Position: AI-Powered Autonomous Weapons Risk Geopolitical Instability and Threaten AI Research

Riley Simmons-Edler [*1 2] **Ryan P. Badman** [*1 2] **Shayne Longpre** [3] **Kanaka Rajan** [1 2]

Abstract

The recent embrace of machine learning (ML) in the development of autonomous weapons systems (AWS) creates serious risks to geopolitical stability and the free exchange of ideas in AI research. This topic has received comparatively little attention of late compared to risks stemming from superintelligent artificial general intelligence (AGI), but requires fewer assumptions about the course of technological development and is thus a nearer-future issue. ML is already enabling the substitution of AWS for human soldiers in many battlefield roles, reducing the upfront human cost, and thus political cost, of waging offensive war. In the case of peer adversaries, this increases the likelihood of "low intensity" conflicts which risk escalation to broader warfare. In the case of non-peer adversaries, it reduces the domestic blowback to wars of aggression. This effect can occur regardless of other ethical issues around the use of military AI such as the risk of civilian casualties, and does not require any superhuman AI capabilities. Further, the military value of AWS raises the specter of an AI-powered arms race and the misguided imposition of national security restrictions on AI research. Our goal in this paper is to raise awareness among the public and ML researchers on the near-future risks posed by full or near-full autonomy in military technology, and we provide regulatory suggestions to mitigate these risks. We call upon AI policy experts and the defense AI community in particular to embrace transparency and caution in their development and deployment of AWS to avoid the negative effects on global stability and AI research that we highlight here.

[*]Equal contribution [1]Department of Neurobiology, Harvard Medical School, Boston, USA [2]Kempner Institute, Harvard University, Cambridge, USA [3]Massachusetts Institute of Technology, Cambridge, USA. Correspondence to: Riley Simmons-Edler <riley_simmons-edler@hms.harvard.edu>.

Proceedings of the 41st International Conference on Machine Learning, Vienna, Austria. PMLR 235, 2024. Copyright 2024 by the author(s).

1. Introduction

With the rise of powerful generative AI models such as GPT-4 and Stable Diffusion, and continued progress in fields such as robotics and reinforcement learning, concerns have grown among both experts and the public about giving AI too much power. Academic concerns have thus far centered on threats in domains such as cybersecurity, biological weapons, disinformation, fraud, and hypothetical rogue artificial general intelligence (AGI) (Future of Life Institute, 2023; Bengio, 2023). Despite these general concerns, there has been relatively little attention given to specific recent developments from military and defense-industry groups, which have already begun to deploy next-generation AI-guided Autonomous Weapon Systems (AWS). Weapons falling under the AWS label have traditionally been either remotely operated (but AI-assisted) or only autonomous within a very narrow scope. However, new fully autonomous (unmanned) AWS in development remove or reduce the human role in the control and decision making process, with the goal of removing humans from the active battlefield en mass. For example, the Pentagon's Replicator program for AI-based weapon "swarms" promises a drastic shift in warfare towards highly autonomous and cooperative AI units within the next few years (Greenwalt, 2023). AWS can serve many battlefield roles, although human-targeting lethal AWS (LAWS) have received most public attention (Taddeo & Blanchard, 2022). Many of these new advanced AI and machine learning (ML) weapons systems are already seeing real-world deployment for the first time in the Ukraine War, and are being designed for every branch of the military and by many nations (Sharma, 2022; Greenwalt, 2023; Pfaff et al., 2023).

Current discussion of ethical use of AWS centers on keeping humans in-the-loop for remote autonomous warfare so that only humans are making lethal decisions, while accepting the idea that remote warfare saves lives (Riesen, 2022; Pfaff et al., 2023). However, we argue that there are fundamental issues caused by removing humans from the battlefield. Human "boots-on-the-ground" can signify a commitment to following the rules of war, improve humanitarian aspects of occupation, and most importantly maintain a human cost to war for aggressor nations that prevents a state of endless

war from being politically feasible (Trapp, 2018). We are concerned the recent embrace of AWS by global militaries is leading to a future where wars are more frequent, with such warfare having negative consequences for global stability even if AWS reduce civilian casualties relative to human soldiers. This new model of AWS-centered warfare will be supported by an increasing fusion of civilian and military AI research that will have devastating effects on research and trust in our field. Thus, we have written this article to make ML researchers aware of the rapid pace of AWS development and the need for action, and to provide community and policy suggestions to help prevent this possible future.

The ethical implications of such AWS capable of selecting and destroying targets without meaningful human control are considerable, and have been discussed by prior work at length (Scharre, 2018; Blanchard & Taddeo, 2022; Brown, 2023). Over the past 2-3 decades many voices have called for these systems to be banned entirely and/or governed by international treaties (Russell, 2023; Bode et al., 2023). We agree that the ideal would be that such systems were not developed, due to moral hazards in their use as well as the issues we highlight here. However, in this paper we focus on the scenario in which a middle ground is achieved: AWS are not banned despite moral reasons they ought to be (Geist, 2016), but manage to comply with basic military ethics and do not increase civilian casualties compared to conventional war (Scharre, 2018; Brose, 2020; Riesen, 2022). This assumption that AWS will not increase civilian casualties is an optimistic one —likely many AWS will fall short of that goal— but it provides the most conservative starting point for our discussion on the indirect negative impacts of AWS. In 2024, major military powers have already committed to making autonomous weapon systems core components of their armed forces, and thus the time to discuss full AWS bans seems increasingly in the past (Greenwalt, 2023; Klare, 2023; Brose, 2020). The conversation now needs to center on responsible development and proliferation of the rapidly improving AWS systems that are already being developed and used by global militaries.

Official statements and announced R&D projects make it clear that the direction of AWS development efforts both short- and long-term is the removal of human soldiers from direct combat roles, to reduce casualties and increase combat effectiveness (Greenwalt, 2023; Pfaff et al., 2023; Bachmann & Grant, 2023). While these goals are reasonable in isolation, a lack of public attention and transparency around the rapid and increasing pace of AWS development and employment risks humanity sleepwalking into an AWS arms race between global powers. We will see these risks in the near future. China and Russia have given 2028-2030 as targets for major automatization of their militaries to begin, and while the USA is set to begin deployment sooner, the country is planning for a long-term transformation (Sharma, 2022; Greenwalt, 2023; Kania, 2020; Warren & Hillas, 2022).

Given these priors, we argue that because highly capable AWS lower the human costs associated with conflict initiation and escalation, they also create a large risk to global geopolitical stability. Critically, this effect worsens the more capable AWS become, even if civilian casualties and collateral damage decrease, and cannot be solved by simply improving the ML systems involved– deliberate policy actions are needed.

In conflicts between powers with large disparities in military strength, invasion or intervention using an AWS-heavy force is politically easier than with an all-human force, since there will be fewer deaths on the aggressor side (Moreau, 2011; Greenwalt, 2023). However, reducing human casualties in a single conflict can be outweighed if the number of conflicts that occur increases. Indeed, the past century suggests that when dominance in a new military technology leads to regional or global hegemony, this does not always translate to greater stability, and can actually increase lower intensity conflicts and terrorism (Benvenisti, 2004; Drezner, 2013; Hegre, 2014). AWS-heavy armies with minimal human battlefield presence may lead to a rise in terrorism, assassination, attacks on civilians, and other methods extending beyond the traditional battlefield (Kwik, 2022). These abhorrent methods provide a way for less powerful nations who lack AWS to deter or retaliate against nations deploying AWS-heavy forces despite their inability to do so through battlefield casualties (Moreau, 2011).

In addition, AWS that function as a substitute for "boots on the ground" are a threat to internal stability and representative government, unless careful safeguards are in place to protect against tyrannical use or sabotage of such systems (Ams, 2023). Further, reduced human battlefield presence makes journalistic transparency and civilian oversight of conflict more difficult, and makes it easier to hide war crimes and the impacts of war from both civilian leadership and the public (Lin-Greenberg, 2020; Ams, 2023).

