# ALEXA, WHOSE FAULT IS IT? AUTONOMOUS WEAPON SYSTEMS INVESTIGATIONS AND THE IMPORTANCE OF A DELIBERATE ACCOUNTABILITY PROCESS

*Major Thomas G. Warschefsky*[*]

*For unto whomsoever much is given, of him shall be much required: and to whom men have committed much, of him they will ask for more.*[1]

## I. Introduction

### A. Hypothetical

It is sometime in the future and the United States (U.S.) military is engaged in a combat operation. During this operation, an Army brigade commander deems it prudent to utilize an autonomous weapon system (AWS)—known as "Weapon X"—to target enemy troops. Weapon X is an aerial platform designed to loiter in a given location while searching for targets, and it is pre-loaded with data to identify and target enemy vehicles, to include armored personnel carriers.[2] On the day in question, the

---

[*] Judge Advocate, United States Army. Presently assigned to United States Army Special Operations Command, Fort Bragg, North Carolina. L.L.M., 2019, The Judge Advocate General's School, United States Army, Charlottesville, Virginia. J.D., 2009, Thomas M. Cooley School of Law; B.A., 2006, Michigan State University. Previous assignments include Officer in Charge, Hohenfels Law Center, Hohenfels, Germany, 2016-2018; Battalion Judge Advocate, Group Support Battalion and 4th Battalion, 7th Special Forces Group (Airborne), Eglin Air Force Base, Florida, 2013-2016; Senior Trial Counsel, U.S. Army Alaska, Joint Base Elmendorf-Richardson, Alaska, 2013; Trial Counsel, 4th Brigade (Airborne), 25th Infantry Division, Joint Base Elmendorf-Richardson, Alaska and Forward Operating Base Salerno, Afghanistan, 2011-2012; Administrative and Operational Law Attorney, U.S. Army Alaska, Joint Base Elmendorf-Richardson, Alaska, 2010-2011. Member of the bar of Michigan. This paper was submitted in partial completion of the Master of Laws requirements of the 67th Judge Advocate Officer Graduate Course.

[1] *Luke* 12:48 (King James); *see also* Stan Lee & Steve Ditko, *Spiderman*, AMAZING FANTASY 15, at 13 (Marvel Entertainment Aug. 1962) ("In this world, with great power there must also come—great responsibility.").

[2] *See generally HARPY Autonomous Weapon for All Weather*, ISRAEL AEROSPACE INDUSTRIES,

commander authorizes Weapon X to deploy to an area where enemy troops may be operating. Although operated in a "human on the loop" capacity, enemy utilization of electromagnetic warfare has greatly restricted the ability of Weapon X to transmit video feed to the command center.[3] As a result, the Soldiers monitoring Weapon X are only able to receive written target analysis conclusions from Weapon X.

At some point after deployment, Weapon X submits a message to the command center: Weapon X has identified an armored personnel carrier and is prepared to strike the target. The Commander has reason to believe armored personnel carriers may be present in the area and, based on this information, allows Weapon X to continue its strike. The target is destroyed. The team later learns the target was a civilian van, and ten children were killed.

In the aftermath, the higher command initiates an administrative investigation into the incident in accordance with Army Regulation 15-6.[4] This investigation examines the commander and those working with the AWS on the date of the incident. It finds that their actions were appropriate based on the information provided by the AWS. Having looked at their actions, the investigation next turns to the AWS itself.

It is at this point that the investigating officer (IO) has difficulty. Despite valiant efforts, the IO has limited experience in computer programming. No individuals within the combat division have the in-depth experience necessary to examine the AWS's designs. Moreover, the system was developed in a collaborative effort between the United States Defense Advanced Research Projects Agency (DARPA) and a private corporation and, although helpful, neither seems particularly motivated to expeditiously provide assistance, as the investigation is coming from well

---

http://www.iai.co.il/2013/36694-16153-en/Business_Areas_Land.aspx (last visited Mar. 14, 2019).

[3] *See* PAUL SCHARRE, ARMY OF NONE: AUTONOMOUS WEAPONS AND THE FUTURE OF WAR 44, 81–82 (2018) (discussing how the option of real-time monitoring of weapon systems is likely to be extremely limited or non-existent if a conflict involving a near-peer with significant capabilities in the electromagnetic spectrum that would allow for disruption of communications links).

[4] *See generally* U.S. DEP'T OF ARMY, REG. 15-6, PROCEDURES FOR ADMINISTRATIVE INVESTIGATIONS AND BOARDS OF OFFICERS (1 Apr. 2016) [hereinafter AR 15-6].

outside their organizational chains of command.[5] With nowhere to turn and the deadline approaching, the IO is forced to conclude that although the commander is not responsible for the deaths of the children, she is unable to determine who—or what—is.

## B. Background

The idea of artificial intelligence (AI) has existed in popular culture since as early as 1920.[6] While some fictional accounts place AI as a great boon to society, others explore its darker side.[7] Today, what was once reserved for the realm of science fiction has entered our everyday lives. Autonomous robotic vacuums clean our houses,[8] and "smart" thermostats control our living environments.[9] Robotic personal assistants, such as Amazon's "Alexa," listen to our day-to-day lives in order to answer questions, play music, or place orders with online retailers,[10] and AI is being tested to drive our cars and pilot commercial airlines.[11] At the same time, the potential of AI has not escaped the watchful eye of militaries throughout the world.

According to Russian President Vladimir Putin, "The one who becomes the leader in [the AI] sphere will be the ruler of the world. When one party's drones are destroyed by drones of another, [that party] will have no other choice but to surrender."[12] Other world powers have taken

---

[5] Both the U.S. C-RAM LPWS and the Israeli Harpy weapon systems were developed in conjunction with private contractors. It is reasonable to assume private business will have heavy involvement in future autonomous weapon systems (AWS) development.

[6] *See, e.g.*, KAREL CAPEK, ROSSUM'S UNIVERSAL ROBOTS (1920).

