LAWS AND LAWYERS: LETHAL AUTONOMUS WEAPONS
BRING LOAC ISSUES TO THE DESIGN TABLE, AND JUDGE
ADVOCATES NEED TO BE THERE.

MAJOR ANNEMARIE VAZQUEZ[*]

I believe there is no deep difference between what can be achieved by a biological brain and what can be achieved by a computer. It therefore follows that computers can, in theory, emulate human intelligence — and exceed it.[1]

I.   Introduction

In August 2017, during a school-year kick-off speech to students in 16,000 schools across Russia, Vladimir Putin announced, "Artificial intelligence [AI] is the future, not only for Russia but for all humankind. Whoever becomes the leader in this sphere will become the ruler of the world."[2] Then a year and a half later, Greg Allen, Chief of Strategy and Communications at the Department of Defense's (DoD) Joint Artificial Intelligence Center (JAIC) reported, "Despite expressing concern on AI arms races, most of China's leadership sees increased military usage of AI

---

[*] Judge Advocate, United States Army. Associate Professor, Contract and Fiscal Law Department, U.S. Army Judge Advocate General's Legal Center and School. B.S., Minnesota State University, 2002; J.D., Mitchell Hamline University School of Law, 2008; Judge Advocate Officer Basic Course, 2010; LL.M., Military Law with National Security Law Concentration, The Judge Advocate General's Legal Center and School, 2019. Career highlights include Defense Counsel, Trial Defense Service Field Office, Fort Bliss, Texas, 2017-2018; Recruiting Officer, Judge Advocate Recruiting Office, Fort Belvoir, Virginia 2015-2017; Brigade Judge Advocate, 108th Air Defense Artillery Brigade, XVIII Airborne Corps, Fort Bragg, North Carolina, 2013-2015; Special Assistant United States Attorney, 1st Infantry Division, Fort Riley, Kansas, 2011-2013; Administrative Law Attorney, 1st Infantry Division, Camp Arifjan, Kuwait, and Fort Riley, Kansas, 2010-2011; Rule of Law Attorney, 1st Infantry Division, Contingency Operating Base Basra, Iraq, 2009-2010; Legal Assistance Attorney, 1st Infantry Division, Fort Riley, Kansas, 2009. Member of the bar of Minnesota, the United States Army Court of Criminal Appeals, and the Supreme Court of the United States.

[1] Professor Stephen Hawking, Speech at the Leverhulme Centre for the Future of Intelligence, Cambridge (Oct. 19, 2016), https://www.cam.ac.uk/research/news/the-best-or-worst-thing-to-happen-to-humanity-stephen-hawking-launches-centre-for-the-future-of.

[2] Tom Simonite, *For Superpowers, Artificial Intelligence Fuels New Global Arms Race*, WIRED (Aug. 8, 2017 7:00 AM), https://www.wired.com/story/for-superpowers-artificial-intelligence-fuels-new-global-arms-race.

as inevitable and is aggressively pursuing it. China already exports armed autonomous platforms and surveillance AI."[3] That same year, Defense Secretary Mark Esper announced on November 5, 2019, that China had exported lethal autonomous drones to the Middle East: "Chinese manufacturers are selling drones advertised as capable of full autonomy, including the ability to conduct lethal targeted strikes."[4] In countering Russian and Chinese pursuit, possession, and export of lethal autonomy the 2018 DoD Artificial Intelligence Strategy emphasized:

> Our adversaries and competitors are aggressively working to define the future of these powerful technologies according to their interests, values, and societal models. Their investments threaten to erode U.S. military advantage, destabilize the free and open international order, and challenge our values and traditions with respect to human rights and individual liberties.[5]

The "powerful technologies" referred to in DoD's AI Strategy and the comments made by Esper, Allen, and Putin refer to lethal autonomous weapons (LAWs),[6] a subset of machines that employ AI. Although there is no internationally agreed-upon definition of LAWs,[7] the DoD defines them as weapons that "can select and engage targets without further intervention by a human operator."[8] These are the "killer robots" referred

---

[3] Gregory C. Allen, *Understanding China's AI Strategy*, CENTER FOR A NEW AMERICAN SECURITY ¶ 4 (Feb. 6, 2019), https://www.cnas.org/publications/reports/understanding-chinas-ai-strategy.

[4] Patrick Tucker, *SecDef: China is Exporting Killer Robots to the Mideast*, DEFENSE ONE (Nov. 5, 2019), https://www.defenseone.com/technology/2019/11/secdef-china-exporting-killer-robots-mideast/161100/.

[5] SUMMARY OF THE 2018 DEP'T OF DEFENSE ARTIFICIAL INTELLIGENCE STRATEGY 17 (2018) [hereinafter DoD AI STRATEGY].

[6] For purposes of the discussion, "Lethal Autonomous Weapons" (LAWs) refer to individual weapons and systems of weapons, including hardware and software, and only those with fully autonomous lethal capabilities, *see infra* Section II. References to LAWs exclude cyber weapons and cyber weapon systems.

[7] Steven Hill & Nadia Marsan, *Artificial Intelligence and Accountability: A Multinational Legal Perspective*, North Atlantic Treaty Organization [NATO] STO-MP-IST-160 (Apr. 16, 2018), https://www.sto.nato.int/publications/STO%20Meeting%20Proceedings/STO-MP-IST-160/MP-IST-160-PP-4.pdf; Christopher Ford & Chris Jenks, *The International Discussion Continues: 2016 CCW Experts Meeting on Lethal Autonomous Weapons*, JUST SECURITY (Apr. 20, 2016), https://www.justsecurity.org/30682/2016-ccw-experts-meeting-laws.

[8] U.S. DEP'T OF DEFENSE, DIR. 3000.09, AUTONOMY IN WEAPON SYSTEMS 13 (21 Nov. 2012) (C1, 8 May 2017) [hereinafter DoDD 3000.09]. The lack of an agreed-up definition for LAWs is evident upon closer look at China's claims of full autonomy in its weapons. The manufacturer of the Blowfish A3 and other Chinese LAWs, Zhuhai Ziyan

to in the media and by organizations dedicated to banning them.[9] Though technology for some LAWs exists,[10] and variants of them have been on the battlefield for decades, fully autonomous lethal systems for offensive use have yet to make their battlefield debut.[11]

In the quest to remain a "leader in this sphere"[12] the United States (U.S.) Congressional and Executive Branches have prioritized research and development of autonomy[13] for military applications. These priorities are evident in the fiscal year 2020 National Authorization Act (FY20

UAV Company, states that though they can organize in a swarm and identify a target autonomously, they do not shoot until a human commands them to do so. Under the DoD's definition, such weapons would not be fully autonomous. *See also* Liu Xuanzun, *Chinese Helicopter Drones Capable of Intelligent Swarm Attacks*, GLOBAL TIMES (May 9, 2019 4:28 PM), http://www.globaltimes.cn/content/1149168.shtml [hereinafter Xuanzun].

[9] *See The Case Against Killer Robots,* HUMAN RIGHTS WATCH (Nov. 19, 2012), https://www.hrw.org/ report/2012/11/19/losing-humanity/case-against-killer-robots [hereinafter *Killer Robots*].

[10] *See, e.g.*, *Tomahawk Cruise Missile*, RAYTHEON, https://www.raytheon.com/capabilities/products/tomahawk (last visited Jan. 23, 2019); MILREM ROBOTICS, THeMIS, https://milremrobotics.com/themis/ (last visited Nov. 20, 2018). *Cf.* PAUL SCHARRE, ARMY OF NONE 129, 266 (2018) [hereinafter SCHARRE, ARMY OF NONE] (Fully autonomous LAWs do not yet exist, but "[a]ll of the tools to build an autonomous weapon that could target people on its own [are] readily available online . . . . Trying to contain the software would be pointless."). Compare to JASON, *Perspectives on Research in Artificial Intelligence and Artificial General Intelligence Relevant to DoD* (Jan. 2017), https://apps.dtic.mil/dtic/tr/fulltext/u2/1024432.pdf [hereinafter JASON] wherein a group of scientific experts examined AI for DoD uses and determined "it is not clear that the existing AI paradigm is immediately amenable to any sort of software engineering validation and verification." *Id.* at 27.

[11] *See, e.g.*, *Aegi: The Shield of the Fleet*, LOCKHEED MARTIN, https://www.lockheedmartin.com/en-us/products/aegis-combat-system.html (last visited Jan. 11, 2019). For recent achievements in operationalizing autonomy see Jen Judson, *Jumping in to Algorithmic Warfare*, DEFENSE NEWS (Sept 5, 2019) [hereinafter Judson], discussing A3I, a networked system of autonomous capabilities developed by Army's Future Vertical Lift cross-functional team. *See also* CONG. RESEARCH SERV., R45178, ARTIFICIAL INTELLIGENCE AND NATIONAL SECURITY 18 (last updated Jan. 30, 2019) [hereinafter CRS, AI AND NAT'L SECURITY] ("The U.S. military does not currently have LAWS in its inventory, although there are no legal prohibitions on the development of LAWS.").

[12] Or "become" a leader in this sphere. Some would argue the United States has lost its lead in the field of artificial intelligence. *See* KAI FU LEE, AI SUPERPOWERS 14-18 (2018); SUMMARY OF THE NAT'L DEF. STRATEGY 3 (2018) [hereinafter NAT'L DEF. STRATEGY]; *but see* Exec. Order No. 13859, 84 Fed. Reg. 31 (Feb. 14, 2019) [hereinafter Exec. Order] ("The United States is the world leader in AI research and development (R&D) and deployment.").

[13] Used here, the term autonomy refers to that which uses machine learning See *infra* Section II.

NDAA), the 2019 National Defense Authorization Act (FY19 NDAA), the President's Executive Order on Maintaining American Leadership in AI, the Pentagon's Third Offset Strategy, the National Defense Strategy, and DoD's AI Strategy.[14] Currently, there are efforts within DoD to facilitate the development of weaponized autonomous platforms, LAWs, capable of operating offensively, beyond human control.[15] At this time, DoD policy, reflected in Department of Defense Directive (DoDD) 3000.09 directs

---

[14] National Defense Authorization Act for Fiscal Year 2020, Pub. L. No. 116-92, §§ 221, 222, 133 Stat. 1198 (2019) [hereinafter FY2020 NDAA] (directing "appropriate entities" in the DoD to review domestic and foreign open source publications to understand adversaries' investments in development of AI; and engaging JASON members to advise on matters involving science, technology, and national security, including methods to defeat existential and technologically-amplified threats to national security."); John S. McCain Nat'l Def. Authorization Act for Fiscal Year 2019, Pub. L. No. 115-232, §§ 238(c)(2)(A)–(B), 238(c)(2)(H), 132 Stat. 1636 (2018) [hereinafter FY2019 NDAA]; Exec. Order, *supra* note 12; Sydney J. Freedberg Jr., *Hagel Lists Key Technologies for US Military; Launches "Offset Strategy,"* BREAKING DEFENSE (Nov. 16, 2014 2:00 PM) (defining Offset Strategy as a "military-industrial term of art for a cluster of technological breakthroughs that can give the United States its edge over potential enemies" and that another example is President Eisenhower's "New" Look which used technology like stealth and computer networks to offset Soviet superiority in numbers). Former Deputy Secretary of Defense, Bob Work, focused the Third Offset Strategy on, among others, autonomous learning systems and network-enabled autonomous weapons. NAT'L DEF. STRATEGY, *supra* note 12, at 9; Bob Work, Deputy Secretary of Defense, Address Before the Center for a New American Security: The Third U.S. Offset Strategy and its Implications for Partners and Allies (Jan. 28, 2015), https://www.defense.gov/Newsroom/Speeches/Speech/Article/606641/the-third-us-offset-strategy-and-its-implications-for-partners-and-allies/; Memorandum from Deputy Sec'y of Army to Chief Management Officer of the Dep't of Defense, et al., subj: Establishment of the Joint Artificial Intelligence Center (June 27, 2018) [hereinafter AI TASK FORCE]; U.S. DEP'T OF ARMY, DIR. 2018-18, ARMY ARTIFICIAL INTELLIGENCE TASK FORCE IN SUPPORT OF THE DEPARTMENT OF DEFENSE JOINT ARTIFICIAL INTELLIGENCE CENTER (2 Oct. 2018) [ARMY DIR. 2018-18]; Yasmin Tadjdeh, *Algorithmic Warfare: Army's AI Task Force Making Strides*, NATIONAL DEFENSE (Oct. 8, 2019) [hereinafter Tadjdeh] (discussing one of AI Task Force's main lines of effort being automated threat recognition and autonomous operational maneuver platforms); *see* U.S. DEP'T OF DEF., DEFENSE SCIENCE BOARD TASK FORCE REPORT, THE ROLE OF AUTONOMY IN DoD SYSTEMS (June 2012) [hereinafter DSB, ROLE OF AUTONOMY] (strongly encouraging the DoD to address the underutilization of autonomy in unmanned systems).

[15] CRS, AI AND NAT'L SECURITY, *supra* note 11, at 13 ("AI is also being incorporated into . . . lethal autonomous weapon systems."); Will Knight, *Military Artificial Intelligence Can be Easily and Dangerously Fooled*, MIT TECHNOLOGY REV. (Oct. 21, 2019), https://www.technologyreview.com/s/614497/military-artificial-intelligence-can-be-easily-and-dangerously-fooled/ ("The Department of Defense's proposed $718 billion budget for 2020 allocates $927 million for AI and machine learning. Existing projects include the rather mundane (testing whether AI can predict when tanks and trucks need maintenance) as well as things on the leading edge of weapons technology (swarms of drones).").

Combatant Commanders to "integrate autonomous and semiautonomous weapon systems into operational mission planning" and identify how LAWs may satisfy operational needs.[16]

So, in a word, LAWs are inescapable. The days of debating whether or not LAWs should be developed are over.[17] Commentators have already shown that fully autonomous lethal weapons are not illegal per se,[18] which is to say that the Law of Armed Conflict (LOAC)[19] does not prohibit their use in all circumstances.[20] Barring an agreed-upon prohibition, States are limited by their own policies, like DoDD 3000.09, and the limitations of the technology itself; the popular concern about robots running amok exaggerates their capabilities.[21] From the United States' perspective, DoDD 3000.09 requires "appropriate levels of human judgment" over autonomous weapons, including those capable of full autonomy. [22]

---

[16] DoDD 3000.09, *supra* note 8, encl. 4, ¶ 10(d)–(e).

[17] CRS, AI AND NAT'L SECURITY, *supra* note 11, at 19 ("Vice Chairman of the Joint Chiefs of Staff, Gen. Paul Selva stated, "I do not think it is reasonable for us to put robots in charge of whether or not we take a human life." But he added that because United States adversaries are pursuing LAWs, the United States must identify its vulnerabilities and address them.); *compare to* CAMPAIGN TO STOP KILLER ROBOTS, https://www.stopkillerrobots.org/learn/ (last visited Nov. 20, 2019) ("Fully autonomous weapons would lack the human judgment necessary to evaluate the proportionality of an attack, distinguish civilian from combatant, and abide by other core principles of the laws of war.").