Beyond the impacts AWS have on global stability, the importance of AWS for warfare will likely lead to major negative impacts on civilian AI research. AWS will become a revolutionary military technology, as did nuclear weapons, mechanized warfare, and others historically, but with important differences in the ease of AWS proliferation and impact on civilian technology development (Stern, 2023; Johnson, 2020; Laird, 2020; Wong et al., 2020). Recent work by Schneider argues that the prevailing military logic—that hegemonic power creates stability—can be highly erroneous when that power is based on technology that relies on a scarce or controlled resource, the capability/vulnerability paradox (Schneider, 2019). In the case of AWS, these key resources are AI experts and knowledge, access to relevant big data, and semiconductor manufacturing—all of which

TOPIC	CHALLENGES	RECOMMENDATIONS
Major power aggression	AWS replacement of human soldiers makes war more domestically palatable. Defenders turn to asymmetric warfare/terrorism for deterrence.	Require significant human battlefield presence, focus on human-AWS teaming over remote and fully autonomous AWS-centered conflict.
Peer conflict escalation	Heavy use of AWS makes conflict initiation easier between major powers, risks AWS arms race.	International transparency about broad capabilities, deployment disclosures for AWS systems.
Conflict transparency	Lack of human presence means accountability in war is harder, war crimes and battlefield underperformance less visible to leaders and the public.	Require detailed public reports on AWS capabilities, deployments, and outcomes. Embed oversight/watchdogs in AWS command centers.
Proliferation	Development and sale of AWS will be widespread, global availability of AWS is inevitable.	Avoid futile AI hardware/software restrictions.
AI researcher restrictions	Military AI needs lead to censorship of civilian research, reduced international collaboration, monitoring and restriction of researchers.	Universities, corporations, governments, etc. establish norms on how much military and civilian research should overlap.
Dual-use AI technology	Many AI algorithms are innately dual-use. Facial recognition, navigation, robotics, etc. Military interest in civilian research is likely to grow.	Improve university ethics oversight and transparency for military-funded AI research, and caution researchers against efforts to weaponize AI.
Risk of over-regulating AI	Public backlash to AWS leads to calls for more limitations on AI research generally, hurting international research community and academic norms.	Avoid restricting basic AI research, regulate explicit AWS research and military-related datasets over general civilian hardware and AI models.
Autonomy transparency	Public data on current AWS are often vague on autonomy levels, definitions of human-in-the-loop, may be more autonomous in practice.	Require governments and AWS manufacturers to clarify the degree of autonomy of AWS. Set international standards for allowed levels of autonomy.

Table 1. Overview of AWS issues raised and policy recommendations in this work.

have historically been dominated by a handful of countries (Chu, 2023; Daniels, 2019; Khan et al., 2021; Farrell & Newman, 2022; Chahal et al., 2020). While these resources are becoming more available globally, governments with a current advantage are likely to try to secure that advantage by preventing their spread. We may see a rise in export restrictions, publication oversight and redaction, and knowledge compartmentalization in the field of AI as a result (Bengio, 2023; Rickli & Ienca, 2021; Warren & Hillas, 2020).

Longer term, we could see AI resources treated analogously to how oil resources have fueled conflict since the advent of industrialized warfare (Schneider, 2019). Indeed, we see hints in current events that this misguided idea is spreading (Templeman & Mastro, 2023; Klingler-Vidra, 2023; Triolo, 2023), and some militaries are already seeking to use civilian AI experts to speed development of domestic AWS (Pfaff et al., 2023). China for example has declared Military-Civil Fusion as foundational to updating its military with AWS capabilities, and major Chinese universities have been roped into AI weapons programs (Konaev et al., 2023; Can & Vieira, 2022; Warren & Hillas, 2021; Fedasiuk, 2021).

To reduce and prevent these outcomes, action by AI researchers, policymakers, and the public will be needed. Below we propose several policies and actions to take in Section 4 and Table 1, and briefly review highlights here. First, the best way to avoid AWS-driven conflicts is for nations to avoid developing or fielding AWS that operate independently rather than in cooperation with humans, especially in large scale or prolonged conflicts. Second, the international community needs to build consensus around AWS autonomy levels, and which levels are acceptable or unacceptable and in what circumstances. Third, there is an urgent need for transparency around AWS capabilities and levels of human oversight in actual use—"Human-in-the-loop" is becoming a meaningless phrase that arms manufacturers and militaries use to reassure the public. Today, "human-in-the-loop" is shifting from implying hands-on remote control to automated targeting where a human presses a "go" button with minimal insight into what the AWS is doing or why (Bode & Watts, 2023; Bode et al., 2023; Boys, 2023; Abraham, 2024). Fourth, universities and companies worldwide must draw lines and declare publicly what level of military-civilian overlap and collaboration on AWS-related research is acceptable (D'Agostino, 2024).

The structure of this article is as follows: Section 2 discusses the state of AWS development and deployment today, as well as near-future directions. Section 3 discusses why limiting AWS proliferation is difficult, and the ineffectiveness of restricting access to ML hardware and expertise to prevent it. Section 4 discusses several approaches to mitigate the risk posed by advanced AWS short of a total ban.

2. Current State of AWS

The past century has seen the development of mechanized, remote-controlled, and semi-autonomous weapons of war as extensively covered in other work (Scharre, 2018; Work, 2021). However, the past 5-10 years have seen unprecedented progress in and acceptance of truly autonomous weapons systems that are increasingly able to hunt and attack targets with minimal or no human oversight. We review the current state of these truly autonomous (i.e. unmanned, uncrewed) weapons systems for land, air and sea in this section to provide context on the likely trajectory AWS development will take. Note that we omit discussion of outer space-based AWS as there is currently too little information publicly available (van Esch et al., 2017). In addition, to our knowledge no work has comprehensively reviewed AWS types across combat domains in an up-to-date way. Current major reviews have focused on AWS progress for only one branch of the military, and/or do not include the most recent developments in this fast-moving field (Sharma, 2022; Castillo, 2022; Longpre et al., 2022; Bode & Watts, 2023).

Note that we do not discuss the impacts of AI on cyber warfare in this work, as it is a complex topic worthy of further study that has distinct impacts compared to directly lethal AWS. Nonetheless it is a rapidly developing field of warfare, with both China and Russia proclaiming it a major topic of military development (Fedasiuk, 2021; Thornton & Miron, 2020; Murdick et al., 2021; Takagi, 2022).

2.1. Airborne AWS

One of the most established and battle-tested AWS types is loitering munitions (LMs) (Bode & Watts, 2023). Loitering munitions include a wide range of designs, but are single-use unmanned aerial combat units which can hover or circle, perform detection and targeting, and then autonomously launch (crash) themselves into targets with lethal payloads. Typically LMs do not require human control or oversight after launch/takeoff, and prosecute their pre-programmed missions autonomously. The main operational LM unit up until approximately 2015 was the Israeli Harpy (Scharre, 2018; Scharre & Horowitz, 2015), which can autonomously detect and destroy radar targets in an area whose precise locations were not known before launch, and can return to base and be re-used if no such targets are found.

Since then, the number of LMs has rapidly expanded, with at least 24 models of varied size and sophistication across 16 countries, and the capabilities of the AI systems used by LMs have increased considerably (Bode & Watts, 2023). Most concerning, the range of targets for LMs has expanded from radar systems to armored vehicles and even specific human personnel that can be identified and targeted by facial recognition without operator involvement. The Kargu-2 developed by Turkey is one of the foremost systems with this capability, and was deployed on battlefields in Libya (Longpre et al., 2022; Bode & Watts, 2023). Other notable models such as the Rafael Spike Firefly and the IAI Rotem L are optimized for targeting in urban environments, and at least 14 other models have specifications suggesting they can autonomously target and follow or execute individual human targets (Bode & Watts, 2023). Both Ukraine and Russia have deployed LMs of various levels of sophistication and autonomy in their current war in large numbers (Bode & Watts, 2023; Cotovio et al., 2024). Notably, the former commander in chief of the Ukrainian military recently highlighted LMs and anti-LM defenses such as jamming as critical to their war effort (Zaluzhnyi, 2023). He also highlighted the Russian Lancet LM, which uses a western-built Nvidia Jetson TX2 ML compute module to autonomously track targets, as being particularly difficult to defend against (Zaluzhnyi, 2023; Faragasso, 2023). A major issue in assessing LMs is that governments and manufacturers are often vague about the capabilities of LM AI-targeting systems, and whether autonomous modes without human oversight have been used operationally (Bode & Watts, 2023; Hambling, 2023a; Sayler, 2021; Longpre et al., 2022).

The other well-established form of AWS is the (reusable) unmanned aerial vehicle (UAV). A major recent announcement in AWS-UAV development is the Pentagon's Replicator Program, which aims to deploy thousands or tens of thousands of UAVs within 18-24 months, across land, air, and sea (Klare, 2023; Greenwalt, 2023). The prime candidates for these programs include budget fighter and bomber drones that can replace traditional crewed aircraft but with a 10-100x lower price tag.

One example system is the XQ-58A Valkyrie developed by Kratos Defense and Security Solutions to be a "loyal wingman" to manned aircraft, defending them and performing offensive actions that would otherwise risk the manned aircraft, first demonstrated in 2019 (Wang et al., 2020; Klare, 2020; 2023; Newdick & Rogoway, 2023). Notably, Kratos has claimed it can produce several hundred Valkyrie units per year at per-unit costs between $2-5 million, cheaper than manned aircraft, most traditional drones like the American Reaper UAV, and even some missiles (Jordan, 2021; Lyu & Zhan, 2022; Boys, 2023). These units can scout, are equipped with weapons for defensive fire, have an operational range of 3000 miles, and can even deploy smaller unmanned aircraft for strike or low-level reconnaissance (Newdick & Rogoway, 2023). Other major candidates are even cheaper, and are intended to be deployed in large numbers using autonomous swarming capabilities, such as Shield AI's V-Bat and Boeing's MQ-28 (Gruszczak & Kaempf, 2023; Saylam & Gündoğdu, 2023; Bachmann & Grant, 2023; Sud, 2020; Castillo, 2022). In addition, unmanned versions of the F-16 have been demonstrated recently, with deployment estimated for 2025 (DARPA, 2023;

Brown, 2023; Kohler, 2024). US allies are developing similar systems, such as the Blue Bear Ghost line of autonomous UAVs (BlueBear, 2018).

