REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
29-10-2021	FINAL	

4. TITLE AND SUBTITLE

Ethics – the Key to Operationalizing AI-Enabled Autonomous Weapons

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Lieutenant Colonel Ross M. Coffey, U.S. Army

Paper Advisor: Dr. Tim Schultz

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Naval War College
686 Cushing Road
Newport, RI 02841-1207

8. PERFORMING ORGANIZATION REPORT NUMBER

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

10. SPONSOR/MONITOR'S ACRONYM(S)

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

12. DISTRIBUTION / AVAILABILITY STATEMENT

Distribution Statement A: Approved for public release; Distribution is unlimited.

13. SUPPLEMENTARY NOTES A paper submitted to the faculty of the NWC in partial satisfaction of the requirements of the EEMT Graduate Certificate Program. The contents of this paper reflect my own personal views and are not necessarily endorsed by the NWC or the Department of the Navy.

14. ABSTRACT

The so-called killer robots have arrived, and artificial intelligence-enabled autonomous weapons stand to be a prominent feature of future war. Against a backdrop of international competitor development of these systems overlaid against international and multinational corporate concern, the National Security Commission on AI's Final Report judges that these types of unmanned weapons can and should be used in ways consistent with international humanitarian law by applying the conditions of human-authorized use and proper design and testing. AI-enabled autonomy and its military applications carry with it the foundational risks in these technologies, and their use in unmanned weapons further challenges militaries seeking legal use within the frameworks of international humanitarian law and Just War Theory. Ethics therefore provides the superior conceptual vehicle to appoint and empower human authorizers and users and to qualitatively establish what constitutes "proper" design and testing. Each of the seven AI worker archetypes established by the DoD's Campaign for an AI Ready Force should apply role-relevant, AI-related ethics to fully realize the conditions established in the Final Report and retain and support the humanity necessary to control the monopoly on violence. The need for ethics education individually and collectively permeates each of the archetypes, and the DoD must recognize the value of public/private partnerships to fully account for these conditions.

15. SUBJECT TERMS

Ethics, Ethics Education, Artificial Intelligence, AI, AI-enabled, Autonomous Weapons, Lethal Autonomous Weapons, Fully Autonomous Weapons

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Dr. Tom Creely EEMT Dir.
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED		58	**19b. TELEPHONE NUMBER** *(include area code)* 401-841-7542

Standard Form 298 (Rev. 8-98)

U.S. NAVAL WAR COLLEGE
Newport, RI

Ethics – the Key to Operationalizing AI-Enabled Autonomous Weapons

By

Ross M. Coffey
Lieutenant Colonel, U.S. Army

Date Submitted: October 29, 2021

A paper submitted to the Faculty of the United States Naval War College Newport, RI, in partial satisfaction of the requirements of the Ethics and Emerging Military Technology Graduate Certificate Program.

Table of Contents

Abstract

The so-called killer robots have arrived, and artificial intelligence-enabled autonomous weapons stand to be a prominent feature of future war. Against a backdrop of international competitor development of these systems overlaid against international and multinational corporate concern, the National Security Commission on AI's Final Report judges that these types of unmanned weapons can and should be used in ways consistent with international humanitarian law by applying the conditions of human-authorized use and proper design and testing. AI-enabled autonomy and its military applications carry with it the foundational risks in these technologies, and their use in unmanned weapons further challenges militaries seeking legal use within the frameworks of international humanitarian law and Just War Theory. Ethics therefore provides the superior conceptual vehicle to appoint and empower human authorizers and users and to qualitatively establish what constitutes "proper" design and testing. Each of the seven AI worker archetypes established by the DoD's Campaign for an AI Ready Force should apply role-relevant, AI-related ethics to fully realize the conditions established in the Final Report and retain and support the humanity necessary to control the monopoly on violence. The need for ethics education individually and collectively permeates each of the archetypes, and the DoD must recognize the value of public/private partnerships to fully account for these conditions.

Introduction – Filling a gap in the Final Report

Artificial intelligence (AI)-enabled autonomous weapons systems are an emergent feature of the modern battlefield. These systems have already been used in recent conflicts in Libya and Nagorno-Karbach, and American allies and adversaries are racing to develop and harness these systems.[1] While the international community is divided on how to approach these developments, and while multinational corporations have levied their own concerns, there is clear evidence the United States will proceed in a direction supporting their maturation and operational employment.

