord of the Rings*, has sold its devices to ten countries, including the U.S.

“I had this belief that the major defense companies didn’t have the right talent or the right incentive structure to invest in things like artificial intelligence, autonomy, robotics,” says Luckey. His company’s drone, called [ALTIUS](https://www.anduril.com/hardware/altius/), is intended to be fired out of a tube and unfold itself, extending its wings and tail; then, steering with a propeller, it acts like a plane capable of carrying a 30-pound warhead. Luckey believes that his approach will result in more AI weapons being built in less time and at a lower cost than could be achieved by traditional defense contractors like McDonnell Douglas.

Anduril, founded in 2017, is also developing the [Dive-LD](https://www.anduril.com/hardware/dive-ld/), a drone that will be used for surveys in littoral and deep water. “It’s an autonomous underwater vehicle that is able to go very, very long distances, dive to a depth of about 6,000 meters (almost 20,000 feet), which is deep enough to go to the bottom of almost any ocean,” says Luckey. Ukraine is already making its own sea drones—essentially jet skis packed with explosives—which have inflicted severe damage on the Russian navy in the Black Sea.

As Anduril’s CEO Brian Schimpf [admits](https://www.youtube.com/watch?v=rQqfjJpnqO8), the introduction of Anduril’s drones to Ukraine has yet to produce any significant results, although he believes that will change. Once they’re launched, these drones will not require guidance from an operator on the ground, making it difficult for the Russians to destroy or disable them by jamming their signals.

“The autonomy onboard is really what sets it apart,” Luckey says. “It’s not a remote-controlled plane. There’s a brain on it that is able to look for targets, identify targets, and fly into those targets.” However, for every innovative weapon system the Ukrainians develop, the Russians counter it with a system that renders it useless. “Technologies that worked really well even a few months ago are now constantly having to change,” says Jacquelyn Schneider, who studies military technology as a fellow at the [Hoover Institution](https://www.hoover.org/), “And the big difference I do see is that software changes the rate of change.”

**The War in Gaza: Lavender**

In their invasion of Gaza, the Israel Defense Forces (IDF) have increasingly relied on a [program supported by artificial intelligence](https://english.elpais.com/technology/2024-02-14/predatory-sparrow-and-other-weapons-of-hybrid-warfare-cheap-fast-undetectable-and-effective.html) to target Hamas operatives, with problematic consequences. According to an April 2024 [report](https://www.972mag.com/lavender-ai-israeli-army-gaza/) by +972 Magazine (an Israeli-Palestinian publication) and Local Call, a Hebrew language news site, the IDF has been implementing a program known as “Lavender,” whose influence on the military’s operations is so profound that intelligence officials have essentially treated the outputs of the AI machine “as if it were a human decision.”

Lavender was developed by the elite Unit 8200, which is comparable to the National Security Agency in the U.S. or the Government Communications Headquarters in the UK.

The Israeli government has defended Lavender for its practicality and efficiency. “The Israeli military uses AI to augment the decision-making processes of human operators. This use is in accordance with international humanitarian law, as applied by the modern Armed Forces in many asymmetric wars since September 11, 2001,” [says](https://english.elpais.com/technology/2024-04-17/lavender-israels-artificial-intelligence-system-that-decides-who-to-bomb-in-gaza.html) Magda Pacholska, a researcher at the TMC Asser Institute and specialist in the intersection between disruptive technologies and military law.

The [data collected](https://www.theguardian.com/world/2024/apr/03/israel-gaza-ai-database-hamas-airstrikes) to identify militants that were used to develop Lavender comes from the more than 2.3 million residents of the Gaza Strip, which was under intense surveillance prior to the Gaza invasion in 2023.

The report states that as many as 37,000 Palestinians were designated as suspected militants who were selected as potential targets. Lavender’s kill lists were prepared in advance of the invasion, launched in response to the Hamas attack of October 7, 2023, which left about 1,200 dead and about 250 hostages taken from Israel. A related AI program, which tracked the movements of individuals on the Lavender list, was called “Where’s Daddy?” Sources for the +972 Magazine report said that initially, there was “no requirement to thoroughly check why the machine made those choices (of targets) or to examine the raw intelligence data on which they were based.” The officials in charge, these sources said, acted as a “rubber stamp” for the machine’s decisions before authorizing a bombing. One intelligence officer who spoke to +972 [admitted](https://www.theguardian.com/world/2024/apr/03/israel-gaza-ai-database-hamas-airstrikes) as much: “I would invest 20 seconds for each target at this stage, and do dozens of them every day. I had zero added value as a human, apart from being a stamp of approval. It saved a lot of time.”