[7] *See, e.g.*, IROBOT (Davis Entertainment 2004); *see also* THE TERMINATOR (Hemdale 1984); THE MATRIX (Warner Bros. 1999); and THE STAR WARS TRILOGY (Lucasfilm 1977, 1980, 1983).

[8] *See, e.g.*, *Roomba Robot Vacuum*, IROBOT, https://www.irobot.com/for-the-home/vacuuming/roomba (last visited Mar. 14, 2019).

[9] *See, e.g.*, *Nest Learning Thermostat*, NEST, https://nest.com/thermostats/nest-learning-thermostat/overview/ (last visited Mar. 14, 2019).

[10] *See, e.g.*, *Echo and Alexa*, AMAZON, https://www.amazon.com/Amazon-Echo-And-Alexa-Devices/b?ie=UTF8&node= 9818047011 (last visited Mar. 11, 2019).

[11] *See, e.g.*, *Our Mission*, WAYMO, https://waymo.com/mission/ (last visited Mar. 14, 2019) (explaining the mission of an autonomous vehicle company).

[12] Russ. President Vladimir Putin, Address to Students at the Beginning of the 2017 School Year (Sep. 1, 2017).

notice of the huge potential of AI as a warfighting tool and are exploring the role autonomous systems will have in the future of combat. This exploration is not merely conceptual. The United States has developed and implemented the Phalanx series of active defense systems (to include the Counter-Rocket Artillery and Mortar or C-RAM) that demonstrate autonomous capabilities.[13] Israel has operationalized the Harpy autonomous drone utilized to hunt and destroy enemy radar stations.[14] Likewise, Russia has publicized their cultivation of autonomous tanks,[15] and China has recently indicated their intent to explore autonomous drone swarms.[16]

While many have recognized the military advantages offered by AWS, many government and non-governmental organizations have taken a negative view of this emerging technology. This has led to a spirited debate on the morality and legality of AWS, with many organizations calling for outright bans.[17] Although many concerns have not stood up to scrutiny, the concern regarding potential inability to assign human blame for collateral damage remains a primary argument for the ban of AWS.[18] As policies are developed on the national and international levels, this concern over lack of human accountability could severely limit the United States' ability to develop autonomous weapon systems and creates the

---

[13] *Counter-Rocket, Artillery, Mortar (C-RAM) Intercept Land-Based Phalanx Weapon System (LPWS)*, U.S. ARMY ACQUISITION SUPPORT CTR., https://asc.army.mil/web/portfolio-item/ms-c-ram_lpws/ (last visited Mar. 14, 2019).

[14] ISRAEL AEROSPACE INDUSTRIES, *supra* note 2.

[15] Daniel Brown, *Russia says it has deployed its Uran-9 robotic tank to Syria—here's what it can do*, BUSINESS INSIDER (May 15, 2018), https://www.businessinsider.com/russia-uran-9-robot-tank-what-can-it-do-syria-2018-5#heres-a-view-from-the-automatic-turret-which-can-detect-and-acquire-targets-on-its-own-up-to-about-four-miles-away-during-the-day-the-operator-however-controls-the-firing-6.

[16] Elsa Kania, *China's Strategic Ambiguity and Shifting Approach to Lethal Autonomous Weapons Systems*, LAWFARE (Apr. 17, 2018), https://www.lawfareblog.com/chinas-strategic-ambiguity-and-shifting-approach-lethal-autonomous-weapons-systems.

[17] *See, e.g.*, *A Growing Global Coalition*, CAMPAIGN TO STOP KILLER ROBOTS, https://www.stopkillerrobots.org/about/ (last visited Mar. 14, 2019) ("The Campaign to Stop Killer Robots is a growing global coalition of 100 international, regional, and national non-governmental organizations...in 54 countries that is working to preemptively ban fully autonomous weapons.") *See also* European Parliament Resolution of 12 September 2018 on Autonomous Weapon Systems, EUR. PARL. DOC. 2018/2752(RSP) (2018) (Adopting "[a]n EU common position on lethal autonomous weapon systems that ensures meaningful human control over the critical functions of weapon systems.").

[18] *See, e.g.*, Tyler D. Evans, *At War with the Robots: Autonomous Weapons Systems and the Martens Claus*e, 41 HOFSTRA L. REV. 697 (2013).

potential to restrict our ability to compete in an ever-changing military environment.[19]

Our ability to use and develop unencumbered AWS requires us to address concerns related to a lack of human accountability in AWS. In order to establish human accountability, we must create a system that allows for efficient and effective investigations into incidents involving AWS and allows for assignment of human responsibility for AWS actions when necessary. After providing a basic understanding of AWS, this article discusses the necessity of accountability within AWS and provides an outline for a deliberate system of responsibility within AWS creation and utilization. This article also identifies the requirement to conduct investigations into AWS incidents and concludes with recommendations for the design and implementation of an AWS investigative system designed to properly assign accountability for AWS incidents.

## II. Understanding Artificial Intelligence and Autonomous Weapon Systems

### A. Artificial Intelligence and Deep Learning

In order to understand issues within AWS investigations, one must first understand some key facets of programming AI. Generally, programing methodologies for AI fall somewhere within a spectrum of practices.[20] On one side of the spectrum, human programmers manually enter code to create a system of logical "decision trees" that a machine must follow. These designers "thought it made the most sense to build machines that reasoned according to rules and logic, making their inner workings transparent to anyone who cared to examine some code."[21] On the other side of the spectrum are programs that:

---

[19] DEF. INNOVATION BD., U.S. DEP'T OF DEF., AI PRINCIPLES: RECOMMENDATIONS ON THE ETHICAL USE OF ARTIFICIAL INTELLIGENCE BY THE DEPARTMENT OF DEFENSE 2, 3 (2019) [hereinafter DIB AI PRINCIPLES], https://media.defense.gov/2019/Oct/31/2002204458/1/1/0/DIB_AI_PRINCIPLES_PRIMARY_DOCUMENT.PDF.