The National Security Commission on AI's 2021 Final Report (referred to hereafter as the Final Report) released to the President and the Congress on March 1, 2021, judges that "provided their use is authorized by a human commander or operator, properly designed and tested AI-enabled and autonomous weapon systems have been and can continue to be used in ways which are consistent with [international humanitarian law]."[2] The report states humans must be involved in decisions regarding life and death in armed conflict but argues human accountability does not always necessitate direct human control throughout the entirety of the engagement process.[3] The Final Report's conditions, though, bear examination. Despite an exhaustive examination of other general AI matters, the 747-page Final Report is comparatively silent on topics of "authorization" and "proper design and testing," neither describing what those are in level of detail matching its other topics nor providing recommendations to explicitly operationalize AI-enabled autonomous weapons. Among the Final Report's eight chapters

[1] Specific examples and citations are provided in Chapter 1.

[2] National Security Commission on Artificial Intelligence, *Final Report* (Arlington, VA: National Security Commission on Artificial Intelligence, 2021), 92, accessed September 13, 2021, https://www.nscai.gov/wp-content/uploads/2021/03/Full-Report-Digital-1.pdf.

[3] *Final Report*, 93.

addressing emerging threats in the AI era, the subject of AI-enabled autonomous weapons is the only one without a corresponding blueprint for action.[4] This professional paper fills that gap.

This paper argues that ethics provides the means to address the Final Report's conditioning the use of AI-enabled autonomous weapons on human-authorized control and proper design and testing, reflecting a common definition of "proper" as "correct according to social or moral rules."[5] Ethics is a forward-looking field that addresses society and morality and answers questions surrounding "should" and for which only qualitative responses are possible. Adopting these conditions as its guiding keywords, this paper establishes a working-level understanding of "proper" design and testing of these types of unmanned weapons and creates a better appreciation of "human authorized use." After all, the ethics envisioned in the nation's founding documents and valued by American society will not be realized if these terms are not fully settled.

The paper opens with a survey of AI-enabled autonomous weapons likely employed in the modern battlefield to establish the current and future security environment and offer challenges for military professionals evaluating their use. After an illumination of AI's foundational risks and their impacts on societies, the paper describes international and multinational corporate concerns with these types of systems to provide a backdrop for the development of the ethical considerations necessary to address the Final Report's judgment and its inherent conditions. The mapping of the necessary legal aspects follows and points towards

[4] The Final Report addresses general testing and evaluation, verification, and validation (TEVV) considerations in its seventh chapter and provides a corresponding blueprint for action. Although this paper recognizes that subject of AI-enabled autonomous weapons and their associated TEVV considerations are inferred from the general considerations, it delves farther in the subject by identifying those measures that specifically create public confidence and trust in AI-enabled autonomous weapons namely through the applications of ethics across the workforce. The paper also recognizes – and builds upon – the Final Report's considerations for the responsible development and fielding of AI in its Appendix C.

[5] "Proper," Merriam-Webster, accessed September 20, 2021, https://www.merriam-webster.com/dictionary/proper.

the exploration of ethical possibilities and a description of the relationship between AI and ethical decision making.

Building upon the security and international environment, risks, and legal aspects, the paper integrates these scene-setting elements by applying ethics to the development and use of AI-enabled autonomous weapons across the communities of humans either in technical or enabling roles. Adopting the seven worker archetypes provided in the NSCAI's AI Workforce Model for Federal Government, it provides the ethical considerations AI experts, AI designers, deployment specialists, end users, non-technical tactical leaders, non-technical strategic leaders, and those in support roles must each apply to fully understand and operationalize "human authorized use" and "proper design and testing."[6] Evaluating consistencies of several of each archetypes' ethical considerations, the paper offers topical, thematic analysis of those considerations. The value of ethics education and public/private partnerships stand as predominate conclusions when attempting to operationally define "human authorized use" and "proper design and testing" of AI-enabled autonomous weapons. To arrive at these conclusions, this paper opens by examining these weapons' concerns and risks.

[6] These seven worker archetypes are further described at Defense Innovation Board, "Campaign for an AI Ready Force," 8, accessed September 5, 2021, https://media.defense.gov/2019/Oct/31/2002204191/-1/-1/0/CAMPAIGN_FOR_AN_AI_READY_FORCE.PDF.