[18] Michael N. Schmitt & Jeffrey S. Thurner, *Out of the Loop: Autonomous Weapons Systems and the Law of Armed Conflict*, 4 HARV. NAT'L SEC. J. 231 (2016) [hereinafter Schmitt, *Out of the Loop*] (Expressing confidence that sophisticated states can determine whether use of LAWs in particular contexts complies with IHL); Kenneth Anderson, Daniel Reisner & Matthew Waxman, *Adapting the Law of Armed Conflict to Autonomous Weapon Systems*, 90 INT'L L. STUD. 386, 387 (2014), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2477095 [hereinafter Anderson, *Adapting the LOAC*].

[19] Law of War (LoW) and Law of Armed Conflict (LOAC) are used interchangeably here, and refer to the international body of law that applies during an armed conflict.

[20] Michael Meier, *Lethal Autonomous Weapons Systems (LAWS): Conducting a Comprehensive Weapons Review*, 30 TEMP. INT'L & COMP. L.J. 119, 126 (2016) [hereinafter Meier, *LAWS Weapons Reviews*]; Charles J. Dunlap et al., *To Ban New Weapons or Regulate Their Use?*, JUST SECURITY (Apr. 3, 2015), https://www.justsecurity.org/21766/guest-post-ban-weapons-regulate-use/ (Urging against emotionally-driven decisions that lead to "unintended consequences of well-intended [weapon] prohibitions."; Christopher M. Ford, *Autonomous Weapons and International Law*, 69 S.C. L. REV. 413, 458-459 (2017) [hereinafter Ford, *Autonomous Weapons*].

[21] Schmitt, *Out of the Loop*, *supra* note 18, at 242.

[22] DoDD 3000.09, *supra* note 8, para. 4a ("Autonomous . . . weapon systems shall be designed to allow commanders and operators to exercise appropriate levels of human judgment over the use of force."); For a discussion of how to give meaning to

Though LAWs are not prohibited under the LOAC, and they can operate lawfully, few commentators discuss pragmatic safeguards for ensuring they actually do operate lawfully when put into operation. The existing legal framework for identifying and addressing potential LOAC concerns in weapons systems is ill-suited to the unique nature of autonomous weapon systems, because of:

- *What* we are pursuing;[23]
- *Where* we are getting it;[24]
- *How* we are acquiring it.[25]

Together, these vulnerabilities set the stage for building risk into LAWs, an already immature and risky technology. While rigorous testing serves a critical role in minimizing these and other risks, it cannot and should not be the cure-all. In a reality where the inevitable trajectory of clashing international interests tosses LAWs into the crucible of armed conflict, the LOAC requires consideration of its tenets during the design of LAWs' "decision-making" models, and in conjunction with those who will be held responsible for employing them: commanders, whose responsibility extends to the foreseeable consequences of their decisions.[26] To this end, selected teams of judge advocates and combat-seasoned commanders, tasked as collaborators and issue-spotters, should be

---

"appropriate levels of human judgment" see Lieutenant Colonel Adam Cook, *Taming Killer Robots*, U.S. AIR FORCE JAG SCHOOL PAPERS, June 2019, at 16, https://www.airuniversity.af.edu/AUPress/ [hereinafter *Taming Killer Robots*]; *see also* Karl Chang, *U.S. Delegation Statement on Human-Machine Interaction*, U.S. MISSION IN GENEVA (Aug. 28, 2018) [hereinafter HUMAN-MACHINE INTERACTION].

[23] Meaning, our inability to understand how a LAW's "black box" of deep neural networks work sets them apart from other weapons. *See* discussion *infra* Section II. *See Taming Killer Robots*, *supra* note 22, at 7, 15 ("[S]imply applying the existing rule of law framework to these fundamentally novel systems is not sufficient to protect against the very real risks . . . Rather additional standards are required.").

[24] In that industry is the most likely source of the component technology of LAWs, and this directly affects *when* DoD becomes involved. Because the black box problem arises during design, testing is only partially effective. *See infra* Section II.

[25] Referring to our use of rapid acquisition authorities to obtain the technology absorbs risk, rather than limits it. *See infra* Section IV.B.3.

[26] *See* discussion *infra* note 83 regarding commander responsibility; *see also* DSB, ROLE OF AUTONOMY, *supra* note 14, at 31-32 (Visualize challenges to autonomy "through the eyes of three key stakeholders: the commander, the operator, and the developer." The commander struggles with understanding how to incorporate autonomy into missions. For the operator, human-machine collaboration is often overlooked during design. And for the developer, "testing and evaluation have few metrics and test beds for verification and validation.").

involved as early as possible in the design and development process of LAWs' learning models.[27]   There are no legal barriers for this involvement, and the current regulatory system allows immediate implementation, limited only by industry's willingness to participate.[28]

In support of this proposition, Section II first defines LAWs, briefly explains the underlying technology, and discusses the "black box" problem, while Section III examines how LAWs' algorithms raise LOAC issues during their development. Section IV describes the current weapons review process and why it is inadequate to mitigate the LOAC issues and risk factors of *what*, *where*, and *how*. Section V explains efforts already in place, where blind spots remain, and what more should be done.

II. Defining Lethal Autonomous Weapons

Autonomy uses artificial intelligence (AI) to mimic human decision-making.[29] Though the U.S. Government has no accepted definition of AI,[30] Section 238 of the FY19 NDAA defines AI as a system that "performs tasks under varying and unpredictable circumstances without significant human oversight, or that can learn from experience and improve performance when exposed to data sets."[31] The DoD further describes autonomous systems as "self-directed toward a goal in that they do not require outside control, but rather are governed by laws and

---

[27] Judge advocates are not the only attorneys well-suited to this task. *See* discussion *infra* Section V; Brigadier General R. Patrick Huston, *The Future JAG Corps: Understanding the Legal Operating Environment*, ARMY LAW., Iss. 1, 2019, at 2-3 ("[J]udge advocates must be positioned to advise coders and developers to ensure LOAC principles are built into emerging technology."); Major Richard J. Sleesman & Captain Todd C. Huntley, *Lethal Autonomous Weapon Systems*, ARMY LAW., Iss. 1, 2019, at 32, 34 ("Since legal issues are likely to arise in development, not just during the use of the weapon system, judge advocates will need to provide legal advice during the development process.").

[28] *See infra*, Part IV.B.2; *see generally*, Nat'l Security Comm'n on Artificial Intelligence, Interim Rep. 32, 45 (Nov. 2019), https://www.epic.org/foia/epic-v-ai-commission/AI-Commission-Interim-Report-Nov-2019.pdf (Discussing challenges and recommendations for DoD's development of artificial intelligence).

[29] In 1955, John McCarthy coined the term "artificial intelligence." *John McCarthy, Standford University 1999 Fellow*, COMPUTER HIST. MUSEUM, https://www.computerhistory.org/fellowawards/hall/john-mccarthy (last visited Jan. 28, 2019).

[30] CRS, AI AND NAT'L SECURITY, *supra* note 11, at 5.

[31] FY2019 NDAA, *supra* note 14, § 238.

strategies that direct their behavior." [32] As stated in the introduction, DoD defines LAWs as weapons that "can select and engage targets without further intervention by a human operator."[33] Upon human deployment, a LAW can identify a target and attack without further human direction, meaning it can operate with a human "out of the loop,"[34] which is a particularly useful capability when operating in a swarm, in communications-denied or degraded areas, when the volume of data exceeds human capacity to review and analyze, or when there is not enough reaction time for human decision-making.[35]

Autonomy is accomplished by algorithms, which are simply "a sequence of instructions telling a computer what to do," or a set of problem-solving processes and rules.[36] These instructions and rules are similar to the decision process a human uses to navigate through traffic to get to work, which can be optimized for different preferred outputs, like the most direct route, the least tolls, the most scenic, or most convenient to a grocery store. Given a decision model, an algorithm predicts the best route. A subcategory of algorithms, called learning algorithms, enable autonomy in LAWs.[37] A learning algorithm looks for patterns within

---

[32] U.S. DEP'T OF DEFENSE, UNMANNED SYSTEMS INTEGRATED ROADMAP FISCAL YEAR 2017–2036 17 (2013) [hereinafter DoD ROADMAP].

[33] DoDD 3000.09, *supra* note 8, at 13. The lack of an agreed-up definition for LAWs is evident upon closer look at China's claims of full autonomy in its weapons. Zhuhai Ziyan UAV Company, the manufacturer of the Blowfish A3 and other Chinese LAWs, states that, though they can organize in a swarm and identify a target autonomously, they do not shoot until a human commands them to do so. Under the DoD's definition, such weapons would not be fully autonomous. Xuanzun, *supra* note 8.

[34] Autonomy in weapons is best described as a spectrum of independence with humans either "in the loop," "on the loop," or "out of the loop." A human "in the loop" must affirmatively act before the weapon can fire. A human "on the loop" is able to intervene prior to firing, much like a supervisor. A human "out of the loop" cannot intervene once the weapon is deployed. Fully autonomous weapons are unique in their ability to *observe* their situations, *orient* themselves by placing those observations in context in time and space, make *decisions*, and then *act* on them. This is the human decision-making cycle coined by Air Force Military Strategist Colonel John Boyd as the "OODA loop." John Boyd, *The Essence of Winning and Losing*, DANFORD (June 28, 1995), http://www.danford.net/boyd/essence.htm.

[35] *See* CRS, AI AND NAT'L SECURITY, *supra* note 11, at 18.

[36] PEDRO DOMINGOS, THE MASTER ALGORITHM: HOW THE QUEST FOR THE ULTIMATE LEARNING MACHINE WILL REMAKE OUR WORLD 1 (2015) [hereinafter DOMINGOS] ("The simplest algorithm is: flip a switch. The state of one transistor is one bit of information: one if the transistor is on, and zero if it is off.").

[37] Dustin A. Lewis, et al., *War Algorithm Accountability*, HARV. LAW SCH. PROGRAM ON INT'L LAW AND ARMED CONFLICT 10 (2016), https://pilac.law.harvard.edu/war-algorithm-accountability-report//#_ftn58 (A "war algorithm" is "any algorithm that is expressed in

inputs (*e.g.*, facial images gathered by its sensors), makes a prediction, and learns from the outcome, continuously improving.[38] Learning algorithms come in different forms and may be referred to as learners, learning systems, agents, or recognizers, depending on the method used to achieve learning and the objective of learning.[39] For this discussion, a LAW's apparatus that enables autonomous "decision-making" will be referred to as a learner.[40] Learners use deep learning and neural networks for unsupervised learning[41] and "mimic the web of neurons in the human

computer code, that is effectuated through a constructed system, and that is capable of operating in relation to armed conflict."); SCHARRE, ARMY OF NONE, *supra* note 4; Yann LeCun, et al., *Deep Learning*, 521 NATURE 436 (2015); Paul Scharre & Michael C. Horowitz, *An Introduction to Autonomy in Weapons Systems* 21 (Ctr. for a New Am. Security, Working Paper, Feb. 2015), https://www.cnas.org/publications/reports/an-introduction-to-autonomy-in-weapon-systems. *Cf.* Chad R. Frost, *Challenges and Opportunities for Autonomous Systems in Space*, *in* FRONTIERS OF ENGINEERING: REPORTS ON LEADING-EDGE ENGINEERING FROM THE 2010 SYMPOSIUM 89-90 (2011), https://www.nap.edu/read /13043/chapter/17 ("An *automated system* doesn't make choices for itself – it follows a script . . . in which all possible courses of action have already been made. . . . By contrast, an *autonomous system* does make choices on its own . . . even when encountering uncertainty or unanticipated events.").

[38] DOMINGOS, *supra* note 36, at 1-2; *but see* CATHY O'NEIL, WEAPONS OF MATH DESTRUCTION Ch 5 (2016) [hereinafter O'NEIL] (Discussing the "pernicious feedback loop" where a model's outputs reinforce biases embedded within the data given it, unbeknownst to those who rely on its predictions.).

[39] RICHARD S. SUTTON & ANDREW G. BARTO, REINFORCEMENT LEARNING: AN INTRODUCTION ¶¶ 1.1-1.3 (2d ed. 2018) [hereinafter SUTTON]; U.S. AIR FORCE OFFICE OF THE CHIEF SCIENTIST, AUTONOMOUS HORIZONS: SYSTEM AUTONOMY IN THE AIR FORCE – A PATH TO THE FUTURE (June 2015), https://www.af.mil/Portals/1/documents/SECAF/Autonomous Horizons.pdf?timestamp=1435068339702 [hereinafter AUTONOMOUS HORIZONS]; GOOGLE, GLOSSARY, https://developers.google.com/machine-learning/glossary/#o (last visited Jan. 28, 2019) [hereinafter GOOGLE].

[40] The term *agent* is widely used when referring to a LAW's decision-making entity. "Agent" is itself a term of art suggesting "agency" or an ability to make decisions and be held accountable for them, a responsibility the U.S. reserves for humans. *See* DoDD 3000.09, *supra* note 8. To avoid confusion with the generic reference to *agent*, the term *learner* is used instead. When referring to "decision-making" capabilities, such capabilities are limited by human coding and training, and so they are similar to, but not the same as, human decisions.

[41] *See* SUTTON, *supra* note 39. Compare to *supervised* learning whereby an algorithm is given a data set with pre-labeled examples, like a grouping of animal photos with the dogs already labeled. The learner learns what dogs are by comparing future unknown samples to the known samples. In *unsupervised* learning, the learner is given an unlabeled data set and must find the patterns itself. Unsupervised learning is better suited to situations where labeled samples are too voluminous, expensive, and/or time intensive to label or acquire. *Id.* ¶¶ 1.1-1.2; Jason Pontin, *Greedy, Brittle, Opaque, and Shallow: The Downsides to Deep Learning*, WIRED (Feb. 2, 2018, 8:00 AM), https://www.wired.com/story/greedy-brittle-opaque-and-shallow-the-downsides-to-deep-

brain" by passing data through layers of filters, looking for patterns until it reaches the output layer, which contains the answer.[42] A programmer sets goals for the learner and may also use reinforcement reward signals to incentivize correct decisions or penalties to deter incorrect decisions, a process called learning or training a model.[43] After achieving its goal, the learner stores its experience to strengthen similar decision-making.[44] Lethal autonomous weapons will likely rely on several different types of learners.[45] For example, one type, known as recognizers, look for patterns within images to classify and predict what the image depicts.[46] Consider an example of a lethal autonomous drone trained to target snipers by a programmer unfamiliar with the LOAC. The programmer learning or training its unsupervised recognizer to identify snipers would give the

---

learning/ [hereinafter Pontin] (suggesting that unsupervised learning offers a path around the limitations of supervised deep learning).

[42] THOMAS H. CORMEN, ET AL., INTRODUCTION TO ALGORITHMS 5 (3d ed. 2009), https://labs.xjtudlc.com/labs/ wldmt/reading%20list/books/Algorithms%20and%20optimization/Introduction%20to%2 0Algorithms.pdf; DOMINGOS, *supra* note 36; GOOGLE, *supra* note 39, https://developers.google.com/machine-learning/glossary/#o (last visited Jan. 28, 2019); *see also* Cade Metz, *In Two Moves, AlphaGo and Lee Sedol Redefined the Future*, WIRED (Mar. 16, 2016, 7:00 AM), https://www.wired.com/2016/03/two-moves-alphago-lee-sedol-redefined-future/. Neural networks may use different types of architectures, and the terminology varies depending on type. CAMERA BASED IMAGE PROCESSING, TUFTS UNIVERSITY (Sept. 26, 2017), https://sites.tufts.edu/ selfdrivingisaac/2017/09/26/camera-based-image-processing/ [hereinafter TUFTS].