China is also making considerable progress in improving the autonomous capabilities of UAVs in the People's Liberation Army-Air Force (PLAAF). The Wing Loong line of unmanned aerial vehicles (UAVs) was declared to be a focus for autonomous upgrades to minimize the need for a remote operator, using 5G, AI, and big data (Liu & Fan, 2023). Looking forward, China recently announced that a significant degree of autonomy (usually translated as "intelligence") must be introduced to the military over the coming decades (Kania, 2020; Warren & Hillas, 2022).

2.2. Ground AWS

The USA and other NATO nations have driven most recent UGV developments of note (Sharma, 2022). In 2017, the US Army declared its mid-term (2021–30) priorities include achieving the ability to perform fully automated convoy operations using unmanned combat and support vehicles that can navigate difficult terrain and carry advanced payloads (U.S. Army Training and Doctrine Command, 2017). The most notable recent unmanned ground vehicle (UGV), already deployed on the Ukrainian battlefield (Milrem Robotics, 2023a), is the Milrem THeMIS unit, which is capable of autonomous targeting and ordinance deployment as well as other support and sensing tasks (Mathiassen et al., 2021; Milrem Robotics, 2023b). Milrem has sought to make the unit more autonomous for faster threat response and to enable coordination in autonomous swarms.

Several other UGV AWS are in late development stages across the US, Europe and India, with major demonstrations or deployments planned (Sharma, 2022). For example, the MAARS Grey Muzzle autonomous weaponized robot developed by QinetiQ that is equipped with machine guns and grenade launchers for direct fire support roles. This system features an onboard AI system for autonomous navigation and waypoint following, but the weapons require human input before firing (QinetiQ, 2023). Legged quadruped UGVs also have shown great promise for use in more complex terrain and environments (Hambling, 2023b; Gilead, 2023). Lastly, Pegasus Intelligence is developing a portfolio of diverse AWS, some of which are amphibious or can transition between flight and ground modes (Pegasus Intelligence, 2023). Looking forward, recent progress in subfields of ML and robotics, such as autonomous vehicle research, has greatly improved UGV capabilities and spurred additional development, with several recent DARPA programs testing self-driving abilities in military technology (Bisht, 2022).

2.3. Naval AWS

Submarine AWS development has been a high priority for navies, as the ocean is an inherent signal jammer, making remote operation difficult or impossible—indeed, modern naval mines and torpedoes can be seen as very simple AWS. One of the most recent notable developments in naval AWS was the 2023 Israel Aerospace Industries (IAI) announcement that its BlueWhale anti-submarine warfare (ASW) uncrewed submarine has participated in NATO exercises (Israel Defense, 2023). BlueWhale ASW can operate for 2-4 weeks with satellite connection and electro-optical sensors for sea targeting, with low operational cost.

Comparable efforts are being made by the US Navy. An array of medium, large and extra-large sized unmanned naval vehicles are planned for undersea and surface operations (O'Rourke, 2023). The largest of these is the Orca XLUUV, with a claimed dry weight in excess of 50 tons. It specializes in Hammerhead mine deployment on the seabed but is reconfigurable depending on needs for combat, surveillance, intelligence, etc. and is capable of carrying multiple types of payloads (O'Rourke, 2023). L3Harris is also developing a line of autonomous surface vessels up to 13 meters long which require no crew for operation and only minimal remote operation for payload delivery (L3Harris, 2023; U.S. Naval Forces Central Command, 2023). These new surface vessels are likely much more capable and autonomous than the smaller semi-autonomous surface vessels which have seen notable use in the Ukraine War (Ozberk, 2023; Santora & Triebert, 2023). For detailed historical progress on autonomous units in the US Navy, see (Castillo, 2022).

China has greatly advanced its claimed unmanned surface and undersea capabilities the past 5 years. The country has announced plans for the world's first AI-run deep sea underwater colony to house and oversee submarine operations for defense and survey in the South China Sea (Chen, 2018). By 2021, an estimated 48 universities and 45 enterprises in China were involved in unmanned (UUV) and autonomous undersea vehicle (AUV) R&D projects (Fedasiuk, 2021). Prototype units include large and small submarines and surface vessels, and military project descriptions mention deep learning-based targeting and sensor analysis, as well as mine deployment (Fedasiuk, 2021; Goldstein, 2019). Disentangling civilian versus military models is increasingly difficult, as the Tianjin Municipal Government for example promotes dozens of AUV and UUV models as being capable of both civilian and military use (Fedasiuk, 2021).

2.4. AWS Systems for Command

Data and intelligence sharing and synthesis are obvious targets for AI enhancement and AWS integration. A number of efforts to build such systems are ongoing (Bihl & Talbert, 2020). Historically these functions have been fragmented

across military branches (Pfaff et al., 2023), but recent initiatives in the US have sought to unify military AI and big data platforms (Allen, 2023). One example system is the Joint All-Domain Command and Control program which attempts to automate and speed up the processing of a large variety of sensor, imaging, and intelligence data across branches of the military (Hoehn, 2022).

A concerning development in this domain recently is Palantir's 2023 demonstration of its Artificial Intelligence Platform (AIP) for Defense, which uses large language models (LLMs) to recommend military command decisions (Mikhailov, 2023; Michel, 2023; Reynolds & Cetin, 2023). Described capabilities include recommendations by the chatbot for deploying specific types of missiles, long range artillery or aerial units in various contexts. The system's performance reportedly deteriorated over time, and intervention was needed to restore accuracy and logic in a demo (Mikhailov, 2023). Scale AI has also recently announced a partnership with the Center for Strategic and International Studies to create LLMs that can perform wargaming, diplomacy, military command advising, cybersecurity analysis, and even help respond to misinformation campaigns (Scale AI, 2024). Finally, while no projects have been announced yet, OpenAI recently lifted their ban on using their products for military purposes in early 2024 (Biddle, 2024).

A major concern in this domain is that as the complexity of AWS systems and geospatial/intelligence data increase, commanders will rely on annotation, analysis and suggestions from AI systems whose output cannot easily be double checked (Pfaff et al., 2023). This concern will likely limit proliferation of advanced AI-based command systems in the short term, but as capabilities improve it may be unwisely disregarded. A recent study has also voiced concern that increased use of AI at the command-level could risk (and may already be) lowering the standards for good tactics and strategy due to how narrow and specialized LLMs inherently are relative to the human brain (Hunter & Bowen, 2022). Supporting the validity of this concern, a 2024 study of military-related LLMs found that LLMs were prone to recommending pro-escalation tactics with unclear motivation and logic, including escalations that provoked arms races and called for nuclear weapons deployment (Rivera et al., 2024). In the same vein, concern has recently been expressed about a lack of human oversight in the Israeli "Lavender" AI system used for classification of civilians versus militants for military targeting purposes in operations in the Gaza Strip (Abraham, 2024). Similar to humans, AI models are not perfect at complex decision making, and more transparency is generally needed in how such systems are used to advise military commanders, particularly regarding what factors are used to set statistical thresholds for lethal decisions.

3. AWS Proliferation and Threats to Academic Research

Presently, most forms of AWS are deployed by few countries, and only in limited numbers. As such, their military significance is relatively low—AWS capabilities are not core to any nation's national defense at present. We anticipate this will change over time, possibly within 5-10 years, with AWS taking a more central role in how war is waged. Further, we expect AWS to proliferate widely, and perform key roles in many armies around the world (Scharre, 2018). If AWS are core to national defense, this means the technologies that power AWS capabilities, machine learning and robotics, are critical defense technologies. This means that developments in those fields have major military implications, which raises the prospect of national security restrictions on AI research and researchers worldwide. In this section, we argue why advanced AWS, once developed, will proliferate widely, and why policies that aim to regulate AI research to prevent AWS proliferation will be both futile and deleterious to the field.

Firstly, regarding proliferation, the technologies needed (for example) to allow a tracked UGV such as the TheMIS to maneuver, detect, and fire on enemy troops with a usefully-high success rate already exist as an extension of methods developed for autonomous driving and computer vision applications (Mathiassen et al., 2021). Likewise, UAVs detecting targets and directing strikes against them without human intervention is already tractable- indeed, multiple loitering munitions advertise some degree of autonomous targeting functionality as described in Section 2. While non-trivial integration work is needed to implement such capabilities for any given AWS, the hardware and software engineering talent required is not scarce or restricted, as evidenced by the multitude of AI and ML products being developed by companies worldwide.