Chapter 1 – AI-enabled autonomous weapons' concerns and risks

AI-enabled autonomous weapons will be used in future war. The presence of AI in those weapons, though, carries with it the societal risks inherent in AI. This chapter provides the backdrop for understanding the technical and international situation by first surveying current and future developments. It evaluates AI's foundational risks and presents international and private concerns associated with these types of weapons. The survey of these developments follows.

AI-enabled, autonomously operated unmanned weapons systems will be a foundational feature of future war.

Unmanned military platforms are everywhere. Unmanned platforms either for non-lethal purposes like surveillance or lethal purposes such as precision attack are being developed and manufactured at reduced costs, and they are no longer the purview of a few, large militaries. Instead, they are being routinely used by militaries around the world and even by non-state actors. These represent a feature of modern warfare that are unescapable next to manned platforms.

Small drones represent both a maturation of unmanned warfare and an evolution towards greater incidence in future war. As opposed to the larger unmanned systems operated by the U.S. Air Force (ex, MQ-1 Predator, MQ-9 Reaper, etc.), smaller drones for employment in tactical roles are likely features of future war. Providing the advantages of reduced cost and increased number, small unmanned systems provide increased options to militaries and political decision makers. Researcher and consultant Michael Blades, reflecting on recent success of Azerbaijan in its seizure of the Nagorno-Karabakh region and its aerial destruction of Armenian soldiers and ground combat vehicles, offers, "You don't need a whole lot of technology, with the small drones and airborne IEDs, you don't need a whole lot of technology to do a lot of

damage."[7] Reports indicate the Taliban have developed and successfully used an improvised unmanned aerial strike capability, clear evidence that even those actors viewed largely as lacking sophisticated military capability are still able to harness the power of unmanned systems.[8]

Beyond the evolution of unmanned systems towards micro-sized platforms, autonomy and AI are also increasingly associated with unmanned platforms. Autonomy, the ability of a machine to complete a task without direct human intervention, and AI, the capability of a machine to make sense of its surroundings and choose actions that maximize the chances of achieving favorable outcomes, seemingly offer benefits to unmanned platforms in that they expand the capacity of unmanned platforms to operate at faster processing speeds. Just as the human performing military tasks has been physically distanced from the task by use of an unmanned platform, autonomy and AI enable the unmanned platform to achieve a cognitive distance from the human operator in that the human need not make every decision for the machine's operation. These two applications create in unmanned platforms the opportunity to operate in manners beyond their original scope of extending the human's reach across the battlespace and reducing physical risk to the human operator.

The subject of lethality adds another layer onto the trajectory of unmanned military technologies applying AI and autonomy. If AI and autonomy can enable the more-rapid operation of an unmanned platform, could these applications enable unmanned platforms to perform lethal tasks? Recent news reporting indicates this has happened. In June 2021, the UN Panel of Experts reported that AI-enabled autonomous weapon systems were employed in Libya where Turkish drones were used by Government of National Accord forces in lethal attacks

[7] Ben Wolfgang, "Drones Have Outsized Impact on Future of War," *Washington Times,* January 12, 2021, https://www.washingtontimes.com/news/2021/jan/12/drones-have-outsized-impact-future-war/.

[8] Fazelminallah Qazizai, "The Drone Unit that Helped the Taliban Win the War," *Newslines Magazine,* September 15, 2021, https://newlinesmag.com/reportage/the-drone-unit-that-helped-the-taliban-win-the-war/.

against Libyan National Army logistics convoys and retreating forces in March 2020. Although the UN did not specify if any of Field Marshall Khalifa Haftar's forces were actually killed, the report identified, "'The lethal autonomous weapons systems were programmed to attack targets without requiring data connectivity between the operator and the munition: in effect, a true 'fire, forget and find' capability…'"[9] These Kargu-2 kamikaze drones, capable of striking ground targets, are equipped with machine learning inherent in their processing systems that allow 20 drones to work together using swarming techniques. Later, a *New York Times* report identified that Iranian nuclear scientist Mohsen Fakhrizadeh perished at the hands of Mossad using "a killer robot machine gun kitted out with artificial intelligence and multiple cameras and capable of firing 600 bullets a minute…"[10] AI technology capably compensated for the time lapse between detection of his car and operated with the precision to kill Fakhrizadeh but leave his wife riding as a passenger unscathed. Although not fully autonomous, this unmanned system delivered lethal effects due to AI-enabling. These might be the first arrival of so-called killer robots on the battlefield.

