REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
04/12/2019	Master's of Military Studies	SEP 2018 - APR 2019

4. TITLE AND SUBTITLE
Autonomous Weapon Systems: A Pragmatic Legal Framework for an Emerging Capability

5a. CONTRACT NUMBER
N/A

5b. GRANT NUMBER
N/A

5c. PROGRAM ELEMENT NUMBER
N/A

6. AUTHOR(S)
Curley, Gregg, F, Major, USMC

5d. PROJECT NUMBER
N/A

5e. TASK NUMBER
N/A

5f. WORK UNIT NUMBER
N/A

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
USMC Command and Staff College
Marine Corps University
2076 South Street
Quantico, VA 22134-5068

8. PERFORMING ORGANIZATION REPORT NUMBER
N/A

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)
Dr. Jill Goldenziel

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
N/A

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited.

13. SUPPLEMENTARY NOTES
This paper (with future publication-specific edits) has been selected for publication in the August 2019 edition of the NATO Legal Gazette

14. ABSTRACT
This paper argues that an autonomous weapon systems (AWS) ban is unlikely, synthesizes disparate proposals to regulate AWS into a pragmatic legal framework, proposes NATO-sponsored non-binding international guidance on AWS development and use, and recommends improvements to US domestic AWS regulation. AWS are already a reality and the technologies and capabilities are advancing rapidly. A robust legal and regulatory framework is essential to ensuring responsible development and employment of these systems while not simultaneously stifling innovation.

15. SUBJECT TERMS
Autonomous Weapon Systems; AWS; Autonomy; Artificial Intelligence; Law of Armed Conflict; LOAC

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USMC Command and Staff College
					19b. TELEPHONE NUMBER *(Include area code)*
Unclass	Unclass	Unclass	UU	67	(703) 784-3330 (Admin Office)

United States Marine Corps
Command and Staff College
Marine Corps University
2076 South Street
Marine Corps Combat Development Command
Quantico, Virginia 22134-5068

MASTER OF MILITARY STUDIES

AUTONOMOUS WEAPON SYSTEMS: A PRAGMATIC APPROACH TO AN EMERGING TECHNOLOGY

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF MILITARY STUDIES

MAJOR GREGG F. CURLEY, UNITED STATES MARINE CORPS

AY 2018-19

Mentor and Oral Defense Committee Member: Jill Goldenziel, J.D., Ph.D.
Approved: Jill Goldenziel PhD
Date: 4/5/19

Oral Defense Committee Member: LtCol Adam M. King, USMC, J.D., LL.M
Approved:
Date: 4/5/19

EXECUTIVE SUMMARY

Title: Autonomous Weapon Systems: A Pragmatic Approach to Emerging Technology

Author: Major Gregg F. Curley, United States Marine Corps

Thesis: The US Government should pursue regulation that allows for the responsible development and employment of Autonomous Weapon Systems. The United States should take the lead in generating international adoption of common-sense regulatory protocols.

Discussion: This paper introduces the topic of autonomous weapon systems (AWS) and then defines AWS. After a brief overview of open-source and currently-employed AWS, this paper summarizes the arguments for and against continued development of the technology. After determining that a ban of AWS is not feasible, this paper covers the laws that AWS will need to comply with. Next this paper details methods in which AWS can achieve compliance with international law. Since AWS do not have specific international regulation, this paper will analyze the United States' domestic regulation of the technology. This paper then identifies a potential liability gap in law and regulation. Finally, this paper recommends changes to United States' domestic regulation and proposes that NATO Supreme Allied Command-Transformation sponsor an international conference to generate a manual that provides guidance on the development and employment of AWS.

Conclusion: There is no current international approach to effectively regulate the inevitable development and use of autonomous weapon systems. Adopting the aforementioned framework will allow for the responsible innovation, adoption, and use of these technologies while adhering to the spirit of the existing LOAC framework. Additionally, pushing for international adoption of this framework has a realistic chance of international consensus and enforcement whereas a ban would be counter-productive and ineffectual.

TABLE OF CONTENTS

DISCLAIMER

THE OPINIONS AND CONCLUSIONS EXPRESSED HEREIN ARE THOSE OF THE INDIVIDUAL STUDENT AUTHOR AND DO NOT NECESSARILY REPRESENT THE VIEWS OF EITHER THE MARINE CORPS COMMAND AND STAFF COLLEGE OR ANY OTHER GOVERNMENT AGENCY. REFERENCES TO THIS STUDY SHOULD INCLUDE THE FOREGOING STATEMENT.

ABSTRACT

This paper synthesizes disparate proposals to regulate AWS into a pragmatic legal framework, proposes NATO-sponsored non-binding international guidance on AWS development and use, and recommends improvements to domestic US regulation of AWS. AWS are already a reality and the technologies and capabilities are advancing rapidly. A significant amount of discussion pertaining to the legal, moral, and ethical implications of AWS exists. However, this scholarship is generally limited to supporting an outright ban or focusing on a narrow aspect of the nascent capability. A robust legal and regulatory framework is essential to ensuring responsible development and employment of these systems while promoting innovation. AWS can be effectively managed with layers of complimentary regulation consisting of the existing weapons approval processes, system-specific restrictions, and significant validation and verification of each system—all based on an underlying ethical framework grounded in the Law of Armed Conflict. Internationally, NATO should develop non-binding guidance on AWS similar to cyber warfare's Tallinn Manual. NATO is ideally positioned to sponsor influential international guidance on AWS because NATO is international in nature, various NATO members have stated opposition to an outright ban, NATO nations currently employ AWS, and NATO has successfully sponsored non-binding influential legal guidance in the past. Domestically, the US already has a fairly robust AWS framework. However, the US should amend portions of Department of Defense Directive 3000.09 (Autonomy in Weapon Systems) and implement a domestic "war tort" model. These US and NATO actions will ensure a viable legal and regulatory framework for AWS development and employment, send a clear signal of reasonable and ethical AWS development to the world, and provide a foundation from which additional international agreement may be generated.

PREFACE

Autonomous weapon systems (AWS) have captured the attention of a wide array of interested parties including governments, private companies, universities, and scholars. While significant scholarship exists on the ethics, legalities, limitations, and potential capabilities of these systems, to the author's knowledge, a pragmatic legal framework for this technology has not been proposed. My intent is to synthesize the existing literature and propose a workable regulatory framework that ensures compliance with international law and does not stifle innovation.

Numerous individuals helped me throughout the research and writing process for this paper. I would like to thank Dr. Jill Goldenziel for her guidance and mentorship. Her breadth of knowledge, expertise, writing ability, flexibility, and eye for detail significantly enhanced the quality of the paper and the recommendations. I would also like to thank Lieutenant Colonel Adam M. King, U.S. Marine Corps; Major David Palacios, U.S. Marine Corps; and Ms. Andrea Hamlin for reviewing this paper and for their technical input. The input from these individuals has proven invaluable and I am grateful to have benefited from their experiences and perspectives. Last, I would like to thank my wife Lauren for her constant support.

While the aforementioned individuals provided invaluable advice during the writing of this paper, the views, opinions, findings, and conclusions expressed in this paper are strictly my own. They are not responsible for any errors or omissions in this paper.

I. Introduction

Autonomous weapon systems (AWS) are already a reality. Significant advances in computers, artificial intelligence, communications, and robotics will only make them more prolific. As a result, humankind is at the precipice of a paradigm shift in the very character of warfare. Many stakeholders have identified significant concerns and proposed various ways to regulate AWS, a challenge that will be a ". . . crossing of a moral Rubicon."[1] The scholarly writing on the topic adequately identifies the legal, ethical, and moral issues inherent in the employment of AWS and even provides some narrowly-scoped solutions. To date, this research has not been amalgamated into a comprehensive and pragmatic framework that will facilitate responsible development and legal employment of AWS without simultaneously stifling military innovation. This paper will provide the background necessary to address AWS, discuss the anticipated benefits and perceived drawbacks to the technology, and explain why an international ban of AWS is unlikely. Next, this paper will address the legal, ethical, and moral framework to which AWS must adhere and then synthesize the disparate proposals into a workable construct. This construct must effectively manage AWS, promote innovation, function domestically, and have a realistic chance of garnering international support (see figure 1).

II. Background

A. Autonomous Weapons Systems

The US Department of Defense (DoD) Directive 3000.09 (Autonomy in Weapon Systems) defines an autonomous weapon system as:

> A weapon system that, once activated, can select and engage targets without further intervention by a human operator. This includes human-supervised [AWS] that are designed to allow human operators to override operation of the weapon system, but can select and engage targets without further human input after activation.[2] (see figure 2)

The DoD definition requires substantial unpacking to ensure a common understanding of these systems. First, United States Army Lieutenant Colonel (LTC) Christopher Ford, an AWS expert deduced, "autonomy is less a technology as it is a capability comprised of multiple technologies."[3] Therefore, any proposed legal framework will have to address the full spectrum of autonomous capabilities across varied and distinct domains, missions, and platforms. Second, many commentators and international organizations differentiate between autonomous weapon systems (AWS) and lethal autonomous weapon systems (LAWS).[4] The distinction between AWS and LAWS is not the byproduct of autonomy or the inadequacy of the current legal framework, but rather the innate human desire to recognize heightened moral, ethical, and legal implications when the loss of human life is a factor. If an autonomous system possesses lethal capabilities, those capabilities will feature prominently in the LOAC analysis, but there is no requirement for a bifurcated regulatory regime. Therefore, recognizing this distinction independent of the Law of Armed Conflict (LOAC) framework is not necessary, the application of the LOAC principles already account for the difference between a non-lethal and a lethal weapon system. This paper will focus on AWS generally.

Autonomy is not binary, it is a capability that exists on a spectrum of varying degrees (see figure 2). A variety of frameworks have been developed to navigate the various levels of semi-autonomy. The simplest framework to conceptualize is based on Colonel John Boyd's ubiquitous observe, orient, decide, act (OODA) loop.[5] A "human-in-the-loop" system is capable of autonomously selecting targets but will only execute once approval from a human operator is granted.[6] A system that will complete a task unless a human intervenes is a "human-on-the-loop" system, and a system that, once activated, a human can no longer intervene is a "human-out-of-the-loop" system.[7] This paper will explore the legal framework applied to AWS as defined

above; that is, weapon systems that have the capability to operate with a human "on-the-loop" or a human "out-of-the-loop" (see figure 2).

AWS are distinguishable from both unmanned systems and automatic weapons. In unmanned systems, the weapon is merely an extension of the operator—albeit an extension that can now employ an astonishing combination of standoff and lethality. While the decision to employ force no longer needs to be co-located with the weapon system, legal, moral, and ethical accountability for unmanned systems falls on the decision-makers and is adequately addressed by the existing regulatory landscape.

Automatic weapons are capable of being triggered without a human decision after employment but are rule-based and passive in nature (e.g. land mines, booby traps, improvised explosive devices, etc.).[8] Automatic weapons follow a programmed script in which every outcome is predetermined by a programmer.[9] In an autonomous weapon system, the script contains unprogrammed improvisation space in which no outcome has been predetermined.[10] Dr. Rebecca Crootof, a leading AWS scholar, succinctly describes the difference between automated weapon systems and AWS: "automated weapon systems merely react to triggers, autonomous weapon systems process information to derive conclusions before responding."[11]

In the near future, AWS will also employ artificial intelligence (see figure 3). Congress has defined artificial intelligence (AI) as: "[a]ny artificial system that performs tasks under varying and unpredictable circumstances, without significant human oversight, or that can learn from their experience and improve their performance. . . . They may solve tasks requiring human-like perception, cognition, planning, learning, communication, or physical action."[12] Current AWS are capable of acting without a human decision-maker, but in the near future they

will also be able to create and then execute their own decision cycles. Any proposed AWS regulatory scheme needs to account for the foreseeable advances in AI technology.

AWS differ from any other weapon systems because eventually they will possess the capability to independently hunt and kill human beings. Appropriate concerns undergird the creation of military-grade apex predators. Even proponents of the technology must recognize the enormous risk inherent in AWS: the systems must differentiate biologically identical targets based on nuanced cultural and behavioral cues. Absent that capability, the machine will turn on its creators. With a common understanding of what constitutes an autonomous weapon system and the instinctual rationale underlying the aversion to this technology, knowledge of the current state of AWS is helpful.

B. Current State of Autonomous Weapon Systems Technology

Many proponents of bans or limitations on AWS incorrectly believe that AWS do not yet exist.[13] AWS have been present on the battlefield for decades, albeit in limited and well-defined roles. Current examples of weapon systems that, under certain circumstances can independently select and engage targets, are:

- The US Phalanx Close-in-Weapon System (CIWS).[14] This system is a radar controlled defensive cannon that protects ships against airborne and surface threats and is capable of operating fully autonomously.[15] Twenty-four allies utilize the phalanx CIWS system and six other nations employ a similar capability;[16]

- US counter-rocket, artillery, and mortar land-based phalanx weapon system (LPWS).[17] This system consists of the radar array (C-RAM) and a kinetic cannon. This system is employed defensively;[18]

- The Israeli Iron Dome is an anti-mortar/missile defense system.[19] This system is the Israeli equivalent to the US's Phalanx/CIWS and LPWS systems;[20]

- The Israeli Harpy Loitering Weapon is an ". . . anti-radar weapon that searches for radars over a wide area and, once it finds them, kamikazes into them."[21] The Harpy can stay aloft for over two hours and operators employing the Harpy do not need to know the specific locations of the enemy radars that will ultimately be targeted;[22]

- South Korea's SGR-A1 (Security Guard Robot). This robot is employed on the 38th parallel and is capable of autonomously locating, targeting, and killing humans that enter the demilitarized zone (DMZ). Despite the fairly extensive precautions and the limited context in which the robot is employed, South Korea still keeps a human in-the-loop;[23]

- The Chinese/Russian PMK-2 encapsulated torpedo mines.[24] These mines can be laid from the air and loiter at depths up to 2000 meters.[25] When a ship or submarine comes within range, the capsule releases a torpedo that tracks and engages the target.[26]

A responsible inventory of the current state of AWS also requires acknowledgement that advanced weapons development is a non-transparent activity. As such, it is a safe assumption that additional autonomous weapon system capabilities already exist and that at least Russia, China, Israel, and Iran have many more in development.[27] The rapid advancements in AI, robotics, and technology, indicate that AWS are a permanent fixture on the modern battlefield and that the role of these systems will only increase in the future.

C. Employment Considerations of Autonomous Weapon Systems

Critics have argued against the continued development of AWS. The prevailing arguments cite negative ramifications stemming from dehumanizing warfare; insufficient moral, ethical, and legal support for the employment of AWS; and fears of a dystopian future wherein humans become subordinate to AWS. Ultimately, each argument against the development of AWS is flawed.