[43] SUTTON, *supra* note 39; GOOGLE, *supra* note 39; *see also* James Le, *12 Useful Things to Know about Machine Learning*, TOWARDS DATA SCIENCE (Jan 26, 2018), https://towardsdatascience.com/12-useful-things-to-know-about-machine-learning-487d3104e28 [hereinafter Le] ("Programming, like all engineering, is a lot of work: we have to build everything from scratch. Learning is a more like farming, which lets nature do most of the work. Farmers combine seeds with nutrients to grow crops. Learners combine knowledge with data to grow programs.").

[44] Loz Blain, *AI Algorithm Teaches a Car to Drive from Scratch in 20 Minutes,* NEW ATLAS (July 5, 2018), https://newatlas.com/wayve-autonomous-car-machine-learning-learn-drive/55340/; U.S. Army Research Laboratory, *Artificial Intelligence Becomes Life-long Learner with New Framework*, SCIENCE DAILY (May 20, 2019), https://www.sciencedaily.com/releases/2019/05/190520115635.htm (Discussing how to avoid "catastrophic loss" in machine learning algorithms by using "backward transfer," in other words, making the learner remember how it completed previous tasks to help it complete new tasks better.).

[45] *See*, *e.g*., Le, *supra* note 43; discussion *infra* Section IV.B.1.

[46] *See* Oleksii Kharkovyna, *The A-Z of AI and Machine Learning: Comprehensive Glossary*, TOWARDS DATA SCIENCE (July 8, 2019), https://towardsdatascience.com/the-a-z-of-ai-and-machine-learning-comprehensive-glossary-fb6f0dd8230; *see also* Salim Chemlal, *A Comprehensive State-Of-The-Art Image Recognition Tutorial*, TOWARDS DATA SCIENCE (July 3, 2019), https://towardsdatascience.com/a-comprehensive-state-of-the-art-image-recognition-tutorial-c34c544d0b4.

drone's software a data set containing images of service members (or combatants, generally), including those exhibiting characteristics associated with snipers. The recognizer would then apply layers of filters to the data to determine what it observed. The recognizer may look for identifying factors like a body in a prone position, camouflaged, motionless, physically isolated from other people, and with a weapon aimed in a particular direction. Each of these features form one layer, or node, and at the output layer, the recognizer would determine whether it was looking at a sniper.[47] Upon reaching an answer, the recognizer would create a model for image classification of snipers.[48] It would then continually refine its model as the recognizer encounters more images. Despite our ability to fine-tune a learner's model, employ reinforcement learning with rewards and penalties, and control the data sets used for training, a learner's decision-making remains opaque.

Evaluating a learner's effectiveness and reliability proves difficult in machine learning because the decision-making occurs within its multiple layers of nodes and neural nets. This creates a "black box" scenario where algorithms create hidden algorithms unknown to software and testing engineers.[49] According to a group of experts, called JASON, tasked with examining AI for DoD uses:

> [T]he sheer magnitude, millions or billions of parameters (i.e. weights/ biases,/etc.), which are learned as part of the training of the net . . . makes it impossible to really understand exactly how the network does

---

[47] TUFTS, *supra* note 42; *see, e.g.*, Waymo Self-Driving Car Image Recognition, *in* Andrew J. Hawkins, *The Google Spinoff has a Head Start in AI, But Can They Maintain the Lead?,* THE VERGE (May 9, 2018), https://www.theverge.com/2018/5/9/17307156/google-waymo-driverless-cars-deep-learning-neural-net-interview [hereinafter Hawkins] (Waymo, a Google sister company, trains self-driving cars using "an automated process and human labelers to train its neural nets.").

[48] One needs little imagination to envision a scenario where a sniper could easily be confused with an injured soldier or civilian hunter. See Section V.A for more discussion on the judge advocate's and commander's roles in refining LAWs' models. *See also* O'NEIL, *supra* note 36, at 18-20 ("A model . . . is nothing more than an abstract representation of some process . . . They tell us what to expect, and they guide our decisions.").

[49] Unlike the black boxes in airplanes known for protecting data so it can become knowable, the "black box" effect within deep learning algorithms obscures it so the information can never be known. *See* Jeff Phillips, *Testing the Unknown: The Real Problem with Autonomous Vehicles,* ELECTRONIC DESIGN (Aug. 9, 2018), https://www.electronicdesign.com/automotive/testing-unknown-real-problem-autonomous-vehicles.

what it does. Thus the response of the network to all possible inputs is unknowable.[50]

Ultimately, not only is testing the network's response to all inputs impossible, but because a learner's decision-making occurs in a black box, evaluators can never know why a learner acts the way it does. "You can't just look inside a deep neural network to see how it works. A network's reasoning is embedded in the behavior of thousands of simulated neurons, arranged into dozens or even hundreds of intricately interconnected layers."[51] The DoD's AI Ethics Principles, which set standards for the use of AI, controls for this limitation.[52] One of the five principles is that AI is "traceable," meaning technicians can examine how the software reached its conclusions. Explainable AI is just that—traceable and knowable—but its early-stage tools are not yet suited for LAWs.[53]

And so the black box problem appears irreconcilable with the requirement for "appropriate levels of human judgment" over LAWs.[54] One may be tempted to suggest rigorous testing will be sufficient but it, too, has limits: "[T]he number of possible input states that such learning systems can be presented with is so large that not only is it impossible to

---

[50] JASON, *supra* note 10, at 28; SCHARRE, ARMY OF NONE *supra* note 10, at 149–50 (One of the greatest challenges in fielding LAWs will be testing them).

[51] Will Knight, *The Dark Secret at the Heart of AI*, MIT TECHNOLOGY REV. (Apr. 11, 2017), https://www.technologyreview.com/s/604087/the-dark-secret-at-the-heart-of-ai/ ("No one really knows how the most advanced algorithms do what they do.").

[52] Patrick Tucker, *The Pentagon's AI Ethics Draft is Actually Pretty Good*, DEFENSE ONE (Oct. 31, 2019), https://www.defenseone.com/technology/2019/10/pentagons-ai-ethics-draft-actually-pretty-good/161005/; *DoD Adopts Ethical Principles for Artificial Intelligence*, U.S. DEP'T OF DEF. (Feb. 24, 2020),
https://www.defense.gov/Newsroom/Releases/Release/Article/2091996/dod-adopts-ethical-principles-for-artificial-intelligence/; AUTONOMOUS HORIZONS, *supra* note 39, at 29 ("The logic and behavior of [machine learning] systems can be quite opaque . . . and often the system developers do not fully understand how the autonomy will behave."). Other entities strive to set standards for the use of AI and echo the need for outcome transparency. *See, e.g.*, the Organization for Economic Cooperation and Development's (OECD) Principles on AI, found at https://www.oecd.org/going-digital/ai/principles/.

[53] Explainability is the algorithms' ability to explain its process and noted the risk that "data training sets could inadvertently introduce errors into a system that might not be immediately recognized or understood by users." CRS, AI AND NAT'L SECURITY, *supra* note 11, at 34.

[54] Some early-stage tools exist to explain AI, but their effectiveness decreases as the machine learning model's complexity increases. *See* Tiernan Ray, *IBM Offers Explainable AI Toolkit, but it's Open to Interpretation*, ZDNET (Aug. 10, 2019), https://www.zdnet.com/article/ibm-offers-explainable-ai-toolkit-but-its-open-to-interpretation/.

test all of them directly, it is not even possible to test more than an *insignificantly small fraction* of them."[55] (emphasis added). If their decision-making models cannot be understood, and cannot be adequately tested, how is a commander to account for the reasonably foreseeable consequences of her decision to use LAWs?[56] Commanders need not rely on faith alone; the black box has windows.

To resolve the black box problem and our inability to adequately test machine learning models, DoD must continue its quest for explainable AI,[57] and in the meantime fully exploit the multiple human touch points occurring across the design timeline that offer critical opportunities for human involvement and understanding.[58] Among them:

- Training decisions, including what data to use;[59]
- Goal selection;[60]
- Choice and weighing of reward and penalty signals;[61]
- Evaluation of the learner's output and its final decision-making model;[62]
- Adjustments to a learner's architecture;[63]

---

[55] Alan Backstrom & Ian Henderson, *New Capabilities in Warfare: An Overview of Contemporary Technological Developments and the Associated Legal and Engineering Issues in Article 36 Weapons Reviews*, 886 Int'l Review of the Red Cross 483, 517 (2012).

[56] U.S. DEP'T OF DEF., DoD LAW OF WAR MANUAL paras. 6.3.1, 6.7.2, *see also* paras. 5.3, 5.3.1, 5.3.2 (Dec. 2016) [hereinafter DoD LoW MANUAL] ("Even when information is imperfect or lacking (as will frequently be the case during armed conflict), commanders and other decision-makers may direct and conduct military operations, so long as they make a good faith assessment of the information that is available to them at that time.").

[57] *See* David Gunning, *Explainable Artificial Intelligence (XAI)*, DEF. ADVANCED RESEARCH PROJECTS AGENCY, https://www.darpa.mil/program/explainable-artificial-intelligence (last visited Jan. 15, 2019) [hereinafter Gunning].

[58] M.L. Cummings, *The Human Role in Autonomous Weapons Design and Deployment*, DUKE UNIVERSITY (2014), https://www.law.upenn.edu/live/files/3884-cummings-the-human-role-in-autonomous-weapons (last visited Jan. 31, 2019).

[59] *See* TUFTS, *supra* note 42; Hawkins, *supra* note 47.

[60] *See* DoD ROADMAP, *supra* note 32, at 46.

[61] SUTTON, *supra* note 39.

[62] SUTTON, *supra* note 39.

[63] Ivan Vasilev, *A Deep Learning Tutorial: From Perceptrons to Deep Network*, TOPTAL, https://www.toptal.com/machine-learning/an-introduction-to-deep-learning-from-perceptrons-to-deep-networks (last visited Jan. 31, 2019); Nick McCrea, *An Introduction to Machine Learning Theory and Its Applications: A Visual Tutorial with Examples,* TOPTAL, https://www.toptal.com/machine-learning/machine-learning-theory-

- Engineering of the machine-operator interface, and how operator adjustments may interact with the learner;[64]
- Integration of recommendations from end users, legal advice and legal reviews into training decisions, goal selection, reinforcement, and evaluation;[65]
- End-user interface options and command decision to employ.

These touch points provide the means for injecting human judgment into a learner even though, when operationalized, a LAW operates fully autonomously, outside human control. In the simplified sniper-targeting drone example above, the drone simply did what its programmer trained it to do by setting reward signals for finding and targeting snipers. The drone's model for making targeting decisions was learner-made, but human-taught. A LAW's decision-making ability is highly dependent upon how its learner's models are programmed and trained,[66] and so the accuracy and reliability of a LAW's performance is directly tied to the human trainers whose inputs, rewards, goals, and adjustments are knowable at the time of programming. But human insight into the black box is fleeting. Once the human touch point passes, that window closes and the model's neural nets run the show, building off training and additional inputs from the environment around it.[67] New windows open as humans interact with the model, but determining which input or adjustment led to a particular output becomes nearly impossible.

---

an-introductory-primer (last visited Jan. 31, 2019); *see* DoDD 3000.09, *supra* note 8, at 6 (Discussing V&V and T&E).

[64] *Compare* DoDD 3000.09, *supra* note 8, at 2-3 (It is DoD's policy for the human-machine interface to: (1) Be readily understandable to trained operators; (2) Provide traceable feedback on system status; and (3) Provide clear procedures for trained operators to activate a deactivate system functions.), *with* the "black box" problem of hidden layers of decision-making in convolutional neural nets. Pontin, *supra* note 41, *and* DARPA's project to create explainable AI. *See* Gunning, *supra* note 57.

[65] *See infra* Part V.A-B.

[66] *See, e.g.*, Nancy Gupton, *The Science of Self-Driving Cars,* FRANKLIN INST., https://www.fi.edu/ science-of-self driving-cars (last visited Jan. 23, 2019) ("By far the most complex part of self-driving cars, the decision-making of the algorithms, must be able to handle a multitude of simple and complex driving situations flawlessly. . . . The software used to implement these algorithms must be robust and fault-tolerant.").

[67] Le, *supra* note 43 ("[M]achine learning is not a one-shot process of building a dataset and running a learner, but rather an iterative process of running the learner, analyzing the results, modifying the data and/or the learner, and repeating.").

Leveraging these windows permits appropriate levels of human judgment and enables commander compliance with the LOAC.

III. Algorithms Raise Legal Issues

The United States is bound by the Law of Armed Conflict, which embodies international treaty law and customary international law.[68] All weapon use must adhere to the LOAC[69] including fully autonomous lethal weapons,[70] which is to say it must comply with the principles of military necessity, humanity, proportionality, distinction, and honor.[71] But ensuring LAWs' programming correctly accounts for the LOAC represents the low bar for legality; layered on top of LOAC requirements are operation-specific rules of engagement, policy considerations, human restraint, and international norms. Applying the LOAC tenets to military operations occurs during planning and execution, when a commander (or servicemember) makes real-time determinations as an operational situation unfolds. But autonomy changes that. The "when" in the decision-making process occurs much earlier. The United States has suggested that LOAC issues are tied to the LAWs' programming,[72]

---

[68] U.S. CONST. art. I, § 8; art. II, § 2; art. III; and art. VI; The Paquete Habana, 175 U.S. 677, 700 (1900) ("International Law is part of our law . . . ."); *see, e.g.*, Geneva Convention for the Amelioration of the Condition of the Wounded and Sick in Armed Forces in the Field of August 12, 1949, art. 63, 1950 U.N.T.S. 32 [hereinafter GC I] ; Michael N. Schmitt & Eric W. Widmar, *On Target: Precision and Balance in the Contemporary Law of Targeting*, 7 J. NAT'L SEC. L. & POL'Y 379, 381. Note this article does not discuss domestic law implications.

[69] U.S. DEP'T OF DEFENSE, DIR. 2311.01E, DoD LAW OF WAR PROGRAM para. 4 (9 May 2006) (certified current as of 22 Feb. 2011) [hereinafter DoD LoW PROGRAM]; DoD LoW MANUAL, *supra* note 56, paras. 4.2, 4.4, 6.5.9.2; DoDD 3000.09, *supra* note 8; U.S. WORKING PAPER, AUTONOMY IN WEAPON SYSTEMS (Nov. 10 2017) [hereinafter AUTONOMY IN WEAPON SYSTEMS].

[70] Protocol Additional to the Geneva Conventions of August 12, 1949, and Relating to the Protection of Victims of International Armed Conflicts, art. 43(2), 50, June 8, 1977, 1125 U.N.T.S. 3 [hereinafter AP I]. Though not a signatory to AP I, the U.S. recognizes several components as reflective of customary international law, and views weapon reviews as a best practice. Refer to discussion in Part IV.A, *infra*, for further discussion of legal reviews of weapons. *See generally* Schmitt, *Out of the Loop*, *supra* note 18 (An outright ban on LAWs is premature).