[20] David Gunning, Defense Advanced Research Projects Agency, Explainable Artificial Intelligence (XAI) Program Update, November 2017, at slide 9, 10 (2017) (published PowerPoint presentation), https://www.darpa.mil/attachments /XAIProgramUpdate.pdf.

[21] Will Knight, *The Dark Secret at the Heart of AI*, 120 MIT TECH. REV. 54, 57 (2017).

> [take] inspiration from biology, and [learn] by observing and experiencing. This mean[s] turning computer programming on its head. Instead of a programmer writing the commands to solve a problem, the problem generates its own algorithm based on example data and a desired output. The machine-learning techniques that would later evolve into today's most powerful AI systems followed the latter path: the machine essentially programs itself.[22]

These machine-learning techniques, known as "neural networks" and "deep learning," present serious considerations in investigations of AWS, centering on the idea that "[n]o one really knows how the most advanced algorithms do what they do."[23] "The computers...have programmed themselves, and they do it in ways we cannot understand. Even the engineers who build these apps cannot fully explain their behavior."[24] Thus, while "[a]lgorithmic transparency means you can see how the decision is reached...you can't with [machine-learning] systems because it's not rule-based software."[25] Indeed, this method of programing is unique enough that some experts take effort to distinguish these machine-learning techniques from other AI systems.[26]

B. Autonomous Weapon Systems

In addition to a fundamental understanding of AI, it is important for one to have a basic definition for and understanding of AWS. While the Department of Defense (DoD) defines AWS as "[a] weapon system that, once activated, can select and engage targets without further intervention by a human operator,"[27] this definition is overly simplistic as it fails to

---

[22] *Id.*
[23] *Id.* at 55.
[24] *Id.* at 56.
[25] David Meyer, *AI Has a Big Privacy Problem and Europe's New Data Protection Law Is About to Expose It,* FORTUNE (May 25, 2018), http://fortune.com/2018/05/25/ai-machine-learning-privacy-gdpr/ (citation omitted).
[26] DIB AI PRINCIPLES, *supra* note 19, at 5 ("When referring to the wider range of considerations, we use the term artificial intelligence (AI); however, where we specifically address machine learning (ML) systems, we refer to ML.").
[27] U.S. DEP'T OF DEF., DIR. 3000.09, AUTONOMY IN WEAPON SYSTEMS 13 (5 Aug. 2017) [hereinafter DODD 3000.09].

adequately distinguish AWS from automated weapons.[28] For example, anti-tank land mines or naval mines that identify appropriate targets based on weight, infra-red, magnetic, or acoustic signature would be in included in this definition of AWS, despite the fact that they have existed for decades.[29] In fact, the DoD recognizes the weakness in its classification by excluding certain items—including mines—from the definition.[30] This is proper because "[i]n contrast to these purely reactive systems, autonomous weapon systems gather and process data from their environment to reach independent conclusions about how to act."[31] As a result, instead of the DoD definition, a better definition of AWS is "a weapon system that, based on conclusions derived from gathered information and preprogrammed constraints, is capable of independently selecting and engaging targets."[32]

Many authorities further the discussion of autonomous systems by considering three sub-categories of weapons with varying levels of autonomous characteristics.[33] First, "semiautonomous weapon systems" utilize automation for many tasks but still require human interface in the target decision process. Thus, while the weapon system itself may identify and classify targets, a human operator remains in the "kill chain" and human authorization is required prior to firing of the weapon. For this reason, semiautonomous weapon systems are often referred to as "human in the loop" systems.[34] Importantly, many experts on AWS, including the

---

[28] Rebecca Crootof, *War Torts: Accountability for Autonomous Weapons*, 164 U. PA. L. REV. 1349, 1367 (2016) [hereinafter Crootof, *War Torts*].

[29] *See, e.g.*, Jon Rabiroff, *U.S. Military Enters New Generation of Sea Mine Warfare*, STARS AND STRIPES (May 9, 2011), https://www.stripes.com/news/u-s-military-enters-new-generation-of-sea-mine-warfare-1.143170. *See also* Anti-Vehicle (Anti-Tank) Mines, Technical Director Geneva International Center for Humanitarian Demining, at slide 18-22 (2002) (published PowerPoint presentation), https://www.gichd.org/fileadmin/GICHD-resources/rec-documents/ERW_AV_AT_Mines.pdf.

[30] DODD 3000.09, *supra* note 27, para. 2b.

[31] Crootof, *War Torts*, *supra* note 28.

[32] Rebecca Crootof, *The Killer Robots Are Here: Legal and Policy Implications*, 36 CARDOZO L. REV. 1837, 1842 (2015).

[33] Crootof, *War Torts*, *supra* note 28. *See also* Michael Press, *Of Robots and Rules: Autonomous Weapons Systems in the Law of Armed Conflict*, 48 GEO. J. OF INT'L L. 1337, 1339–1342 (2017); SCHARRE, *supra* note 3, at 44.

[34] SCHARRE, *supra* note 3, at 44.

DoD, do not include semiautonomous weapon systems in their definition of AWS.[35]

The next category refers to systems that involve human supervision of the weapon but do not require human permission to act. Known as "human on the loop" systems, or "supervised autonomous weapon systems," these systems act largely of their own accord, but in a supervised manner. Although humans monitor these systems and remain available to react in real time should a mishap be identified, their permission is not needed for the AWS to act.[36]

Finally, "fully autonomous weapon systems," or "human off the loop" systems, operate in a manner entirely without human intervention.[37] These systems would be deployed and have the ability to search for, identify, categorize, and carry out an attack without further human involvement.[38]

## III. Accountable Artificial Intelligence

### A. Accountability Concerns

When fused with deep-learning AI, the concept of AWS leads to many concerns regarding lack of accountability. As the Campaign to Stop Killer Robots contends:

> The use of fully autonomous weapons would create an accountability gap as there is no clarity on who would be legally responsible for a robot's actions: the commander, programmer, manufacturer, or robot itself? Without accountability, these parties would have less incentive to ensure robots did not

---

[35] DODD 3000.09, *supra* note 27, at 14 (Defining a semiautonomous weapon system as "[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." Fire and Forget munitions are included in this definition.).
[36] SCHARRE, *supra* note 3, at 45.
[37] *Id.* at 46.
[38] *Id.* at 81–82.