It was already known that the Lavender program made errors in 10 percent of the cases, meaning that a fraction of the individuals selected as targets might have had no connection with Hamas or any other militant group. The strikes generally occurred at night while the targeted individuals were more likely to be at home, which posed a risk of killing or wounding their families as well.

A [score](https://english.elpais.com/technology/2024-04-17/lavender-israels-artificial-intelligence-system-that-decides-who-to-bomb-in-gaza.html) was created for each individual, ranging from 1 to 100, based on how closely he was linked to the armed wing of Hamas or Islamic Jihad. Those with a high score were killed along with their families and neighbors despite the fact that officers reportedly did little to verify the potential targets identified by Lavender, citing “efficiency” reasons. “This is unparalleled, in my memory,” said [one intelligence officer](https://www.theguardian.com/world/2024/apr/03/israel-gaza-ai-database-hamas-airstrikes) who used Lavender, adding that his colleagues had more faith in a “statistical mechanism” than a grieving soldier. “Everyone there, including me, lost people on October 7. The machine did it coldly. And that made it easier.”

The IDF had previously used another AI system called “The Gospel,” which was described in a [previous investigation](https://www.972mag.com/mass-assassination-factory-israel-calculated-bombing-gaza/) by the magazine, as well as in the Israeli military’s own [publications](https://www.israeldefense.co.il/node/57256#google_vignette), to target buildings and structures suspected of harboring militants. “The Gospel” draws on [millions of items of data](https://www.defenseone.com/ideas/2024/04/ai-revolution-already-here/395722/), producing target lists more than 50 times faster than a team of human intelligence officers ever could. It was used to [strike](https://www.theguardian.com/world/2023/dec/01/the-gospel-how-israel-uses-ai-to-select-bombing-targets) 100 targets a day in the first two months of the Gaza fighting, roughly five times more than in a similar conflict there a decade ago. Those [structures of political or military significance](https://www.asahi.com/ajw/articles/15209132) for Hamas are known as “power targets.”

**Weaknesses of AI Weapons**

If an AI weapon is autonomous, it needs to have the capacity for accurate perception. That’s to say, if it mistakes a civilian car for a military target, its response rate isn’t relevant. The civilians in the car die regardless. In many cases, of course, AI systems have excelled at perception as AI-powered machines and algorithms have become refined. When, for instance, the Russian military conducted a test of 80 UAVs simultaneously flying over [Syrian battlefields](https://ndupress.ndu.edu/Media/News/News-Article-View/Article/2846375/ai-is-shaping-the-future-of-war/#endnote-180) with unified visualization, then Russian Defense Minister Sergei Shoigu compared it to a “semi-fantastic film” that revealed all potential targets.

But problems can creep in. In designing an AI weapon, developers first need access to data. Many AI systems are trained using data that has been labeled by an expert system (e.g., labeling scenes that include an air defense battery), usually a human. An AI’s image-processing capability won’t function well when given images [that are different from its training set](https://www.nature.com/articles/d41586-019-03013-5)—for example, pictures produced [where lighting conditions are poor, that are at an obtuse angle, or that are partially obscured.](https://www.theverge.com/2017/4/12/15271874/ai-adversarial-images-fooling-attacks-artificial-intelligence) AI [recognition systems](https://mwi.westpoint.edu/artificial-intelligence-future-warfare-just-not-way-think/) don’t understand what the image is; rather, they learn textures and gradients of the image’s pixels. That means that an AI system may correctly recognize a part of an image but not its entirety, which can result in misclassification.

To better [defend AI](https://www.theverge.com/2017/4/12/15271874/ai-adversarial-images-fooling-attacks-artificial-intelligence) against deceptive images, engineers subject them to “adversarial training.” This involves feeding a classifier adversarial images so it can identify and ignore those that aren’t going to be targeted. Research by Nicolas Papernot, a graduate student at Pennsylvania State University, shows that a system, even bolstered by adversarial training, may be ineffective if overwhelmed by the sheer number of images. Adversarial images take advantage of a feature found in many AI systems known as “[decision boundaries](https://www.theverge.com/2017/4/12/15271874/ai-adversarial-images-fooling-attacks-artificial-intelligence).”

These boundaries are the invisible rules that instruct a system whether it is perceiving a lion or a leopard. The objective would be to create a mental map with lions in one sector and leopards in another. The line dividing these two sectors—the border at which a lion becomes a leopard or leopard a lion—is known as the decision boundary. Jeff Clune, who has also studied adversarial training, remains dubious about such classification systems because they’re too [arbitrary](https://www.theverge.com/2017/4/12/15271874/ai-adversarial-images-fooling-attacks-artificial-intelligence).“All you’re doing with these networks is training them to draw lines between clusters of data rather than deeply modeling what it is to be [a] leopard or a lion.”