[71] DoD LoW MANUAL, *supra* note 56, ch II; AP I, *supra* note 70, pt. IV, § 1, ch. II, arts. 45 and 51(3); Protocol Additional to the Geneva Conventions of August12, 1949, and Relating to the Protection of Victims of Non-International Armed Conflicts (Protocol II), art. 13, June 8, 1977, 1125 U.N.T.S. 609 [hereinafter AP II].

[72] AUTONOMY IN WEAPON SYSTEMS, *supra* note 69, at 2; *see also* Schmitt, *Out of the Loop*, *supra* note 18, at 273 ("Given the technological advances likely to be embedded in

meaning a learner's training must enable its later use to conform to the LOAC.[73]

A. Legal Issues Under the Law of Armed Conflict

Among the LOAC issues raised by LAWs are the bedrock principles of distinction and proportionality. Distinction simply means only proper military objectives are made the subject of attack.[74] A commander using a LAW must reasonably believe that the learner can distinguish between its intended target and those it must avoid. If used to select and engage targets autonomously, the LAW must be able to distinguish between combatants and non-combatants, and between military objectives and civilian objects.[75] In conflicts where adversaries clearly indicate their military membership, like wearing a recognizable military uniform and openly bearing arms, a particular combatant's targetable status would be readily apparent to a LAW.[76] But where adversaries and civilians are outwardly indistinguishable, a combatant's targetable status must be determined by other less visible clues, like past behavior and intent. For LAWs, interpreting body language and context pose significant hurdles, though not insurmountable.[77] Yet, to be used lawfully, a commander must

autonomous weapons . . . [l]awyers conducting the reviews will need to work closely with computer scientists and engineers to obtain a better appreciation for the measures of reliability and the testing and validation methods used on the weapons.").

[73] U.S. WORKING PAPER, HUMAN-MACHINE INTERACTION IN THE DEVELOPMENT, DEPLOYMENT, AND USE OF EMERGING TECHNOLOGIES IN THE AREA OF LETHAL AUTONOMOUS WEAPONS SYSTEMS, para. 31 (Aug. 28, 2018) [hereinafter HUMAN-MACHINE INTERACTION] ("Current artificial intelligence systems often use processes that are opaque to the human operators of the systems. This lack of understandability and transparency hinders trust and accountability and undermines the commander's ability to use LAWs properly."). Compounding the transparency problem are biases introduced into the teaching or training of the algorithm. *See* Ayanna Howard & Jason Borenstein, *The Ugly Truth About Ourselves and Our Robot Creations: The Problem of Bias and Social Inequity,* SCIENCE AND ENGINEERING ETHICS 1521–1536 (Oct. 2018) (Algorithms "find patterns within datasets that reflect implicit biases and, in so doing, emphasize and reinforce these biases as global truth."). After all, "Models are opinions embedded in mathematics." O'NEIL, *supra* note 36, at 21.

[74] DoD LoW MANUAL, *supra* note 56, para. 2.5; AP I, *supra* note 70, pt. IV, § 1, ch. II, arts. 45 and 51(3); AP II, *supra* note 71, art. 13.

[75] *Id.*

[76] *See* TUFTS, *supra* note 42.

[77] *Killer Robots*, *supra* note 9; Pontin, *supra* note 41 (Determining when civilians may be targeted because they directly participate in hostilities poses additional challenges); Kalev Leetaru, *Why Machine Learning Needs Semantics Not Just Statistics,* FORBES (Jan. 15, 2018, 11:03 AM), https://www.forbes.com/sites/kalevleetaru/2019/01/15/why-

reasonably believe that a LAW can distinguish between correct and incorrect targets and behave predictably even when circumstances change after the LAW's mission commences. If not, the commander's choice to employ the LAW in that particular circumstance would be unlawful.

To comply with proportionality, a commander must ensure an attack's likely collateral damage is not excessive in relation to the concrete military advantage expected to be gained.[78] After making a proportionality determination, the commander using his LAW must reasonably believe its effects will conform to his estimation of damage.[79] But, like the principle of distinction, the LAW's programming and training has been conducted and tested long before the facts of the commander's engagement present themselves. So, the commander must be able to predict with reliability how the LAW will behave. Unlike servicemembers, whose training and decision-making are relatively transparent, LAWs' deep learning models are opaque. The commander cannot know how it was trained or how it will make decisions given situation-specific, real-time facts.[80]

---

machine-learning-needs-semantics-not-just-statistics/#35673b0377b5 (When humans assess a situation, they do so by finding meaning and interrelationships between the various components while machine learning looks for correlations and patterns); *see also* DoD LoW Manual, *supra* note 56, para. 5.8 (Civilians directly participating in hostilities (DPH) forfeit protection from being attacked); Nils Melzer, Int'l Comm. of the Red Cross, Interpretive Guidance on the Notion of Direct Participation in Hostilities under International Humanitarian Law 70 (2009), https://www.icrc.org/en/doc/assets/ files/other/icrc-002-0990.pdf (The U.S. view on what it means to DPH is broader than others' interpretations.).

[78] DoD LoW Manual, *supra* note 56, para. 2.4.1.2. Collateral damage, including civilian death, is permissible so long as it is not "excessive" when balanced against the advantage gained by the underlying attack. *See also* AP I *supra* note 70, art. 51(5)(b) (Reiterating that indiscriminate attacks include those resulting in excessive civilian damage relative to the military advantage).

[79] "Military advantage" algorithms could shift this calculation to the LAW, presenting a programmer with the added feat of understanding and training a model on highly contextual decisions long before a conflict exists. Proposed solutions include making the collateral damage threshold adjustable, or very conservative. Nevertheless, such algorithms add complexity to an already difficult problem. *See* Schmitt, *Out of the Loop, supra* note 18, at 255-57.

[80] *See* O'Neil, *supra* note 36, at 20-21 ("[M]odels are, by their very nature, simplifications. No model can include all of the real world's complexity or the nuance of human communication. Inevitably, some important information gets left out . . . A model's blind spots reflect the judgments and priorities of its creators. . . . [M]odels, despite their reputation for impartiality, reflect goals and ideology. . . . Our own values and desires influence our choices, from the data we choose to collect to the questions we ask.").

This disparity can be overcome to an extent with rigorous testing and evaluation and operator training,[81] but ultimately, commander confidence requires well-trained LAWs—training which occurs during design. Thus, experts familiar with advising commanders on LOAC issues in military operations must be present during design to help equip LAWs' learners with lawful and reliable parameters when their models are trained.

The DoD has determined that principles like distinction and proportionality are complicated and weighty enough to assign co-located legal advisors to deployed combat and combat support units.[82] The DoD also mandated combatant commanders to obtain legal reviews of all plans, policies, directives, and rules of engagement for LOAC compliance.[83] In an operational environment, a commander's decision-making is reactive to real time circumstances, informed by battlefield experience and accompanying legal advice and judgment, among other information.[84] This is in sharp contrast to LAW's decision-making learners, which are trained and tested by human programmers—likely non-DoD.[85] The

---

[81] Other forms of algorithms are more transparent but do not offer the problem-solving advantages of deep learning, so testing and evaluation and operator training are critical to overcoming the black-box problem. *See* DoDD 3000.09, *supra* note 8, para. 4; Autonomy in Weapon Systems, *supra* note 69, at 2.

[82] DoD LoW Program, *supra* note 69, at 7-8. The LoW Program sets requirements for the military's compliance with the LOAC. To this effect, all levels of command must have qualified legal advisors available to advise on the law of war.

[83] DoD LoW Program, *supra* note 69, para. 5.11.8. This paper does not discuss responsibility or accountability, though generally acknowledges that the U.S. would bear responsibility for LOAC violations caused by use of LAWs. *See* Human-Machine Interaction, *supra* note 73, at 8; Autonomy in Weapon Systems, *supra* note 69, at 2, 4 ("[P]ersons are responsible for their individual decisions to use weapons with autonomous functions . . . it is for individual human beings . . . to ensure compliance with [the law of war] when employing any weapon or weapons system, including autonomous or semi-autonomous weapons systems."); DoD LoW Manual, *supra* note 56, para. 18.3 (Individual members of the armed forces must comply with the law of war.); Int'l Law Comm'n on the Work of Its Fifty-Third Session, *Responsibility of States for Internationally Wrongful Acts*, U.N. Doc. A/56/83, arts. 1, 4 (2001) ("Every internationally wrongful act of a State entails the international responsibility of that State.").

[84] DoD LoW Manual, *supra* note 56, para. 5.3 (Commanders must make good-faith assessments of the information available to them at the time, even if imperfect or lacking.); DoD LoW Program, *supra* note 69, at 7-8.

[85] *See* discussion *infra* Section IV.B.2; *see also* Nat'l Def. Strategy, *supra* note 12, at 3 ("[M]any technological developments will come from the commercial sector . . . ."); Gregory C. Allen & Taniel Chan, *Artificial Intelligence and National Security,* Bulletin of the Atomic Scientists (Feb. 21, 2018) [hereinafter Allen] ("There are multiple Silicon Valley and Chinese companies who each spend more annually on AI R&D than the entire United States government does on R&D for all of mathematics and computer

combat learner's decision-making models evolved and took shape in the hands of engineers long before the commander received it and likely lacked meaningful legal guidance during its training.   Thus the key difference between addressing a commander's real-time LOAC challenges, and addressing a LAW's LOAC challenges is not tied to who makes the decisions, but when they are made.

B. Issues Arise During Programming

In November 2017, the U.S. submitted to the Convention on Certain Weapons[86] Group of Governmental Experts (GGE),[87] "Weapons that use autonomy in target selection and engagement seem unique in the degree to which they would allow consideration of targeting issues during the weapon's development."[88] Restated:

> [I]f it is possible to program how a weapon will function in a potential combat situation, *it may be appropriate to consider the law of war implications of that programming*. In particular, it may be appropriate for weapon designers and engineers to consider measures to reduce the likelihood that use of the weapon will cause civilian casualties.[89] (emphasis added)

Such commentary reflects the U.S. view that LAWs must be designed in accordance with the LOAC, not that LAWs must themselves make legal decisions.[90] To emphasize that point, the U.S. offered that "it might be

---

science combined."); WORK WITH US, DEF. INNOVATION UNIT, https://www.diu.mil/work-with-us/companies (last visited Jan. 24, 2019). Though technology for LAWs will likely come from outside the DoD, the Army Research Lab (ARL) includes in its 2015–2035 research strategy a variety of AI research priorities for use in military operations, ARL, https://www.arl.army.mil/www/default.cfm?page=2401 (last visited Jan. 11, 2019).

[86] 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 (CCW), Nov. 10, 1980, 1342 U.N.T.S. 137, https://treaties.un.org/pages/viewdetails.aspx?src=treaty&mtdsg_no=xxvi-2&chapter=26&lang=en.

[87] To learn more about the Group of Governmental Experts and its inception by the Fifth Review Conference of the High Contracting Parties to the CCW, *see* Hayley Evans and Natalie Salmanowitz, *Lethal Autonomous Weapons: Recent Developments,* LAWFARE (Mar. 7, 2019, 3:28 PM), https://www.lawfareblog.com/lethal-autonomous-weapons-systems-recent-developments.

[88] AUTONOMY IN WEAPON SYSTEMS, *supra* note 69, at 4.

[89] *Id*., para. 8.

[90] *Id*., paras. 11-13.

appropriate to consider whether it is possible to program or build mechanisms into the weapon that would reduce the risk of civilian casualties."[91] In effect, this means the U.S. acknowledges that, although law of war issues typically arise within a particular military operation in real time, the unique character of autonomy bends the timeline for when such issues should be considered back to the point of programming.

In its August 28, 2018 submission to the GGE, the U.S. again emphasized the need to consider LOAC principles like distinction, proportionality, humanity, and military necessity when deciding whether to "develop or deploy an emerging technology in the area of lethal autonomous weapons systems."[92] In its January 2019 report on AI and National Security, the Congressional Research Service (CRS) reported that "domain adaptability" presents challenges for militaries when "systems developed in a civilian environment are transferred to a combat environment," and that these failures are exacerbated when AI systems are deployed at scale.[93] Thus the critical juncture for training an autonomous system's learners to stay within the bounds of the LOAC lies squarely during design when the goals and parameters that guide a learner's decisions are set.[94] The design timeframe varies by the particular aspect of technology being developed, and so determining when a judge advocate's involvement is timely must consider how the risks associated with autonomy render the current system for reviewing LOAC compliance in weapon systems inadequate.

IV. Current Process for Mitigating LOAC Issues is Inadequate

A. Legal Reviews for Weapons

When an agency contemplates buying a weapon, whether building one from scratch or adapting a commercially available variant, the current process requires at least one legal review and, for developmental weapons, an earlier legal review prior to full-scale engineering.[95] As outlined in

---

[91] *Id.*, para. 14.

[92] HUMAN-MACHINE INTERACTION, *supra* note 73, at 4-6.

[93] CRS, AI AND NAT'L SECURITY, *supra* note 11, at 33.

[94] The term "design" is used generically and not tied to acquisition process definitions.

[95] AP I, *supra* note 70; DOD LAW OF WAR MANUAL, *supra* note 56, at 337 (Discussing requirement for legal reviews of weapons); U.S. DEP'T OF DEFENSE, DIR. 5000.01, THE DEFENSE ACQUISITION SYSTEM para. E1.1.15 (May 12, 2003) (C2, 31 Aug. 2018) [hereinafter DoDD 5000.01]; U.S. DEP'T OF ARMY REG. 27-53, LEGAL REVIEW OF

DoDD 3000.09, the acquisition of LAWs requires two legal reviews: a preliminary legal review prior to formal development, and another legal review prior to fielding.[96] But these reviews examine a weapon's legality too narrowly and too belatedly. [97]   Apart from providing "weapons reviews," as they are referred to in shorthand, attorneys scrutinize weapons and weapon systems from many angles, like during an acquisition, for example, but only weapons reviews address potential LOAC concerns.[98]

When conducting a weapons review, the legal advisor receives a requirements document, a general description of the weapon, a description of the mission, the desired terminal ballistic effects of the weapon, along with tests and lab studies, if included.[99] The attorney's review focuses on if the weapon is "illegal *per se*," [100] that is, whether the weapon is prohibited for all uses, including when the U.S. has agreed to a prohibition. The review also considers "whether the weapon is 'inherently indiscriminate,' i.e., if the weapon is capable, under any set of circumstances and, in particular, the intended concept of employment, of being used in accordance with the principles of distinction and proportionality."[101] Distilled further, if a weapon is not prohibited, if it can be aimed, and if its effects can be limited, it would pass legal review.[102]

WEAPONS AND WEAPON SYSTEMS paras. 6(a), (e) (Sept. 23, 2019) [hereinafter AR 27-53] ("[D]evelopment and procurement of weapons and their intended use in armed conflict shall be consistent with the obligations assumed by the United States Government under all applicable treaties, with customary international law . . .").