Arguments Against Continued Development of Autonomous Weapon Systems

The first significant concern with AWS is that removing humans from various aspects of the battlefield will increase the likelihood of war. China and the non-governmental organization Human Rights Watch have both expressed concern that AWS will lower the threshold for war.[28] The argument is predicated on one nation being technologically superior to another to such a degree that the domestic cost of war, in lives and material, is minimal. This argument relies on a

flawed assumption. Every bilateral relationship between nations is not solely an economic equation wherein each nation has a threshold price point below which war will automatically be conducted. If this were the state of reality, power disparities would dictate that powerful nations should already prey on less powerful ones.[29] The dehumanization argument ignores the impact of deterrence, alliances, the international order, human decency, and the myriad other factors present in a society's decision to go to war.[30]

The argument that employment of AWS is morally, ethically, or legally unsupportable collapses as soon as it is tested for validity. Prohibiting development of AWS without allowing for distinctions that recognize the context and manner in which autonomy is employed, leads to suboptimal moral and ethical outcomes. At times, LPWS systems must respond faster than human cognition is capable of perceiving, processing, and reacting. Inserting a human into the LPWS decision loop in those instances negates the utility of the system. Prohibiting the autonomous functions of time-sensitive systems on moral or ethical grounds necessitates the immoral and unethical decision to incur needless death and destruction. It is not morally and ethically superior to permit death and injury from incoming mortars simply because the LPWS system cannot function timely and effectively when a human remains "in-the-loop." As an autonomous weapon advances to the point that it has been validated and verified in certain circumstances as providing superior compliance with the LOAC principles relative to humans, employing humans in these circumstances would be the immoral and unethical option. Superior compliance with the LOAC principles necessarily means fewer military deaths, fewer civilian casualties, and less collateral damage. Last, as the legal "mirror thesis" of law posits, law will adapt and change to reflect the "intellectual, social, economic, and political climate of its time."[31]

When national survival becomes contingent on the development, adoption, and use of AWS, the laws associated with the technology will evolve to accommodate the technology.

More than any other factor, western science fiction depictions—the Terminator Effect[1]—appear to drive opposition to the development and use of AWS.[32] Fear that machines may become self-aware and operate independently of all human input is still premature. A Phalanx CIWS system cannot realistically become sentient and commandeer a destroyer, at least not in the near future. While AWS are a long way from that capability, the rate of technological advance, particularly with AI, means that the singularity is closer than it may appear. The fleeting window of time that exits now, provides an opportunity to develop a cogent international AWS framework, which will better enable addressing more difficult and complex systems in the future. Any proposed framework needs to take into account the significant power inherent in autonomy and AI and be flexible enough to adapt in lockstep with technologies. However, autonomous-capable systems are present now, and responsible discussion on the topic requires setting dystopian concerns aside with the realization that, where practicable, humans will remain "in-the-loop" for the foreseeable future.[33]

Arguments for Continued Development of Autonomous Weapon Systems

Arguments in favor of AWS anticipate fewer civilian and military deaths, lower human and material costs in war, superior compliance with the LOAC principles, and a recognition of the inevitability of these systems. Arguably, a benefit to the employment of AWS is fewer military and civilian deaths—at least on an individual engagement basis. The potential for these systems to better comply with the principles of the LOAC is not mere speculation; it is both an

[1] The "Terminator Effect" is a reference to a science-fiction action franchise that has spanned almost 40 years. In the franchise, an artificial intelligence network utilizes robots, called Terminators, to exterminate the human race. The dystopian fears exploited by the franchise are the same ones supporting an international ban on AWS.

inevitability and proposed prerequisite to employment. Removing human factors from various tasks in warfare will lead to more precise outcomes. AWS will be quicker, more accurate, and more effective than humans at an increasing number of battlefield tasks. An autonomous weapon system without a human in the loop will be unimpeded by human factors—emotions, biological limitations, or survival instincts—that inject additional risk into warfare. Additionally, each battlefield task completed by an autonomous weapon system is one that will no longer require risking the lives of servicemembers.

Countries that do not develop autonomous capabilities will be at a military disadvantage, making continued development of these systems an inevitability. The nuclear arms race illustrates this paradigm very well. When the US was the only nation that possessed nuclear technology, all other nations were dependent on the benevolence and judgment of the US not to employ those weapons. Once two nations possessed nuclear weapons, survival of all parties became the impetus to refrain from using nuclear weapons.[34] Rightly or wrongly, only one country has employed nuclear weapons against an adversary and did so when the technology disparity presented a significant military advantage. To maintain parity, it is clear that nations will need to pursue AWS or risk being at the mercy of those that do. The International Committee of the Red Cross recognizes that AWS attract "considerable interest and research funding so such weapons may well be a feature of warfare in the future."[35] Over thirty nations employ or are currently developing autonomous weapon technologies and "[s]tate and non-state actors will certainly pursue such technology since the barriers to entry are much lower, with greater tactical advantages readily available."[36] Despite the current proliferation of AWS, many in the international community are calling for an outright ban.[37]

D. International Ban of Autonomous Weapon Systems

One proposed response to the anticipated problems of AWS is to institute an international ban. Human Rights Watch, the International Committee for Robot Arms Control, and over fifty other non-governmental organizations have advocated for a ban on AWS.[38] One thousand experts and thought leaders, including famous physicist Stephen Hawking, entrepreneur Elon Musk, and Apple co-founder Steve Wozniak, have also advocated for an outright ban on AWS.[39] A ban is problematic for a few reasons. With the currently existing AWS, a ban would require thirty nations to forfeit validated missile defense systems, or require a ban to have exceptions broad enough they would effectively render a ban meaningless. Second, a ban on these systems is predicated on the beliefs 1) that enforcement is possible and 2) the risks of non-compliance are greater than the risks associated with compliance.

Regulating AWS is unlike regulating weapons of mass destruction. Nuclear, chemical, and biological weapons have unique characteristics (e.g. precursor materials; large quantities of rare materials; and specific technologies and equipment for creation, storage, and protection) that render those weapons amenable to international inspection and enforcement. A coercive yet viable inspection and enforcement program targeting autonomy would be impossible. Autonomy is a capability comprised of many technologies.[40] No country would grant the transparency required for effective inspections, and no international agency has the manpower required for enforcement. Setting aside the impossibility of inspection and enforcement, empirical evidence suggests that an outright ban of AWS could lead to worse outcomes.

In their article encouraging open dialogue on AWS, Judge Advocates LTC Reeves and Major Johnson draw on history to explain an apparent contradiction: an outright ban on a nascent weapon system may actually lead to *more* casualties.[41] The theory holds that as new warfighting

technology develops, responsible and thoughtful dialogue has the potential to foster appropriate and complementary advances in technology, law, and tactics, whereas an outright ban stifles advances in those areas.[42] As more advanced AWS are employed in warfare (an inevitability) a ban will constitute an opportunity cost—time lost in developing technology, law, and tactics.

In 1899, a five-year international ban of balloon-launched projectiles led to significantly more civilian death and destruction during World War II (WWII).[43] Had the ban never been imposed, appreciably better outcomes vis-à-vis the principle of humanity may have been achieved.[44] The outright ban on aerial bombardment effectively tolled all technological development and responsible dialogue on the employment of aerial bombardment.[45] When Allied participation in WWII aerial bombardment became necessary to counter Axis aggression, effective aerial bombardment required indiscriminate obliteration and fire-bombing tactics to generate effects.[46] The technology, applicable legal framework, and tactics were orders of magnitude behind where they could have been if development of the technology and constructive dialogue of the capability had continued unabated.[47] While aerial bombardment technology took almost ninety years to reasonably comply with the LOAC principles, the five years of development lost as a result of the ban translated to avoidable civilian death and destruction in WWII.[48]

The distinction between successful and unsuccessful bans hinge on the difference between a capability and a means.[49] Generally, successful bans prohibit a means but not a capability. A ban on a munition amounts to a nation accepting inefficiency in certain areas in return for the benefits such a regulatory scheme provides their forces (e.g. the banned weapons will not be used against their forces or civilians). Nations that agree to such a ban do not forfeit the ability to utilize a capability, but rather forfeit the ability to use a specific type of munition

(means). The successful bans on hollow-point rounds, glass rounds, poisoned rounds, chemical, and biological weapons adhere to this capability/means distinction.[50] Nations accept inefficiencies inherent in inferior means to accomplish objectives as a matter of comity and humanity while preserving the overall capability. An indiscriminate ban on aerial bombardment attempted to eliminate a capability and it failed. AWS represent a capability with different degrees of autonomy; different processes employing autonomy; different autonomous functions, means, and missions; across all domains and platforms. The largest obstacle to a ban is highlighted by the inherent inability to answer the operative question, "ban what?" Until that question can be definitively answered in the narrower context of means, a ban will fail.

The United States has officially stated its opposition to a ban on AWS at the United Nations, "[r]ather than trying to stigmatize or ban such emerging technologies in the area of lethal autonomous weapon systems, States should encourage such innovation that furthers the objectives and purposes of the Convention."[51] Currently, twenty-six countries support a ban and five, including France, Israel, Russia, United Kingdom, and the United States, outright oppose one.[52] Without those five nations, an effective international ban is unlikely to be enacted. While nominally supporting a ban, China has hedged by officially stating on the record at a UN meeting on AWS, "there should not be any pre-set premises or prejudged outcome which may impede the development of [artificial intelligence] technology."[53]

Last, civil-military considerations related to autonomy will also drive adoption of AWS. There is an inflection point at which market forces require businesses to automate. This point occurs when the cost to automate is comparable to the cost of labor and the quality and quantity of output can equal or exceed that of a human workforce. Businesses that do not automate at this inflection point will lose profits and market share to those that do. Autonomous vehicles are

already a reality on the battlefield. Autonomous vehicles are capable of following logistics trains and unmanned helicopters are capable of delivering supplies.[54] Domestically, autonomous cars are on the horizon with millions of autonomous miles logged and active testing programs in Silicon Valley, CA; Phoenix, AZ; and Pittsburgh, PA.[55] In short order, an outright ban of AWS would be wholly incompatible with a society that promotes private sector automation and allows automated systems to assume more and more domestic and non-combat battlefield tasks. Such a modern society will demand the use of AWS to spare the blood of its youth.

III. Autonomous Weapon Systems and Current Law

A. Weapons Law

Article 36 of Additional Protocol I to the Geneva Convention requires nations to conduct a legal review of new weapons, means, or methods of warfare to ensure that their employment will not be prohibited by international law.[56] This review is concerned primarily with two things: avoiding unnecessary suffering and preventing weapons that are indiscriminate or unlimited in scope.[57] Weapons violate this Article when they inflict damage beyond what is necessary for a military objective (e.g. hollow-point projectiles, poisoned weapons, glass projectiles, etc.).[58] Indiscriminate weapons are incapable of being used in a manner in which the proponent can reasonably distinguish between civilian and military targets (e.g. chemical weapons).[59] Biological weapons are an example of weapons that are unlimited in scope: once unleashed, the effects cannot be controlled. As a result, they are unlawful under Article 36.[60]

In the context of Article 36, the distinguishing feature of AWS is the autonomy—the methods and processes by which the AWS select and authorize target engagement—not the means with which AWS engage those targets. As a result, an autonomous system that utilizes chemical, biological, or glass projectiles would be *per se* illegal; whereas, an autonomous system

that utilizes an internationally accepted munition would not be precluded by Article 36. Significant advances in AI might eventually pose some issues with regard to limiting the scope of an autonomous system once deployed (e.g. an autonomous weapon system that independently and continuously selects and engages targets).[61] Technology has not advanced to this point, but with the proliferation of AI this capability is not as distant as it may seem. To ensure continued Article 36 compliance in this regard, every autonomous-capable system regardless of munition, should have human over-rides to ensure control over the scope of employment. Additionally, common-sense safeguards in AWS architecture and a margin of error discussed below will minimize future risk. In sum, AWS will have some Article 36 implications but Article 36 will not serve as a bar to the development and use of AWS.

B. Law of Armed Conflict Principles

Any autonomous weapon system will need to comply with the Law of Armed Conflict. Legal scholars Gregory and Diana Noone note that, "[n]o academic or practitioner is stating anything to the contrary. . . . Simply put, no one would agree to any weapon that ignores LOAC obligations."[62] The LOAC principles are codified in Additional Protocol I (AP I) to the Geneva Conventions and, although not ratified by the US, the US does consider significant portions of the protocols, including the principles, customary international law.[63] These principles are necessity, humanity, proportionality, and distinction.[64]

Under the necessity principle "[a]ttacks shall be limited strictly to military objectives. . . . military objectives are limited to those objects which by their nature, location, purpose . . . offers a definite military advantage."[65] To be lawful, any destruction or seizure of property must be required by the military dictates of the situation.[66] Humanity pertains to civilians and requires: "[i]n the conduct of military operations, constant care shall be taken to

spare the civilian population, civilians, and civilian objects."[67] Of vital import to this principle, combatants are required to take all feasible precautions to limit the injury and suffering of civilians.[68] Proportionality recognizes the tension between necessity and humanity and the reality that war is messy. Proportionality prohibits "an attack which may be expected to cause incidental loss of civilian life, injury to civilians, damage to civilian objects, or a combination thereof, which would be excessive in relation to the concrete and direct military advantage anticipated."[69] Proportionality is a subjective determination made by the cognizant commander. In a war crimes context, the proportionality decision is subject to a reasonableness standard.[70] Distinction requires safeguards to ensure the military nature of targets and parties. AP (1) 48 states, "[i]n order to ensure respect for and protection of the civilian population and civilian objects, the Parties to the conflict shall at all times distinguish between the civilian population and combatants and between civilian objects and military objectives and accordingly shall direct their operations only against military objectives."[71]

Prior to approval for autonomous use, an autonomous weapon system capable of operating with a human "on-the-loop" or a human "outside-of-the-loop" should verify and validate compliance with the LOAC at levels superior to humans under the same circumstances. Theoretically, a machine can be made more compliant with the LOAC than a human. There are two complementary means of accomplishing this compliance. The first strategy limits the situations and scenarios in which a weapon system can/will operate autonomously, thereby minimizing the risk of a violation of the LOAC. Second, the algorithms utilized need to produce an error rate within the employment criteria that is lower than the human error rate under the same conditions. Essentially, the context in which the system is used and the programmed script must be sufficiently restrictive to ensure any improvised outcome will be compliant with the

LOAC. Once an autonomous weapon system has demonstrated superior compliance relative to humans, the machine will have achieved *de facto* compliance with the LOAC. *De facto* compliance recognizes that a subjective judgment (e.g. the decision to employ force) can be determined objectively by an autonomous system provided a sufficient number and combination of criteria are met. To determine if a violation of the LOAC occurred, a reasonableness standard will be applied by cognizant tribunals.[72] An inverse relationship between the likelihood of an autonomous weapon committing a violation of the LOAC and the reasonableness of the employment of the system exists. The narrower the employment criteria, the more restrictive the algorithms, and the more rigorous the verification and validation of the system, the more reasonable it is to employ the autonomous system.