Secondly, ML software proliferates easily and widely and is cheap to develop relative to many military procurement programs (Warren & Hillas, 2020; Bengio, 2023). The key breakthroughs enabling advanced AWS now rather than in past decades come from the field of machine learning, where major advances in algorithms have enabled new robotic applications such as warehouse logistics and remote survey and monitoring UAVs and UGVs (Schedl et al., 2021; del Cerro et al., 2021). In the near future, academic and industry researchers aim to expand that frontier to include human-assistant robots capable of navigating complex indoor environments (Fan et al., 2020), dexterous object manipulation (Andrychowicz et al., 2020; Nagabandi et al., 2020), and even surgery (Saeidi et al., 2022), with much effort going into developing such systems. These are worthy applications, but many of these algorithms are inherently "dual-use" in that they can be applied to AWS development

by defense contractors (Hoijtink & Planqué-van Hardeveld, 2022; Gómez de Ágreda, 2020). While it is a good thing that most of the ML software ecosystem remains open source and freely available, this does mean that the core software needed for AWS development is also readily available.

This rapid proliferation of ML talent and open access ML code, while a net positive for humanity, has unfortunately led to recent calls by several top ML researchers for strict regulation, restriction, censorship, or requiring security clearances for many topics in ML (Bengio, 2023; Schaake, 2021; European Parliament News, 2023; Henshall, 2023; O'Brien, 2023). Given how vague the criteria for what needs to be regulated are, a paranoid response that censors ML research could have devastating effects on ML progress and global collaboration, and would not limit the proliferation of AWS. AWS effectiveness and proliferation will likely be limited mostly by the quality and scope of the military datasets available to train them, datasets which have different constraints and goals than civilian datasets (Chahal et al., 2020). Thus, protecting and securing trained AWS models and relevant datasets is likely a more feasible goal than security compartmentalization of AI research. We also note that the opportunity to prevent proliferation through locking down the field of ML has largely passed—ML expertise and codebases have spread worldwide, and retroactively clamping down on that knowledge would not be sufficient to prevent substantial AWS development from occurring.

One major form these misplaced efforts to restrict knowledge take is immigration and research visa restrictions on civilian scientists. Indeed, such visa bans have already been put in place for some Chinese researchers by the USA (Hawkins, 2024). Such bans are not just ineffective for the reasons given above, but are pointlessly cruel to the people who are targeted, and can backfire and increase proliferation rather than reduce it. As a historical example, Qian Xuesen was a Chinese-American rocket scientist who was unjustly expelled from the USA during the early Cold War when rocket science suddenly became a field of critical national security interest, and who as a result became the father of the Chinese rocketry program under coercion after his deportation from the USA (BBC News, 2020).

Beyond visa restrictions, attempts to stem the proliferation of potentially dangerous AI capabilities have so far focused on blocking hardware exports to select countries, most notably the US block on GPU exports to China. However, while economically disruptive (Triolo, 2023; Allen, 2022), blocking access to top-of-the-line GPUs and other compute capacity will not limit the proliferation of AWS. In China's example this is partially due to the involvement of domestic academic and tech infrastructure in AWS development, including domestic semiconductor fab technology (Chu, 2023; Warren & Hillas, 2022; Konaev et al., 2023).

However AWS development, like most robotics currently, does not require high end compute capacity. While much attention has accrued to LLMs such as ChatGPT which run on large clusters of expensive and scarce cutting-edge GPUs, most robotics and other edge ML research needs only modest hardware (Chen & Ran, 2019). These algorithms and the chips they run on are most relevant to AWS development, because both civilian and military robotics must reckon with limited volume, mass, power, and time budgets for on-robot compute. The chips suitable for these applications must be both small and low power, running models much smaller than state of the art LLMs, and will overlap heavily with those used for common civilian applications.

As a concrete example, Shield AI, who produce the cutting-edge V-Bat AWS platform capable of autonomous swarming behavior, lists the on-platform compute requirements for autonomy as follows: "AI Performance: 100 TOPS, GPU: 1024-core Nvidia Ampere architecture GPU + 32 Tensor Cores, CPU: 8-core Arm® Cortex®-A78AE v8.2 64-bit CPU 2MB L2 + 4MB L3" (Shield AI, 2023). These specifications correspond to Nvidia's Jetson Orin NX edge compute module, which debuted in 2023 on Samsung's 8nm process from 2018, and is representative of the current standard for embedded GPU compute. The performance of this module falls far below the most recent performance-based export restrictions on GPUs the US government has imposed, with the threshold for export control at 4800 TOPS, over 48 times more powerful than the Jetson Orin NX's 100 TOPS (Department of Commerce, 2023). By comparison, Chinese semiconductor manufacturer SMIC advertises a 7nm-class process, and multiple Chinese silicon design companies have GPU accelerator chips in production or announced, making the production of a Chinese domestic chip equivalent to the Jetson Orin NX likely in the near future.

As AWS technology proliferates and becomes increasingly critical to militaries worldwide, we will likely see nations make use of academic and corporate research talent to aid AWS development. The example of China's fusion of military and civilian infrastructure and institutions is unfortunately likely to become a common model for developing states interested in AWS weapons systems (Seraphin & Miles, 2023). Maintaining separate well-funded military and civilian research spheres to preserve responsible management of dual-use AI technology is morally admirable, but a significant luxury. Several major high ranking Chinese universities have already been blacklisted by the US for contributing to AI weapons development (Konaev et al., 2023), but it is likely universities will be co-opted into AWS development in many other countries over the next decade. At minimum, it is likely that government funding for academic AI research will increasingly concentrate on dual-use topics that can be applied to AWS development.

Similarly, it is likely that companies with significant civilian AI expertise and infrastructure will see pressure and incentives to contribute to defense contracts and develop technology that is of dual-use interest to the military (Reynolds & Cetin, 2023; Biddle, 2024). If the threat posed by AI and AWS continues to be equated in rhetoric to that of nuclear weapons, countries may start restricting ML researchers similar to how nuclear scientists have had their travel restricted and their research on some topics blocked from publication, despite the ineffectiveness of such policies (Daniels, 2019).

4. Policy Recommendations

4.1. Historical AWS Policy Efforts

To begin this section, we will survey a few major events in AWS regulation and policy over the past two decades. Awareness and concern about AWS began to increase after 2001, when the US Congress ordered the Pentagon to make one-third of their "deep strike" aircraft fleet unmanned by 2010, and one-third of their ground combat vehicles unmanned by 2015 (U.S. Congress, 2001). These timelines proved overly aggressive, but sparked a number of efforts to regulate AWS development—In 2010, the International Committee for Robot Arms Control called for regulations on remote-controlled and autonomous weapons (Markoff, 2010). Then, in 2013, the Campaign to Stop Killer Robots started as a coalition of NGOs attempting to push for global bans on lethal AWS (Stop Killer Robots, 2021). In 2014, the phrase "meaningful human control" (MHC) was coined by Article 36, which has since become a core principle of AWS ethics and regulation (Article 36, 2014).

Calls for regulation have not been limited to NGOs—A number of national and international efforts to regulate AWS have been seen in recent years. As of 2020, at least 30 countries have called for a ban on lethal AWS, with many others expressing some degree of concern about AWS (Stauffer, 2020). In the USA, the Block Nuclear Launch by Autonomous Artificial Intelligence Act was introduced in 2023 (Markey, 2023). In 2023 60 countries, including China and the USA, issued a call to action in the Hague towards more responsible use of AI for military applications (Sterling, 2023). Also in 2023, the UN passed a resolution calling for strict regulation on AWS, and for humans to stay in control of lethal decision-making in war (UN, 2023). Lastly, while not directly targeting AWS, the EU has moved to regulate AI risks, first in 2021 and more recently in 2024, with the EU AI Act passed as the first major regulation on artificial intelligence, though its implications for AWS are currently unclear (European Parliament News, 2023).

4.2. International Policy Recommendations

Here, we will we present several policy recommendations that can help to reduce the negative impacts of AWS, and exhort both ML and policy researchers to work to address these issues and to advocate for their implementation where possible. Although evidence suggests that the widespread use of AWS is inevitable, the manner in which this comes to pass will play a large role in determining the risks they pose to global stability and to academic freedoms.

The first policy recommendation we make is that nations can mitigate the negative effects of AWS on conflict risk by maintaining a ban on human-independent use of AWS. That is, AWS should not be deployed into combat without a meaningful human battlefield presence alongside them, whether using human-AWS teaming processes or as part of separate operations in the same battlespace. While the widespread deployment of AWS is likely inevitable, there is still time to influence where and how it is deployed, and human-independent deployment is not technically trivial.

There are two avenues to preventing human-independent use: AWS manufacturers should not develop systems designed to operate in an environment free of friendly humans, and policymakers should not field independent combat formations composed entirely of AWS. On the technical front, requiring strong human-in-the-loop measures goes a long way, as does limiting the use of generalist AI models in AWS. The AI research community should be wary of and resist attempts to use large pretrained models in AWS development—These models are likely more useful for building generalist AWS capable of human-independent operations compared to more specialized AWS, which can be developed using smaller and more specialized ML models. Further, AWS using specialist models will likely prove easier to develop to maturity and more easily fit into existing military command systems and doctrine, so this approach is also less risky from a military procurement perspective.