[96] DoDD 3000.09, *supra* note 8. Note that in addition to legal reviews, other types of review may be triggered, but only weapons reviews as discussed herein look at LOAC. *See, e.g.*, U.S. DEP'T OF DEFENSE, DIR. 2060.1, IMPLEMENTATION OF, AND COMPLIANCE WITH, ARMS CONTROL AGREEMENTS (Jan. 9, 2001) (C2, Aug. 31, 2018).

[97] *See also Taming Killer Robots*, *supra* note 22.

[98] Army Federal Acquisition Regulation (AFARS) 5101.602-2-90 (*Revised* July 20, 2018) [hereinafter AFARS] (Requires contracting officers to obtain legal reviews and consider advice of legal counsel throughout the acquisition process.). Though the AFARS does not include a list of actions requiring legal review, the Air Force Federal Acquisition Regulation Supplement (AFFARS) does, in 5301.602-2 (*Revised* May 25, 2018). For example, when using or applying unique or unusual contract provisions; when actions are likely to be subject to public scrutiny or receive higher-level agency attention; and when issues dealing with licensing, technical data rights, and patents arise.

[99] AR 27-53, *supra* note 95, para. 6c(2); *see* Meier, *LAWS Weapons Reviews*, *supra* note 20.

[100] DoD LoW MANUAL, *supra* note 56, para. 6.2.2 (Questions considered in legal review of weapons); AUTONOMY IN WEAPON SYSTEMS, *supra* note 69, at 2.

[101] DoD LoW MANUAL, *supra* note 56, para. 6.2.2 (Questions considered in legal review of weapons); AUTONOMY IN WEAPON SYSTEMS, *supra* note 69, at 2.

[102] Though not discussed here, modifications to the weapons review process could provide another avenue for enhanced legal advisor involvement. For example, weapons

Under DoDD 3000.09, only autonomous weapons that use autonomy in new ways trigger (seemingly) additional requirements.[103] The drone in the sniper example above would have been subjected to senior official approval before formal development, and senior official approval again before fielding.[104] Although DoDD 3000.09 directs rigorous verification and validation (V&V) and testing and evaluation (T&E), from a legal perspective, none of the enhanced measures mandated by DoDD 3000.09 actually require any additional legal scrutiny beyond that already directed by Army Regulation (AR) 27-53 and DoDD 5000.01 for all weapons.[105] All new weapons, whether autonomous or not, may receive a legal review before full-scale development, and must receive one prior to fielding.[106] This means lethal, fully autonomous weapons used in ways never before seen in combat receive the same level of legal scrutiny as the L5 "Ribbon Gun," a one-time contender to replace the Army's tried and true M4 carbine.[107] But lethal autonomous weapons are not M4s; LAWs are

---

reviews for LAWs could examine how the software works and whether it would run afoul of the requirement for an operator to be able to limit its effects. *Taming Killer Robots*, *supra* note 22*; see* Meier, *LAWS Weapons Reviews*, *supra* note 20.

[103] DoDD 3000.09, *supra* note 8. Category 4c(1) through 4c(3) weapons are human-supervised, used for self-defense (as opposed to offensive use), or are non-lethal. These categories require no additional legal review beyond that required of DoDD 5000.01 for any weapon. DoDD 3000.09 para. 4d states that autonomous weapons intended for use in a manner that falls outside paragraphs 4c(1)-(3) (e.g., fully autonomous lethal weapons for offensive use) require approval of Under Secretary of Defense for Policy (USD(P)); the Under Secretary of Defense for Acquisition, Technology, and Logistics (USD(AT&L)); and the Combined Joint Chiefs of Staff before formal development and again before fielding in accordance with the guidelines in Encl. 3, DoDD 5000.01, and DoDI 5000.02 (now DoDI 5000.02T). In circular fashion, the level of scrutiny required by Encl. 3, DoDD 5000.01, DoDI 5000.02 and DoDI 5000.02T is no different than for defensive and non-lethal categories of autonomous weapons.

[104] DoDD 3000.09, *supra* note 8, para. 4d, encl. 3, para. 1a(5). The senior official review prior to formal development is intended to "ensure that military, acquisition, legal, and policy expertise is brought to bear before new types of weapons systems are used." HUMAN-MACHINE INTERACTION, *supra* note 73, at 5. Specifically, the review looks at whether five criteria are satisfied. None of the five criteria raise the legal bar. Instead they only mandate a "preliminary legal review," which follows the same rules as DoDD 5000.01.

[105] DoDD 5000.01, *supra* note 95, para. E1.1.15 requires legal review of "the intended acquisition of weapons or weapon systems." *Cf.* paras. 1(a)(5) and 1(b)(6) of DoDD 3000.09, *supra* note 8, encl. 3; AR 27-53, *supra* note 95, para. 8.

[106] AR 27-53, *supra* note 95.

[107] Tom Roeder, *The Army May Have Found its Next Rifle in a Colorado Garage,* THE GAZETTE, *reprinted in* TASK & PURPOSE, https://taskandpurpose.com/army-rifle-ribbon-gun/ (last visited Dec. 6, 2018). The Forward Defense Munitions L5 can shoot up to five rounds of 6mm caseless ammunition at once triggered by an electromagnetic actuator.

characterized by their software, which receives no scrutiny under the current weapons review process.[108] Even if it did, the gates for weapons reviews occur so late in the acquisition process that any LOAC issues arising during design would long have been set and obscured within a LAW's algorithmic black box.

B. Risk Factors Unique to Lethal Autonomous Weapons

The major risk factors rendering the current process for identifying LOAC compliance concerns fall into three categories: what we are pursuing, where we are getting it, and how we are getting it. The following discussion addresses each.

1. *What we are pursuing.*

As discussed in Section II, the technology that enables autonomy in LAWs presents significant obstacles to understanding how it works, even for the experts who create it. The greatest obstacles to fielding LAWs is the inability to test and evaluate them because combat presents near-infinite possibilities for LAWs' decision-making.[109]

The black box problem means we cannot know how a learner's model makes decisions, what biases may be trained into the model, how it set about achieving its goals, how the built-in parameters affected its decision-making, and so on. What limited opportunities exist to observe the structure and contents of the black box, the human touch points, exist when

---

See FORWARD DEFENSE MUNITIONS, https://www.fdmunitions.com/ (last visited Nov. 21, 2019).

[108] AR 27-53, *supra* note 95, para. 6(b)(1). Software and computer applications that do not directly or indirectly cause death or inflict injury to persons, facilities, or property are excluded from the definition of weapons or weapon systems, and therefore are not subject to review under AR 27-53.

[109] AUTONOMOUS HORIZONS, *supra* note 39 (Regarding testing complex autonomous systems, "Traditional methods [of testing] fail to address the complexities associated with autonomy software. . . . There are simply too many possible states and combinations of states to be able to exhaustively test each one."); *see also* SCHARRE, ARMY OF NONE, *supra* note 10, at 8 (Bradford Tousley, Director of the Tactical Technology Office, DARPA stating, "[T]he technology for autonomy and the technology for human-machine integration and understanding is going too far surpass our ability to test it."); SCHARRE, ARMY OF NONE, *supra* note 10, at 287 (paraphrasing Christof Heyns, Professor of Human Rights Law, former United Nations Special Rapporteur on extra judicial, summary or arbitrary executions from 2010-2016, "He felt it was impossible for programmers to anticipate ahead of time all of the unique circumstances surrounding a particular use of force, and thus no way for an algorithm to make a fully informed contextual decision.").

the model is trained. For the attorneys conducting weapons reviews, the aperture of these already narrow windows is further constricted by time and distance. Relatively far into the process, the LAW's legal reviewer receives from the developer or acquiring agency a prepared batch of information.[110] With only the provided documentation, testing, and lab results, the legal advisor must learn how the LAW operates well enough to opine as to its legality.

Even if weapons reviews examined software capabilities,[111] the information provided must somehow be comprehensive enough to identify issues buried deep within the learner's model *at the points in time* humans imbued the model with injects of human judgment. The attorneys conducting the weapons reviews are separated by time and distance to such a degree that a written request for a weapons review and accompanying enclosures simply cannot produce a picture of how the model was built. Unless a legal advisor versed in the weapons review process participated at key points in a model's training,[112] and could enhance and explain information provided in the request for a weapons review, paper is simply insufficient to capture what must be glimpsed in person.[113]

2. *Where we are obtaining the technology.*

Compounding our inability to adequately test LAWs, the research and development of their underlying technology occurs in scattered pockets,

---

[110] Army, Navy, and Air Force weapons reviewers' offices are housed within the Pentagon. Telephone interviews of Michael Meier, Special Assistant for Law of War Matters, U.S. Army Office of The Judge Advocate General (Feb. 5, 2019, Oct. 23, 2019) [hereinafter Interview, Meier]; Telephone interview with Aaron Waldo, Lieutenant Commander, Head of Maritime Law, U.S. Navy (Nov. 19, 2018) [hereinafter Interview, Waldo]; Telephone interview of William Toronto, Major, Chief, Operations and Int'l Law Division, Judge Advocate Office, U.S. Air Force (Dec. 18, 2018) [hereinafter Interview, Toronto].

[111] *See* AR 27-53, *supra* note 95, para. 6(b)(1).

[112] The recommendation is not persistent shadowing, but rather collaborative involvement at agreed-upon points in time based on the expertise of those involved. See discussion *infra* Section V.

[113] *See also* Michael C. Horowitz, *The Promise and Peril of Military Applications of Artificial Intelligence,* BULLETIN OF THE ATOMIC SCIENTISTS (Apr. 23, 2018), https://thebulletin.org/2018/04/the-promise-and-peril-of-military-applications-of-artificial-intelligence/ ("AI systems deployed against each other on the battlefield could generate complex environments that go beyond the ability of one or more systems to comprehend, further accentuating the brittleness of the systems and increasing the potential for accidents and mistakes.").

some within DoD but the vast majority outside DoD.[114] A LAW will not arrive to the Pentagon's front steps fully formed and ready for purchase.[115] Thus DoD will most likely acquire various AI-enabled component technologies from multiple internal and external sources,[116] often without knowing how they may ultimately be used, and then layering those on top of other AI technologies.[117] Absent access to the design table, we are limited to testing upon acquisition (or seeking to acquire) the technology.

---

[114] *See* NAT'L DEF. STRATEGY, *supra* note 12; Allen, *supra* note 85. For example, almost all of the technology for the "Architecture, Automation, Autonomy and Interfaces" capability, or A31, a product of Army Futures Command's (AFC) Future Vertical Lift Cross-Functional Team (CFT), came from small businesses and academia. Judson, *supra* note 11. And, Carnegie Mellon's Robotics Institute and National Robotics Engineering Center has partnered with the AFC's AI Task Force. Tadjdeh, *supra* note 14; Lieutenant Colonel Alan M. Apple, *Government Communication with Industry*, ARMY LAW., Iss. 3, 2019, at 44.

[115] *See* Anderson, *supra* note 18, at 388. And they should not. Though enabling a learner to identify potential targets is too removed from the commander's decision to engage them to amount to an inherently governmental function (IGF) this issue requires more discussion. The notion of an IGF is an evolving one, but at its core sets apart activities that are so completely interwoven with the sovereign nature of the U.S. that they may only be performed by federal government personnel. Included among the list of IGFs is "all combat." Combat is a bright line IGF, but even activities closely associated with an IGF may become one. Programming and training an unsupervised learner to distinguish between combatants and non-combatants inches toward what combat is all about, though falls short of specifically choosing targets. *See* 10 U.S.C. § 2330a (2012); Policy Letter 11-01, Office of Federal Procurement Policy, subj.: Performance of Inherently Governmental and Critical Functions, app. A, para. 4, app. B, paras. 5-1(a)(2), 5-1(a)(1)(ii)(B) (Sept. 12, 2012); Federal Activities Inventory Reform Act of 1998 (FAIR ACT), Pub. L. No. 105–270, § 5, 112 Stat. 2382 (1998); U.S. DEP'T OF DEF., INSTR. 1100.22, POLICY AND PROCEDURES FOR DETERMINING WORKFORCE MIX (Apr. 12, 2010) (C1, Dec. 1, 2017); *see also* DoDD 3000.09, *supra* note 8, para. 4d (Requiring high level DoD approval prior to formal development of new autonomous weapon technology.).

[116] *See* JESSE ELLMAN, LISA SAMP, & GABRIEL COLL, CTR. FOR STRATEGIC & INT'L STUDIES, ASSESSING THE THIRD OFFSET STRATEGY 14 (Mar. 2017) [hereinafter ELLMAN].

[117] DEFENSE INNOVATION BOARD, AI PRINCIPLES: RECOMMENDATIONS ON THE ETHICAL USE OF ARTIFICIAL INTELLIGENCE BY THE DEPARTMENT OF DEFENSE 37 (Oct. 31, 2019) [hereinafter AI PRINCIPLES] ("While some AI applications will be stand-alone solutions, many of the Department's efforts include layering AI solutions."); GARY SHEFTICK, U.S. ARMY, AI TASK FORCE TAKING GIANT LEAPS FORWARD (Aug. 13, 2019), https://www.army.mil/article/225642/ai_task_force_taking_giant_leaps_forward [hereinafter SHEFTICK] ("While the Army AI Task Force didn't necessarily sponsor that work [on fully autonomous cars and disaster clean-up robots] we're befitting from it. . . . We're not starting from zero. . . . That's what's allowing us to go so fast when it comes time to build out a new sensor package for automated recognition. We're able to put those systems together, because they've already solved those problems."); *see also* Le, *supra* note 43 ("In the Netflix prize, teams from all over the world competed to build the best video recommender system. As the competition progressed, teams found that they

Even if the technology is generated internally, or is industry-developed and internally refined, convincing researchers, scientists, engineers, and developers that collaborating with an attorney in the early stages of designing a learner is actually *beneficial* may require a colossal culture shift in how the role of the attorney, and attorneys themselves, are viewed. This institutional recoiling could hamper any willingness to identify projects raising possible LOAC issues in order to avoid bringing attorneys into the design process, allowing those projects to slip through the cracks until they arrive at the required weapons review gate, too late for preventative legal involvement. Operating within the status quo, to the extent it excludes judge advocates from the design process, results in a detriment to the effective and lawful use of LAWs.

3. *How we are acquiring the technology.*

As discussed above, fully autonomous lethal weapons do not yet exist, but some capabilities do.[118] Over time, machine learning capabilities will be layered together with other autonomous capabilities, and then fitted to a physical platform, punctuated throughout by iterations of testing, modifying, and refining the technology specifically for DoD's needs. Along the way, DoD will look to industry for its technology, expertise, and resources to partner with DoD's own technology, expertise, and resources to create the first LAWs.[119] To effectuate this exchange, DoD will follow an acquisition strategy, or combination of strategies. Numerous strategies exist, but the traditional process follows the DoDD 5000-series, starting with DoDD 5000.01.[120] The "5000-series," for what are now called major capability acquisitions, has been derided as slow, ineffective, expensive, risk-averse, and cumbersome for industry and the DoD alike, making it a less attractive route for rapid development, production, and fielding of emerging technologies like LAWs.[121] If DoD wished to develop a LAW from start to finish on its own, including research, development, testing, and evaluation (RDT&E), prototyping,

---

obtained the best results by combining their learners with other teams', and merged into larger and larger teams. The winner and runner-up were both stacked ensembles of over 100 learners, and combining the two ensembles further improved the results. Doubtless we will see even larger ones in the future.").