Narrow Employment Criteria

Narrowing the context and manner in which AWS are employed can significantly increase the likelihood that an autonomous system will comply with the spirit and intent of the LOAC.[73] For instance, employing an autonomous system only in self-defense and against inanimate targets (such as the Phalanx CIWS, LPWS, and Iron Dome systems) eliminate almost all humanitarian, proportionality, necessity, and distinction concerns. Destroying imminent aerial military threats (identified by speed, direction, radar signature, etc.)[74] in self-defense does not generally expose the proponents of those AWS to war crime liability. Tight employment parameters applied to other autonomous capabilities may not effectively generate near total compliance with the LOAC as they do in missile defense systems, but employment parameters can complement sophisticated LOAC algorithms to generate *de facto* compliance. An example of this complementary construct is South Korea's SGR-A1 robot. First, the robot is defensive in nature—it guards the DMZ between North and South Korea against human incursions with lethal and non-lethal munitions.[75]

The DMZ is a militarized hellscape 160 miles long and 2.5 miles wide consisting of a significant military presence, land mines, barbed wire, watch towers, obstacles, and signs.[76] Employing the robot in this particular context and manner significantly reduces the likelihood that the autonomous system will engage innocent civilians, non-military targets, or cause disproportionate destruction. While the employment criteria render violations of the LOAC less likely, the SGR-A1 autonomous system still requires additional algorithmic safeguards to ensure the system does not violate the LOAC.[77]

LOAC Algorithms

To complement sufficiently narrow employment constraints, AWS programming must generate compliance with the LOAC at a rate equal to or better than humans under the same circumstances. Proponents of a ban argue that the LOAC principles inherently require human judgment, and therefore a machine will never be able to comply.[78] These arguments fail to recognize the concept of *de facto* compliance. Humans are fallible and make mistakes. These mistakes translate to an error rate. Once an error rate in a given scenario is quantified, an autonomous weapon system's performance can be measured against humans. If the autonomous weapon's error rate is less than humans under similar circumstances, it is reasonable to use the system; if the error rate of the autonomous system is more than a human, employment is unreasonable. In fact, if an autonomous weapon generates a lower error rate than a human operator, the moral imperative is to employ the system.

Proportionality, necessity, and humanity are principles that can often be reduced to quantifiable algorithms. Proportionality is already a mathematical equation that is executed by humans in targeting cells. A commander sets numerical values on a military target and numerical values on collateral damage. When the value of the target exceeds the acceptable amount of

collateral damage, it is permissible to prosecute the target. Provided the values assigned to the target and the collateral damage are reasonable, this principle is met. An autonomous weapon system algorithm would simply compute the predetermined values in the proportionality analysis faster and more accurately than a human. This principle also would require an autonomous system have a real-time update capability, whereby a commander can update the subjective values of targets and collateral damage as often as necessary. If an autonomous system with those capabilities has not malfunctioned, liability for a proportionality violation would fall to the commander that assigned unreasonable values on the military target and/or the collateral damage, not with the autonomous system. Preprogrammed military targets and self-defense algorithms will ensure that necessity is met. The Israeli Harpy Missile system is an example of a system pre-programmed to only destroy military targets. The missile will only attack transmitting radars that meet set criteria.[79] This constraint ensures that any target, while not identifiable at the deployment of the system, is military. Effectively, the pre-programmed script narrows any improvised autonomous action by the Harpy system to targets that satisfy the necessity principle. Similarly, the factors vital to the humanity principle can often be quantified and programmed for optimal results. For example, AWS can be programed to strike a target based on both pattern of life data and real-time assessments to minimize civilian casualties. These are the very same considerations used now for non-AWS strikes. Additionally, the ability to remove emotions and assume more risk prior to prosecuting targets will lead to better humanitarian outcomes.[80]

Commentators have called distinction, "the greatest hurdle to the legal deployment of AWS."[81] The perceived difficulty of this hurdle stems from fact that technology has not yet achieved the capability to appropriately distinguish between combatants and civilians in most

scenarios. This reality does require real-time human decision making in most cases, but that may change as technology advances. While certainly challenging, it is feasible that over time narrow and complementary employment constraints, sufficiently robust sensors, AI technology, and appropriate algorithms will be able to achieve *de facto* distinction over broader employment scenarios.

De facto distinction is best illustrated via analogy. A South Korean human sentry tasked with guarding the DMZ between North and South Korea must comply with the principle of distinction.[82] As a baseline, the likelihood that a non-combatant would disregard all posted warnings and attempt to navigate military obstacles and mines in the DMZ is low; therefore, the simple presence of someone in the DMZ already provides the sentry with significant information that aids in the distinction calculus. Next, if an individual is in the DMZ, wearing a North Korean military uniform, carrying a firearm, and does not have his arms raised, a viable case for distinction is satisfied and the decision to engage the target is likely reasonable.[83] The reasonableness of the engagement does not change if the engagement is the result of human judgment or an algorithm. This statistical capability is *de facto* distinction—stacking a sufficient number of required conditions prior to engagement that the autonomous weapon system has a demonstrated error rate lower than a human.

While the South Korean sentry/autonomous system example only paired a narrow employment envelope with three requisite conditions, additional distinction criteria and parameters could be programmed to increase the ability to distinguish combatants and lower the error rate. The number of if/then statements that can be programmed into AWS are limited only by the capability of the sensors. However, for humans, the limitation is the tension between the personal risk to the warfighter and the number of conditions that must be met prior to the to the

employment of lethal force. While far from perfect, rules of engagement try to navigate this tension between the acceptable amount of risk and the application of lethal force. AWS—free of the strictures of self-preservation, fear, revenge, emotion, and biological constraints—can assume far more risk than what would be acceptable to impose on, or expect of, a human prior to the employment of lethal force (e.g. enhanced escalation of force procedures, voice commands, non-lethal ammunition, de-escalation procedures, etc.). Theoretically, an autonomous weapon system could execute an extremely complex decision tree comprised of thousands of if/then statements in a fraction of a second. Once an autonomous system has demonstrated the capability to outperform humans, *de-facto* distinction has been achieved and a human "in-the-loop" is no longer necessary for compliance with this principle of the LOAC.

For an autonomous system to be employed, validation and verification of the system should confirm that, when employed as designed, the system is superior to a human operator in adhering to the LOAC principles. While discussion of potential testing protocols and strategies is beyond the scope of this paper, it is important to note that the design, execution, and verification of valid autonomous system tests will be a complex and difficult task.[84] However, the difficulty in designing and implementing effective testing will not absolve the sponsor of liability under the LOAC or remove the requirement to verify and validate the efficacy of the system. Today, AWS technology is not advanced enough to outperform humans in most applications or across broad scenarios. A regulatory scheme should be in place before technology advances to the point that broader *de facto* compliance is possible.[85]

Underlying Ethical Architecture

The underlying architecture is a system of constraints, restraints, and defaults to which AWS algorithms and AI must comply.[86] In his book, *Governing Lethal Behavior in Autonomous*

Robots, Ronald Arkin proposes a detailed and layered ethical architecture for autonomous weapons.[87] Three parts of his ethical decision matrix apply at the operating system level: the ethical governor, the ethical behavior control, and the ethical adaptor.[88] The ethical governor requires an autonomous system to execute non-lethal decision loops for validity after engagement criteria are met but prior to prosecuting the target.[89] This safeguard ensures that a viable non-lethal option does not exist prior to engagement—ensuring that lethality is a last resort. The second element of the ethical decision matrix is an ethical behavior control that limits a lethal response to inside an acceptable ethical framework.[90] The third is an ethical adaptor that permits AI to create a more restrictive ethical framework but never authorizes expansion.[91] Essentially, AI may employ additional "learned" criteria prior to engaging a target, but AI may never disregard pre-set parameters to expand permissible decision space or change a system's initial charter.

The ethical behavior control provides the largest opportunity for ensuring ethical and legal employment of AWS. The platform-specific behavior control systems should be hashed out by the military, experts, ethicists, and other stakeholders on a case-by-case basis. However, some common-sense ethical behavior controls should be included in the underlying architecture of all AWS such as a "do not engage default" that must be affirmatively overridden by precise compliance with all engagement criteria.[92] This default should also be executed whenever the system malfunctions, suffers damage, or a sensor breaks.[93] Pre-programmed self-destruct, self-deactivation, or self-neutralization mechanisms should also be included.[94] Ethical behavior controls are an area ripe for international dialogue, codification, and agreement. Even in the absence of international agreement, the United States should consider implementing domestic regulation requiring ethical behavior controls in all AWS.

Control Measures

Once an autonomous system is approved for use, there must be clearly defined parameters under which the autonomous system is verified and validated. Responsible use of AWS will then include control measures that provide a margin of safety. A margin of safety further narrows the employment window and reduces risk.[95] These additional parameters will be customized to each autonomous system based on the individual system's design and functions. For instance, limiting a geographic maneuver box, capping the amount of time an autonomous system may operate independently, limiting payloads, limiting fuel, and withholding approval authority to a higher commander are all reasonable controls that could be placed on an autonomous weapon system to ensure a margin of safety.[96] This concept is also consistent with current practice. The US has universal safety rules (e.g. "never point a weapon at anything you do not intend to shoot") but also has weapon-specific safety criteria to fill gaps in the general rule created by the particularities of a weapon (e.g. check the back-blast area before using a shoulder-fired rocket).

C. US DoD Directive 3000.09 (Autonomy in Weapon Systems)

DoD Directive 3000.09, last updated May 8, 2017, currently implements many constraints and restraints relative to the development of AWS. First and foremost, humans must be "on-the-loop" for AWS that provide defense of manned installations and platforms—humans must be "in-the-loop" for all other AWS.[97] Next, all systems must be verified and validated through a rigorous testing and evaluation process.[98] The employment of these systems will be limited to a reasonable period of time, and (3) safeguards are required to prevent unanticipated consequences including an adversary hijacking the system.[99] To ensure an autonomous weapon system works as intended, system hardware must have appropriate user interfaces, user-friendly controls, pertinent safeties, anti-tamper measures, clear activation/deactivation procedures, and traceable

feedback capabilities.[100] DoDD 3000.09 also ensures AWS employment complies with, ". . . law of war, applicable treaties, weapon system safety rules, and applicable rules of engagement."[101] Additionally, offensively-employed AWS must be designed to disengage when communications are degraded.[102] An autonomous system designed to function outside the parameters of the Directive requires Under Secretary of Defense approval at both the development and fielding stages.[103] Last, the regulation addresses sales and transfers of AWS technology.[104]

IV. Accountability

Accountability is another essential safeguard against violations of the LOAC. Accountability enables punishment and promotes deterrence, two interrelated concepts that shape the decisions and behavior of individual actors. Some legal theorists posit that violations of the LOAC by AWS do not present accountability problems.[105] They cite unanimous consent among lawyers that, ". . . anyone who commits a LOAC violation should be held accountable (i.e. in [an] AWS scenario that may be the system programmer) and anyone in a superior/command position who knew or should have known about the violation may be held accountable as well."[106] This is a logical leap. This position is correct in every case where intent and/or negligence on the part of a stakeholder exists. The programmer that intentionally programs malicious code into the system, the commander that intentionally employs the autonomous system outside the verified scenarios, and the negligent autonomous system "on-the-loop" supervisor that did not intervene when he had a duty to do so, all provide a clear and direct path to legal liability for violations of the LOAC. Deeper analysis reveals a potential gap in accountability that occurs when 1) there is a violation of the LOAC; 2) the violation is the result of an unforeseen autonomous weapon system malfunction; and 3) no stakeholder has the requisite *mens rea* for personal accountability.[107] The commander, the deployer, the programmer, the contractor, and the

manufacturers of the various sensors have all been proposed as individuals that could or should shoulder responsibility in the event of such an autonomous system malfunction.[108]

A. Commanders

Many have proposed holding the commander responsible if an autonomous weapon system "goes rogue."[109] This option is suboptimal and would hold a commander responsible for actions over which he had no control simply by nature of his command position. If an autonomous system is able to pass the approval crucible and is employed within the defined parameters and approved scenario, the autonomous system would have demonstrated superior compliance with the LOAC principles relative to humans. If commanders are forced to assume liability for the employment of AWS, they are incentivized *not* to employ the system despite its validated superiority. This perverse incentive structure is a moral temptation: a tension that exists when there is a right thing to do, but competing interests provide justification not to do it.[110] If an autonomous system provides better compliance with the LOAC, the "right" thing to do is to utilize the system. Holding commanders criminally or administratively responsible when they do not act intentionally or negligently has the second-order effects of less compliance with the LOAC principles and stifling military innovation. Additionally, such a liability scheme is not consistent with customary interpretations of command responsibility.[111]

B. Contractors

Others have suggested domestic product liability law fill the accountability gap.[112] This proposal is problematic. Carried to its logical conclusion, no company would manufacture weapons of war if it could then be held liable for the use of those weapons. Nations need weapons for survival and modern necessity dictates those weapons be produced by industry. The solution to this tension is to grant domestic immunity for weapons contractors—precisely the state of the law in

the US. In *Boyle v. United Technologies Corp.*,[113] the US Supreme Court held "liability for design defects in military equipment cannot be imposed, pursuant to state law."[114] The government contractor defense holds that liability is not appropriate when: "(1) the United States approved reasonably precise specifications; (2) the equipment conformed to those specifications; and (3) the supplier warned the United States about the dangers in the use of the equipment that were known to the supplier but not to the United States."[115] Imposing liability on contractors for AWS employment would stifle military innovation and create a system in which a nation would be incapable of defending itself. Since autonomy is a capability that is applicable across myriad domains, platforms, and munitions, an exception to domestic product liability for AWS employment will necessarily be so broad that it would subsume the general rule. Imposing liability on contractors leads to the inevitability of industry withdrawing from weapons production or exorbitant costs[116]—untenable outcomes for any nation.

C. Programmers

Programmers who do not act intentionally or negligently pose two issues for accountability. The first issue is the government contractor defense. As a subset of contractors, programmers also enjoy the protection of the government contractor legal defense.[117] Second, programming is now generally done in teams. These teams effectively dilute individual liability to the point that there is no personal accountability to be had for malfunctions resulting from unintentional programming errors that were not detected in the testing stage.

D. Potential Liability Gap

Who should be responsible if an autonomous weapon system acts outside its prescribed parameters and commits a war crime? Suppose the commander was not negligent, the deployer was following lawful orders; the programmer, manufacturer, and the developers did not act

negligently or intentionally and are effectively immune from civil liability under various legal doctrines.[118] In such a scenario, every conceivable stakeholder will lack the requisite *mens rea* for a war crime.[119]

Assuming full compliance with approval processes, directives, and reviews; verification and validation; and proper employment, the potential for a malfunction that leads to an unintended violation of the LOAC, while minimized, still exists. Proponents of AWS do not anticipate perfect systems, just better ones. Therefore when, *not if*, a violation of the LOAC occurs despite proper implementation of all safeguards, there would be no party to hold responsible under current liability frameworks. The remaining entity that can be held responsible for violations of the LOAC by AWS is the state. There is a split between scholars as to whether international current international law defaults to state liability or if there is a liability gap. Some scholars have determined that the law eventually defaults to state liability through at least two avenues.[120] Although these default paths to state liability are fraught with jurisdictional issues.[121] Other scholars argue that the more difficult cases when no person or entity acts intentionally or negligently, create an "accountability gap."[122] Whether this area of the law only needs jurisdictional reform or an entirely new accountability mechanism, it is an area of US and international law ripe for clarification and specificity.

V. Recommendations

A. Proposed Changes to Department of Defense Directive 3000.09 Autonomy in Weapon Systems (see appendix 1 for DoDD 3000.09 with proposed changes)

A regulatory framework that synergistically limits the decision-space for AWS (and in the future AWS with incorporated AI) will help ensure the responsible development of these systems (see figure (3) for a graphic depiction of systematically limited decision-space). An updated DoDD

3000.09 can serve as a strategic messaging tool to the rest of the world clearly communicating the United States' stance on AWS.[123] Additionally, the Directive can serve as a model that can be utilized as a baseline for determining points of international agreement on many aspects of AWS.