> endanger civilians, and victims would be left unsatisfied that someone was punished for the harm they experienced.[39]

While this potential lack of transparency causes distrust for some, those concerns are misplaced. To understand this, one must briefly dissect how the concepts of explainability and responsibility relate to accountability of AWS.

Explainability in AI seeks to solve the problem that "[c]ertain algorithms act as a 'black box,' where it is impossible to determine how the output was produced..."[40] The argument holds that "[b]y exposing the logic behind a decision, explanation can be used to prevent errors and increase trust."[41] Nevertheless, while explainability in AI is an important feature (and one that considerable resources are being leveraged to solve),[42] it is not required to establish accountability. An illustration of this is provided by the widespread use of animals in the military, such as working dogs.[43]

B. (Un)Explainable AI

In many ways, military working dogs act in a semiautonomous or fully autonomous manner.[44] Like AWS, military working dogs possess a significant amount of autonomy but "[t]heir independence is tempered through extensive training; [and] their propensity for unpredictable action

---

[39] *The Problem*, CAMPAIGN TO STOP KILLER ROBOTS, https://www.stopkillerrobots.org/learn/#problem (last visited Mar. 14, 2019).

[40] Chamith Fonseka, *Hold Artificial Intelligence Accountable*, HARV. U. SCI. IN THE NEWS (Aug. 28, 2017), http://sitn.hms.harvard.edu/flash/2017/hold-artificial-intelligence-accountable/.

[41] Finale Doshi-Velez & Mason Kortz, *Accountability of AI Under the Law: The Role of Explanation* 2 (Berkman Klein Ctr. Working Grp. on Explanation and the Law, Berkman Klien Ctr. for Internet and Soc'y Working Paper, 2017), http://nrs.harvard.edu/urn-3:HUL.InstRepos:34372584.

[42] *See generally* Gunning, *supra* note 20.

[43] Linda Crippen, *Military Working Dogs: Guardians of the Night*, U.S. ARMY NEWS (May 23, 2011), http://www.army.mil/article/56965/Military_Working_Dogs__Guardians_of_the_ Night.

[44] Major Charles T. Kirchmaier, *Unleashing the Dogs of War: Using Military Working Dogs to Apprehend Enemy Combatants*, ARMY LAW., Oct. 2006, at 4; *see also* Aiden Warren and Alek Hillas, *Lethal Autonomous Weapons Systems: Adapting to the Future of Unmanned Warfare and Unaccountable Robots*, 12 YALE J. OF INT'L AFF. 71, 75–79 (2017).

is addressed through limited use." [45] Despite their autonomous characteristics, the legal analysis of animals in armed conflict is limited to Protocol II of the Convention on Certain Conventional Weapons, which prohibits the use of animal-borne booby-traps or other devices.[46] This should lead one to consider "[w]hat then, would happen if an animal combatant were to take an action that resulted in what seemed to be a serious violation of international humanitarian law?"[47]

To remedy this, some remove explainability from the equation and suggest an analysis based on the responsibility of the human handlers.[48] Indeed, as there are no requirements under international law to attribute explainability for the actions of animals in warfare, examining responsibility of associated humans is a logical method of ensuring accountability.

### C. Human Responsibility

Likewise, accountability in AWS should focus less on explainability and more on human responsibility. The assignment of human responsibility can be premised on the fact that just as military working dogs are not truly autonomous since they rely on a handler to operate, AI will never be completely autonomous. Indeed, "[n]o entity—and for that matter, no person—is capable enough to be able to perform competently in every task and situation. On the other hand, even the simplest machine can seem to function 'autonomously' if the task and context are sufficiently constrained." [49] Put differently, "there exist no fully autonomous systems, just as there are no fully autonomous soldiers, sailors, airmen or Marines."[50] Given this understanding, one can begin to envision how AWS responsibility can be established. Much like military

---

[45] Rebecca Crootof, *Autonomous Weapons Systems and the Limits of Analogy*, 9 HARV. NAT'L SECURITY J. 51, 78 (2018) [hereinafter Crootof, *Limits of Analogy*].

[46] Protocol on Prohibitions or Restrictions on the Use of Mines, Booby-Traps, and Other Devices (Protocol II) art. 7(1), Oct. 10, 1980 S. TREATY DOC. No 105-1, 2048 U.N.T.S. 133 (amended May 3, 1996). *See also* Crootof, *Limits of Analogy*, *supra* note 46, at 77.

[47] Crootof, *Limits of Analogy*, *supra* note 46, at 77.

[48] Karsten Nowrot, *Animals at War: The Status of 'Animal Soldiers' Under International Humanitarian Law*, 40 HIST. SOC. RES. 128, 142 (2015).

[49] Robert R. Hoffman, *The Seven Deadly Myths of Autonomous Systems*, 28 IEEE INTELLIGENT SYS. 1541, 1545 (2013).

[50] DEF. SCI. BD., U.S. DEP'T OF DEF., TASK FORCE REPORT: THE ROLE OF AUTONOMY IN DOD SYSTEMS 23 (2012), https://fas.org/irp/agency/dod/dsb/autonomy.pdf.

parachute riggers annotate responsibility for each phase of the parachute packing and inspection process,[51] the AWS design and implementation process should annotate and designate human responsibility for the phases of AWS creation and use.[52] In other words, human responsibility for AWS must be *traceable*.[53]

To determine when and where traceable human responsibility may be interjected in AWS, it is helpful to consider the defense acquisition framework, which is utilized for the procurement of defense materials.[54] Under this framework, acquisition of an item follows one of six acquisition pathways, based on the particular item to be procured and the urgency of the need.[55] Although the terminology used for the phases of various acquisition pathways differs, two of the phases discussed in the Major Capability Acquisition pathway provide an outline to discuss traceable human responsibility in AWS that can be translated to other acquisition strategies.