Our second policy recommendation is that policy experts develop consensus on levels of functional autonomy in AWS, and what constitutes a sufficiently autonomous AWS to require regulation (UNODA, 2023; Qiao-Franco & Bode, 2023). Clear capability and autonomy levels will help to avoid tensions and arms-race dynamics between competing nations as AWS become central to defense planning. Our ideal would be a full international ban on AWS that make a lethal decision without direct human involvement at the time of the weapon firing, a line that has already been crossed in combat in the Ukraine War (Hambling, 2023a) and by loitering munitions in multiple conflicts (Bode & Watts, 2023). Short of this ideal, policymakers need to lead military R&D by deciding where to draw the line before systems that cross it are developed and deployed in combat, which makes rolling back to lower autonomy levels much harder.

The third policy suggestion we make is for urgent improvements to transparency and oversight regarding planned and deployed AWS capabilities. In particular, we need clarity on what military tasks are acceptable to automate, and with what level of human supervision, both on paper and in the field. For example, there is (hopefully) wide consensus that nuclear arsenals should not be turned over to fully autonomous systems exempt from human oversight (Tamburrini, 2023; Matthew Robert, 2021). Unfortunately, given recent enthusiasm for autonomous weapons as well as recognition that contexts like signal jamming and underwater warfare inherently benefit from human-out-of-the-loop autonomy, anything less dangerous than autonomous nuclear weapons seems to be under consideration at the moment (Greenwalt, 2023; Pfaff et al., 2023; O'Rourke, 2023). All branches of the US military are already investing in command-augmenting or command-replacing AI systems (Hoehn, 2022; Mikhailov, 2023), but the aims and scope of these projects are vague at best when it comes to the degree of human removal from analysis and targeting. We support measures to mitigate the worst of these risks, and hope legislation can be pushed to create oversight processes for autonomy in AWS for conventional munitions.

In addition, we note the value of transparency in AWS deployment as battlefields become more and more dominated by machines. In such a scenario, there will by default be fewer humans present to report on war time events. Past AWS-led military excursions have been under-reported to the public and increasingly cross lines that prior political and military statements said would never be crossed (Hernandez, 2021; Bergman & Fassihi, 2021; Hambling, 2023a; Abraham, 2024). As AWS come to require less human involvement, covert AWS campaigns and secret AWS programs may be known to fewer military and intelligence personnel as well. This situation may lead to cover-ups of war crimes, as well as insufficient civilian assessment of new systems, leading decision makers to overestimate their capabilities (Pfaff et al., 2023). Providing access to independent watchdogs and journalists should be mandatory for advanced AWS use. In addition, a robust reporting process which maximizes public transparency should be implemented to assess the outcomes and effectiveness of these AI-based weapons, and to prevent the conduct of war from being further hidden from the public interest.

4.3. Academic Policy and Action Recommendations

Lastly, we present several actions for academics and academic institutions to take to address the issues we describe in this work, beginning with an overview of recent actions researchers have taken on AWS regulation. Most recent academic initiatives have centered on open letters or petitions seeking to regulate or ban AWS (Future of Life Institute, 2016; Amoroso & Tamburrini, 2020). For example, in 2017 a Clearpath Robotics co-founder published an open letter calling for the UN to ban AWS, with 116 founders of robotics and AI companies from 26 countries signing on (Bogdon, 2017). Similarly, in 2018 AI researchers from 30 countries wrote an open letter protesting a new department at KAIST that merged academic and military research (Walsh, 2018). In 2018, Google employees' protest led to the company canceling a Pentagon contract that could have led to more accurate drone strike technology (Crofts & van Rijswijk, 2020). Another Google employee protest has started recently, opposing their Project Nimbus (Perrigo, 2024). Furthermore, we note that global public opinion is favorable for AWS regulation currently. A 2020-2021 public opinion survey found that approximately 61 percent of adults across 28 countries oppose the use of fully autonomous weapons, with India being the only country surveyed with majority support for AWS (Ipsos, 2021).

Looking forward, AI research may increasingly result in dual-use hardware and software useful for either weapons systems or civilian applications, and expanding military funding for AWS risks more recruitment of AI/ML researchers and institutions onto AWS development projects. An immediate way to address this issue is to treat military funding similar to how industry funding is treated by many universities. Currently, many universities require substantial ethics training in and provide administrative oversight of industry funding sources. These measures seek to prevent bias in research or negative effects on the public's perceptions of the role of the university in society (Bickford, 2004; Robertson, 2011; Fabbri et al., 2018; Larrick et al., 2022; Bero, 2022). Military funding is not currently handled with the same degree of oversight and awareness (D'Agostino, 2024). This measure is simple to implement, and can prevent the most ethically questionable AWS research projects and defend academic independence.

We emphasize that this is not a call for universities to swear off military funding altogether. Defense-funded academic research that benefits military applications diffusely or is disconnected from combat technology is neither new nor a major threat to academic freedom. However, the rise of AWS will likely make academic research central to new AI projects for defense and national security that directly lead to more deadly new weapon systems on short timescales. Given this trend, military funding should be handled with the same awareness and caution that industry funding is currently handled with. Clear institutional rules around which military projects are acceptable and which endanger academic independence are needed (D'Agostino, 2024).

To conclude, unconstrained AWS development poses a range of risks to academics and nations worldwide, many of which have not been properly examined. It is urgent that we begin to reckon with these threats to our field and society.

Acknowledgements

We are grateful to Samantha Hubner for reviewing and providing feedback on drafts of this work, and for helpful discussion.

Impact Statement

This position paper is intimately concerned with the ethical and societal impacts of machine learning. Thus, in some sense the paper itself is an impact statement about global militaries' recent embrace of ML/AI-based autonomous weapons (AWS), as well as the ML community's growing advertent and inadvertent contributions to AWS advancements. Regarding the impacts of this paper itself, we hope that they will be both significant and positive, spurring awareness, constructive discussion, and ultimately action from the global ML community and beyond to mitigate the negative effects of AWS. That said, we do not expect one paper alone to do this. It is our hope that this work will encourage further analysis and debate on the impacts of AWS by ML and policy researchers worldwide as well as the public, ultimately leading to action by policymakers in many nations. We do not anticipate likely negative impacts from this work (such as encouraging ML researchers to become more, rather than less, active in advancing AWS technology), as AWS are already being recklessly developed and deployed. Furthermore in our opinion, ignoring or limiting awareness of active AWS research carries much greater risks to society than open and honest discussion, and as such it is appropriate and necessary for the field to discuss these issues in a public way.

Last, some may argue that AWS will become essential for national security, and thus that it is unethical to discourage nations from developing AWS because they are needed to maintain military deterrence. More broadly, one could ask under what extremes may highly-autonomous AWS development and production be viewed as a necessary evil, and discouraging it therefore unethical? As a hypothetical example, an aggressive world power may refuse to limit fully-autonomous, human-free AWS development, and over years or decades their technology could progress to the point where it is impossible to counter using conventional forces, human-controlled AWS, or nuclear deterrence. In this particular situation we concede that fully-autonomous AWS development with a deterrence mentality, mirroring modern nuclear deterrence and tenuous geopolitical equilibriums, may pragmatically become a necessity. However, as outlined in this paper, while AWS battlefield deployments have started in preliminary ways, AWS are currently not the core of any nation's military, and global public and institutional opinions are largely in favor of substantial regulation of some form. As such, this hypothetical is far from guaranteed to come to pass. However, we do not believe that the contributions of this paper are contingent on taking a position regarding the above hypothetical. Our work concerns the impacts of AWS on conflict risk and academic freedom, which will occur regardless of whether AWS are under strict human control or are highly autonomous, so long as they are critical to warfare and replace the human battlefield element. The goal of this work is to open discussion on the impacts of AWS among AI researchers and ground it in prior work, and we expect that the ethics of meaningful human control will be a major topic of future debate in the community, but our contribution does not rely on taking an ethical position on the matter.

References

Abraham, Y. 'Lavender': The AI machine directing Israel's bombing spree in Gaza. *+972 Magazine*, April 2024. Accessed: 2024-4-5.

Allen, G. C. Choking off China's access to the future of AI. *Center for Strategic and International Studies*, October 2022.

Allen, G. C. Six questions every DOD AI and autonomy program manager needs to be prepared to answer. *Center for Strategic and International Studies*, May 2023.

Amoroso, D. and Tamburrini, G. Autonomous weapons systems and meaningful human control: Ethical and legal issues. *Current Robotics Reports*, 1(4):187–194, December 2020.

Ams, S. Blurred lines: the convergence of military and civilian uses of AI & data use and its impact on liberal democracy. *International Politics*, 60(4):879–896, August 2023.