Nuclear Interface

The obvious and missing safeguard from DoDD 3000.09 is a categorical prohibition on any AWS/nuclear interface. AWS should never carry, control, respond to, or target nuclear weapons. The consequences of nuclear weapons are so grave that a human, preferably multiple humans, should remain in nuclear decision loops into perpetuity.[124] Placing this simple safeguard into the regulation also sends a favorable international message. Intentions are never certain in international relations. Prohibiting nuclear/AWS interface is one intention that should unequivocally be broadcast, and ideally, reciprocated.[125] The United States—with its nuclear triad (submarine-launched, land-based and air-delivered nuclear warheads)—would not be required to forfeit second-strike capability[126] with the inclusion of a nuclear prohibition in the Directive.[127] The upside to this action is that international actors with less or no nuclear diversity and therefore higher risk may be more likely to agree to an AWS/nuclear interface ban if the US has already taken that step.

Memorializing

AWS must be capable of recording and storing the external stimuli and objective criteria relied on to carry out the autonomous functions of the systems. While this seems similar to traceable feedback,[128] the clause should be strengthened and clarified—if a system is capable of operating without a human in-the-loop, humans need to be able to evaluate the efficacy of the loop. Such a recording of inputs will provide for continuous process improvement, accountability, and justification. Additionally, a recording will also allow for the reconstruction of accidents, a

vitally important capability for assessing liability for malfunctions. Last, and most importantly, recording will ensure these issues do not reoccur.

Stated Policy Preference on Levels of Autonomy

The current Directive does not enumerate an autonomy hierarchy. The Directive should be amended to clearly articulate a policy preference that requires a human "in-the-loop" when practicable, "on-the-loop" when feasible, and only "out-of-the-loop" when an autonomous weapon system will not be able to function effectively under either of the other two modes. Simply because the capability to remove human decision making from the battlefield may quickly become a reality in broader scenarios, does not mean it should be the reality without carefully weighing the alternatives.

Ethical Behavior Control Requirement

DoDD 3000.09 should also include a requirement for an underlying ethical architecture as discussed above. The Directive should merely require the presence of an architecture as a pre-condition to approval of all autonomous weapons. The actual architecture will change as the understanding and regulation of the autonomous capability grows.

Approval Authorities

Last, the Directive should require the AWS approval process assign employment authorization levels commensurate with the missions, capabilities, and risks of each autonomous weapon system. Withholding approval authorities for weapon employment at higher-level commanders with the responsibility, perspective, and experience commensurate with the risks involved is a common safeguard utilized to balance equities appropriately. The viability of this safeguard and the appropriate level for its execution should be considered for each AWS.

B. Indemnification

Addressing the shortcomings in accountability law is more complex. Dr. Crootof has proposed an innovative way to ensure a clear path of accountability in the context of AWS. Dr. Crootof proposed "war torts."[129] A war tort would apply when no individual party or entity has the requisite mental state required for criminal liability. In this narrowly-defined category, the country employing the autonomous system would be strictly liable for damages incurred as a result of a violation of the LOAC. A war tort construct serves to indemnify the commander when an autonomous weapon system is employed as designed and approved. Alternative AWS liability frameworks try to impute liability where none exists, reach the same conclusion by default through complex analysis rife with jurisdictional issues, or simply accept the accountability gap. Absent a war tort regime and the accompanying commander indemnification, a perverse incentive exists to utilize warfighting means leading to sub-optimal outcomes under the LOAC in an effort to avoid personal liability for the unforeseeable actions of a machine. A war tort system serves to absolve a commander of liability if technology that enhances compliance with the LOAC is utilized as designed. Indemnification also reinforces robust national verification and validation processes and procedures, encourages reasonable constraints on the employment of AWS, and provides monetary compensation for damages to victims and their families. A war tort regime will also incentivize contractors to ensure their systems behave as intended—or suffer loss of contracts, clawed-back profits, and other economic damages. A war tort regime would be designed to supplement, not supplant the existing international war crime system.[130] Given the potential benefits of AWS and accountability concerns, there is a realistic chance that many nations will agree to implement a war tort regime as a matter of comity. Additionally, the United States and allies could utilize soft power over time to promote an international war tort

regime. The flow chart in figure 1 demonstrates the limited circumstances in which an accountability concerns could arise. A war tort regime would provide a limited mechanism of accountability to victims and a clear path to AWS liability.

C. International Approach

Three NATO nations oppose a ban of AWS, the US, UK, and France; therefore, NATO is an international entity that has a realistic opportunity to develop well-reasoned principles and meaningful international consensus on development and employment of AWS.[131] With an AWS ban unlikely, international consensus on many facets of AWS is still be feasible. There have been multiple efforts at producing non-binding manuals on the application of International Humanitarian Law to different aspects of warfare—for example the Tallinn Manual, San Remo Manual, and the Manual on International Law Applicable to Air and Missile Warfare.[132] NATO Allied Command Transformation is uniquely situated to charter and develop a similar manual incorporating the legal, moral, and ethical landscape of AWS. Such a manual would go beyond the US executive fiat in DoDD 3000.09 and provide additional room for policy, intent, explanation, and scenarios. A non-binding manual that fleshes out the thorny legal issues associated with AWS would provide a blueprint for the responsible international development and employment of AWS without stifling innovation, investment, and employment of these systems. In time, validated concepts developed as part of this manual may become customary international law and/or be codified in treaties.

VI. Conclusion

The potential benefits and advantages of AWS are significant enough that international consensus on a ban is untenable and, similar to the 1899 ban on aerial bombardment, may be counter-productive long-term. Responsible development and employment of these systems

require a disciplined and reasoned approach to AWS development, as well as slight changes to US policy, US law, and international law to ensure war crime accountability.[133] Nations must require AWS to comply with standard weapons law. Second, prior to approving an autonomous weapon system for use, that system must have demonstrated superior compliance with the LOAC relative to a human operator under similar conditions. Next, the underlying software architecture, regardless of the AWS' platform and capabilities, must have similar fail-safe attributes including a default to restraint, canceling the mission, and/or self-destructing. On top of that architecture, weapon system-specific control measures should further promote compliance with LOAC, such as preprogrammed maneuver boxes, time limits, payload restrictions, and elevated approval authorities. Service regulations and international legal frameworks must criminalize employment of AWS outside of the validated design parameters. All of these measures are capable of unilateral US implementation but are ripe for international review, non-binding guidance, and/or regulation.[134] Last, when an autonomous system violates principles of the LOAC and no individual liability attaches, the country employing the system should be held strictly liable for the damages under a formalized war tort regime.

The proposed legal framework for AWS recognizes the inevitability of AWS and the potential benefits of this technology. The framework ensures compliance with customary international law and encourages responsible AWS development without stifling innovation. Most importantly, this proposal has the potential to garner both domestic and international backing as it complements long-standing principles of war that already enjoy near universal support.

Illustrations

Figure 1

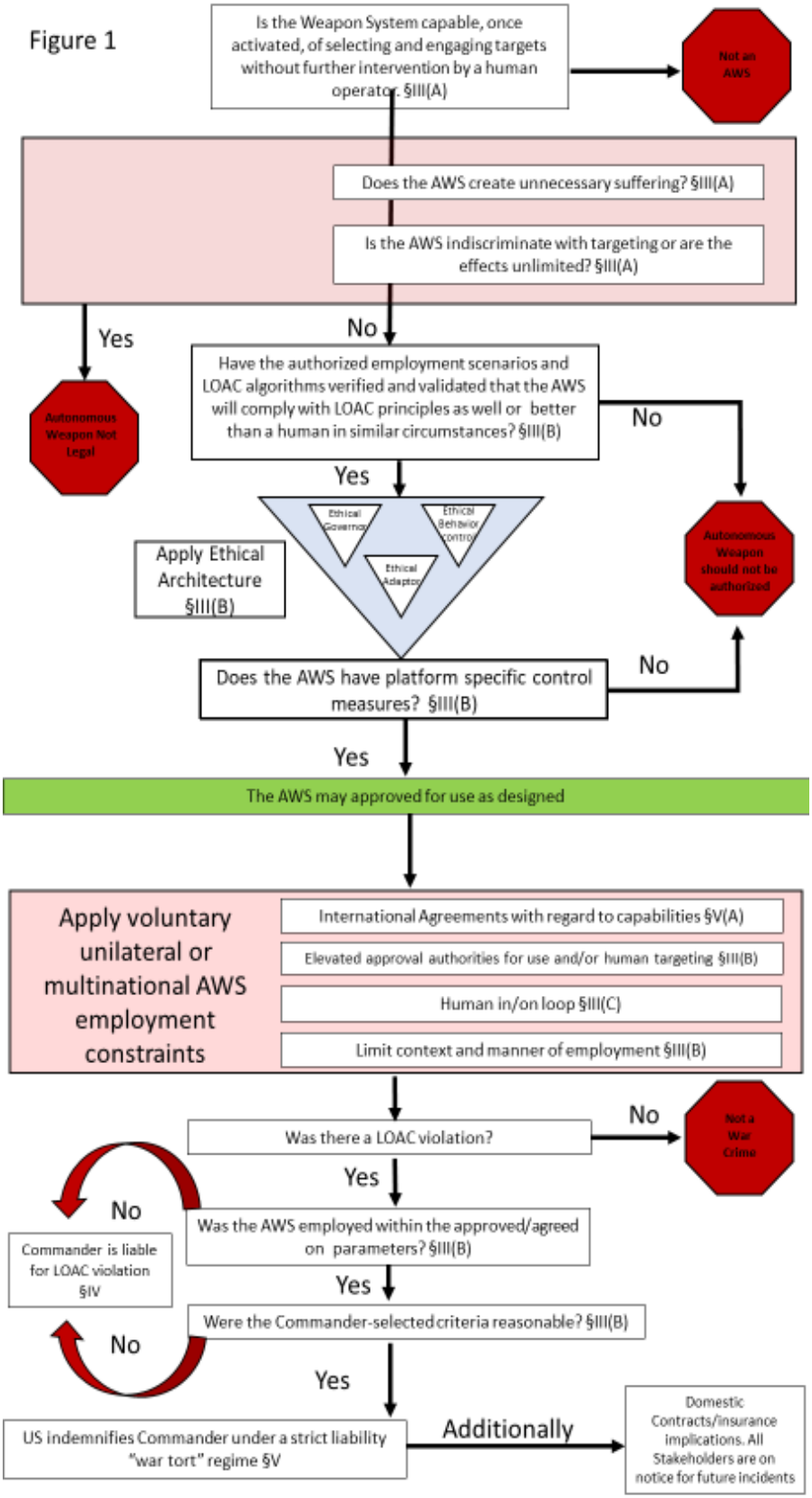

Is the Weapon System capable, once activated, of selecting and engaging targets without further intervention by a human operator. §III(A)
Not an AWS
Does the AWS create unnecessary suffering? §III(A)
Is the AWS indiscriminate with targeting or are the effects unlimited? §III(A)
Yes
Autonomous Weapon Not Legal
No
Have the authorized employment scenarios and LOAC algorithms verified and validated that the AWS will comply with LOAC principles as well or better than a human in similar circumstances? §III(B)
No
Autonomous Weapon should not be authorized
Yes
Apply Ethical Architecture §III(B)
Ethical Governor
Ethical Behavior Control
Ethical Adaptor
Does the AWS have platform specific control measures? §III(B)
No
Yes
The AWS may approved for use as designed
Apply voluntary unilateral or multinational AWS employment constraints
International Agreements with regard to capabilities §V(A)
Elevated approval authorities for use and/or human targeting §III(B)
Human in/on loop §III(C)
Limit context and manner of employment §III(B)
Was there a LOAC violation?
No
Not a War Crime
Yes
No
Commander is liable for LOAC violation §IV
Was the AWS employed within the approved/agreed on parameters? §III(B)
Yes
No
Were the Commander-selected criteria reasonable? §III(B)
Yes
US indemnifies Commander under a strict liability "war tort" regime §V
Additionally
Domestic Contracts/insurance implications. All Stakeholders are on notice for future incidents