Turkey and Iran are not the only countries experimenting with AI-enabled autonomous weapons systems: global competitor China has sought cruise missile technology that will incorporate AI and automation.[11] Beyond human operation, senior designers have been eyeing fire-and-forget modes that will permit cruise missiles to ostensibly make terminal guidance decisions formerly the exclusive purview of the human operator. Russia is similarly

[9] Metin Gurcan, "Turkish Drone Sets off International Buzz over 'Killer Robots'," *Al-Monitor,* June 8, 2021, https://www.al-monitor.com/originals/2021/06/turkish-drone-sets-international-buzz-over-killer-robots.

[10] Rachel Sharp, "Revealed: How Iran's Top Nuclear Scientist Was Assassinated by a Killer AI Machine Gun That Allowed Sniper Based 1,000 Miles Away to Fire 15 Bullets After Disguised Spy Car Had Pinpointed His Location," *DailyMail.com,* September 20, 2021, https://www.dailymail.co.uk/news/article-10004269/How-Irans-nuclear-scientist-assassinated-killer-robot-machine-gun-kitted-AI.html.

[11] Reuters Staff, "China Eyes AI for New Cruise Missiles," *Reuters,* August 19, 2016, https://www.reuters.com/article/us-china-defence-missiles-idUSKCN10U0EM.

experimenting with swarming technologies: its weapons groups are exploring applying AI to unmanned platforms expressly designed to operate in swarming mode.[12] Other Russian weapons developers are working on guided missiles that will use AI to "analyze the aerial and radio-radar situation and determine its direction, altitude and speed" at operational ranges of 7,000 kilometers, effectively bringing Seattle, Los Angeles, Denver, Chicago, Philadelphia, New York, and Boston within range of Siberian-launched AI-enabled autonomous weapons.[13] Iran is planning to obtain capabilities to field and deploy fully autonomous suicide drones and other AI-enabled weapons such as missiles and robots by 2024.[14]

The United States and its allies and partners are also experimenting with the military applications of AI and autonomy in weapon systems. There is the recognition among the Department of Defense's futures community that humans may not be able to make decisions as fast as needed on the future battlefield, decision space that is further reduced in the event swarming technologies become reality. Recognizing AI should be safely controlled by humans as per the Final Report's conditions, operating in future war may rely on decisions being made by AI. General John Murray, Commanding General of the U.S. Army's Futures Command, posed, ""When you are defending against a drone swarm, a human may be required to make that first decision, but I am just not sure any human can keep up."[15] The United Kingdom is moving

[12] Tom O'Conner, "Russia's Military Challenges U.S. and China by Building a Missile That Makes Its Own Decisions," *Newsweek,* July 20, 2017, https://www.newsweek.com/russia-military-challenge-us-china-missile-own-decisions-639926.

[13] O'Connor, "Russia's Military."

[14] David Freedman, "US Is Only Nation with Ethical Standards for AI Weapons. Should We Be Afraid?" *Newsweek,* September 15, 2021, https://www.newsweek.com/2021/09/24/us-only-nation-ethical-standards-ai-weapons-should-we-afraid-1628986.html.

[15] Matthew Cox, "To Defeat Enemy Drone Swarms, Troops May Have to Take a Back Seat to Machines, General Says," *Military.com,* January 25, 2021, https://www.military.com/daily-news/2021/01/25/defeat-enemy-drone-swarms-troops-may-have-take-back-seat-machines-general-says.html#:~:text=To%20Defeat%20Enemy%20Drone%20Swarms%2C%20Troops%20May%20Have,Multinational%20Simulation%20Center%2C%20Grafenwoehr%2C%20Germany%2C%20in%20March%2C%202018.

to develop AI-enabled autonomous airplanes that will be able to target and shoot enemy aircraft.[16] Dubbed the "loyal wingman," the United Kingdom's Spirit AeroSystems is building test vehicles for envisioned pairing with British Typhoon and American F-35 Joint Strike Fighters by the conclusion of the 2020s.[17] Test flights might be possible as early as 2023.[18] Beyond the aforementioned Mossad unmanned sniper capability, Israel's Harpy drone is "programmed to fly to a particular area, hunt for specific targets, and then destroy them using a high-explosive warhead nicknamed 'Fire and Forget.'"[19] Not only are unmanned systems ubiquitous, the future possibility of AI-enabled autonomous weapons is a real one.