Andrychowicz, O. M., Baker, B., Chociej, M., Józefowicz, R., McGrew, B., Pachocki, J., Petron, A., Plappert, M., Powell, G., Ray, A., Schneider, J., Sidor, S., Tobin, J., Welinder, P., Weng, L., and Zaremba, W. Learning dexterous in-hand manipulation. *Int. J. Rob. Res.*, 39(1):3–20, January 2020.

Article 36. Key areas for debate on autonomous weapons systems. *Memorandum for delegates at the Convention on Certain Conventional Weapons (CCW) Meeting of Experts on Lethal Autonomous Weapons Systems (LAWS)*, 2014.

Bachmann, S.-D. D. and Grant, R. V. The need for an Australian regulatory code for the use of artificial intelligence (AI) in military application. *American University National Security Law Brief*, 13(2):2, 2023.

BBC News. Qian Xuesen: The man the US deported - who then helped China into space. *BBC*, October 2020.

Bengio, Y. My testimony in front of the U.S. Senate - the urgency to act against AI threats to democracy, society and national security. https://yoshuabengio.org/2023/07/25/my-testimony-in-front-of-the-us-senate/, July 2023. Accessed: 2023-12-7.

Benvenisti, E. The US and the use of force: Double-edged hegemony and the management of global emergencies. *Eur. J. Int. Law*, 15(4):677–700, September 2004.

Bergman, R. and Fassihi, F. The scientist and the A.I.-assisted, remote-control killing machine. *The New York Times*, September 2021.

Bero, L. Industry influence on research: A cycle of bias. *The Commercial Determinants of Health*, pp. 185–C19, October 2022.

Bickford, A. Maintaining integrity in industry-sponsored research. *Virtual Mentor*, 6(11), November 2004.

Biddle, S. OpenAI quietly deletes ban on using ChatGPT for "military and warfare". *The Intercept*, 2024. Accessed: 2024-1-12.

Bihl, T. J. and Talbert, M. Analytics for Autonomous C4ISR within e-Government: a research agenda. In *HICSS*, pp. 1–10, 2020.

Bisht, I. S. DARPA to field test autonomous military Off-Roader. https://www.thedefensepost.com/2022/01/18/darpa-test-auto-off-roader/, January 2022. Accessed: 2023-12-19.

Blanchard, A. and Taddeo, M. Jus in bello necessity, the requirement of minimal force, and autonomous weapons systems. *Journal of Military Ethics*, 21(3-4):286–303, October 2022.

BlueBear. BlueBear - A Saab company. https://bbsr.co.uk/, July 2018. Accessed: 2024-1-3.

Bode, I. and Watts, T. F. A. Loitering munitions and unpredictability: Autonomy in weapon systems and challenges to human control. *Center for War Studies*, 2023.

Bode, I., Huelss, H., Nadibaidze, A., Qiao-Franco, G., and Watts, T. F. A. Prospects for the global governance of autonomous weapons: comparing Chinese, Russian, and US practices. *Ethics Inf. Technol.*, 25(1):5, February 2023.

Bogdon, C. Clearpath founder signs open letter to UN for ban on autonomous weapons. https://clearpathrobotics.com/blog/2017/08/clearpath-founder-signs-open-letter-un-ban-autonomous-weapons/, 2017. Accessed: 2024-3-22.

Boys, J. D. Predator's progress: the bureaucratic challenges to the Clinton administration's development and deployment of unmanned aerial vehicles (1993-2001). *Intell. Natl. Sec.*, 38(4):538–557, June 2023.

Brose, C. *The Kill Chain: Defending America in the Future of High-Tech Warfare*. Hachette Books, April 2020.

Brown, A. Ethics, autonomy, and killer drones: Can machines do right? *Comparative Strategy*, 42(6):731–746, November 2023.

Can, M. and Vieira, A. The Chinese Military-Civil Fusion strategy: A state action theory perspective. *Int. Spect.*, 57 (3):85–102, July 2022.

Castillo, V. A. Future utilization of Unmanned Aerial Vehicles within the United States Navy. *Transportation and Engineering Management, California State University Maritime Academy*, 2022.

Chahal, H., Fedasiuk, R., and Flynn, C. Messier than oil: Assessing data advantage in military AI. *CSET Issue Brief*, 2020. Accessed: 2024-1-3.

Chen, J. and Ran, X. Deep learning with edge computing: A review. *Proc. IEEE*, 107(8):1655–1674, August 2019.

Chen, S. Beijing plans an AI Atlantis for the South China Sea without a human in sight. *South China Morning Post*, November 2018. Accessed: 2024-3-22.

Chu, M.-C. M. China's defence semiconductor industrial base in an age of globalisation: Cross-strait dynamics and regional security implications. *Journal of Strategic Studies*, pp. 1–26, 2023.

Cotovio, V., Sebastian, C., and Goodwin, A. Ukraine's AI-enabled drones are trying to disrupt Russia's energy industry. so far, it's working. *CNN*, April 2024.

Crofts, P. and van Rijswijk, H. Negotiating 'evil': Google, Project Maven and the corporate form. *Law, Technology and Humans*, 2(1):75–90, January 2020.

D'Agostino, S. Does military AI research at universities 'benefit humanity'? https://www.insidehighered.com/news/tech-innovation/artificial-intelligence/2024/01/31/does-military-ai-research-universities, January 2024. Accessed: 2024-4-22.

Daniels, M. Controlling knowledge, controlling people: Travel restrictions of U.S. scientists and national security. *Diplomatic History*, 43(1):57–82, January 2019.

DARPA. ACE program's AI agents transition from simulation to live flight. https://www.darpa.mil/news-events/2023-02-13, 2023. Accessed: 2023-12-6.

De Vynck, G. The U.S. says humans will always be in control of AI weapons. but the age of autonomous war is already here. *The Washington Post*, July 2021.

del Cerro, J., Cruz Ulloa, C., Barrientos, A., and de León Rivas, J. Unmanned aerial vehicles in agriculture: A survey. *Agronomy*, 11(2):203, January 2021.

Department of Commerce. Implementation of additional export controls: Certain advanced computing items. *Federal Register*, 88(205), 2023.

Drezner, D. W. Military primacy doesn't pay (nearly as much as you think). *Int. Secur.*, 38(1):52–79, July 2013.

European Parliament News. EU AI Act: first regulation on artificial intelligence. https://www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence, August 2023. Accessed: 2024-1-5.

Fabbri, A., Lai, A., Grundy, Q., and Bero, L. A. The influence of industry sponsorship on the research agenda: A scoping review. *Am. J. Public Health*, 108(11):e9–e16, November 2018.

Fan, T., Long, P., Liu, W., and Pan, J. Distributed multi-robot collision avoidance via deep reinforcement learning for navigation in complex scenarios. *Int. J. Rob. Res.*, 39 (7):856–892, June 2020.

Faragasso, S. Russian Lancet-3 kamikaze drone filled with foreign parts: Western parts enable Russian Lancet-3 drone to have advanced targeting and anti-jamming capabilities. *Institute for Science and International Security*, December 2023.

Farrell, H. and Newman, A. L. Weak links in finance and supply chains are easily weaponized. http://dx.doi.org/10.1038/d41586-022-01254-5, May 2022. Accessed: 2023-12-16.

Fedasiuk, R. Leviathan wakes: China's growing fleet of autonomous undersea vehicles. https://cimsec.org/leviathan-wakes-chinas-growing-fleet-of-autonomous-undersea-vehicles/, 2021. Accessed: 2024-1-27.

Future of Life Institute. Autonomous weapons open letter: AI & robotics researchers. https://futureoflife.org/open-letter/open-letter-autonomous-weapons-ai-robotics/?cn-reloaded=1., February 2016. Accessed: 2024-3-22.

Future of Life Institute. Pause giant AI experiments: An open letter. https://futureoflife.org/open-letter/pause-giant-ai-experiments/, March 2023. Accessed: 2024-1-27.

Geist, E. M. It's already too late to stop the AI arms race—we must manage it instead. *Bull. At. Sci.*, 72(5):318–321, September 2016.

Gilead, A. Israel recruits robot dogs for Gaza fighting. https://en.globes.co.il/en/article-israel-recruits-robot-dogs-for-gaza-fighting-1001465142, December 2023. Accessed: 2024-4-5.

Goldstein, L. J. Chinese sea mines are threatening the U.S. Navy. https://nationalinterest.org/blog/buzz/chinese-sea-mines-are-threatening-us-navy-80251, 2019. Accessed: 2024-1-27.

Gómez de Ágreda, Á. Ethics of autonomous weapons systems and its applicability to any AI systems. *Telecomm. Policy*, 44(6):101953, July 2020.

Greenwalt, W. C. DOD's Replicator program: Challenges and opportunities. *American Enterprise Institute*, October 2023.

Gruszczak, A. and Kaempf, S. (eds.). *Routledge Handbook of the Future of Warfare*. Routledge, 1 edition, September 2023.