Figure 2

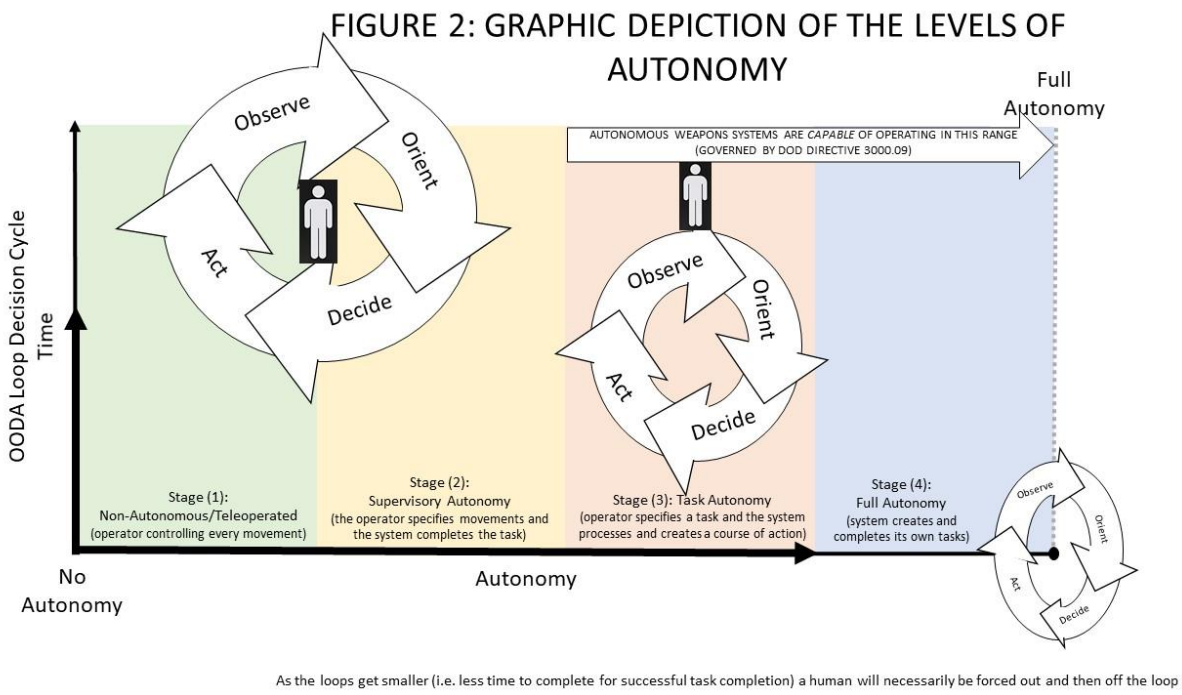
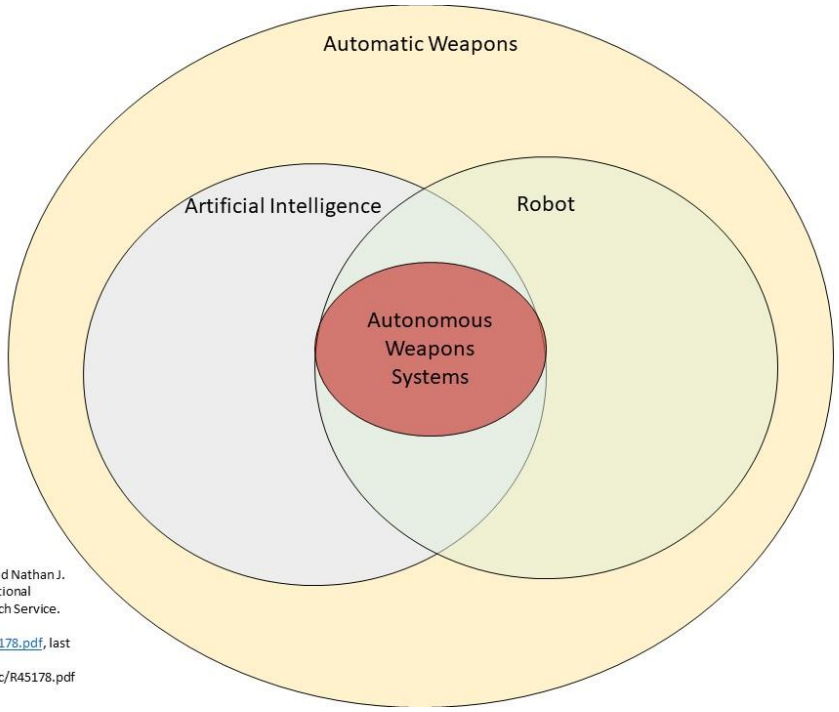
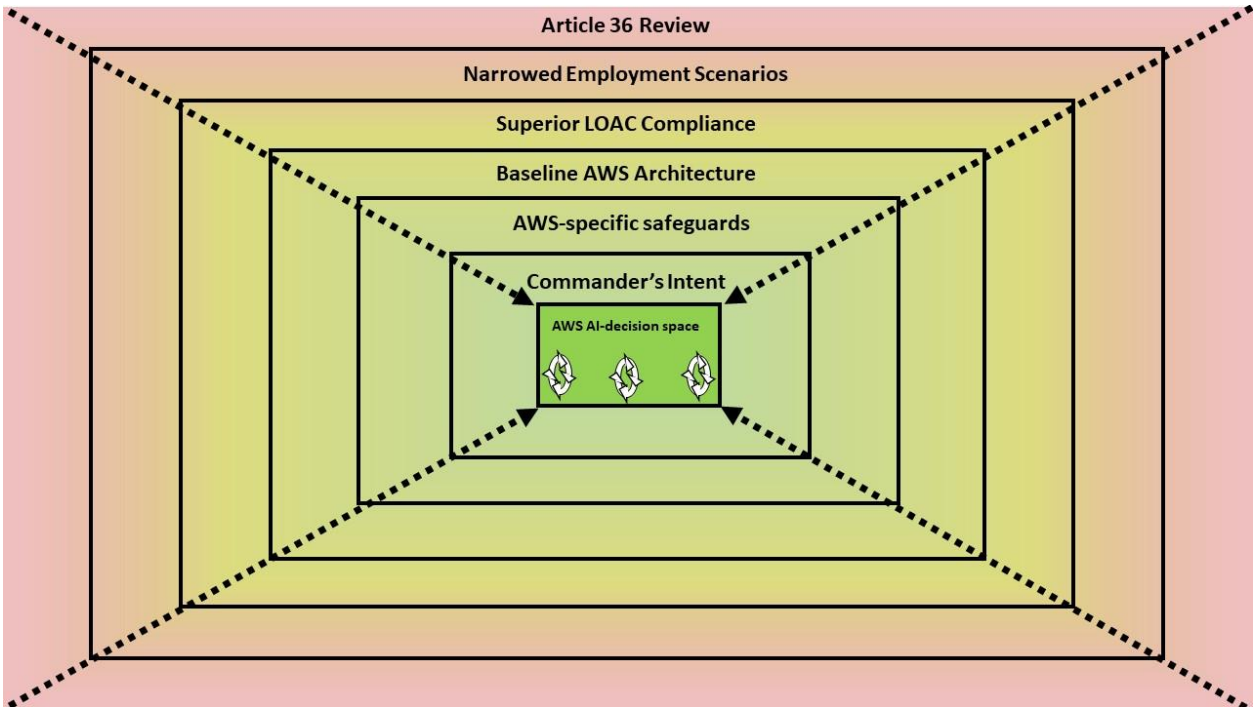


Figure 3



Adapted from Daniel S. Hoadley and Nathan J. Lucas, Artificial Intelligence and National Security. The Congressional Research Service. April 26, 2018 available at https://fas.org/sgp/crs/natsec/R45178.pdf, last retrieved December 20, 2018.https://fas.org/sgp/crs/natsec/R45178.pdf

Figure 4


Article 36 Review
Narrowed Employment Scenarios
Superior LOAC Compliance
Baseline AWS Architecture
AWS-specific safeguards
Commander's Intent
AWS AI-decision space

[1] Rebecca Crootof, *War Torts: Accountability for Autonomous Weapons*, University of Pennsylvania Law Review, May 2016, 14.

[2] Department of Defense Directive 3000.09, "Autonomy in Weapon Systems," May 8, 2017, 13 [hereinafter DoDD 3000.09].

[3] Lieutenant Colonel Christopher M. Ford, *Autonomous Weapons and International Law*, University of South Carolina Law Review 69 S. Car. Law Rev. 413, 416.

[4] Daniel S. Hoadley and Nathan J. Lucas, *Artificial Intelligence and National Security*, The Congressional Research Service, April 26, 2018, https://fas.org/sgp/crs/natsec/R45178.pdf, 12.

[5] *See* John R. Boyd, *Destruction and Creation (PDF)*, September 3, 1976, U.S. Army Command and General Staff College, https://globalguerrillas.typepad.com/JohnBoyd/Destruction%20and%20Creation.pdf.

[6] See Paul Scharre, *Autonomous Weapons and Operational Risk*, Washington: Center for a New American Security, 2016, https://search-proquest-com.lomc.idm.oclc.org/docview/1834992075?accountid=14746, 43; and Amitai Etzioni PhD, and Oren Etzioni PhD, "Pros and Cons of Autonomous Weapons Systems," *Military Review* 97 (3) 2017, https://search-proquest-com.lomc.idm.oclc.org/docview/1922376987?accountid=14746, 78 (citing Bonnie Docherty, *Losing Humanity: The Case against Killer Robots* (Cambridge, MA: Human Rights Watch, 19 November 2012), https://www.hrw.org/report/2012/11/19/losing-humanity/case-against-killer-robots, 2).

[7] See Scharre, *Autonomous Weapons*, 43; and Etzioni & Etzioni, 78 (citing Docherty, *Losing Humanity*, 2).

[8] Hoadley and Lucas, *Artificial Intelligence*, 4.

[9] Ford, *Autonomous Weapons*, 420.

[10] Ford, *Autonomous Weapons*, 420.

[11] Rebecca Crootof, *The Killer Robots are Here: Legal and Policy Implications*, 36 Cardozo L. Rev. 1837, 12.

[12] U.S. Congress, House, FUTURE of Artificial Intelligence Act of 2017, HR 4625, 115th Cong., introduced in House December 12, 2012, https://www.congress.gov/bill/115th-congress/house-bill/4625/related-bills?q=%7B%22search%22%3A%5B%22H.+R.+83%22%5D%7D12/12/2017; and U.S. Congress, Senate, FUTURE of Artificial Intelligence Act of 2017, HR 4625, 115th Cong., introduced in Senate December 12, 2012, https://www.congress.gov/bill/115th-congress/house-bill/4625/related-bills?q=%7B%22search%22%3A%5B%22H.+R.+83%22%5D%7D12/12/2017.

[13] See Kelly Cass, *Autonomous Weapons and Accountability: Seeking Solutions in the Law of War*, 48 Loy. L.A. L. Rev. 1017, Spring 2015, 5; Bradan T. Thomas, *Autonomous Weapon Systems: The Anatomy of Autonomy and the Legality of Lethality*, 37 Hous. J. Int'l L. 235, 2; Benjamin Kastan, *Autonomous Weapons Systems; A Coming Legal "Singularity"?*, 2013 U. Ill. J.L. tech. & Pol'y 45, 17-18; Shane R. Reeves, and William J. Johnson, *Autonomous Weapons: Are You Sure these are Killer Robots? Can we Talk about it?*, The Army Lawyer, https://search-proquest-com.lomc.idm.oclc.org/docview/1540957074?accountid=14746, 25; China, *Position Paper*, Group of Governmental Experts of the High Contracting Parties to the Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to Be Excessively Injurious or to Have Indiscriminate Effects, April 9-13; and Gregory P. Noone and Diana C. Noone, *The Debate over Autonomous Weapons Systems*, Case Western Reserve Journal of International Law 47, no. 1 (Spring 2015): 29, http://scholarlycommons.law.case.edu/jil/vol47/iss1/6/, 35.

[14] Brian K. Hall, "Autonomous Weapons Systems Safety," *Joint Force Quarterly: JFQ* (86) https://search-proquest-com.lomc.idm.oclc.org/docview/1916950387?accountid=14746, last retrieved December 17, 2018, 89; and Crootof, *War Torts*, 15.

[15] The Phalanx CIWS system usually has a human in-the-loop; however, when it is in "casualty mode," the system is capable of fully autonomous operation. See Kastan, *Autonomous Weapons Systems*, 5.

[16] See Phalanx Close-in Weapon System, *Raytheon*, https://www.raytheon.com/capabilities/products/phalanx, last retrieved January 17, 2019; and Crootof, *War Torts*, 15.

[17] Hall, *Autonomous Weapons Systems Safety*, 89; and Scharre, *Autonomous Weapons*, 43.

[18] Counter-Rocket, Artillery, Mortar (C-RAM) Intercept Land-Based Phalanx Weapon System (LPWS), US Army, https://asc.army.mil/web/portfolio-item/ms-c-ram_lpws/, last retrieved January 17, 2019.

[19] Crootof, *The Killer Robots*, 15.

[20] Crootof, *The Killer Robots*, 15.

[21] Scharre, *Autonomous Weapons*, 19.

[22] Scharre, *Autonomous Weapons*, 19.

[23] Christopher P. Toscano, *"Friend of Humans": An argument for Developing Autonomous Weapons Systems*, 8 J. Nat'l Security L. & Pol'y 189, 9.

[24] Crootof, *War Torts*, 15.

[25] Scott C. Truver, Taking Mines Seriously: Mine Warfare in China's Near Seas, *Naval War College Review*, Volume 65, 2012, https://digital-commons.usnwc.edu/cgi/viewcontent.cgi?referer=https://www.google.com/&httpsredir=1&article=1429&context=nwc-review, 12.

[26] See Paul Scharre, *Autonomy, "Killer Robots, " and Human Control in the Use of Force--Part I*, JUST SECURITY, July 9, 2014, *http://justsecurity.org/12708/autonomy-killer-robots-human-control-force-part*.

[27] Toscano, *Friend*, 4.

[28] See China, *Position Paper*; and Bonnie Docherty, "We're Running Out of Time to Stop Killer Robot Weapons," *Human Rights Watch*, April 11, 2018, https://www.hrw.org/news/2018/04/11/were-running-out-time-stop-killer-robot-weapons.

[29] Toscano, *Friend*, 29.

[30] Toscano, *Friend*, 29.

[31] Brian Z. Tamanaha, *"Law and Society,"* St. John's University School of Law, Legal Studies Research Paper Series, [Draft], available at The Social Science Research Network Electronic Paper Collection http://ssrn.com/abstract=13452042, last retrieved March 16, 2019.

[32] Toscano, *Friend*, 2.

[33] Toscano, *Friend*, 9.

[34] Lawrence Freedman, The First Two Generations of Nuclear Strategists, In *Makers of Modern Strategy from Machiavelli to the Nuclear Age*, Ed. By Peter Paret (Princeton, NJ: Princeton University, 1986), 738-739.

[35] Toscano, *Friend*, 11 (citing ICRC Resource Center, *Autonomous Weapons: States Must Address Major Humanitarian Ethical Challenges* (Sept. 2, 2013), http://www.icrc.org/eng/resources/documents/faq/q-and-aautonomous-weapons.htm).

[36] Toscano, *Friend*, 12.

[37] See Dan Smith, "Stephen Hawking, Elon Musk Warn of 'Third Revolution in Warfare' with Autonomous Weapons," *ABC Premium News*, Jul 28, 2015, https://search-proquest-com.lomc.idm.oclc.org/docview/1699087903?accountid=14746, December 18, 2018; and William D. Hood, "Autonomous Weapons Systems: What Commanders Should Know," *The Marine Corps Gazette*, March 2015, 43-44.

[38] Hood, *What Commanders Should Know*, 43-44.

[39] Smith, Third Revolution. This source indicates that the supporters of this ban have not distinguished between lethal and non-lethal autonomous weapon systems.

[40] Ford, *Autonomous Weapons*, 416.

[41] Reeves and Johnson, *Can we Talk about it?*, 27.

[42] Reeves and Johnson, *Can we Talk about it?*, 27.

[43] See Reeves and Johnson, *Can we Talk about it?*, 28 (Twenty four countries agreed to the original 1899 ban. A subsequent 1907 ban encountered significantly more resistance and was narrower in scope).

[44] Declaration (IV,1), to Prohibit, for the Term of Five Years, the Launching of Projectiles and Explosives from Balloons, and Other Methods of Similar Nature, *The Hague*, 29 July 1899, https://ihl-databases.icrc.org/ihl/INTRO/160?OpenDocument.

[45] Reeves and Johnson, *Can we Talk about it?*, 28.

[46] These techniques were most famously applied to London, Dresden, and Tokyo.

[47] Reeves and Johnson, *Can we Talk about it?*, , 28-29.

[48] Reeves and Johnson, *Can we Talk about it?*, , 29.

[49] See Law of Armed Conflict Deskbook, 5th Ed. *The Judge Advocate General's Legal Center and School,* 2017, http://www.loc.gov/rr/frd/Military_Law/pdf/LOAC-Deskbook-2015.pdf, 152 [hereinafter, Deskbook], (hollow-point bullets, frangible rounds, and fragmentation rounds); Deskbook 151, (glass rounds); Deskbook, 159, (chemical weapons); and Deskbook 20, (biological weapons).

[50] See Deskbook, pp 20, 151, 152, and 159.

[51] United States of America, "Humanitarian Benefits of Emerging Technologies in the Area of Lethal Autonomous Weapon Systems," *Group of Governmental Experts of the High Contracting Parties to the Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons which May be Deemed to be Excessively Injurious or to Have Indiscriminate Effects*, April 9-13, 2018, 6.

[52] Tucker Davey, "Lethal Autonomous Weapons: An Update from the United Nations," April 30, 2018, https://futureoflife.org/2018/04/30/lethal-autonomous-weapons-an-update-from-the-united-nations/.

[53] China, *Position Paper*, 2.

[54] See Ryan Felton, "Lockheed Martin's Autonomous Military Vehicles Aim To Save Lives In A Different Way," *FoxtrotAlpha*, February 18, 2017, available at https://foxtrotalpha.jalopnik.com/lockheed-martins-autonomous-military-vehicles-aim-to-sa-1792128164; and K-Max, Lockheed Martin, available at https://www.lockheedmartin.com/en-us/products/k-max.html, last retrieved March 16, 2019.

[55] See Aaron Aupperlee, "5 Reasons Pittsburgh is Still Tops in Autonomous Vehicles," *The Pittsburgh Tribune-Review*, July 21, 2017, http://www.govtech.com/fs/5-Reasons-Pittsburgh-is-Still-Tops-in-Autonomous-Vehicles.html; and Ryan Randazzo, "Waymo announces 'Waymo One,' but self-driving ride service isn't public — yet," *Arizona Republic*, December 5, 2018, https://www.azcentral.com/story/money/business/tech/2018/12/05/waymo-one-launches-self-driving-car-service-arizona/2114688002/.