To begin, the Engineering and Manufacturing Development phase of the Major Capability Acquisition pathway offers three opportunities for establishment of responsibility. The first opportunity is when program requirements are set, evaluated, and approved. While establishing formal responsibility during this phase of an acquisition may be unnecessary for traditional weapon systems,[56] AWS program requirements will require much greater detail as they encroach on decisions that have been

---

[51] *See generally* U.S. Dep't of Army, DA Form 3912, Army Parachute Log Record (1 June 1979); U.S. DEP'T OF ARMY, REG. 59-4, JOINT AIRDROP INSPECTION RECORDS, MALFUNCTION INVESTIGATIONS, AND ACTIVITY REPORTING {OPNAVINST 4630.24D; AFJ 13 210(I); MCO 13480.1C} (8 Apr. 2008) (RAR 23 June 2009) [hereinafter AR 59-4].

[52] A complete discussion on legalities of imputing civilian contractor liability for potential Law of War violations resulting from AWS use is outside the scope of this paper. This issue could be resolved by ensuring the "persons responsible" for key portions of the AWS acquisition process are members of the military.

[53] DIB AI PRINCIPLES, *supra* note 19, at 8 ("DoD's AI engineering discipline should be sufficiently advanced such that technical experts possess and appropriate understanding of the technology, development process, and operational methods of its AI systems, including transparent and auditable methodologies, data sources, and design procedure and documentation.").

[54] *See generally* U.S. DEP'T OF DEF., DIR. 5000.02, OPERATION OF THE DEFENSE ACQUISITION FRAMEWORK (23 January 2020) [hereinafter DODD 5000.02].

[55] *Id.* at 9.

[56] For example, the requirement that a precision guided munition be able to strike a given location with a high degree of accuracy does not necessitate a complex analysis of the Law of War to be incorporated into the design of the munition.

traditionally made on the battlefield. Specifically, requirements must include the ability for an AWS to comply with law of war principles, such as distinction,[57] proportionality,[58] and military necessity[59] during operations.[60] Because this ability to comply with law of war principles is an essential task, forming the backbone of lawful AWS use, it is critical that responsibility is established for this portion of the AWS procurement process.

The second opportunity for responsibility within the Engineering and Manufacturing Development phase is found in the design and production of the item.[61] At this time of the acquisition process, a designated individual should attest to the accuracy of the computer programming utilized to achieve the specific AWS requirement. As these requirements will include compliance with law of war principles, this person must be able to attest to the accuracy with which the AWS complies with these requirements.

Third, responsibility should be designated in the testing and validation portion of the Engineering and Manufacturing Development phase of the Major Capability Acquisition pathway.[62] While methods of testing weapons systems are generally well established, designating responsibility at this stage will ensure testing and validation utilize the best available

---

[57] U.S. DEP'T OF DEF., DOD LAW OF WAR MANUAL para. 2.5 (May 2016) [hereinafter LAW OF WAR MANUAL]. *See also* Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of International Armed Conflicts, adopted June 8, 1977, 1125 U.N.T.S. 3, entered into force December 7, 1978, art. 48, 51(4) [hereinafter Protocol I].

[58] LAW OF WAR MANUAL, *supra* note 57, para.2.4. *See also* Protocol I, art. 51(5)(b).

[59] LAW OF WAR MANUAL, *supra* note 57, para. 2.2

[60] *But see* HUM. RTS. WATCH & INT'L HUM. RTS. CLINIC, HARV. L. SCH., LOSING HUMANITY: THE CASE AGAINST KILLER ROBOTS 30–36 (2012) (arguing that it will be impossible for AWS to comply with the Laws of War). It is the author's opinion that these arguments are conclusory and subject to challenge as technology advances. Autonomous weapon systems are likely able to conduct—at a minimum—a conservatively accurate analysis of an engagement that complies with these principles. For example, an AWS could be designed such that it only targets enemy tanks firing in the open, located on the enemy side of the forward line of troops, where there are no living objects within a given safety radius of the target.

[61] DODD 5000.02, *supra* note 54, at 11.

[62] *Id.*

efforts to examine the unique characteristics of an AWS prior to its validation as a weapons system.[63]

Lastly, a system of responsibility must include the final stage of the procurement process: Deployment of the AWS.[64] As with conventional weapons, this phase must assign responsibility for utilization of an AWS to commanders and individual end-users of the item. Although establishing a chain of responsibility along these constructs is arduous, it is necessary to take these deliberate actions in order to ultimately provide the structure to allow accountability of AWS through investigations.

## IV. Investigative Considerations

### A. Requirement to Investigate

It can be expected that accountability for AWS will be established through investigations, as inquiries into use of force by the U.S. military take place in formal and informal manners on a regular basis. By policy, U.S. military forces must evaluate "the overall effectiveness of employing joint force targeting capabilities during military operations."[65] Known as a "Combat Assessment," these inquiries into the effects of a targeting operation include conducting a Battle Damage Assessment (BDA) which determines, among other things, if a strike resulted in "unintentional or incidental injury or damage to persons or objects that would not be lawful military targets in the circumstances ruling at the time."[66] Unwarranted or unexpected collateral damage identified in the BDA (or identified by other sources such as reports from media) often becomes the driver of follow-on investigations.[67]

Although "[u]nder the current state of IHL (International Humanitarian Law), there is no express requirement placing states under