Large datasets are [often labeled by companies](https://www.mturk.com/) that [employ manual methods.](https://cloud.google.com/ai-platform/data-labeling/docs) Obtaining and sharing datasets is a challenge, especially for an organization that prefers to classify data and restrict access to it. A military dataset may contain images produced by thermal-imaging systems, for instance, but unless this dataset is shared with developers, an AI weapon wouldn’t be as effective. For example, AI devices that rely on chatbots limited to hundreds of words might not be able to completely [replace a human](http://arxiv.org/abs/1706.03762) with a much larger vocabulary.

AI systems are also hampered by their inability to [multitask](https://mwi.westpoint.edu/artificial-intelligence-future-warfare-just-not-way-think/). A human can identify an enemy vehicle, decide on a weapon system to employ against it, predict its path, and then engage the target. An AI system can’t duplicate these steps. At this point, a system trained to identify a T-90 tank most likely would be unable to identify a Chinese Type 99 tank, despite the fact that they are both tanks and both tasks require image recognition. Many researchers are trying to solve this problem by [working to enable systems](https://doi.org/10.1109/ICDSP.2018.8631696) to [transfer their learning,](https://doi.org/10.1109/TMM.2018.2869276) but such systems are years away from production.

Predictably, adversaries will try to take advantage of these weaknesses by fooling image recognition engines and sensors. They may also try mounting cyberattacks to evade intrusion detection systems or feed altered data to AI systems that will supply them with false requirements.

**U.S. Preparedness**

The U.S. Department of Defense has been more partial to contracting for and [building hardware](https://mwi.westpoint.edu/artificial-intelligence-future-warfare-just-not-way-think/) than to implementing new technologies. All the same, the Air Force, in cooperation with Boeing, [General Atomics, and a company called Kratos](https://www.twz.com/air/kratos-has-a-hypersonic-drone-in-the-works), is developing AI-powered drones. The Air Force is also testing pilotless XQ-58A Valkyrie experimental aircraft [run by artificial intelligence](https://www.eglin.af.mil/News/Article-Display/Article/3480769/ai-successfully-pilots-xq-58-aircraft-at-eglin/). This next-generation drone is a prototype for what the Air Force hopes can become a potent supplement to its fleet of traditional fighter jets. The objective is to give human pilots a swarm of highly capable robot wingmen to deploy in battle. The Valkyrie is not autonomous, however. Although it will use AI and sensors to identify and evaluate enemy threats, it will still be up to pilots to decide whether or not to strike the target.

Pentagon officials may not be deploying autonomous weapons in battle yet, but they are [testing and perfecting weapons](https://www.thenation.com/article/world/artificial-intelligence-us-military/) that will not rely on human intervention. One example is the Army’s [Project Convergence.](https://crsreports.congress.gov/product/pdf/IF/IF11654/6) In a test, conducted as part of the project, held in August 2020 at the Yuma Proving Ground in Arizona, the Army used a variety of air- and ground-based sensors to track simulated enemy forces and then process that data using AI-enabled computers at a base in Washington state. Those computers, in turn, issued fire instructions to ground-based artillery at Yuma. “This entire sequence was supposedly accomplished within 20 seconds,” the Congressional Research Service later [reported](https://crsreports.congress.gov/product/pdf/IF/IF11654/6).

In a [U.S. program](https://www.diu.mil/replicator) known as the Replicator initiative, the Pentagon said it planned to mass-produce thousands of autonomous drones. However, no official policy has condoned the use of autonomous weapons, which would allow devices to decide whether to strike a target without a human's approval.

The Navy has an AI equivalent of Project Convergence called “Project Overmatch.” In the words of Adm. Michael Gilday, chief of naval operations, this is [intended](https://sgp.fas.org/crs/natsec/R46725.pdf) “to enable a Navy that swarms the sea, delivering synchronized lethal and nonlethal effects from near-and-far, every axis, and every domain.” Very little has been revealed about the project.

About 7,000 analysts employed by the National Security Agency (NSA) are trying to integrate AI into its operations, according to General Timothy Haugh, who serves as the NSA Director, U.S. Cyber Command Commander, and Chief of the Central Security Service. General Haugh has [disclosed](https://www.eurasiantimes.com/raven-sentry-how-us-military-weaponized/) that as of 2024, the NSA is engaged in 170 AI projects, of which 10 are considered critical to national security. “Those other 160, we want to create opportunities for people to experiment, leverage, and compliantly use,” he says.