[56] The US is not a signatory to Additional Protocol I but recognizes Article 36 as customary international law and conducts legal reviews of all new weapons systems. Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of International Armed Conflicts, Art. 36, June 8, 1977, 1125 U.N.T.S. 3 [hereinafter Additional Protocol I].

[57] See Int'l Comm. of the Red Cross, Weapons that May Cause Unnecessary Suffering or Have Indiscriminate Effects (1973), *http://www.loc.gov/rr/frd/Military_Law/pdf/RC-Weapons.pdf*.

[58] Toscano, *Friend*, 17.

[59] Deskbook, 137.

[60]Additional Protocol I, Art. 36, June 8, 1977, 1125 U.N.T.S. 3.

[61] See the plot to W. D. Richter, *Stealth*, DVD, Directed by Rob Cohen, Los Angeles, Columbia, July 29, 2005.

[62] Noone & Noone, *The Debate*, 29.

[63] Deskbook, 71.

[64] See Deskbook, 133-162.

[65] Additional Protocol I, Art. 52(2).

[66] Additional Protocol I, Art. 52(2).

[67] Additional Protocol I, Art. 57(1).

[68] Additional Protocol I, Art. 57(1).

[69] Additional Protocol I, Arts. 51(5)(b), 57(2)(a)(iii).

[70] Deskbook, 148.

[71] Additional Protocol I, Art. 48.

[72] Deskbook, 148.

[73] Ford, *Autonomous Weapons*, 429.

[74] Robert H. Stoner, "R2D2 with Attitude: The Story of the Phalanx Close-In Weapons," *NavWeaps*, http://www.navweaps.com/index_tech/tech-103.php, last retrieved January 18, 2019.

[75] Etzioni & Etzioni, 79.

[76] See Demilitarized Zone, Encyclopedia Britannica, https://www.britannica.com/place/demilitarized-zone-Korean-peninsula, last retrieved January 18, 2019.

[77] The SGR-1 is capable of firing non-lethal ammunition, can distinguish a human with his or her hands in the air, and given that there is time for a human to exercise discretion, retaining a human in the loop; and Toscano, *Friend*, 9.

[78] Noone & Noone, *The Debate*, 26; and International Committee for Robot Arms Control, Berlin Statement, October 2010, https://www.icrac.net/statements/.

[79] Israel Aerospace Industries, *Harpy NG*, http://www.iai.co.il/Sip_Storage//FILES/5/41655.pdf, last retrieved February 12, 2019.

[80] Frank Sauer, "Stopping 'Killer Robots': Why Now Is the Time to Ban Autonomous Weapons Systems," *Arms Control Today*, October 2016, 9.

[81] Kastan, *Autonomous Weapons Systems*, 14.

[82] Additional Protocol I, Art. 48.

[83] Additional Protocol I, Art. 48

[84] Ford, *Autonomous Weapons*, 457-460.

[85] See "AI is a rare case where I think we need to be proactive in regulation than be reactive."—Elon Musk (Catherine Clifford, 9 of the most jaw-dropping things Elon Musk said about robots and AI in 2017, *CNBC*, December 18, 2017, https://www.cnbc.com/2017/12/18/9-mind-blowing-things-elon-musk-said-about-robots-and-ai-in-2017.html).

[86] Isaac Asimov's Three Laws of Robotics espoused in his science fiction book, *I, Robot*, provide an example of a simplistic but broadly-applicable underlying ethical architecture: 1) A robot may not injure a human being or,

through inaction, allow a human being to come to harm. 2) A robot must obey orders given it by human beings except where such orders would conflict with the First Law. 3) A robot must protect its own existence as long as such protection does not conflict with the First or Second Law. See Isaac Asimov, *I, Robot*, (Garden City, NY: Doubleday, 1950), 40.

[87] Ronald C. Arkin, Governing Lethal Behavior in Autonomous Robots (2009), 125.

[88] Arkin, 125.

[89] Arkin, 125.

[90] Arkin, 125.

[91] Arkin, 125.

[92] See Kastan, *Autonomous Weapons Systems*, note 71, 8; and Toscano, *Friend*, 20.

[93] "Do not engage" defaults are consistent with US military practice vis-à-vis non-autonomous weapon systems. For example, an observer can send an artillery mission as "do not load." This means that required data is generated, but the weapon is not loaded and cannot fire. To fire the mission, the observer must send the message "cancel do not load," and the mission becomes active.

[94] Crootof, *The Killer Robots*, 43.

[95] Margin of safety, "the margin required in order to ensure safety; in engineering the margin of safety is the factor of safety (strength of the material divided by the anticipated stress) minus one." Ronald A. Beaulieu, "*Margin of Safety Definition and Examples used in Safety Basis Documents and the USQ Process,"* https://www.osti.gov/servlets/purl/1134068/, last retrieved February 13, 2019.

[96] See e.g. International Committee for Robot Arms Control, *Berlin Statement*, October 2010, https://www.icrac.net/statements/.

[97] DODD 3000.09 4 a and 4 c 2 (a) & (b).

[98] DODD 3000.09 4 a 1.

[99] DODD 3000.09 4 a 1 (a), (b), & (c).

[100] DODD 3000.09 4 a 2 (a) & (b) and 3 (a), (b), (c).

[101] DODD 3000.09 4 b.

[102] DODD 3000.09 4 c 1.

[103] DODD 3000.09 4 d.

[104] DODD 3000.09 4 e.

[105] Noone & Noone, *The Debate*, 30-31.

[106] Noone & Noone, *The Debate*, 31.

[107] Crootof, *War Torts*, 9.

[108] Crootof, *War Torts*, 14.

[109] Toscano, *Friend*, 28.

[110] Rebecca J. Johnson, Moral Decision Making, October 1, 2015, Slide 8 Notes, Lecture delivered October 9, 2018.

[111] See Deskbook, 185 (Citing *U.S. v. Tomoyuki Yamashita*, "The commander's personal dereliction must have contributed to or failed to prevent the offense" and *The United States of America vs. Wilhelm von Leeb, et al.,* "Military subordination is a comprehensive but not conclusive factor in fixing criminal responsibility . . . A high commander cannot keep completely informed of the details of military operations of subordinates . . . He has the right to assume that details entrusted to responsible subordinates will be legally executed . . . There must be a personal dereliction. That can only occur where the act is directly traceable to him or where his failure to properly supervise his subordinates constitutes criminal negligence on his part. In the latter case, it must be a personal neglect amounting to a wanton, immoral disregard of the action of his subordinates amounting to acquiescence. Any other interpretation of international law would go far beyond the basic principles of criminal law as known to civilized nations.")

[112] See Kastan, *Autonomous Weapons Systems*, 19.

[113] *Boyle v. United Technologies Corp*., 487 U.S. 500 (1988).

[114] *Boyle*, 512.

[115] *Boyle*,512.

[116] *Boyle*,512.

[117] See Cass, *Autonomous Weapons*, 23 citing Int'l Comm. of the Red Cross, Report of the ICRC Expert Meeting on "*Autonomous Weapons Systems*: Technical, Military, *Legal* and Humanitarian Aspects', 26-28 March 2014, Geneva 2 (2014), 8 (out of the four *legal* regimes listed, the only one applicable to programmers acting unintentionally is product liability which is subject to the same restrictions as contractor liability).

[118] Kastan, *Autonomous Weapons Systems*, 22.

[119] See Crootof, *War Torts*, 24, note 155, *citing e.g.*, Rome Statute, *supra* note 13, art. 30(1) ("[A] person shall be criminally responsible and liable for punishment . . . only if the material elements are committed with intent and knowledge."); *see also* *Prosecutor v. Blaškić*, Case No. IT-95-14-T, Trial Chamber Judgment, P 152 (Int'l Crim. Trib. for the Former Yugoslavia Mar. 3, 2000), *http://www.icty.org/x/cases/blaskic/tjug/en/bla-tj000303e.pdf* [*https://perma.cc/4FG6-WZRE*] ("[T]he *mens rea* constituting all the [grave breaches of the Geneva Conventions] includes both guilty intent and recklessness which may be likened to serious criminal negligence.").

[120] Ford, *Autonomous Weapons*, 475.

[121] Crootof, *War Torts*, 28.

[122] Crootof, *War Torts*, 4.

[123] Freedman, *Nuclear Strategists*, 754.

[124] Ashley Deeks, Noam Lubell, & Daragh Murray, Daragh, "Machine Learning, Artificial Intelligence, and the Use of Force by States," 10 J. Nat'l Security L. & Pol'y __ (forthcoming 2019).

[125] Freedman, *Nuclear Strategists*, 754.

[126] Freedman, *Nuclear Strategists*, 753.

[127] U.S. Department of Defense, *Nuclear Posture Review*, Office of the Secretary of Defense Washington, DC: Pentagon, February 2018, 42.

[128] DODD 3000.09 4 a (3) (b).

[129] Crootof, *War Torts*, 28.

[130] Crootof, *War Torts*, 28.

[131] Elizabeth Minor, Prohibiting Autonomous Weapons Systems, OpenDemocracy, London, 23 Apr 2015, 2.

[132] See Tallinn Manual on the International Law Applicable to Cyber Warfare (Michael N. Schmitt ed., 2013); San Remo Manual on International Law Applicable to Armed Conflicts at Sea (Louise Doswald Beck ed.,1995); and Program on Humanitarian Policy and Conflict Research, Manual on International Law Applicable to Air and Missile Warfare (2009).

[133] Crootof, *War Torts*, 28.

[134] Crootof, *The Killer Robots*, 44.

Department of Defense

DIRECTIVE

NUMBER 3000.09
~~November 21, 2012~~
Proposed Changes
Incorporating Change 1, May 8, 2017

USD(P)

SUBJECT: Autonomy in Weapon Systems

References: See Enclosure 1

1. PURPOSE. This Directive:

a. Establishes DoD policy and assigns responsibilities for the development and use of autonomous and semi-autonomous functions in weapon systems, including manned and unmanned platforms.

b. Establishes guidelines designed to minimize the probability and consequences of failures in autonomous and semi-autonomous weapon systems that could lead to unintended engagements.

c. Establishes domestic accountability measures in the event that an autonomous weapons system, while operating autonomously, commits a war crime.

2. APPLICABILITY. This Directive:

a. Applies to:

(1) OSD, the Military Departments, the Office of the Chairman of the Joint Chiefs of Staff and the Joint Staff (CJCS), the Combatant Commands, the Office of the Inspector General of the Department of Defense, the Defense Agencies, the DoD Field Activities, and all other organizational entities within the DoD (hereinafter referred to collectively as the "DoD Components").

(2) The design, development, acquisition, testing, fielding, and employment of autonomous and semi-autonomous weapon systems, including guided munitions that can independently select and discriminate targets.

(3) The application of lethal or non-lethal, kinetic or non-kinetic, force by autonomous or semi-autonomous weapon systems.

b. Does not apply to autonomous or semi-autonomous cyberspace systems for cyberspace operations; unarmed, unmanned platforms; unguided munitions; munitions manually guided by the operator (e.g., laser- or wire-guided munitions); mines; or unexploded explosive ordnance.

3. DEFINITIONS. See Glossary.

4. POLICY. It is DoD policy that:

a. Autonomous and semi-autonomous weapon systems shall be designed to allow commanders and operators to exercise appropriate levels of human judgment over the use of force. Appropriate levels of judgment require a human in the loop where possible, on the loop where practicable, and out of the loop only when time/space restrictions would prevent a weapon system from performing as designed with a human in a real-time decision loop.

(1) Systems will go through rigorous hardware and software verification and validation (V&V) and realistic system developmental and operational test and evaluation (T&E) in accordance with the guidelines in Enclosure 2. Training, doctrine, and tactics, techniques, and procedures (TTPs) will be established. Autonomous and semi-autonomous weapon systems that have been verified and validated may be employed by cognizant commanders in accordance with the established procedures and guidelines. A commander that properly employs an approved autonomous or semi-autonomous system will not be liable for a malfunction of that system. These measures will ensure that autonomous and semi-autonomous weapon systems:

(a) Function as anticipated in realistic operational environments against adaptive adversaries.

(b) Complete engagements in a timeframe consistent with commander and operator intentions and, if unable to do so, terminate engagements or seek additional human operator input before continuing the engagement.

(c) Are sufficiently robust to minimize failures that could lead to unintended engagements or to loss of control of the system to unauthorized parties.

(d) Enhance compliance with the law of armed conflict relative to a human operator.

(2) Consistent with the potential consequences of an unintended engagement or loss of control of the system to unauthorized parties, physical hardware and software will be designed with appropriate:

(a) Safeties, anti-tamper mechanisms, and information assurance in accordance with DoD ~~Directive~~ *Instruction* 8500.01~~E~~ (Reference (a)).

(b) Human-machine interfaces and controls.

(c) Self-destruct, self-destruction, or self-neutralization protocols.

(d) Ethical behavior controls.

(e) A margin of safety.

(3) In order for operators to make informed and appropriate decisions in engaging targets, the interface between people and machines for autonomous and semi-autonomous weapon systems shall:

(a) Be readily understandable to trained operators.

(b) Provide traceable feedback on system status.

(c) Provide clear procedures for trained operators to activate and deactivate system functions.

b. Persons who authorize the use of, direct the use of, or operate autonomous and semi-autonomous weapon systems must do so with appropriate care and in accordance with the law of war, applicable treaties, weapon system safety rules, and applicable rules of engagement (ROE).

c. The approval authority for the specific employment of approved autonomous and semi-autonomous weapons will be set at a level commensurate with the risk and capabilities of the weapon system. This authority may not be delegated further without approval by the Secretary of Defense.

d. Autonomous and semi-autonomous weapon systems intended to be used in a manner that falls within the policies in subparagraphs 4.c.(1) through 4.c.(3) will be considered for approval in accordance with the approval procedures in DoD Directive 5000.01 (Reference (b)), DoD Instruction 5000.02 (Reference (c)), and other applicable policies and issuances.

(1) Semi-autonomous weapon systems (including manned or unmanned platforms, munitions, or sub-munitions that function as semi-autonomous weapon systems or as subcomponents of semi-autonomous weapon systems) may be used to apply lethal or non-lethal, kinetic or non-kinetic force. Semi-autonomous weapon systems that are onboard or integrated with unmanned platforms must be designed such that, in the event of degraded or lost communications, the system does not autonomously select and engage individual targets or specific target groups that have not been previously selected by an authorized human operator.

(2) Human-supervised autonomous weapon systems may be used to select and engage targets, with the exception of selecting humans as targets, for local defense to intercept attempted time-critical or saturation attacks for:

(a) Static defense of manned installations.

(b) Onboard defense of manned platforms.

(3) Autonomous weapon systems may be used to apply non-lethal, non-kinetic force, such as some forms of electronic attack, against materiel targets in accordance with DoD Directive 3000.*0*3*E* (Reference (d)).

e. Autonomous or semi-autonomous weapon systems intended to be used in a manner that falls outside the policies in subparagraphs 4.c.(1) through 4.c.(3) must be approved by the Under Secretary of Defense for Policy (USD(P)); the Under Secretary of Defense for Acquisition, Technology, and Logistics (USD(AT&L)); and the CJCS before formal development and again before fielding in accordance with the guidelines in Enclosure 3, References (b) and (c), and other applicable policies and issuances.

f. International sales or transfers of autonomous and semi-autonomous weapon systems will be approved in accordance with existing technology security and foreign disclosure requirements and processes, in accordance with ~~Directive-Type Memorandum 11-053~~ *DoD Directive 5111.21* (Reference (e)).

g. Autonomous or semi-autonomous weapon systems without a human in the loop are categorically prohibited from carrying, employing, controlling, responding to, or targeting nuclear weapons or being designed specifically to do so.