---

[63] *See generally*: U.S. DEP'T OF ARMY, REG. 73-1, TEST AND EVALUATION POLICY (16 Nov. 2016).
[64] DODD 5000.02, *supra* note 54, at 11.
[65] JOINT CHIEFS OF STAFF, JOINT PUB. 3-60, JOINT TARGETING app. D, sec. 1.a (28 Sep. 2018) [hereinafter JP 3-60].
[66] *Id.* app. D, sec. 1.a.5.
[67] U.S. DEP'T OF DEF., DIR. 2311.01E, DOD LAW OF WAR PROGRAM para. 3.2 (22 Feb. 2011) [hereinafter DODD 2311.01E] (requiring investigation of "[a]ll possible...violation(s) of the law of war, for which there is credible information").

a duty to investigate all strikes resulting in civilian losses,"[68] it is widely accepted that states are required to prevent and prosecute grave breaches of IHL.[69] "In order to discharge the obligation to prosecute those who commit grave breaches, a state must *ipso facto* conduct credible investigations that could, if warranted, lead to prosecutions."[70] Further, some argue that investigations into breaches that amount to less than grave breaches of IHL can "be deduced from articles 1 and 146 of [the Fourth Geneva Convention] as well as from articles 1 and 87(3) of [Additional Protocol] I."[71] This theory is based on the assertion that "IHL creates an obligation to penalize all kinds of breaches and not only those which qualify as grave."[72] The obligation to penalize, when combined with the requirement that "[i]n all circumstances the accused person shall benefit by safeguards of proper trial and defence,"[73] suggests some form of proper and credible investigation must be carried out to account for other than grave breaches of IHL.

In this regard, U.S. policy is clear. The DoD requires all "possible, suspected, or alleged violation[s] of the law of war, for which there is credible information...[be] reported promptly, investigated thoroughly, and, where appropriate, remedied by corrective action."[74] Analysis must also determine if incidents are classified as war crimes.[75] Indications of

---

[68] Michal Drabik, *A Duty to Investigate Incidents Involving Collateral Damage and the United States Military's Practice*, 22 MINN. J. INT'L L. ONLINE 15, 19 (2013).

[69] *Rule 158 Prosecution of War Crimes*, INT'L COMMITTEE OF THE RED CROSS, https://ihl-databases.icrc.org/customary-ihl/eng/docs/v1_rul_rule158 (last visited Mar. 13, 2019) ("States must investigate war crimes allegedly committed by their nationals or armed forces, or on their territory, and, if appropriate, prosecute the suspects. They must also investigate other war crimes over which they have jurisdiction and, if appropriate, prosecute the suspects."); *see also How 'grave breaches" are defined in the Geneva Conventions and Additional Protocols*, INT'L COMMITTEE OF THE RED CROSS, https://www.icrc.org/en/doc/resources/documents/faq/5zmgf9.htm (last visited Mar. 13, 2019).

[70] Brendan Groves, *Civil-Military Cooperation in Civilian Casualty Investigations: Lessons Learned from the Azizabad Attack*, 65 A.F. L. Rev. 1, 41 (2010).

[71] Drabik, *supra* note 68, at 19 n. 10.

[72] *Id.*

[73] Convention (IV) Relative to the Protection of Civilian Persons in Time of War, art. 146, Aug. 12, 1949, 75 U.N.T.S 287. *See also* Drabik, *supra* note 68, at 19 n. 10.

[74] DODD 2311.01E, *supra* note 67, paras. 3.2, 4.4.

[75] *See* 18 U.S.C. § 2441 (2006) (defining "war crimes" as grave breaches of IHL); *see also* LAW OF WAR MANUAL, *supra* note 57, para. 18.9.5 ("The term 'war crime' has been used in different ways in different contexts. In contemporary parlance, the term 'war crime' is most often used to mean serious violations of the law of war.").

war crimes typically "[require] that higher authorities receiving an initial report request a formal investigation by the cognizant military criminal investigative organization."[76] These organizations consist of trained professional investigators, such as Army Criminal Investigative Command (CID) or Navy Crime Scene Investigators (NCIS), who operate under unique authorities and regulations.[77] In situations that may not rise to the level of war crimes, investigation of reportable incidents is commonly accomplished through the military departments' and services' administrative investigative processes.[78] Both administrative investigations and criminal investigations face unique issues when investigating AWS incidents.

B. Centrally Managed Investigations

To account for unique considerations in AWS investigations, information sharing must be improved. Under current methods of conducting administrative investigations, IOs are appointed, conduct investigations, and their findings and recommendations are approved by an authority who also considers any recommendations they may have.[79] The investigation is then maintained on file for a period of years.[80] While this technique of categorizing and storing information is useful for the less complex situations that might give rise to an administrative investigation, it does not offer the ability for units to readily share problems that are experienced across military formations—let alone amongst military branches.[81] Similarly, military criminal investigations are managed at localized levels, and while information sharing is much more efficient than in administrative investigations,[82] it can be improved upon for purposes of managing information related to AWS investigations.

---

[76] LAW OF WAR MANUAL, *supra* note 57, para. 18.13.

[77] *See, e.g.*, U.S. MARINE CORPS, MCTP 10-10F, MILITARY POLICE OPERATIONS para. 4-7 (2 May 2016); *see also* U.S. DEP'T OF ARMY, REG. 195-2, CRIMINAL INVESTIGATION ACTIVITIES para. 3-3a(6) (9 June 2016).

[78] LAW OF WAR MANUAL, *supra* note 57, para. 18.13.2.

[79] *See, e.g.*, AR 15-6, *supra* note 4, secs. II and III; *see also* DEP'T OF THE NAVY, JAGINST 5800.7F, MANUAL OF THE JUDGE ADVOCATE GENERAL (JAGMAN) ch. II (26 June 2012).

[80] AR 15-6, *supra* note 4, para. 3-19 ("The approval authority will keep the original and a digital copy of the final report of proceedings on file for a period of not less than 5 years.").