Where will military professionals be challenged to use AI-enabled autonomous weapons?

The Final Report sees AI as a comparative advantage in any field to those actors and agencies who can harness its power.[20] Recognizing that "no comfortable historical reference captures the impact of AI on national security," the report illuminates its inherent nature as an encompassing aspect of future technology, drawing itself upon Thomas Edison's determination of electricity's power and his expectation it would drastically alter life and society.[21] Indeed, AI will affect future societies' computing, communications, robotics, and electronics.

AI will also affect future war. The Final Report states, "In this new kind of warfare, traditional confines of the battlefield will be expanded through AI-enabled micro-targeting, disinformation, and cyber operations."[22] It recognizes warfare's attributes like its scale, tempo,

---

[16] Andrew Chuter, "British Shell Out Seed Funding for 'Loyal Wingman' Combat Drone," *Defense News,* January 25, 2021, https://www.defensenews.com/global/europe/2021/01/25/british-shell-out-seed-funding-for-loyal-wingman-combat-drone/.

[17] Chuter, "British Shell Out."

[18] Chuter.

[19] Kai-Fu Lee, "The Third Revolution in Warfare," *The Atlantic,* September 11, 2021, https://www.theatlantic.com/technology/archive/2021/09/i-weapons-are-third-revolution-warfare/620013/.

[20] *Final Report*, 2.

[21] *Final Report,* 7.

[22] *Final Report*, 79.

and velocity will be reshaped, as well as "the relationships service members have with machines; the persistence with which the battlefield can be monitored; and the discrimination and precision with which targets can be attacked."[23] Future war will not be human versus the machine; rather, and compared to historical war and human versus human, future war will likely feature humans' machines versus humans' machines and human-machine teams versus human-machine teams. Thus, those more capable of harnessing the power of AI will gain tactical advantage over their adversaries. Decisions might also need to be taken at machine speed.

For its possible benefits, though, AI has inherent drawbacks. AI can magnify cyber intrusions or disinformation campaigns. It can strengthen authoritarian regimes and fuel extremism. The potential of digital manipulation is concurrently increased as AI is furthered, and due to its ubiquitous spread, the power of AI is not only the purview of legitimate authorities, but it is also exploitable by non-state actors. These vulnerabilities must be addressed if tactical advantage is to be translated into strategic success, and ethics provides a vehicle to do so.

Unmanned weapons, while novel from a technological sense, in some degree merely represent an evolution of technology. Melvin Kranzberg analyzed the complex interrelationship technology has, and has made upon, human history, postulating several laws governing this interrelationship. Two are particularly instructive when considering these types of weapons and reflect both the seemingly positive aspects and inherent vulnerabilities just explored.

Kranzberg's First Law provides technology is neither good nor bad; nor is it neutral. Relating this to the AI embedded in military systems, it purports that AI, like any technology,

[23] *Final Report.*

offers trade-offs between "various 'good's' and possible 'bad's.'"[24] Enabling lethal weapon systems with AI-derived autonomy further reflects Kranzberg's First Law, posing several challenges for the military ethicist. Autonomous systems should know when enough is enough, or in a technical sense, when sufficient effect has been achieved against its directed target. The extent AI will perform battle damage assessment, particularly in terms of assessing if the military and political objectives were achieved, is not yet established. Recognition when objectives have been achieved and lethal force, while technically authorized under the rules of engagement, is no longer required must similarly be defined. Adaptation to the rules of engagement is a necessary element of these autonomous systems, although how fast that can occur must be determined. AI-enabled autonomy might increase military effectiveness in the short term, although its long-term effects are not yet fully known.

Kranzberg's Third Law is also germane to the subject of AI-enabled autonomous weapon systems: Technology comes in packages, big and small. This law identifies that in any one macro technology, there are multiple micro technologies each with their own set of design and ethical considerations. Kranzberg provides, "final [products are] composed of many separate elements brought together in a system that could not function without every single one of the components."[25] This law is particularly appropriate to autonomous AI-enabled weapon systems. Not only are the weapon systems technologies in and of themselves, their internal guidance and decision-making mechanisms are also their own technologies.

Applying this Third Law to these systems raises challenges with the transferability of the technology. Consider scenarios where the US found it operating in new and unfamiliar

[24] Melvin Kranzberg, "Technology and History: 'Kranzberg's Laws,'" *Technology and Culture* 27, no 3 (July 1986): 545.