RESPONSIBILITIES. See Enclosure 4.

RELEASABILITY. ~~UNLIMITED~~ ***Cleared for public release***. This Directive is ~~approved for public release and is~~ available on the ~~Internet from the~~ DoD Issuances Website at http://www.dtic.mil/whs/directives.

5. EFFECTIVE DATE. This Directive:~~-~~*is effective November 21, 2012.*

a. ~~Is effective November 21, 2012.~~

b. ~~Must be reissued, cancelled, or certified current within 5 years of its publication in accordance with DoD Instruction 5025.01 (Reference (f)). If not, it will expire effective November 21, 2022 and be removed from the DoD Issuances Website.~~

Name
Deputy Secretary of Defense

Enclosures
1. References
2. V&V and T&E of Autonomous and Semi-Autonomous Weapon Systems
3. Guidelines for Review of Certain Autonomous or Semi-Autonomous Weapon Systems
4. Responsibilities

Glossary

ENCLOSURE 1

REFERENCES

~~(a) DoD Directive 8500.01E, "Information Assurance (IA)," October 24, 2002~~
(a) DoD Instruction 8500.01, "Cybersecurity," March 14, 2014
(b) DoD Directive 5000.01, "The Defense Acquisition System," May 12, 2003
(c) DoD Instruction 5000.02, "Operation of the Defense Acquisition System," ~~December 8, 2008~~ *January 7, 2015, as amended*
(d) ~~DoD Directive 3000.3, "Policy for Non-Lethal Weapons," July 9, 1996~~
(d) DoD Directive 3000.03E, "DoD Executive Agent for Non-Lethal Weapons (NLW), and NLW Policy," April 25, 2013
(e) ~~Directive Type Memorandum (DTM) 11-053, "Technology Security and Foreign Disclosure (TS&FD) Processes," January 9, 2012~~ *DoD Directive 5111.21, "Arms Transfer and Technology Release Senior Steering Group and Technology Security and Foreign Disclosure Office," October 14, 2014*
(f) ~~DoD Instruction 5025.01, "DoD Directives Program," September 26, 2012~~
(~~g~~*f*) DoD Directive 2311.01E, "DoD Law of War Program," May 9, 2006*, as amended*
(~~h~~*g*) DoD Directive 1322.18, "Military Training," January 13, 2009*, as amended*

ENCLOSURE 2

V&V AND T&E OF AUTONOMOUS AND SEMI-AUTONOMOUS WEAPON SYSTEMS

To ensure autonomous and semi-autonomous weapon systems function as anticipated in realistic operational environments against adaptive adversaries and are sufficiently robust to minimize failures that could lead to unintended engagements or to loss of control of the system, in accordance with subparagraph 4.a.(1) above the signature of this Directive:

a. Systems will go through rigorous hardware and software V&V and realistic system developmental and operational T&E, including analysis of unanticipated emergent behavior resulting from the effects of complex operational environments on autonomous or semi-autonomous systems.

b. After initial operational test and evaluation (IOT&E), any further changes to the system will undergo V&V and T&E in order to ensure that critical safety features have not been degraded.

(1) A regression test of the software shall be applied to validate critical safety features have not been degraded. Automated regression testing tools will be used whenever feasible. The regression testing shall identify any new operating states and changes in the state transition matrix of the autonomous or semi-autonomous weapon system.

(2) Each new or revised operating state shall undergo integrated T&E to characterize the system behavior in that new operating state. Changes to the state transition matrix may require whole system follow-on operational T&E, as directed by the Director of Operational Test and Evaluation (DOT&E).

(3) If a system does not perform as anticipated, it is to be removed from the inventory, updated, and the V&V process must start over. Failure to remove a system from inventory after obtaining information that a system has not perform as designed will expose the cognizant commander to liability.

ENCLOSURE 3

GUIDELINES FOR REVIEW OF CERTAIN AUTONOMOUS OR SEMI-AUTONOMOUS WEAPON SYSTEMS

1. Autonomous or semi-autonomous weapon systems intended to be used in a manner that falls outside the policies in subparagraphs 4.c.(1) through 4.c.(3) above the signature of this Directive must be approved by the USD(P), USD(AT&L), and CJCS before formal development and again before fielding.

a. Before a decision to enter into formal development, the USD(P), USD(AT&L), and CJCS shall ensure:

(1) The system design incorporates the necessary capabilities to allow commanders and operators to exercise appropriate levels of human judgment in the use of force.

(2) The system is designed to complete engagements in a timeframe consistent with commander and operator intentions and, if unable to do so, to terminate engagements or seek additional human operator input before continuing the engagement.

(3) The system design, including safeties, anti-tamper mechanisms, intended context and manner of employment, and information assurance in accordance with Reference (a), addresses and minimizes the probability or consequences of failures that could lead to unintended engagements or to loss of control of the system.

(4) Plans are in place for V&V and T&E to establish system reliability, effectiveness, and suitability under realistic conditions, including possible adversary actions, to a sufficient standard consistent with the potential consequences of an unintended engagement or loss of control of the system.

(5) A preliminary legal review of the weapon system has been completed, in coordination with the General Counsel of the Department of Defense (GC, DoD) and in accordance with References (b) and (c), DoD Directive 2311.01E (Reference (~~g~~f)), and, where applicable, Reference (d).

b. Before fielding, the USD(P), USD(AT&L), and CJCS shall ensure:

(1) System capabilities, human-machine interfaces, doctrine, TTPs, and training have demonstrated the capability to allow commanders and operators to exercise appropriate levels of human judgment in the use of force and to employ systems with appropriate care and in accordance with the law of war, applicable treaties, weapon system safety rules, and applicable ROE.

(2) Sufficient safeties, anti-tamper mechanisms, and information assurance in accordance with Reference (a) have been implemented to minimize the probability or consequences of failures that could lead to unintended engagements or to loss of control of the system.

(3) V&V and T&E assess system performance, capability, reliability, effectiveness, and suitability under realistic conditions, including possible adversary actions, consistent with the potential consequences of an unintended engagement or loss of control of the system.

(4) Adequate training, TTPs, and doctrine are available, periodically reviewed, and used by system operators and commanders to understand the functioning, capabilities, and limitations of the system's autonomy in realistic operational conditions.

(5) System design and human-machine interfaces are readily understandable to trained operators, provide traceable feedback on system status, and provide clear procedures for trained operators to activate and deactivate system functions.

(6) A legal review of the weapon system has been completed, in coordination with the GC, DoD, and in accordance with References (b), (c), (~~g~~*f*), and, where applicable, Reference (d).

2. The USD(P), USD(AT&L), and CJCS may request a Deputy Secretary of Defense waiver for the requirements outlined in section 1 of this enclosure, with the exception of the requirement for a legal review, in cases of urgent military operational need.

ENCLOSURE 4

RESPONSIBILITIES

1. USD(P). The USD(P) shall:

a. Provide policy oversight for the development and employment of autonomous and semi-autonomous weapon systems.

b. In coordination with the USD(AT&L) and CJCS, review and consider for approval weapon systems submitted in accordance with paragraph 4.d. above the signature of this Directive.

c. Review, as necessary, the appropriateness of guidance established in accordance with this Directive given the continual advancement of new technologies and changing warfighter needs.

d. Approve the DoD position on international sales or transfers of autonomous and semi-autonomous weapon systems in accordance with existing technology security and foreign disclosure requirements and processes.

2. USD(AT&L). The USD(AT&L) shall:

a. Provide principal oversight responsibility for the establishment and enforcement of standards for testing, safety and reliability, hardware and software V&V, anti-tamper mechanisms, and information assurance in accordance with Reference (a), for autonomous and semi-autonomous weapon systems in order to minimize the probability and consequences of failures that could lead to unintended engagements or to loss of control of the system.

b. Provide principal oversight responsibility for the establishment of science and technology and research and development priorities for autonomy in weapon systems, including the development of new methods of V&V and T&E.

c. Oversee adequate developmental testing of autonomous and semi-autonomous weapon systems to assess the risk of failures that could lead to unintended engagements or to loss of control of the system.

d. In coordination with the USD(P) and CJCS, review and consider for approval weapon systems submitted in accordance with paragraph 4.d. above the signature of this Directive.

3. UNDER SECRETARY OF DEFENSE FOR PERSONNEL AND READINESS (USD(P&R)). The USD(P&R) shall, consistent with DoD Directive 1322.18 (Reference (~~h~~*g*)), oversee and provide policy for:

a. Individual military training programs for the Total Force relating to autonomous and semi-autonomous weapon systems.

b. Individual and functional training programs for military personnel and the collective training programs of military units and staffs relating to autonomous and semi-autonomous weapon systems.

4. DOT&E. The DOT&E shall:

a. Provide principal oversight responsibility for the development of realistic operational T&E standards for semi-autonomous and autonomous weapon systems, including standards for T&E of any changes to the system following IOT&E, in accordance with subparagraph 4.a.(1) above the signature of this Directive and Enclosure 2.

b. Evaluate whether semi-autonomous and autonomous weapon systems under DOT&E oversight have met sufficient V&V and T&E in realistic operational conditions, including potential adversary action, in order to minimize the probability and consequences of failures that could lead to unintended engagements or to loss of control of the system to unauthorized parties.

5. GC, DoD. The GC, DoD, shall, in accordance with References (b), (c), (~~g~~*f*), and, where applicable, Reference (d), provide for guidance in and coordination of legal reviews of weapon systems submitted in accordance with paragraph 4.d. above the signature of this Directive.

6. DEPARTMENT OF DEFENSE CHIEF INFORMATION OFFICER (DoD CIO). The DoD CIO, shall monitor, evaluate, and provide advice to the Secretary of Defense regarding information assurance for autonomous and semi-autonomous weapon systems, in accordance with subparagraph 4.a.(2)(a) above the signature of this Directive and Reference (a).

7. ASSISTANT TO THE SECRETARY OF DEFENSE FOR PUBLIC AFFAIRS (ATSD(PA)). The ATSD(PA) shall coordinate and approve guidance on public affairs matters concerning autonomous and semi-autonomous weapon systems and their use.

8. SECRETARIES OF THE MILITARY DEPARTMENTS; COMMANDER, U.S. SPECIAL OPERATIONS COMMAND (USSOCOM); AND THE HEADS OF THE DEFENSE AGENCIES AND DoD FIELD ACTIVITIES. The Secretaries of the Military Departments; the Commander, USSOCOM; and the Heads of the Defense Agencies and DoD Field Activities shall:

a. Develop and implement employment concepts, doctrine, experimentation strategies, TTPs, training, logistics support, V&V, anti-tamper mechanisms, physical hardware and software-level safeties, information assurance in accordance with Reference (a), and

developmental and operational T&E appropriate for autonomous and semi-autonomous weapon systems.

(1) Design autonomous and semi-autonomous weapon systems in such a manner as to minimize the probability and consequences of failures that could lead to unintended engagements or to loss of control of the system.

(2) Perform rigorous and realistic developmental and operational T&E and V&V, including T&E of any changes to the system following IOT&E, in accordance with subparagraph 4.a.(1) above the signature of this Directive and Enclosure 2.

(3) Design autonomous and semi-autonomous weapon systems with sufficient safeties, anti-tamper mechanisms, and information assurance in accordance with subparagraph 4.a.(2) above the signature of this Directive and Reference (a).

(4) Design human-machine interfaces for autonomous and semi-autonomous weapon systems to be readily understandable to trained operators, provide traceable feedback on system status, and provide clear procedures for trained operators to activate and deactivate system functions, in accordance with subparagraph 4.a.(3) above the signature of this Directive.

(5) Certify that operators of autonomous and semi-autonomous weapon systems have been trained in system capabilities, doctrine, and TTPs in order to exercise appropriate levels of human judgment in the use of force and employ systems with appropriate care and in accordance with the law of war, applicable treaties, weapon system safety rules, and applicable ROE.

(6) Establish and periodically review training, TTPs, and doctrine for autonomous and semi-autonomous weapon systems to ensure operators and commanders understand the functioning, capabilities, and limitations of a system's autonomy in realistic operational conditions, including as a result of possible adversary actions.

b. Ensure that legal reviews of autonomous and semi-autonomous weapon systems are conducted in accordance with References (b), (c), (~~g~~*f*) and, where applicable, Reference (d). Legal reviews should ensure consistency with all applicable domestic and international law and, in particular, the law of war.

c. Consider for support only those autonomous and semi-autonomous weapon systems that are technically feasible and that conform to this Directive. Submit to the USD(P), USD(AT&L), and CJCS for review, in accordance with paragraph 4.d. above the signature of this Directive, any autonomous or semi-autonomous weapon system intended to be used in a manner that falls outside the policies in subparagraphs 4.c.(1) through 4.c.(3) above the signature of this Directive before a decision to enter into formal development and again before fielding of any such system.

9. CJCS. The CJCS shall:

a. Advise the Secretary of Defense on the capability needs and employment of autonomous and semi-autonomous weapon systems.

b. Assess military requirements for autonomous and semi-autonomous weapon systems, including applicable key performance parameters and key system attributes.

c. Develop and publish joint doctrine, as appropriate, to incorporate emerging capabilities of autonomous and semi-autonomous weapon systems.

d. In coordination with the USD(P) and USD(AT&L), review and consider for approval autonomous weapon systems submitted in accordance with paragraph 4.d. above the signature of this Directive.

10. COMMANDERS OF THE COMBATANT COMMANDS. The Commanders of the Combatant Commands shall:

a. Use autonomous and semi-autonomous weapon systems in accordance with this Directive and in a manner consistent with their design, testing, certification, operator training, doctrine, TTPs, and approval as autonomous or semi-autonomous systems.

b. Employ autonomous and semi-autonomous weapon systems with appropriate care and in accordance with the law of war, applicable treaties, weapon system safety rules, and applicable ROE, in accordance with paragraph 4.b. above the signature of this Directive.

c. Ensure that weapon systems are not employed or modified to operate in a manner that falls outside the policies in subparagraphs 4.c.(1) through 4.c.(3) above the signature of this Directive without specific approval in accordance with paragraph 4.d. above the signature of this Directive.

d. Integrate autonomous and semi-autonomous weapon systems into operational mission planning.

e. Through the CJCS, identify warfighter priorities and operational needs that may be met by autonomous and semi-autonomous weapon systems.