[81] *See id.* para. 3-19 (discussing filing of investigations at the local level); s*ee also id.* app. C-4, para. b(7) (indicating the approval authority's permission is required to release the investigation outside the organization).

[82] *U.S. Army Crime Records Center*, U.S. ARMY CRIM. INVESTIGATION COMMAND, https://www.cid.army.mil/crc.html (last visited Mar. 13, 2019).

With AWS platforms likely to become ubiquitous across military formations,[83] central management of AWS is key to identifying common issues that may manifest within individual AWS platforms. In turn, this will assist in AWS accountability and traceability by allowing compilation of data from AWS across the military.[84] For example: analysis of multiple false identifications of weather radar stations as anti-aircraft batteries may help AWS designers to explain, and solve, the problem of AWS returning false identifications. While this input- and output-based analysis of AWS is not the single answer, allowing this form of examination is a step toward ensuring accountability of AWS.[85]

Luckily, the concept of centrally managed investigations is not foreign to the U.S. military. While not as technologically in depth as AWS, airdrop operations routinely involve coordination between multiple branches of the military, utilizing aircraft and complex parachute delivery systems.[86] By ensuring "proper analysis to improve existing procedures and technology as rapidly as possible," [87] the services maintain a joint regulation laying out combined duties and responsibilities. Under this joint regulation, the individual services are required to conduct an internal malfunction investigation in the event of a malfunction during an airborne operation.[88] Once complete, these investigations are forwarded to a centralized directorate who publishes "all reported malfunction/incident

---

[83] *See, e.g.*, COUNTER ROCKET, ARTILLERY, MORTAR, (C-RAM), https://www.msl.army.mil/Pages/C-RAM/default.html (last visited Mar. 13, 2019) (describing the C-RAM, a defense weapon with autonomous characteristics that has been adapted from the Navy's Phalanx Weapon System).

[84] It is reasonable to assume a certain amount of modularity will occur between AWS and non-weaponized artificial intelligence (AI) items in the military inventory. For example, the computer program operating an autonomous tank may share programing with the computer system operating an autonomous fuel truck. As a result, it would be advantageous to implement centrally managed investigations to all AI platforms.

[85] S. Wachter, S, B. Mittelstadt, B., & L. Floridi, *Transparent, Explainable, and Accountable AI for Robotics*, SCI. ROBOTICS (May 31, 2017), https://discovery.ucl.ac.uk/id/eprint/10038294/1/Wachter_Transparent_explainable_accountable_AI.pdf ("Inscrutability in AI challenges calls for transparency. Mechanisms not reliant on full interpretability, including pre-deployment certification and algorithmic auditing, require further development to ensure transparency and accountability in opaque systems. It remains to be seen whether such "black box" approaches that assess inputs and outputs will comply with legal requirements.").

[86] AR 59-4, *supra* note 51, para. 1-5.

[87] *Id.* para. 1-5.

[88] *Id.* paras. 1-4, 3-3, ch. 4.

activity data for review and analysis during the triannual airdrop malfunction and safety analysis review board meeting."[89]

Investigations into AWS incidents should follow a format similar to airborne malfunction operations. While there is no need for micromanagement of individual service or command investigations, it is important that data on AWS incidents be compiled in a centralized location where it can be appropriately analyzed to allow improvements in AWS design. In addition to improving AWS and increasing explainability of AWS, centrally managed investigations will solve another issue present in AWS investigations by allowing subsequent investigations and incorporation of experts into the AWS investigation process.

C. Incorporating Experts

As demonstrated by the hypothetical at the beginning of this article, traditional investigative methods are not well positioned to examine the complex technology and multiple levels of government and private organizations that will have interplay in AWS incidents. Although current administrative investigative regulations require appointment of IOs "best qualified by reason of their education, training [and] experience...[and allow for appointing authorities to designate] assistant IOs...to provide special technical knowledge..."[90] the sheer complexity of AWS will likely result in the inability of anyone other than a true expert to understand technological questions posed by AWS. For this reason, AWS investigations must allow for the incorporation of technological experts into the investigative process to ensure results are credible and can support accountability by providing a reliable basis for necessary criminal or adverse administrative actions.[91]

While criminal investigations have successfully integrated experts into the investigative process for some time,[92] incorporation of experts into administrative investigations is less common.[93] Fortunately, best

---

[89] *Id*. paras. 1-5, 1-6.

[90] AR 15-6, *supra* note 4, para. 2-3.

[91] DoDD 2311.01E, *supra* note 67, paras. 3.2, 4.4; s*ee also* Groves, *supra* note 70.

[92] DEF. FORENSIC SCI. CTR., https://www.cid.army.mil/dfsc-usacil.html (last visited Mar. 13, 2019).

[93] AR 15-6, *supra* note 4, app. C-3, para. 3e(4) (providing the following as the sole guidance on incorporating experts in the investigative process: "It may be necessary or advisable to interview experts having specialized understanding of the subject matter of

practices can be derived from time-tested methods that allow for integration of technically complex concerns into investigative processes such as aircraft accident investigations.

With the invention of powered flight in 1903, complex mechanical and engineering issues quickly became apparent to the public.[94] By 1928, the need for aeronautic accident investigations was recognized, and Congress passed the Air Commerce Act giving the U.S. Department of Commerce the mandate to investigate the causes of aircraft accidents.[95] They do so through the present-day National Transportation Safety Board (NTSB).[96] Today, the NTSB employs approximately 400 full-time employees between its headquarters in Washington, D.C., and four regional field offices.[97] Through combined efforts with the Federal Aviation Administration, the NTSB has successfully conducted more than 132,000 investigations into the complex issues presented by aircraft accidents.[98]

To effectively conduct investigations of aviation incidents (and other public transportation incidents), the NTSB utilizes investigators in "Go Teams" who remain "[o]n call 24 hours a day, 365 days a year...[and are prepared to] travel through the country and to every corner of the world to investigate significant accidents."[99] Importantly, due to the fact that "[a]viation accidents are...usually the culmination of a sequence of events, mistakes, and failures,"[100] the NTSB supplements their own internal experts with a "party system" of investigations.