[118] SHEFTICK, *supra* note 117; Judson, *supra* note 11; Tadjdeh, *supra* note 14.

[119] *See* discussion *infra* Section IV.B.2.

[120] *See* DoDD 5000.01 *supra* note 95.

[121] *See* Bill Greenwalt, *Build Fast, Effective Acquisition: Avoid the System We've Got*, BREAKING DEFENSE (Apr. 25, 2014), https://breakingdefense.com/2014/04/build-fast-effective-acquisition-avoid-the-system-weve-got/ (Claiming the acquisition system "is really so bad we just need to figure out how to get around it most of the time. . .").

and full-scale production, it would likely follow the 5000-series framework for a major capability acquisition.[122] This scenario seems unlikely given that the lion's share of research and development for the LAWs' enabling technology will occur outside DoD's purview.[123]

Other more flexible pathways exist and that flexibility makes them more attractive for acquiring cutting edge technology. For example, Section 804 of the FY16 NDAA established Middle Tier Acquisitions (MTA) for two categories: rapid prototyping and rapid fielding of emerging military needs.[124] They are intended to be completed quickly and are therefore exempt from the most cumbersome aspects of the 5000-series.[125] Rapid prototyping requires operational capability within five years from requirement, and rapid fielding means production within six months and complete fielding within five years of a validated requirement.[126] Another authority flows from Section 2447d of the FY17

---

[122] *See* U.S. Dep't of Defense, Instr.5000.02T, Operation of the Defense Acquisition System 15 (Jan. 17, 2015) (C5, Oct. 32, 2019) [hereinafter DoDI 5000.02T] (Model 6: Hybrid Program B (Software Dominant)). Effective January 23, 2020, the 2015 version of DoDI 5000.02 was renumbered to DoDI 5000.02T (transition), and remains in effect until content is removed, cancelled, or transitioned to a new issuance, and the new DoDI 5000.02 cancels it. *See* U.S. Dep't of Defense, Instr. 5000.02, Operation of the Adaptive Acquisition Framework 3 (23 Jan. 2020). Other acquisition pathways could feed into a major capability acquisition program at different points. *See* Moshe Schwartz, Cong. Research Serv., RL34026, Defense Acquisitions: How DoD Acquires Weapon Systems and Recent Efforts to Reform the Process (2014).

[123] *See* discussion *infra* Section IV.B.2.

[124] Nat'l Def. Authorization Act for Fiscal Year 2016, Pub. L. No. 114–92, § 804, 129 Stat. 726, 882 (2015) [hereinafter FY2016 NDAA]; *see also* U.S. Dep't of Defense, Instr. 5000.80, Operation of the Middle Tier of Acquisition (MTA) (Dec. 30, 2019) [hereinafter DoDI 5000.80]; Daniel E. Schoeni, *Still Too Slow for Cyber Warfare: Why Extension of the Rapid Acquisition Authority and the Special Emergency Procurement Authority to Cyber are Half Measures*, 46 Pub. Cont. L.J. 833, 841 (2017) (Discussing the 2005 NDAA, which empowered the Secretary of Defense to waive any provision of acquisition law or regulation that unnecessarily impedes the rapid acquisition of urgently needed equipment.).

[125] To streamline the process, MTAs are not subject to the Joint Capabilities Integration and Development System (JCIDS) or the programmatic requirements of DoDI 5000.02 or DoDI 5000.02T. DoDI 5000.80, *supra* note 124; *see* Combined Joint Chiefs of Staff, Instr. 5123.01H, Charter of the Joint Requirements Oversight Council (JROC) and Implementation of the Joint Capabilities Integration and Development System (JCIDS) (Aug. 31, 2018) (The purpose of which is to enable the JROC to execute its statutory duties to identify, prioritize, and fill capability gaps.); *see generally* DoDI 5000.02T, *supra* note 122, para. 5a(4), 5b.

[126] FY2016 NDAA, *supra* note 124, § 804. For a good snapshot of Section 804 MTA and other streamlined acquisition pathways see Pete Modigliani, et al., *Middle Tier*

NDAA, which permits non-competitive follow-on production contracts or other transactions for prototype projects when the project "addresses a high priority warfighter need or reduces the costs of a weapon system."[127] Section 2447d also grants Service Secretaries transfer authority, which means they can transfer available procurement funds to pay for low-rate initial production.[128]

Despite its reputation, the 5000-series has its own efficiencies. Department of Defense Directive 5000.71 enables combatant commands to request processing of urgent operational needs, which means a validated request sees a fielded solution within two years.[129] This process may be used in conjunction with Section 806 MTA.[130] Section 806's Rapid Acquisition Authority (RAA) authority used together with DoDD 5000.71 enables warfighter needs to be fulfilled exceptionally quickly.[131]

Acquisition and Other Rapid Acquisition Pathways, MITRE (Mar. 2019), https://aida.mitre.org/wp-content/uploads/2019/03/Middle-Tier-and-Rapid-Acquisition-Pathways-8-Mar-19.pdf.

[127] National Defense Authorization Act for Fiscal Year 2017, Pub. L. No. 114-328, § 2447d 130 Stat. 2259 (2016) (10 U.S.C. 2447d) [hereinafter FY2017 NDAA] ("Mechanisms to speed deployment of successful weapon system component or technology prototypes").

[128] FY2017 NDAA, *supra* note 127, § 2447d(b).

[129] U.S. DEP'T OF DEFENSE, DIR. 5000.71, RAPID FULFILLMENT OF COMBATANT COMMANDER URGENT OPERATIONAL NEEDS (24 Aug. 2012) (C1, 31 Aug. 2018) (Rapid Fulfilment of Combatant Commander Urgent Operational Needs) [hereinafter DoDD 5000.71]; see also DoDI 5000.02T, supra note 122, encl. 13 (Urgent Capability Acquisition); U.S. DEP'T OF DEFENSE, INSTR. 5000.81, URGENT CAPABILITY ACQUISITION (31 Dec. 2019). Contract vehicles may add additional efficiencies to the various available rapid acquisition pathways, like Other Transaction Authority (OTA) under 10 U.S.C. § 2371, Pub. L. No. 85-- 568, 72 Stat. 426 (1958), or industry engagement and streamlining efforts like the Commercial Solutions Opening pilot program authorized in § 879 of FY2017 NDAA, *supra* note 127; and the Small Business Innovation Research Program under 15 U.S.C. § 638; *see also* U.S. SMALL BUS. ADMIN. OFFICE OF INV. AND INNOVATION, SMALL BUSINESS INNOVATION RESEARCH (SBIR) AND SMALL BUSINESS TECHNOLOGY TRANSFER (STTR) PROGRAM POLICY DIRECTIVE (May 2, 2019), https://www.sbir.gov/sites/default/files/SBIR-STTR_Policy_Directive_2019.pdf [hereinafter SBIR/STTR].

[130] Bob Stump National Defense Authorization Act for Fiscal Year 2003, Pub. L. No. 107-314, § 806(c), as amended (10 U.S.C. 2302 note).

[131] Memorandum from Under Secretary of Defense (Acquisition, Technology and Logistics) to Secretaries of the Military Departments, subj.: Acquisition Actions in Support of Joint Urgent Operational Needs (JUONs) attachment, para. 4 (Mar. 29, 2010) (The goal in using SECDEF's Rapid Acquisition Authority is to achieve contract award within 15 days.). Two of the 5000-series' most efficient acquisition pathways, DoDI 5000.02T Model 4, Accelerated Acquisition Program, and DoDD 5000.71/ DoDI 5000.81 for Urgent Capability Acquisitions, have very short timelines and lower dollar thresholds, making them unlikely for start-to-finish development of LAWs, though could

Though not an acquisition pathway, the DoD may also pursue and adapt commercial technology derived from Independent Research and Development (IR&D) under 10 U.S.C. § 2372.[132] Independent Research and Development envisions DoD adapting research and development conducted in the commercial sector for defense purposes.[133] Under Section 2372, DoD reimburses contractor expenses for research and development conducted outside of the department's control and without direct DoD funding.[134] Projects must have potential interest to the DoD, and include those that improve U.S. weapon system superiority and promote development of critical technologies.[135]

With all that flexibility and speed, one may wonder where in the process weapons reviews fall. Each acquisition pathway follows its own procedural rules and allows for varying degrees of overlap with other pathways,[136] but the only one that dictates when weapons reviews must be conducted is the 5000-series. The 2019 version of Army Regulation (AR) 27-53 contemplates rapid acquisition strategies and acquisition of emerging technology and attempts to bridge the gap by requiring a weapons review pre-development for weapons or weapon systems sought through a rapid acquisition process.[137] Acknowledging the importance of early reviews, AR 27-53, paragraph 6g requires preliminary legal reviews for pre-acquisition category projects, like advanced concept technology demonstrations, rapid fielding initiatives, and general technology development and maturation projects when the technology is "intended to be used . . . in military operations of any kind."[138]

---

be used effectively once the technology matures. DoDI 5000.02T, *supra* note 122, model 4; DoDD 5000.71, *supra* note 129, para. 4.2a; *see generally*, DoDI 5000.81, *supra* note 129.

[132] 10 U.S.C. § 2372 (1990); Defense Federal Acquisition Regulation Supplement (DFARS) 231.205-18 [hereinafter DFARS].

[133] 10 U.S.C. § 2372. Given the private sector's investment in AI research and development, the technology for LAWs will likely come from industry, as opposed to from within the federal government. *See* discussion *supra* notes 85, 114, 117. SBIR/STTR contracts are another variant on how DoD partners with industry to rapidly develop emerging technology. *See* SBIR/STTR, *supra* note 129.

[134] 10 U.S.C. § 2372; *see also* Federal Acquisition Regulation (FAR) 31.205-18 [hereinafter FAR] (definition of IR&D).

[135] DFARS, *supra* note 132, 231.205-18.

[136] *See, e.g.*, DoDI 5000.80, *supra* note 124.

[137] AR 27-53, *supra* note 95, para. 6h.

[138] AR 27-53, *supra* note 95, para. 6g.

Refocusing the issue, the 5000-series is the least likely path for acquiring LAWs' technology because it is notoriously slow and rigid, but the governing DoD policy on LAWs, DoDD 3000.09, points to the DoDD 5000-series framework for the timing of weapons reviews within the acquisition process. Yet, as discussed in Section III.B, fluid timing of judge advocate involvement is a crucial element to mitigating LOAC issues. This is problematic.

Because LOAC issues raised by LAWs' algorithms arise when the learners are trained, the current acquisition process, regardless of pathway, renders weapons reviews either too late, too narrow, or too disconnected from the various human touch points that allow consideration of targeting issues during the weapon's development.[139] Those human touch points offer crucial windows for appropriate levels of human judgment to be incorporated into LAWs' algorithmic models and their training—judgment tempered by legal counsel similar to that which commanders receive during military operations.[140] Fortunately, no regulatory hurdles prevent an enhanced legal advisor role, but hesitancy from industry could.

V. Building on Current Efforts to Address Blind Spots

The current legal framework allows for broadening the scope of judge advocate involvement. The services have taken steps to involve judge advocates earlier on in a weapon's development, even before the weapon or its technology enters the acquisition process. But these efforts are just the first steps and blind spots remain. The following section touches on the permissive character of regulations governing legal advisor involvement, some efforts to expand the scope of current legal advisor

---

[139] *See also Killer Robots*, *supra* note 9, *citing* INT'L COMM. OF THE RED CROSS, A GUIDE TO THE LEGAL REVIEW OF NEW WEAPONS, MEANS AND METHODS OF WARFARE 23, http://e-brief.icrc.org/wp-content/uploads/ 2016/09/12-A-Guide-to-the-Legal-Review-of-New-Weapons.pdf (last visited Jan. 31, 2019) [hereinafter ICRC GUIDANCE] ("[R]eviews should take place at the stage of the conception/design of the weapon . . . .").
[140] This article does not discuss implications of LAWs' creating their own problem-solving algorithms and how LAW's own self-modification may impact the need for review once deployed. The DoD uses agile contracting methods for other software-dependent applications, which allow for continuous iterations of updates and modifications while maintaining operator employability. Judge advocate/commander teams should also participate in that process. *See* U.S. DIGITAL SERV., DIGITAL SERV. PLAYBOOK, https://playbook.cio.gov (last visited Dec. 2, 2019) (Explaining how the U.S. Digital Service approaches software development, and the emphasis on agile practices (play 4)).

involvement, where vulnerabilities remain, and how to use existing resources to address them.

A. Getting the right people in the right place.

Generally, those seeking legal advice in carrying out DoD business may readily obtain it.[141] The issue is not a lack of legal advisors but not knowing how to or being unwilling to use them. Figuring out how judge advocates add value during design and training of the LAWs' enabling technology opens doors of possibilities but remains an unanswered question, partially because LAWs' technology only exists in incomplete fragments, and partially because lawyer involvement in the earliest stages only occurs on an ad hoc basis, if at all.[142]

Within the acquisition arena, attorneys play important roles throughout the process, but are not tasked with reviewing LOAC concerns in weapon systems.[143] For instance, when an acquisition is contemplated, legal advisors located within requiring agencies prepare acquisition packages, provide support to contracting units reviewing proposed solicitations, participate as members of acquisition teams offering legal and non-legal counsel, and offer legal advice to source selection decision authorities.[144] Within the Army, many of those attorneys are not co-

---

[141] *See, e.g.*, OFFICE OF THE GENERAL COUNSEL, https://www.hhs.gov/about/agencies/ogc/index.html (last visited Nov. 25, 2019); OFFICE OF THE ARMY GENERAL COUNSEL, https://ogc.altess.army.mil/ (last visited Nov. 25, 2019); THE ARMY JUDGE ADVOCATE GENERAL'S CORPS, https://www.goarmy.com/jag/about.html (last visited Nov. 25, 2019).

[142] Interview, Meier, *supra* note 110; *see* discussion *infra* note 174, Telephone interview of Graham Todd, Chief, Operations Policy and Planning, Judge Advocate's Office, U.S. Air Force (Dec. 18, 2018) [hereinafter Interview, Todd]; *see also*, DSB, ROLE OF AUTONOMY, *supra* note 14, at 18 (Recommending that "technologists and designers get direct feedback from the operators, [and] the Military Services should schedule periodic, on-site collaborations that bring together academia, government and not-for-profit labs and industry and military operators to focus on appropriate challenge problems.").

[143] Responsibility for that rests solely with the Special Assistant for Law of War Matters, U.S. Army Office of The Judge Advocate General (OTJAG).