Hambling, D. Ukraine's AI drones seek and attack Russian forces without human oversight. *Forbes Magazine*, October 2023a.

Hambling, D. U.S. marines test robot dog armed with a rocket launcher. *Forbes Magazine*, November 2023b.

Hawkins, A. Chinese students in US tell of 'chilling' interrogations and deportations. *The Guardian*, April 2024.

Hegre, H. Democracy and armed conflict. *J. Peace Res.*, 51 (2):159–172, March 2014.

Henshall, W. The heated debate over who should control access to AI. *Time*, 2023.

Hernandez, J. A military drone with a mind of its own was used in combat, U.N. says. *NPR*, June 2021.

Hoehn, J. R. Joint All-Domain Command and Control (JADC2). https://sgp.fas.org/crs/natsec/IF11493.pdf, 2022. Accessed: 2023-12-17.

Hoijtink, M. and Planqué-van Hardeveld, A. Machine learning and the platformization of the military: A study of Google's machine learning platform TensorFlow. *Int Polit Sociol*, 16(2), June 2022.

Hunter, C. and Bowen, B. E. We'll never have a model of an AI major-general: Artificial intelligence, command decisions, and kitsch visions of war. *Journal of Strategic Studies*, pp. 1–31, 2022.

Ipsos. Global survey highlights continued opposition to fully autonomous weapons. https://www.ipsos.com/en-us/global-survey-highlights-continued-opposition-fully-autonomous-weapons, February 2021. Accessed: 2024-3-22.

Israel Defense. IAI's BlueWhale ASW makes debut in dynamic messenger 2023 exercise. https://www.israeldefense.co.il/en/node/59756, October 2023. Accessed: 2023-11-27.

Johnson, J. Artificial intelligence, drone swarming and escalation risks in future warfare. *The RUSI Journal*, 165 (2):26–36, February 2020.

Jordan, J. The future of unmanned combat aerial vehicles: An analysis using the three horizons framework. *Futures*, 134:102848, December 2021.

Kania, E. B. "AI weapons" in China's military innovation. https://www.brookings.edu/wp-content/uploads/2020/04/FP_20200427_ai_weapons_kania_v2.pdf, 2020. Accessed: 2024-1-3.

Khan, S. M., Mann, A., and Peterson, D. The semiconductor supply chain: Assessing national competitiveness. https://cset.georgetown.edu/wp-content/uploads/The-Semiconductor-Supply-Chain-Issue-Brief.pdf, 2021. Accessed: 2023-12-16.

Klare, M. T. The coming of automated warfare. *Curr. Hist.*, 119(813):9–14, January 2020.

Klare, M. T. Pentagon plans mass autonomous weapons deployment. *Arms Control Today*, 53(8):36–37, October 2023.

Klingler-Vidra, R. The microchip industry would implode if China invaded Taiwan, and it would affect everyone. *The Conversation*, June 2023.

Kohler, C. USAF test pilot school and DARPA announce breakthrough in aerospace machine learning. https://www.edwards.af.mil/News/Article-View/Article/3744695/usaf-test-pilot-school-and-darpa-announce-breakthrough-in-aerospace-machine-lea/, April 2024. Accessed: 2024-4-20.

Konaev, M., Fedasiuk, R., Corrigan, J., Lu, E., Stephenson, A., Toner, H., and Gelles, R. U.S. and Chinese Military AI Purchases. *CSET Data Brief*, August 2023. Accessed: 2023-11-27.

Kwik, J. Mitigating the risk of autonomous weapon misuse by insurgent groups. *Laws*, 12(1):5, December 2022.

L3Harris. L3Harris forward deploys autonomous maritime capabilities to deliver Manned-Unmanned Teaming. https://www.l3harris.com/newsroom/editorial/2023/04/l3harris-forward-deploys-autonomous-maritime-capabilities-deliver-manned, April 2023. Accessed: 2023-12-6.

Laird, B. The risks of autonomous weapons systems for crisis stability and conflict escalation in future U.S.-Russia confrontations. https://www.rand.org/pubs/commentary/2020/06/the-risks-of-autonomous-weapons-systems-for-crisis.html, June 2020. Accessed: 2023-11-27.

Larrick, B. M., Dwyer, J. T., Erdman, J. W., D'Aloisio, R. F., and Jones, W. An updated framework for industry funding of food and nutrition research: Managing financial conflicts and scientific integrity. *J. Nutr.*, 152(8): 1812–1818, August 2022.

Lin-Greenberg, E. Allies and artificial intelligence: Obstacles to operations and Decision-Making (spring 2020). *Texas National Security Review*, 2020.

Liu, X. and Fan, W. China's Wing Loong drones to see more applications, become more intelligent: chief designer. *Global Times*, 2023. Accessed: 2024-4-6.

Longpre, S., Storm, M., and Shah, R. Lethal autonomous weapons systems & artificial intelligence: Trends, challenges, and policies. *MIT Science Policy Review*, 3:47–56, August 2022.

Lyu, C. and Zhan, R. Global analysis of active defense technologies for unmanned aerial vehicle. *IEEE Aerospace and Electronic Systems Magazine*, 37(1):6–31, January 2022.

Markey, E. J. Block Nuclear Launch by Autonomous Artificial Intelligence Act of 2023, May 2023.

Markoff, J. War machines: Recruiting robots for combat. *The New York Times*, November 2010.

Mathiassen, K., Baksaas, M., Græe, S. A., Mentzoni, E. A., and Nielsen, N. H. Making the Milrem Themis UGV ready for autonomous operations. In *Unmanned Systems Technology XXIII*, volume 11758, pp. 221–240. SPIE, April 2021.

Matthew Robert, L. *The rise of killer robots: The impact on Autonomous Weapon Systems (AWS) on the United States' Nuclear Deterrence*. PhD thesis, Stellenbosch : Stellenbosch University, February 2021.

Michel, A. H. AI-Assisted warfare. *MIT Technology Review*, 126(5):46+, 2023.

Mikhailov, D. Artificial intelligence integration as a strategic imperative for national security. *ACM Technology Policy Council*, March 2023.

Milrem Robotics. Milrem Robotics partners with the Ukrainian Defence industry to forge next-generation multi-domain robotic systems. https://milremrobotics.com/milrem-robotics-partners-with-the-ukrainian-defence-industry-to-forge-next-generation-multi-domain-robotic-systems/, December 2023a. Accessed: 2023-12-19.

Milrem Robotics. THeMIS. https://milremrobotics.com/defence/, 2023b. Accessed: 2023-12-6.

Moreau, D. M. Unmanned ground vehicles support of irregular warfare: A Non-Lethal approach. *United States Marine Corps University*, 2011.

Murdick, D., Chou, D., Fedasiuk, R., and Weinstein, E. The public AI research portfolio of China's security forces. https://cset.georgetown.edu/wp-content/uploads/CSET-The-Public-AI-Research-Portfolio-of-Chinas-Security-Forces-1.pdf, 2021. Accessed: 2024-1-27.

Nagabandi, A., Konolige, K., Levine, S., and Kumar, V. Deep dynamics models for learning dexterous manipulation. In Kaelbling, L. P., Kragic, D., and Sugiura, K. (eds.), *Proceedings of the Conference on Robot Learning*, volume 100 of *Proceedings of Machine Learning Research*, pp. 1101–1112. PMLR, 2020.

Newdick, T. and Rogoway, T. Marine XQ-58 Valkyries will be electronic warfare platforms for F-35s. https://www.thedrive.com/the-war-zone/marine-xq-58-valkyries-will-be-electronic-warfare-platforms-for-f-35, May 2023. Accessed: 2023-12-6.

O'Brien, M. AI's future could be 'open-source' or closed. tech giants are divided. *AP News*, December 2023. Accessed: 2024-1-5.

O'Rourke, R. Navy large unmanned surface and undersea vehicles: Background and issues for Congress. *Congressional Research Service*, 2023.

Ozberk, T. Ukraine's Magura V5 USV on the stage at DSEI 2023. https://www.navalnews.com/event-news/dsei-2023/2023/09/ukraines-magura-v5-usv-on-the-stage-at-dsei-2023/, 2023. Accessed: 2023-12-6.

Pegasus Intelligence. Autonomous: Unmanned systemsacross land, sea and air domains. https://pegasusintelligence.com/autonomous/, 2023. Accessed: 2023-12-19.

Perrigo, B. Exclusive: Google workers revolt over $1.2 billion contract with Israel. *Time*, 2024.

Pfaff, C. A., Lowrance, C. J., Washburn, B. M., and Carey, B. A. *Trusting AI: Integrating Artificial Intelligence Into the Army's Professional Expert Knowledge*. US Army War College (USAWC) Press, 2023.

Qiao-Franco, G. and Bode, I. Weaponised artificial intelligence and Chinese practices of Human–Machine interaction. *Chinese Journal of International Politics*, 16(1): 106–128, March 2023.