---

the investigation, if the information may be helpful to the appointing authority in making a final determination.").

[94] A BRIEF HISTORY OF THE FAA, https://www.faa.gov/about/history/brief_history/ (last visited Feb. 7, 2019).

[95] HISTORY OF THE NAT'L TRANSP. BD., https://www.ntsb.gov/about/history/Pages/default.aspx (last visited Feb. 7 2019).

[96] *Id.*

[97] NTSB CAREERS, https://www.ntsb.gov/about/employment/Pages/Careers.aspx (last visited Feb. 7, 2019); *see also* FEDERAL AVIATION ADMIN., AVIATION SAFETY WORKFORCE PLAN 2018–2017, at 23 (2018).

[98] HISTORY OF THE NAT'L TRANSP. BD., *supra* note 95.

[99] *Id.*

[100] Clinton V. Oster Jr, et al., *Analyzing Aviation Safety: Problems, challenges, opportunities*, 43 RES. IN TRANSP. ECON. 148, 151 (2013) ("Take a very simple example of an engine failure during takeoff where the crew then fails to take the needed actions to land the plan safely with the result of an accident. Had the engine not failed, there would not have been an accident. Had the crew responded to the engine failure quickly and properly, there would not have been an accident.").

Under this methodology, the NTSB designates federal, state, or local government agencies, as well as organizations or corporations with expertise, to actively participate in the investigation.[101] This results in the NTSB investigative process including smaller working groups comprised of true subject matter experts in various fields relevant to the given investigation.[102] Through the use of internal and external experts, the NTSB is able to effectively investigate complex accident scenarios and arrive at scientifically accurate results.

In order to ensure scientifically sound investigations into complex situations, AWS investigations should incorporate experts into the investigative process in a manner similar to the NTSB. While expert integration may be feasible at the local level in certain situations,[103] the ability to employ and contract with experts in the AI field is best handled at a central location. By establishing central management of AWS investigations, the DoD can build the structure necessary to employ internal experts and coordinate for outside expertise when needed. This, in turn, will inform investigations that comply with international and DoD requirements and provide human accountability for AWS actions.

## V. Bringing It Together: An AWS Investigative Model

While there is no need to reinvent the time-tested methods utilized by military services to conduct administrative investigations, the unique factors that present themselves in AWS investigations require a modified process to ensure accountability for AWS is properly established. Adopting the Joint Airdrop Malfunction/Incident Investigation methodology, individual services should be allowed to conduct initial

---

101 *The Investigative Process*, NAT'L TRANSP. SAFETY BD., https://www.ntsb.gov/investigations/process/pages/default.aspx (last visited Mar. 14, 2019).

102 *See id.*

103 *See, e.g.*, MANUAL FOR COURTS-MARTIAL, UNITED STATES, R.C.M. 703(d) (2019) ("When the employment at Government expense of an expert witness or consultant is considered necessary by a party, the party shall, in advance of employment of the expert, and with notice to the opposing party, submit a request to the convening authority to authorize the employment and to fix the compensation for the expert. The request shall include a complete statement of reasons why employment of the expert is necessary and the estimated cost of employment.").

AWS investigations utilizing their respective investigative methods.[104] However, like Joint Airdrop Investigations, the DoD should direct that specific questions be answered at this phase.[105]

First, initial unit-level investigations should address responsibility at the command and end-user level to determine if the utilization of AWS was in compliance with law of war requirements. Because a key driver of this analysis includes the command's understanding of what the AWS should have done, documentation of this expectation is key. Having established the command's expectation of the AWS, initial unit-level investigations should next document the actual actions of the AWS, highlighting any deviation from the expected action. Finally, the initial unit-level investigation should document the outcome from the AWS actions.

Utilizing the hypothetical scenario presented at the beginning of this article as an example, a unit-level investigation would determine the commander appropriately used the AWS, as he believed the AWS had properly identified an enemy vehicle. Investigation would also determine that the AWS misidentified a school bus as an enemy vehicle resulting in the death of civilians. Having reached this conclusion, the AWS investigation would be forwarded to the centrally managed AWS investigation database.

With the end-user analysis complete by the unit, experts at the centrally managed location would then begin to analyze the other stages of responsibility in the AWS creation process. By adopting the NTSB model for utilization and incorporation of experts, AWS investigators would have access to experts from other government agencies and private business to assist with the investigation as needed. Utilizing the facts provided in the unit-level investigation and by conducting analysis of the AWS in question, the experts would attempt to identify the point of failure within the AWS and, if identified, examine why testing and evaluation did not predict and prevent the AWS failure.

With a scientifically accurate investigation complete, investigators would then examine the actions of individuals in designated positions of responsibility during the creation of the AWS. Finally, investigators and

---

[104] AR 59-4, *supra* note 51, para. 1-5.
[105] *Id.* app. B.

commanders would be able to examine the accountability of individual persons and, if necessary, take appropriate punitive or administrative actions utilizing existing methods and command structures.

## VI. Conclusion

By allowing assignment of human responsibility for AWS actions through efficient and effective investigations, the U.S. military can ensure its ability to use and develop AWS without unnecessary restrictions. Designing actionable solutions to AWS accountability issues will allow the United States to remain competitive in an ever-changing military environment, while simultaneously ensuring that the moral and legal concerns surrounding AWS use are addressed. Although it remains to be seen whether "[t]he one who becomes the leader in this sphere will be the ruler of the world,"[106] one can be certain that AI and AWS offer great power. And "[i]n this world, with great power there must also come—great responsibility."[107]

---

[106] Putin, *supra* note 12.

[107] Lee & Ditko, *supra* note 1; s*ee also* Luke 12:48, *supra* note 1.