[144] AFARS, *supra* note 98, 5101.602-2-90. Contracting vehicles are as varied as acquisition pathways. Depending on the type of contract vehicle used, attorneys may be involved more or less. For example, the Other Transaction Authority Guide recommends securing "the early participation of subject matter experts such as legal counsel" when awarding an OTA and throughout the OTA process. OFF. OF UNDER SEC'Y OF DEF. FOR ACQUISITION AND SUSTAINMENT, OTHER TRANSACTION AUTHORITY GUIDE sect. II(D)(1) (Nov. 2018), https://www.dau.mil/tools/t/Other-Transactions-(OT)-Guide [hereinafter OTA Guide]; *see, e.g.*, CONTRACT & FISCAL LAW DEP'T, THE JUDGE ADVOCATE GEN.'S

located with the agency they support; rather they belong to a contracting support unit (e.g. Contracting Support Brigades), a Staff Judge Advocate's Office, or within Army Material Command. Despite their involvement as legal advisors, these attorneys' roles are not especially intended for spotting design or operational issues associated with the LOAC, and they are not physically co-located in the places most likely to encounter them.[145] Their roles in refining requirements for a LAW would be more concerned with accurately describing what the LAW needs to be able to *do*, not *how* the LAW must do it.[146] An attorney assisting with refining an agency's needed capabilities for a LAW could simply include a requirement for LOAC compliance. But the complexity of translating what that actually means—and threading LOAC compliance through programmer, evaluator, and operator—lends itself poorly to simple insertion as a contractual requirement.[147] Furthermore, downstream attorneys reviewing performance of that requirement are as ill-equipped as the weapons reviewer to spot potential flaws or operational defects in how a programmer trained a model to function within the LOAC. Recent efforts to modernize how the Army acquires emerging technology and advances certain types of technology set the stage for an expanded judge advocate role.

The Army's hub of innovation and cutting-edge research resides within Army Futures Command (AFC), headquartered in Austin, Texas,

---

LEGAL CTR. & SCH., U.S. ARMY, CONTRACT ATTORNEYS DESKBOOK, at 2-13 to 2-24 (2018) (Sample Contract Review Checklist).

[145] Although the substance of an attorney's legal advice could be affected by the nature of the procurement, and their role could evolve to include spotting potential LOAC issues. *See* FAR, *supra* note 134, 1.102-3; AFARS, *supra* note 98; *see* Major Andrew S. Bowne, U.S. Air Force, *Innovation Acquisition Practices in the Age of AI*, ARMY LAW., Iss. 1, 2019, at 75 (discussing different roles acquisitions attorneys can play, including understanding the technical possibilities of AI and the ethical and legal implications of such acquisitions).

[146] As applied to LAWs, this would mean its firing feature would receive legal scrutiny, but not the software programming upon which it operates. *See* Vincent Boulanin, *SIPRI Insights On Peace And Security No. 2015/1, Implementing Article 36 Weapon Reviews In The Light Of Increasing Autonomy In Weapon Systems* 2 (Nov. 2015), https://www.sipri.org/sites/ default/files/files/insight/SIPRIInsight1501.pdf; DoD LoW MANUAL, *supra* note 56, para. 6.2.2 (Questions considered in legal review of weapons).

[147] Which is not to say it is unimportant. A requiring agency's legal advisors (civilian or uniformed) offer tremendous value in the requirements development phase of LAWs. Though the critical timing for spotting LOAC issues in machine learning models occurs early on, ensuring an agency's requirements adequately capture its needs regarding LOAC compliance is no less significant.

with offices scattered throughout the U.S.[148] Judge advocates and civilian attorneys working within AFC already advise its cross-functional teams (CFTs), CCDC research labs, the Artificial Intelligence Task Force (AI TF) and its Applications Lab.[149] The breadth of legal advice they offer remains in its nascent stages, but could include early issue-spotting across the spectrum of legal topics, including LOAC issues.[150] This is one of the locations within the Army most likely to encounter the technology for LAWs in its earlier stages, either by virtue of the Army's own internal research and development, or resulting from some variety of Army-Industry partnership.[151] The judge advocates and civilian attorneys within AFC and it subordinate units may be dispatched outside of AFC, including upon industry request, wherever their presence is needed. [152] The vulnerability resides in the assumption that AFC (and sister service equivalents) is an omniscient entity, when AFC is but one agency within DoD with limited resources and capable of seeing only those projects that fall within its broad reach.

To standardize efforts on this issue, DoD should promulgate a consistent, uniformly applicable policy requiring the employment of judge advocates in service of identifying LOAC issues in LAWs. The judge advocate/commander teams should be situated within AFC but mobile and readily available to whomever needs them. Recalling the sniper-targeting drone example from Section II.B, the programmer unfamiliar with the LOAC would doubtlessly also be unfamiliar with its prohibition on targeting those who are *hors de combat*, meaning they are "out of the

---

[148] ARMY FUTURES COMMAND, https://www.army.mil/futures#org-about. Though there are other pockets of innovation outside of AFC. *See*, *e.g*. U.S. ARMY RAPID CAPABILITIES AND CRITICAL TECHNOLOGIES OFFICE, https://rapidcapabilitiesoffice.army.mil/ (last visited 8 Feb. 2020); DEFENSE DIGITAL SERVICE, https://dds.mil/ (last visited Feb. 8, 2020).

[149] Interview with Darren Pohlmann, Lieutenant Colonel, Deputy Staff Judge Advocate, Army Futures Command (Nov. 13, 2019) [hereinafter Interview, Pohlmann]. Legal advisors support other federal and non-federal entities conducting research in this area as well, like DARPA and the Massachusetts Institute of Technology (MIT).

[150] *Id*. These attorneys are able to work with the Army's OTJAG's Special Assistant for Law of War Matters, Mr. Michael Meier, who is solely responsible for conducting the Army's weapons reviews.

[151] *See* ARMY FUTURES COMMAND, https://www.army.mil/futures#org-about (last visited Nov. 25, 2019) ("Army Futures Command leads a continuous transformation of Army modernization in order to provide future warfighters with the concepts, capabilities and organizational structures they need to dominate a future battlefield.")

[152] Interview, Pohlmann, *supra* note 149.

fight."[153] It takes little imagination to envision a scenario where a sniper exhibits the same qualities as an unconscious soldier lying motionless aside his weapon, a civilian hunter awaiting a clear shot, or a medic rendering aid to a fallen comrade. Each may appear to be laying in a prone position, camouflaged, motionless, isolated, and aiming a weapon in a particular direction, yet only the sniper would be a valid target.

Training a learner's model to identify the nuances of what makes the sniper's legal status different—and thus subject to attack—requires both a firm understanding of the law that governs when one is out of the fight and the characteristics, behavior and tactics employed by one who is fairly in it. Put another way, the model must set a sniper apart from a teenager hiding with a paintball gun. The experienced operational commander (or former operator) would understand these characteristics and be able to articulate them so a programmer could train the model to search for and recognize them. The judge advocate versed in dispensing operational advice would complement the commander's tactical expertise with legal perspective, thus adding dimension and detail to the programmer's understanding, ergo the model's understanding, of the LOAC. Lethal autonomous weapons' models are simply extensions of humans' prediction and problem-solving models; they both need multiple sources of "expertise" in developing their decision-making. The entity within the Army with attorneys best-situated to team up with commanders and offer their expertise at the critical time is AFC.

While judge advocates offer the advantages of training, experience, and education,[154] they are not the only attorneys able to provide such

---

[153] When *hors de combat*, a person must be treated humanely and may not be subject to attack. The phrase is used in Common Article 3 of all four 1949 Geneva Conventions, *see e.g.*, GC I, *supra* note 68, art. (3)(1), though not defined until 1977 in AP I, *supra* note 70, art. 41. The LoW MANUAL defines the following persons as *hors de combat* when they "abstain from any hostile act and do not attempt to escape: persons in the power of an adverse party; persons not yet in custody, who have surrendered; persons rendered unconscious or otherwise incapacitated by wounds, sickness, or shipwreck; and persons parachuting from aircraft in distress." DoD LoW MANUAL, *supra* note 56, para. 5.9.

[154] Determining which judge advocates, DoD civilian attorneys, and commanders are best suited to advise warrants additional discussion, though being possessed of significant deployment/operational experience applying the LOAC, and familiarity with the concepts of machine learning would be necessary. The FY2020 NDAA directs the Secretary of Defense to develop an education strategy for servicemembers in "relevant occupational fields on matters relating to artificial intelligence." The curriculum includes topics on software coding, artificial intelligence decisionmaking via machine learning and neural

support. The DoD abounds with highly capable civilian attorneys and those with prior service as judge advocates across all services. Their expertise and experience with military operations and the acquisitions process provides a valuable resource. On the issue of whether the attorney must be conversant in coding, a familiarity with the concepts would be desirable, but the emphasis should be instead on collaborating with the various experts designing the technology, which requires communication and interpersonal skills and a well-rounded support network as much as anything else.[155]

B. Doing the right things.

The role of the judge advocate/commander or operator team should be in assisting the engineers, scientists, and programmers build LOAC durability into the deep learning algorithms' architecture, leveraging the human touch points, so that when a commander or operator manipulates the LAWs' various capabilities and constraints, whatever machinations take place within the black box also stay within the bounds of the LOAC.[156] As a practical matter, a LAW is useless unless a commander can reliably control it. Knowing that she or he is accountable for the foreseeable consequences of its behavior, a commander contemplating using a LAW that she or he does not understand, would simply bench it.[157] Outwardly, a commander experiences a model's training through its performance and the LAW's operator interface, which is the means by which the commander "makes informed and appropriate decisions in

---

networks, and ethical issues. The practice of law should be a "relevant occupational field." FY2020 NDAA, *supra* note 14, § 256.

[155] *See e.g.*, MISSION COMMAND DEVELOPMENT INTEGRATION DIRECTORATE (CDID) BATTLE LAB, https://usacac.army.mil/organizations/mccoe/cdid (last visited Nov. 25, 2019), which supports AFC and other agencies.

[156] Le, *supra* note 43 ("There is no sharp frontier between designing learners and learning classifiers; rather, any given piece of knowledge could be encoded in the learner or learned from data. So machine learning projects often wind up having a significant component of learner design, and practitioners need to have some expertise in it.").

[157] DSB, ROLE OF AUTONOMY, *supra* note 14, at 11 ("[A]utonomous systems present a variety of challenges to commanders, operators and developers . . . these challenges can collectively be characterized as a lack of trust that the autonomous functions of a given system will operate as intended in all situations."); *see also* Michael Meier, *Lethal Autonomous Weapons Systems – Is It the End of the World as We Know It – or Will We Be Just Fine? in* COMPLEX BATTLESPACES: THE LAW OF ARMED CONFLICT AND THE DYNAMICS OF MODERN WARFARE (Christopher M. Ford & Winston S. Williams eds., 2018).

engaging targets."[158] Thus, the interface provides a critical means for the commander to set mission-specific parameters on the LAW. A recent study by the Combat Capabilities Development Command (CCDC) Army Research Lab (ARL) examined what it takes for a human to trust a robot.[159] The research team found that soldiers reported lower trust after seeing a robot commit an error, even when the robot explained the reasoning behind its decisions. The lack of trust endured, even when the robot made no more errors. The heart of the issue is trust, which means those responsible for designing LAWs' deep learning models must not only have a keen awareness of commanders' real-time operational needs but also how to translate those needs through the operator interface into a LOAC-resilient model.[160]

To this end, like in the sniper-targeting drone example discussed above, a judge advocate/commander team would provide real-world operational scenarios, offer insights on the interplay between targeting decisions and the LOAC, theater-specific rules of engagement and policy considerations, and explore how different options built into the operator interface could control for varying levels of risk. The team could also assist with ensuring the machine and human share the same objective, and that they are able to adjust in unison as circumstances change.[161] Related to this concept is understanding each other's "lanes" or in other words, the machine and human knowing the limitations of the others' decision capabilities, and how that may change as objectives change.[162] Integrating operational realities into a learner's model means they must be taught, and what better teachers than those who bear the responsibility in real life?

---

[158] DoDD 3000.09, *supra* note 8, para. 4a(3).

[159] ARMY RESEARCH LABORATORY PUBLIC AFFAIRS, WHEN IT COMES TO ROBOTS, RELIABILITY MAY MATTER MORE THAN REASONING (Sept. 25, 2019), https://www.army.mil/article/226681/when_it_comes_to_robots_reliability_may_matter_more_than_reasoning.

[160] *See* ELLMAN, *supra* note 116, at 16 (Lieutenant General (Ret.) Robert Schmidle, USMC, and former Principal Deputy Director of the Office of Cost Assessment and Program Evaluation (CAPE) emphasized this point: "[I]f you want decision-makers to trust the algorithms you need those decision-makers to be involved in, and capable of understanding, the development of those algorithms, because they are not going to necessarily be involved in the real-time decisions that the algorithms would make.").

[161] CRS, AI AND NAT'L SECURITY, supra note 11, at 35 (Referring to the concept of *goal alignment*).

[162] CRS, AI AND NAT'L SECURITY, supra note 11, at 35 (Referring to the concept of *task alignment*).

The DoD has mandated legal advice for all operational decision-makers and offers in-theater judge advocates to dispense it, but judge advocates offer more. They bring to the table critical thinking skills and a diversity of thought that is important to the collaborative process, and is exactly what they offer commanders in operational settings.[163] Viewing the lawyers as teammates as opposed to ivory tower gate keepers maximizes the skill set they possess. Providing the same access for researchers, programmers, and engineers as the military offers operational commanders just means the judge advocate's place of duty changes; their advice is required at the design table, not just while deployed.[164]

C. At the right time.

Ensuring the right people are in the right place at the right time hinges on when DoD gets its first opportunity to examine autonomous technology. If the first opportunity comes as part of the acquisition process, the Federal Acquisition Regulation (FAR) and its supplements, applicable to the vast majority of acquisitions, permit early and ongoing legal involvement beyond legal reviews.[165] As discussed above, for developmental weapons or weapon systems, AR 27-53 provides that initial reviews may be made at the earliest possible stage, and pre-acquisition technology projects intended for military use must receive a preliminary legal review.[166]

---

[163] Telephone Interview with William Gamble, General Counsel, Defense Digital Service (Oct. 3, 2019).

[164] Additional safeguards have been proposed. Meier, *LAWS Weapons Reviews*, *supra* note 20; Ford, *Autonomous Weapons*, *supra* note 20, at 26; *Killer Robots*, *supra* note 9 (recommending involving lawyers early in the development process); ICRC GUIDANCE, *supra* note 139 (suggesting need to conduct reviews during concept and design); Larry Lewis, *Insights for the Third Offset*, CNA (Sept., 2017), https://www.cna.org/CNA_files/PDF/DRM-2017-U-016281-Final.pdf; Larry Lewis *Redefining Human Control*, CNA (Mar. 2018), https://www.cna.org/CNA_files/PDF/DOP-2018-U-017258-Final.pdf.

[165] FAR, *supra* note 134; DFARS, *supra* note 132, 203.170 (discussing legal review pre-award of contract); AFARS, *supra* note 98, 5101.602-2-90.

[166] AR 27-53, *supra* note 95, paras. 6e, 6g. Worth noting, though outside the scope of this discussion, is the requirement to obtain a weapons review if a weapon or weapon system changes after fielding "such that it is no longer the same system or capability described in the legal review request." This includes substantial changes to its intended use or anticipated effects. *Id*. at para. 6f. This further complicates the black box problem discussed *supra* Section II if para. 6f is interpreted to mean that LAWs must come equipped with a mechanism to determine when its machine learning models have changed to such a degree that they are no longer the same system or capability, or that

The Navy counterpart to AR 27-53, Secretary of the Navy Instruction 5000.2E, requires that potential acquisition or development of weapons receives a legal review during "the program decision process."[167] The Air Force equivalent, Air Force Instruction 51-401, requires a legal review "at the earliest possible stage in the acquisition process, including the research and development stage."[168] But in practice, across all services, actual legal advisor involvement more closely aligns with the baseline requirements discussed in Section IV.A,[169] meaning early involvement of legal advisors to spot LOAC issues rarely occurs.