GLOSSARY

PART I. ABBREVIATIONS AND ACRONYMS

ATSD(PA)	Assistant to the Secretary of Defense for Public Affairs
CJCS	Chairman of the Joint Chiefs of Staff
DoD CIO	Department of Defense Chief Information Officer
DOT&E	Director of Operational Test and Evaluation
GC, DoD	General Counsel of the Department of Defense
IOT&E	initial operational test and evaluation
ROE	rules of engagement
T&E	test and evaluation
TTP	tactics, techniques, and procedures
USD(AT&L)	Under Secretary of Defense for Acquisition, Technology, and Logistics
USD(P)	Under Secretary of Defense for Policy
USD(P&R)	Under Secretary of Defense for Personnel and Readiness
USSOCOM	U.S. Special Operations Command
V&V	verification and validation

PART II. DEFINITIONS

These terms and their definitions are for the purpose of this Directive.

automated regression testing. A type of regression testing that uses testing tools and repeatable test scripts.

autonomous weapon system. A weapon system that, once activated, can select and engage targets without further intervention by a human operator. This includes human-supervised autonomous weapon systems that are designed to allow human operators to override operation of

the weapon system, but can select and engage targets without further human input after activation.

electronic attack. Division of electronic warfare involving the use of electromagnetic energy, directed energy, or antiradiation weapons to attack personnel, facilities, or equipment with the intent of degrading, neutralizing, or destroying enemy combat capability and is considered a form of fires.

failures. An actual or perceived degradation or loss of intended functionality or inability of the system to perform as intended or designed. Failures can result from a number of causes, including, but not limited to, human error, human-machine interaction failures, malfunctions, communications degradation, software coding errors, enemy cyber attacks or infiltration into the industrial supply chain, jamming, spoofing, decoys, other enemy countermeasures or actions, or unanticipated situations on the battlefield.

human-supervised autonomous weapon system. An autonomous weapon system that is designed to provide human operators with the ability to intervene and terminate engagements, including in the event of a weapon system failure, before unacceptable levels of damage occur.

operating state. A variable or vector reflecting the status of the system.

operator. A person who operates a weapon system.

regression testing. A type of software testing that seeks to uncover new deficiencies (i.e., regressions) in the existing functional and non-functional areas of a system created by changes to the software, including enhancements, patches, emergency transports, or configuration changes.

semi-autonomous weapon system. A weapon system that, once activated, is intended to only engage individual targets or specific target groups that have been selected by a human operator. This includes:

Semi-autonomous weapon systems that employ autonomy for engagement-related functions including, but not limited to, acquiring, tracking, and identifying potential targets; cueing potential targets to human operators; prioritizing selected targets; timing of when to fire; or providing terminal guidance to home in on selected targets, provided that human control is retained over the decision to select individual targets and specific target groups for engagement.

"Fire and forget" or lock-on-after-launch homing munitions that rely on TTPs to maximize the probability that the only targets within the seeker's acquisition basket when the seeker activates are those individual targets or specific target groups that have been selected by a human operator.

state transition matrix. A matrix that characterizes the ability of a system to transition from one operating state to another.

target selection. The determination that an individual target or a specific group of targets is to be engaged.

unintended engagement. The use of force resulting in damage to persons or objects that human operators did not intend to be the targets of U.S. military operations, including unacceptable levels of collateral damage beyond those consistent with the law of war, ROE, and commander's intent.

unmanned platform. An air, land, surface, subsurface, or space platform that does not have the human operator physically onboard the platform.

BIBLIOGRAPHY

Arkin, Ronald C. Governing Lethal Behavior in Autonomous Robots (2009).

Asimov, Isaac. *I, Robot*. Garden City, NY: Doubleday. 1950.

Aupperlee, Aaron . "5 Reasons Pittsburgh is Still Tops in Autonomous Vehicles." *The Pittsburgh Tribune-Review*. July 21, 2017. http://www.govtech.com/fs/5-Reasons-Pittsburgh-is-Still-Tops-in-Autonomous-Vehicles.html.

Beaulieu, Ronald A. "Margin of Safety Definition and Examples used in Safety Basis Documents and the USQ Process." https://www.osti.gov/servlets/purl/1134068/. Last retrieved February 13, 2019.

Beck, Louise D Ed. San Remo Manual on International Law Applicable to Armed Conflicts at Sea (1995).

Boyd, John R. (3 September 1976). *Destruction and Creation (PDF)*. U.S. Army Command and General Staff College.

Boyle v. United Technologies Corp., 487 U.S. 500 (1988).

Carlton, Harlye. On Stable Ground: Remotely Operated Unmanned Ground Vehicles Enhancing Department of Defense Compliance with the Law of Armed Conflict. 2018. (unpublished, on file with author).

Cass, Kelly. Autonomous Weapons and Accountability: Seeking Solutions in the Law of War. 48 Loy. L.A. L. Rev. 1017, Spring 2015.

Clifford, Catherine. "9 of the most jaw-dropping things Elon Musk said about robots and AI in 2017." *CNBC*. December 18, 2017. https://www.cnbc.com/2017/12/18/9-mind-blowing-things-elon-musk-said-about-robots-and-ai-in-2017.html.

Counter-Rocket, Artillery, Mortar (C-RAM) Intercept Land-Based Phalanx Weapon System (LPWS). *US Army*. https://asc.army.mil/web/portfolio-item/ms-c-ram_lpws/. Last retrieved January 17, 2019.

Crootof, Rebecca. The Killer Robots are Here: Legal and Policy Implications. 36 Cardozo L. Rev. 1837, 1854, 12.

Crootof, Rebecca. Article: War Torts: Accountability for Autonomous Weapons. University of Pennsylvania Law Review. May 2016.

Davey, Tucker. Lethal Autonomous Weapons: An Update from the United Nations. April 30, 2018. https://futureoflife.org/2018/04/30/lethal-autonomous-weapons-an-update-from-the-united-nations/.

Declaration (IV,1), to Prohibit, for the Term of Five Years, the Launching of Projectiles and Explosives from Balloons, and Other Methods of Similar Nature. The Hague, 29 July 1899. https://ihl-databases.icrc.org/ihl/INTRO/160?OpenDocument.

Deeks, Ashley, Lubell, Noam, & Murray, Daragh. *Machine Learning, Artificial Intelligence,and the Use of Force by States.* 10 J. Nat'l Security L. & Pol'y __ (forthcoming 2019).

Demilitarized Zone. Encyclopedia Britannica. https://www.britannica.com/place/demilitarized-zone-Korean-peninsula. Last retrieved January 18, 2019.

Docherty, Bonnie. *Losing Humanity: The Case against Killer Robots* (Cambridge, MA: Human Rights Watch. November 19, 2012, 2, https://www.hrw.org/report/2012/11/19/losing-humanity/case-against-killer-robots.

Docherty, Bonnie. We're Running Out of Time to Stop Killer Robot Weapons. *Human Rights Watch*. April 11, 2018. https://www.hrw.org/news/2018/04/11/were-running-out-time-stop-killer-robot-weapons.

Etzioni, Amitai, PhD. and Oren Etzioni PhD. 2017. "Pros and Cons of Autonomous Weapons Systems." *Military Review* 97 (3): 72-81. https://search-proquest-com.lomc.idm.oclc.org/docview/1922376987?accountid=14746.

Faggell, Daniel. (All) Elon Musk Artificial Intelligence Quotes – A Catalogue of His Statements. November 29, 2018. https://emerj.com/ai-future-outlook/elon-musk-on-the-dangers-of-ai-a-catalogue-of-his-statements/.

Felton, Ryan. "Lockheed Martin's Autonomous Military Vehicles Aim To Save Lives In A Different Way." *FoxtrotAlpha.* February 18, 2017. Available at https://foxtrotalpha.jalopnik.com/lockheed-martins-autonomous-military-vehicles-aim-to-sa-1792128164.

Ford, Lieutenant Colonel Chrisopher M. Autonomous Weapons and International Law. *University of South Carolina Law Review*. 69 S. Car. Law Rev. 413.

Freedman, Lawrence. The First Two Generations of Nuclear Straegists. In *Makers of Modern Strategy from Machiavelli to the Nuclear Age*. Ed. By Peter Paret (Princeton, NJ: Princeton University, 1986).

Hall, Brian K. 2017. "Autonomous Weapons Systems Safety." *Joint Force Quarterly: JFQ* (86): 86-93. https://search-proquest-com.lomc.idm.oclc.org/docview/1916950387?accountid=14746.

Hoadley, Daniel S. and Lucas, Nathan J. Artificial Intelligence and National Security. The Congressional Research Service. April 26, 2018.

Hood, William D. Autonomous Weapons Systems: What Commanders Should Know. *The Marine Corps Gazette*. March 2015. 99.

International Committee for Robot Arms Control. Berlin Statement. October 2010. https://www.icrac.net/statements/.

Int'l Comm. of the Red Cross. *Weapons that May Cause Unnecessary Suffering or Have Indiscriminate Effects* (1973). Available at http://www.loc.gov/rr/frd/Military_Law/pdf/RC-Weapons.pdf.

Int'l Comm. of the Red Cross. Report of the ICRC Expert Meeting on "*Autonomous Weapons Systems*: Technical, Military, *Legal* and Humanitarian Aspects'. 26-28 March 2014, Geneva 2 (2014).

Israel Aerospace Industries. "Harpy NG." http://www.iai.co.il/Sip_Storage//FILES/5/41655.pdf. Last retrieved February 12, 2019.

Jai Galliot, *Military Robots: Mapping the Moral Landscape* (Routledge 2015).

K-Max, Lockheed Martin, available at https://www.lockheedmartin.com/en-us/products/k-max.html, last retrieved March 16, 2019.

Kastan, Benjamin Autonomous Weapons Systems; A Coming Legal "Singularity"?, 2013 U. Ill. J.L. tech. & Pol'y 45.

Kidder, Rushworth, M. *How Good People Make Tough Choices*. Rev. Ed. 2009, Harper (2003).

Law of Armed Conflict Deskbook. 5th Ed. *The Judge Advocate General's Legal Center and School.* 2017. http://www.loc.gov/rr/frd/Military_Law/pdf/LOAC-Deskbook-2015.pdf.

Losing Humanity: The Case Against Killer Robots, HUM. RTS. WATCH (Nov. 2012), http://www.hrw.org/sites/default/files/reports/arms1112ForUpload_0_0.pdf.

Minor, Elizabeth. Prohibiting Autonomous Weapons Systems. *OpenDemocracy*. London. 23 Apr 2015.

Noone, Gregory P. and Noone, Diana C. "The Debate over Autonomous Weapons Systems." *Case Western Reserve Journal of International Law* 47, no. 1 (Spring 2015): 29. http://scholarlycommons.law.case.edu/jil/vol47/iss1/6/.

Parakilas, Jacob and Xenia Wickett. 2017. *Transatlantic Rifts: Managing the use of Autonomous Weapons Systems*: Chatham House: The Royal Institute of International Affairs. https://search-proquest-com.lomc.idm.oclc.org/docview/1873244640?accountid=14746.

Phalanx Close-in Weapon System. *Raytheon.* https://www.raytheon.com/capabilities/products/phalanx. Last retrieved January 17, 2019.

Program on Humanitarian Policy and Conflict Research, Manual on International Law Applicable to Air and Missile Warfare (2009).

"Prohibiting Autonomous Weapons Systems." 2015. *OpenDemocracy*, Apr 23. https://search-proquest com.lomc.idm.oclc.org/docview/1677847866?accountid=14746.

Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of International Armed Conflicts, art. 35(2), June 8, 1977, 1125 U.N.T.S. 3.

Putin: Leader in Artificial Intelligence Will Rule the World. *CNBC*. September 4, 2017. https://www.cnbc.com/2017/09/04/putin-leader-in-artificial-intelligence-will-rule-world.html.

Rogers, Logan Jay. Case Note: Legal Judgment Day for the Rise of the Machines; a National Approach to Regulating Fully Autonomous Weapons. 56 Arizona Law Review 1257. 2014.

Reeves, Shane R. and William J. Johnson. 2014. "Autonomous Weapons: Are You Sure these are Killer Robots? can we Talk about it?" *The Army Lawyer*: 25-31. https://search-proquest-com.lomc.idm.oclc.org/docview/1540957074?accountid=14746.

Scharre, Paul. 2016. *Autonomous Weapons and Operational Risk*. Washington: Center for a New American Security. https://search-proquest-com.lomc.idm.oclc.org/docview/1834992075?accountid=14746.

Sauer, Frank. 2016. "Stopping 'Killer Robots': Why Now is the Time to Ban Autonomous Weapons Systems." *Arms Control Today* 46 (8): 8-13. https://search-proquest-com.lomc.idm.oclc.org/docview/1829055999?accountid=14746.

Scharre, Paul. *Autonomy "Killer Robots," and Human Control in the Use of Force--Part I*. JUST SECURITY. July 9, 2014. *http://justsecurity.org/12708/autonomy-killer-robots-human-control-force-part*.

Schmidt, Michael N. Ed. Tallinn Manual on the International Law Applicable to Cyber Warfare. Cambridge University Press, 2013.

Smith, Dan. 2015. "Stephen Hawking, Elon Musk Warn of 'Third Revolution in Warfare' with Autonomous Weapons." *ABC Premium News*, Jul 28. https://search-proquest-com.lomc.idm.oclc.org/docview/1699087903?accountid=14746.

Tamanaha, Brian Z. “Law and Society.” St. John’s University School of Law, Legal Studies Research Paper Series [Draft]. Available at The Social Science Research Network Electronic Paper Collection http://ssrn.com/abstract=13452042. Last retrieved March 16, 2019.

Randazzo, Ryan. “Waymo announces 'Waymo One,' but self-driving ride service isn't public — yet.” *Arizona Republic*. December 5, 2018. https://www.azcentral.com/story/money/business/tech/2018/12/05/waymo-one-launches-self-driving-car-service-arizona/2114688002/.

Ricter, W. D. *Stealth.* DVD. Directed by Rob Cohen. Los Angeles, Columbia. July 29, 2005.

Stoner, Robert H. “R2D2 with Attitude: The Story of the Phalanx Close-In Weapons.” *NavWeaps*. Available at http://www.navweaps.com/index_tech/tech-103.php, last retrieved January 18, 2019.

The Tallinn Manual Process. NATO Cooperative Cyber Defence Centre of Excellence. https://ccdcoe.org/tallinn-manual.html.

Thomas, Bradan T. “Autonomous Weapon Systems: The Anatomy of Autonomy and the Legality of Lethality.” 37 Hous. J. Int'l L. 235.

Toscano, Christopher P. “Friend of Humans”: An argument for Developing Autonomous Weapons Systems. 8 J. Nat’l Security L. & Pol’y 189.

Truver, Scott C. Taking Mines Seriously: Mine Warfare in China’s Near Seas. *Naval War College Review*. Vol 65. 2012. https://digital-commons.usnwc.edu/cgi/viewcontent.cgi?referer=https://www.google.com/&httpsredir=1&article=1429&context=nwc-review, 12.

U.S. Congress. House. FUTURE of Artificial Intelligence Act of 2017. HR 4625, 115th Cong. Introduced in House December 12, 2012. https://www.congress.gov/bill/115th-congress/house-bill/4625/related-bills?q=%7B%22search%22%3A%5B%22H.+R.+83%22%5D%7D12/12/2017

U.S. Congress. Senate. FUTURE of Artificial Intelligence Act of 2017. HR 4625, 115th Cong. Introduced in Senate December 12, 2012. https://www.congress.gov/bill/115th-congress/house-bill/4625/related-bills?q=%7B%22search%22%3A%5B%22H.+R.+83%22%5D%7D12/12/2017.

United States of America. “Humanitarian Benefits of Emerging Technologies in the Area of Lethal Autonomous Weapon Systems.” *Group of Governmental Experts of the High Contracting Parties to the Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons which May be Deemed to be Excessively Injurious or to Have Indiscriminate Effects*. April 9-13, 2018.

United States Department of Defense Directive 3000.09. “Autonomy in Weapon Systems.” 8 May 2017.

U.S. Department of Defense, *Nuclear Posture Review*, Office of the Secretary of Defense Washington, DC: Pentagon, February 2018.