At present, though, AI is still regarded as a supplement to conventional platforms. AI is also envisioned as playing [four additional roles](https://mwi.westpoint.edu/artificial-intelligence-future-warfare-just-not-way-think/): automating planning and strategy; fusing and interpreting signals more efficiently than humans or conventional systems can do; aiding space-based systems, mainly by collecting and synthesizing information to counter hypersonics; and enabling next-generation cyber and information warfare capabilities.

**Ethics of AI Use**

Although the use of autonomous weapons has been a subject of debate for decades, few observers expect any international deal to establish new regulations, especially as the U.S., China, Israel, Russia, and others race to develop even more advanced weapons. “The geopolitics makes it impossible,” [says](https://www.thenation.com/article/world/artificial-intelligence-us-military/) Alexander Kmentt, Austria’s top negotiator on autonomous weapons at the UN. “These weapons will be used, and they’ll be used in the military arsenal of pretty much everybody.”

Despite such challenges, [Human Rights Watch](https://www.hrw.org/news/2024/05/06/submission-united-nations-secretary-general-autonomous-weapons-systems) has called for “the urgent negotiation and adoption of a legally binding instrument to prohibit and regulate autonomous weapons systems.” It has launched the [Campaign to Stop Killer Robots](https://www.stopkillerrobots.org/), which the human rights organization says has been joined by more than 270 groups and 70 countries. Even though the controversy has centered around autonomous weapons, Brian Schimpf, CEO of AI drone manufacturer Anduril, has another [perspective](https://www.knkx.org/2024-07-09/he-created-oculus-headsets-as-a-teenager-now-he-makes-ai-weapons-for-ukraine). He says AI weapons are “not about taking humans out of the loop. I don’t think that’s the right ethical framework. This is really about how we make human decision-makers more effective and more accountable [for] their decisions.”

All the same, autonomous AI weapons are already under development. Aside from the ethics of relying on a weapon to make life-and-death decisions, there is a problem with AI itself. Errors and miscalculations are relatively common. Algorithms underlying the operations of AI systems are capable of making mistakes—“[hallucinations](https://en.wikipedia.org/wiki/Hallucination_%28artificial_intelligence%29)”—in which seemingly reasonable results turn out to be entirely illusory. That could have profound implications for deploying AI weapons that operate with deeply flawed instructions undetectable by human operators. In a particularly dystopian scenario, an adversary might substitute robot generals for human ones, forcing the U.S. to do the same, with the result that AI systems may be pitted against one another on the battlefield with unpredictable and possibly catastrophic consequences.

Dr. [Elke Schwarz](https://www.qmul.ac.uk/research/featured-research/the-ethical-implications-of-ai-in-warfare/) of Queen Mary University of London views the AI weapon dilemma through a theoretical framework that relies on political science and empirical investigations in her consideration of the ethical dimensions of AI in warfare. She believes that the integration of AI-enabled weapon systems facilitates the objectification of human targets, leading to heightened tolerance for collateral damage. In her view, automation can “weaken moral agency among operators of AI-enabled targeting systems, diminishing their capacity for ethical decision-making.” The bias towards autonomous systems may also encourage the defense industry to rush headlong into funding military AI systems, “influencing perceptions of responsible AI use in warfare.” She urges policymakers to take risks into account before it’s too late.

“(T)he effect of AI is much, much more than the machine gun or plane. It is more like the shift from muscle power to machine power in the last Industrial Revolution,” [says](https://www.asahi.com/ajw/articles/15209132) Peter Singer, a professor at Arizona State University and a strategist and senior fellow at the U.S. think tank New America, who has written extensively about AI and warfare. “I believe that the advent of AI on the software side and its application into robotics on the hardware side is the equivalent of the industrial revolution when we saw mechanization.” This transformation raises new questions “of right and wrong that we weren’t wrestling with before.” He advocates setting “frameworks to govern the use of AI in warfare” that should apply to those people who are working on the design and use.

One of the issues Singer calls [“machine permissibility”](https://www.asahi.com/ajw/articles/15209132) is what the machine should be allowed to do apart from human control. He calls attention to a second issue “that we’ve never dealt with before,” which is “machine accountability.” “If something happens, who do we hold responsible if it is the machine that takes the action? It’s very easy to understand that with a regular car, it’s harder to understand that with a so-called driverless car.” On the battlefield, would the machine be held responsible if the target was mistaken or if civilians were killed as a result?