In an effort to integrate judge advocates earlier into the process *pre-acquisition*, the Air Force includes judge advocates as members of cross-functional acquisition teams, advising within an assigned portfolio, like F-15s, Cyberspace, or Intelligence, Surveillance, and Recognizance (ISR).[170] Air Force judge advocates also provide direct legal support to the research labs. Of the ten research lab directorates, three have in-house legal counsel and the remaining satellite locations receive support from a nearby legal office.[171] If LOAC-specific issues arise, servicing legal advisors send them through their channels to a single office at the Air Force Judge Advocate's Office (AF JAO).[172]

In the Navy, the judge advocates performing weapons reviews engage in outreach with program managers, educating them about their responsibilities to get legal reviews and involve legal advisors in the

---

their intended use changes beyond that described in the initial request for a weapons review.

[167] U.S. DEP'T OF NAVY, SEC'Y OF NAVY INSTR. 5000.2E, DEP'T OF THE NAVY IMPLEMENTATION AND OPERATION OF THE DEFENSE ACQUISITION SYSTEM AND THE JOINT CAPABILITIES INTEGRATION AND DEVELOPMENT SYSTEM 1.6.1 (Sept. 1, 2011) [hereinafter SECNAVINST 5000.2E].

[168] U.S. DEP'T OF AIR FORCE, INSTR. 51-401, THE LAW OF WAR para. 2.5.2.1 (Aug. 3, 2018).

[169] Interview, Meier, *supra* note 110; Interview, Waldo, *supra* note 110; Interview, Toronto, *supra* note 110. Note that the Navy conducts legal reviews for Marine Corps' weapons. SECNAVINST 5000.2E, *supra* note 167; Telephone interview with Joe Rutigliano, Branch Head, International and Operational Law Branch, Judge Advocate Division, U.S. Marine Corps (Feb. 11, 2019).

[170] Interview of Andrew Bowne, Major, U.S. Air Force, Associate Professor, Contract and Fiscal Law Dep't, The Judge Advocate Gen.'s Legal Ctr. and Sch., in Charlottesville, Va. (Jan. 27, 2019) [hereinafter Interview, Bowne].

[171] Telephone interview with Jonathan Compton, Attorney Advisor, Headquarters, Air Force Research Lab, Wright-Patterson Air Force Base (Dec. 12, 2018).

[172] Telephone interview with Jonathan Compton, Attorney Advisor, Headquarters, Air Force Research Lab, Wright-Patterson Air Force Base (Dec. 12, 2018); Interview, Toronto, *supra* note 96.

acquisition process.[173] Legal advisors are also physically located in or near some research labs, though their support does not envision addressing LOAC concerns.[174] For all services, unless the researchers, programmers, and engineers know to ask, LOAC issues may well go unnoticed until it is too late to fix them.[175] A DoD policy could change that.

As discussed in Section IV.B.2, DoD's first opportunity to examine autonomous technology will likely arise from outside DoD. This scenario leads to the greatest challenge and most promising solution to mitigating the various risk factors bearing on LAWs and the LOAC: access. Specifically, whether industry is willing to bring DoD into its design process.

The DoD has been directed to engage with industry. In his March 2018 memorandum the Deputy Secretary of Defense encouraged cooperation with industry: "While we must always be mindful of our legal obligations, they do not prevent us from carrying out our critical responsibility to engage with industry." [176] Congress goes beyond encouragement and directs the DoD to "accelerate the development and fielding of artificial intelligence capabilities [and to] ensure engagement with defense and private industries."[177] In Section 238(c)(2)(H) of the FY2019 NDAA, Congress states that designated officials "shall work with appropriate officials to develop appropriate ethical, legal, and other

---

[173] Interview, Waldo, *supra* note 110.

[174] *Id.*

[175] Two examples of ad hoc requests for judge advocate support illustrate the need and value added by involving legal advisors early in the research, development, testing, and evaluation process. Example 1: Researcher from an Air Force Research Lab asked about legality of biological research, which required higher level review prior to proceeding. The question only came up because researcher thought to ask. Interview, Toronto, *supra* note 110. Example 2: Scientist at research lab asked for legal support during software development testing. Legal advisor went to lab for a few days, observed, exchanged feedback on designing the algorithms, and how the machine would behave. This request only came up because scientist thought to ask, but the collaborative process was already in place and ongoing between the requirements owner and the lab. Interview, Todd, *supra* note 142.

[176] *See* Memorandum from Deputy Secretary of Defense to the Secretaries of the Military Departments, et al., subj.: Engaging with Industry (Mar. 2, 2018) [hereinafter Engaging with Industry].

[177] FY2019 NDAA, *supra* note 14, §§ 238(c)(2)(A)-(B), 238(c)(2)(H); Engaging with Industry, *supra* note 176, at 2 ("The Department's policy continues to be that representatives at all levels of the Department have frequent, fair, even, and transparent dialogue with industry on matters of mutual interest . . ."); NAT'L DEF. STRATEGY, *supra* note 12. Other recent initiatives support this endeavor. *See* AI TASK FORCE, *supra* note 14; DOD AI STRATEGY, *supra* note 5.

policies for the Department governing the development and use of artificial intelligence enabled systems and technologies in operational situations." [178]   (emphasis added).   Industry engagement is not only permitted, it is mandated.[179]

Though DoD may desire industry engagement, that willingness is not necessarily mutual. Barriers include mistrust of DoD, more lucrative and less cumbersome options elsewhere, resistance to supporting DoD's mission, lack of awareness about opportunities to work with DoD, and lack of understanding how to access those opportunities.[180] The DoD has taken strides to address the latter four concerns by creating an approachable physical presence in tech hubs like the Army Applications Lab in Capitol Factory, Austin, Texas, SOFWERX in Tampa, Florida, the Air Force's AFWERX innovation hubs in Washington, D.C., Las Vegas, and Austin, and the AI Lab in Pittsburgh, Pennsylvania. It has also expanded opportunities for quick turnaround payoffs with on-the-spot contracts awarded during industry engagement events, like the Air Force's Pitch Days, the Navy's Small Business Innovation and Small Business Technology Transfer (SBIR/STTR) and NavalX, and the Army's Innovation Days. [181] Reverse Industry Days foster transparency and encourage communication by offering industry a chance to share its

---

[178] FY2019 NDAA, *supra* note 14, § 238(c)(2)(H); *see also* DoD LoW PROGRAM, *supra* note 69; NAT'L DEF. STRATEGY, *supra* note 12 (Prioritizing investment in advanced autonomous systems).

[179] Hesitancy for increased legal advisor participation may stem from ethical objections to DoD members helping private companies develop new technology or for fear of giving one company an unfair competitive advantage over another. 5 C.F.R. § 2635.702 (1997); U.S. DEP'T OF DEF., DIR. 5500.07-R, JOINT ETHICS REGULATION 3-209 (Nov. 17, 2011); FAR, *supra* note 134, 9.505(2)(b), 3.104-4(a) (One of the main principles for avoiding conflicts in acquisitions is preventing unfair competitive advantage . . ."). However, the Office of Federal Procurement Policy's "myth-busting" series allays such fears. OFFICE OF MGM'T & BUDGET, OFFICE OF FEDERAL PROCUREMENT POLICY, https://www.whitehouse.gov/omb/ management/office-federal-procurement-policy/ (last visited July 8, 2019).

[180] CONG. RESEARCH SERV., R45521, DEP'T OF DEF. USE OF OTHER TRANSACTION AUTHORITY 16 (Feb. 22, 2019); FUTURE OF LIFE INSTITUTE, OPEN LETTER ON AUTONOMOUS WEAPONS (July 28, 2015), https://futureoflife.org/open-letter-autonomous-weapons/?cn-reloaded=1.

[181] Brenda Marie Rivers, *Will Roper: Air Force Expanding 'Pitch Day' Across US*, EXECUTIVE GOV (Mar. 18, 2019), https://www.executivegov.com/2019/03/will-roper-air-force-expanding-pitch-day-across-us/; U.S. DEP'T OF THE NAVY, NAVY SBIR/STTR HOME, https://www.navysbir.com (last visited Feb. 8, 2020); U.S. DEP'T OF THE NAVY, NAVALX, https://www.secnav.navy.mil/agility/Pages/default.aspx (last visited 8 Feb. 2020); U.S. DEP'T OF THE ARMY, ARMY APPLICATIONS LABORATORY, https://aal.army/innovationdays/ (last visited Feb. 8, 2020).

practices and lessons learned with the military to improve its processes to secure more industry collaboration.[182]

Pitch Days, Innovation Days, Industry and Reverse Industry Days, flexible acquisition strategies discussed in Section IV.B.3, and ease of access to DoD's storefront-type locations help nudge forward industry-DoD cooperation. But the intractable problem remains; fostering trust within industry that DoD's participation during design does not equate to giving away the crown jewels. For many companies, guarding the inner workings of their processes and technology is the same as guarding the viability of the company itself. Allowing an unknown government employee to observe, poke, prod, and question is simply unthinkable. Overcoming that intransigence means taking consistent, measured steps to incentivize access.

This can and should be accomplished from many angles. Among them, tying design process access to money by making it a condition of contract or other transaction award, with an emphasis on those agreements that entail researching, designing, and developing autonomous capabilities that could later be used in a LAW.[183] As seen in DoD technology challenges, like the Defense Advanced Research Projects Agency's (DARPA) robotics challenge, commercial start-ups placed a premium in "establishing themselves as the market standard" far and above their own investments in their technology. Commercial firms are willing to trade technology, or access to it, in exchange for notoriety and DoD adoption.[184] Another is to start with small successes, sending judge advocates to participate in isolated lower-threat projects. Judge advocates already

---

[182] U.S. Air Force District of Washington Contracting Officers, Reverse Industry Day, at slide 2 (Apr. 8, 2019), https://govtribe.com/opportunity/federal-contract-opportunity/reverse-industry-day-fa701419reverseindustryday#; U.S. Army Contracting Command Orlando, Reverse Industry Day Engagement, at slides 4-5 (Nov. 14, 2017), https://acc.army.mil/contractingcenters/acc-orl/palt_files/industryref/ReverseIndustryDayEngagementRIDENov2017.pdf.

[183] Contracting professionals must abide by rules designed to avoid conflicts and unfair competitive advantage. FAR, *supra* note 134, 15.201, 15.306. Additionally, the Defense Trade Secrets Act of 2016, Pub. L. 114-153, 130 Stat. 376 (2016) and Procurement Integrity Act, 41 U.S.C. § 2101 et seq. (2011); FAR, *supra* note 134, 3.104, limit disclosure of protected industry information. Given their current responsibilities and training, judge advocates are especially well-matched to the task of protecting the proprietary information of the companies they engage and ensuing their interactions are "fair, even, and transparent." Engaging with Industry*, supra* note 175, at 2; *see also* Memorandum from The Judge Advocate General to Judge Advocate Legal Services Personnel, subj.: Guidance for Strategic Legal Engagements (Sept. 8, 2016).

[184] ELLMAN, *supra* note 116, at 14.

support industry outreach efforts, as discussed in Section V.A. Literally, they are physically present when private sector innovators hawk their creations hoping for a deal with DoD.[185] Leveraging that presence with training, a strong support network, and a clear objective (access to the design process) advances DoD's interests for early involvement in the design of learners whose future calling may be within a LAW.

Most importantly, DoD needs a clear and consistent policy, announced to all potential industry partners, that its objective in pursuing machine learning autonomy is to actually be able to use it, which means minimizing the risk that vulnerabilities, indiscernible during testing, are smuggled inside the black boxes we buy. And to achieve that, the policy should encourage industry to invite judge advocate/commander teams as collaborators and facilitators as early as possible to identify and prevent possible LOAC issues before they arise. Whenever feasible, when DoD contemplates acquiring machine learning technology, the request for proposals should include a requirement that DoD gets the intellectual property (IP) and data necessary for weapons reviews.[186] The potential contractor and DoD could negotiate a special license for the pertinent data required for the sole and express purpose of conducting weapons reviews, accounting for the need to recertify the license as the learner modifies itself over time.

These efforts could avoid costly delays in later acquisition stages, provide the private developers a means to keep their valuable IP and data rights yet allow DoD the access it needs to help engender trust and reliability for the end user, and prevent mishaps and other operational challenges during operation.[187]

VI. Conclusion

The complexity of how LAWs' enabling technology learns, combined with its industry origins and unpredictable uses, and the rapid, risk-

---

[185] Interview with Melissa Fowler, Major, Acquisitions Program Attorney, U.S. Air Force (Dec. 4, 2019) (Discussing SBIR contracts, and participation as a judge advocate in Pitch Day, Nov. 5, 2019, San Francisco, CA).

[186] Interview with Janet Eberle, Lieutenant Colonel, Special Counsel, SAF/GCQ, U.S. Air Force (Nov. 12, 2019).

[187] *See, e.g.* DSB, ROLE OF AUTONOMY, *supra* note 14, sect. 1.4.3 (Discussing challenges encountered by unmanned systems operators resulting from rapid deployment of prototype and developmental capabilities and the pressures of conflict.).

absorbing acquisition pathways employed to obtain it require adjusting the current process for identifying and addressing potential LOAC issues in weapon systems. Though weapons reviews serve an important and necessary function, and rigorous testing will ferret out many of the problems, they should not be the only safeguards against the unique LOAC issues posed by autonomy in weapon systems. Relying solely on weapons reviews and ad hoc requests for legal support fails to consider how autonomy transforms battlefield LOAC concerns into laboratory LOAC concerns, and ignores the limitations of arms-length legal reviews. Because no legal barriers exist to judge advocates' enhanced participation in the design process, the DoD should take immediate action to incentivize the use of judge advocate/commander teams by commercial developers working on machine learning capabilities, and DoD organizations should be required to request it. Project managers, cross-functional team members, DoD employees engaging with industry, and anyone participating in projects to design machine learning models for DoD applications should be empowered to identify those human touchpoints when a judge advocate should be present. Lethal autonomous weapons will be commanders' tools, intended to assist them achieve mission success, and judge advocates trusted legal advisors. As the military prepares for LAWs to assume their inevitable place in formation, changing the fundamental nature of war, [188] leveraging the judge advocate's historical role as combat advisors is the right place to start.[189]

---

[188] Aaron Mehta, *AI Makes Mattis Question "Fundamental" Beliefs about War*, C4ISRNET (Feb. 17, 2018), https://www.c4isrnet.com/intel-geoint/2018/02/17/ai-makes-mattis-question-fundamental-beliefs-about-war/.

[189] FREDERIC L. BORCH, JUDGE ADVOCATES IN COMBAT: ARMY LAWYERS IN MILITARY OPERATIONS FROM VIETNAM TO HAITI (2001).