QinetiQ. MAARS weaponized robot. https://www.qinetiq.com/en-us/capabilities/robotics-and-autonomy/maars-weaponized-robot, 2023. Accessed: 2023-12-19.

Reynolds, I. and Cetin, O. A. War is messy. AI can't handle it. *Bulletin of the Atomic Scientists*, 2023. Accessed: 2024-1-12.

Rickli, J.-M. and Ienca, M. The security and military implications of neurotechnology and artificial intelligence. In Friedrich, O., Wolkenstein, A., Bublitz, C., Jox, R. J., and Racine, E. (eds.), *Clinical Neurotechnology meets Artificial Intelligence: Philosophical, Ethical, Legal and Social Implications*, pp. 197–214. Springer International Publishing, Cham, 2021.

Riesen, E. The moral case for the development and use of autonomous weapon systems. *Journal of Military Ethics*, 21(2):132–150, April 2022.

Rivera, J.-P., Mukobi, G., Reuel, A., Lamparth, M., Smith, C., and Schneider, J. Escalation risks from language models in military and diplomatic decision-making. *arXiv:2401.03408*, January 2024.

Robertson, C. The money blind: How to stop industry bias in biomedical science, without violating the first amendment. *American Journal of Law and Medicine*, 358, 2011.

Russell, S. AI weapons: Russia's war in Ukraine shows why the world must enact a ban. http://dx.doi.org/10.1038/d41586-023-00511-5, February 2023. Accessed: 2023-11-27.

Saeidi, H., Opfermann, J. D., Kam, M., Wei, S., Leonard, S., Hsieh, M. H., Kang, J. U., and Krieger, A. Autonomous robotic laparoscopic surgery for intestinal anastomosis. *Sci Robot*, 7(62):eabj2908, January 2022.

Santora, M. and Triebert, C. Ukraine hits a distant Russian ship, showing reach of naval drones. *The New York Times (Digital Edition)*, pp. NA, August 2023.

Saylam, S. and Gündoğdu, F. K. Decision for the next unmanned aerial vehicle considering regional operational requirements based on Z-Fuzzy Best-Worst method. In *2023 10th International Conference on Recent Advances in Air and Space Technologies (RAST)*, pp. 1–6. IEEE, June 2023.

Sayler, K. M. Emerging military technologies: Background and issues for Congress. Technical report, Congressional Research Service, October 2021.

Scale AI. Scale AI announces partnership with Center for Strategic and International Studies. https://scale.com/blog/scale-csis-partnership-announcement, 2024. Accessed: 2024-1-13.

Schaake, M. The European Commission's Artificial Intelligence Act. *Stanford Univeristy Human-Centered Artificial Intelligence (BHAI) Issue Brief*, 2021.

Scharre, P. *Army of None: Autonomous Weapons and the Future of War*. W. W. Norton & Company, April 2018.

Scharre, P. and Horowitz, M. C. An introduction to autonomy in weapon systems. *JSTOR: Center for a New American Security*, 2015.

Schedl, D. C., Kurmi, I., and Bimber, O. An autonomous drone for search and rescue in forests using airborne optical sectioning. *Sci Robot*, 6(55), June 2021.

Schneider, J. The capability/vulnerability paradox and military revolutions: Implications for computing, cyber, and the onset of war. *Journal of Strategic Studies*, 42(6): 841–863, September 2019.

Seraphin, A. and Miles, W. Toward a balanced approach: Bridging the military, policy, and technical communities. *Ethics Int. Aff.*, 37(3):272–286, March 2023.

Sharma, S. Unmanned ground vehicles: Global developments and future battlefield. *Manohar Parrikar Institute for Defence Studies and Analyses Issue Brief*, April 2022.

Shield AI. V-BAT teams. https://shield.ai/v-bat-teams/, September 2023. Accessed: 2023-12-7.

Stauffer, B. Stopping killer robots: Country positions on banning fully autonomous weapons and retaining human control, 2020. URL https://www.hrw.org/report/2020/08/10/stopping-killer-robots/country-positions-banning-fully-autonomous-weapons-and.

Sterling, T. U.S., China, other nations urge 'responsible' use of military AI. *Reuters*, February 2023.

Stern, J. AI is like ... nuclear weapons? *The Atlantic*, March 2023.

Stop Killer Robots. Our vision and values. https://www.stopkillerrobots.org/vision-and-values/, 2021. Accessed: 2024-3-22.

Sud, A. UAVs and counter UAVs technologies in the world and the indigenous capability. https://media.neliti.com/media/publications/327303-uavs-and-counter-uavs-technologies-in-th-e794317d.pdf, 2020. Accessed: 2023-12-8.

Taddeo, M. and Blanchard, A. A comparative analysis of the definitions of autonomous weapons systems. *Sci. Eng. Ethics*, 28(5):37, August 2022.

Takagi, K. Artificial intelligence and future warfare. https://www.hudson.org/defense-strategy/artificial-intelligence-future-warfare, 2022. Accessed: 2024-1-27.

Tamburrini, G. Nuclear weapons and the militarization of AI. In *Nuclear Risks and Arms Control - Problems and Progresses in the Time of Pandemics and War*, pp. 147–158. Springer International Publishing, 2023.

Templeman, K. A. and Mastro, O. Deepening US-Taiwan cooperation through semiconductors. *Book chapter in Silicon Triangle, edited by Diamond, L., Ellis J.O. Jr., and Schell, O.*, July 2023.

Thornton, R. and Miron, M. Towards the 'third revolution in military affairs'. *The RUSI Journal*, 165(3):12–21, April 2020.

Trapp, K. N. Boots (on the ground). In Hohmann, J. and Joyce, D. (eds.), *International Law's Objects*. Oxford University Press, December 2018.

Triolo, P. Technology crossroads: innovation in China's telecommunications and high-performance computer sectors threatened by US stranglehold on semiconductors. *Asian Security*, 19(2):143–158, May 2023.

UN. First committee approves new resolution on lethal autonomous weapons, as speaker warns 'an algorithm must not be in full control of decisions involving killing'. https://press.un.org/en/2023/gadis3731.doc.htm, 2023. Accessed: 2024-3-22.

UNODA. Lethal autonomous weapon systems (LAWS) – UNODA. https://disarmament.unoda.org/the-convention-on-certain-conventional-weapons/background-on-laws-in-the-ccw/, 2023. Accessed: 2023-12-6.

U.S. Army Training and Doctrine Command. The U.S. army robotic and autonomous systems strategy. https://www.tradoc.army.mil/wp-content/uploads/2020/10/RAS_Strategy.pdf, 2017.

U.S. Congress. 106th Congress Public Law 398. https://www.govinfo.gov/content/pkg/PLAW-106publ398/html/PLAW-106publ398.htm, 2001. Accessed: 2024-3-25.

U.S. Naval Forces Central Command. Unmanned surface vessel transits Strait of Hormuz with U.S. Coast Guard. https://www.navy.mil/DesktopModules/ArticleCS/Print.aspx?PortalId=1&ModuleId=2892&Article=3367448, 2023. Accessed: 2023-12-6.

van Esch, P., Northey, G., Striluk, M., and Wilson, H. Autonomous weapon systems: Is a space warfare manual required? *Computer Law & Security Review*, 33(3):382–389, June 2017.

Walsh, T. Open letter to Professor Sung-Chul Shin, president of KAIST. https://www.cse.unsw.edu.au/~tw/ciair//kaist.html, 2018. Accessed: 2024-3-22.

Wang, H., Cheng, H., and Hao, H. The use of unmanned aerial vehicle in military operations. In *Man-Machine-Environment System Engineering*, pp. 939–945. Springer Singapore, 2020.

Warren, A. and Hillas, A. Friend or frenemy? The role of trust in human-machine teaming and lethal autonomous weapons systems. *Small Wars & Insurgencies*, 31(4): 822–850, May 2020.

Warren, A. and Hillas, A. Red robots: Chinese military modernization and perceptions on lethal autonomous weapons system. *CIAR*, 15(1):9–56, 2021.

Warren, A. and Hillas, A. 'Xi Jinping Thought': Lethal autonomous weapons systems and military modernization with Chinese characteristics. *The Journal of International Relations, Peace Studies, and Development*, 7(1):6, 2022.

Wong, Y. H., Yurchak, J. M., and Button, R. W. *Deterrence in the Age of Thinking Machines*. RAND Corporation, 2020.

Work, R. O. A short history of weapon systems with autonomous functionalities. *Center for a New American Security*, 2021.

Zaluzhnyi, V. Modern positional warfare and how to win in it. https://athenalab.org/wp-content/uploads/2023/11/ZALUZHNYI_FULL_VERSION-2.pdf, 2023.

A. Global Regions with Alleged AWS Deployment Discussed in This Work

Figure 1. Locations of alleged deployment of fully or near fully autonomous AI-powered weapon systems that were discussed in this work. The Turkey, Azerbaijan and Armenia deployments are further discussed by (De Vynck, 2021).