[25] Kranzberg, "Technology and History," 549.

environments. Autonomous systems engineered for lethal effect against peer competitors in the South China Sea or in the Baltic states might also need to effectively and ethically operate if American interests were threatened by asymmetric adversaries in Central America or Sub-Saharan Africa.[26] Kranzberg's Third Law holds that any one of the micro technologies might limit the effectiveness of the system, challenging views of ethical use in an environment where suboptimal performance can and should be expected.

Adaptability of the technology is another challenge when applying Kranzberg's Third Law. The value of manned systems is the ability to recognize when the environment and operating area are novel and to determine if - and how - their capabilities are appropriate. Manned systems, or those unmanned systems under the direct control of a human operator, can, with proper education and training of its human operators, make ethical determinations in a timely manner. While AI can and does aid in decision-making particularly when faced with large data sets, it's not clear AI can operate in the same manner when presented with novel environments, challenging AI-enabled autonomous machines expressly designed to create lethal effects.

Further, given the complexity inherent in autonomous AI-enabled weapons, any one of the micro technologies might pose an adaptability challenge threatening the entire system. For example, how would autonomous systems withstand changes in environment and operating area? How might adaptability of AI-enabled systems be exploited by adversaries? Given the US' global commitment, could AI-enabled autonomous systems demonstrate the same adaptability that manned systems provide? And at what cost?[27] Mapping the implications of Kranzberg's

[26] If this results in an inherent operational limitation, then to what degree should the Department of Defense invest in AI-enabled autonomous systems? This represents a reverse salient.

[27] This represents another reverse salient.

Third Law stating "technology comes in packages, big and small" to AI in general and AI-enabled lethal weapon systems yields both challenges with transferability and adaptability of the systems. The questions posed by Kranzberg's First and Third Law provide the backdrop of the AI enabling future autonomous weapons systems and its possible advantages and disadvantages in future war.

Foundational risks inherent in AI further complicate the development and use of AI-enabled autonomous weapons.

Beyond the military challenges and advantages and disadvantages described above, foundational societal risks in AI further complicate the subject. Reliance on AI in matters of public safety and governance bring about other risks, risks that further erode rights of private citizens and their trust in government. The risks of racial bias, misapplication, and spoofing present ethical dilemmas when attempting to use data to address societal problems.

Bias towards race represents a tangible risk in AI. The Government Accountability Office's June 2021 accountability framework recognized that while bias is not specific to AI, it has the possibility to worsen those biases that already exist. It provides, "Biases arise from the fact that AI systems are created using data that may reflect preexisting biases or social inequities."[28] One example involves predictive health care models that excluded patients' races yet consistently assigned lower risk scores to Black patients, resulting in an under-identification of that demographic's health needs despite having similar conditions.[29]

The same exists in policing systems using AI to predict future crime risk: "the systemic biases can be perpetuated and amplified as police departments use biased predictions to make

[28] Government Accountability Office, *Artificial Intelligence: An Accountability Framework for Federal Agencies and Other Entities*, GAO-21-519SP (Washington, DC: GAO, June 2021), 13, https://www.gao.gov/assets/gao-21-519sp.pdf.

[29] Government Accountability Office, *Artificial Intelligence,* 23.

tactical policing decisions."[30] Police departments are adopting, and in some cases aggressively adopting, AI-enabled computer algorithms to make risk assessments towards committing future crime and placing at risk public confidence in law enforcement and, more broadly, America's democratic institutions. Provided that computers could predict those defendants with a greater likelihood of committing future, more violent crime, this could result in a more equitable justice system. But the computers often get it wrong. Consider the case chronicled by Julia Angwin and her co-authors analyzing software used by the Fort Lauderdale police department to predict future criminals, comparing the cases of a black woman and a white man both apprehended for $80 petty thefts.[31] Despite the white man having served five years in prison on felony charges, and despite the black woman having only misdemeanors on her record, the computer program predicting future crime at the time of their bookings rated the black woman as a high risk and the white man as a low risk. Two years later, though, the white man was convicted on felony-level theft charges and is presently serving an eight-year term whereas the black woman has not been charged with any new crimes. Similar cases of wildly inaccurate predictions involving blacks and whites have been documented in Kentucky and Arizona.[32]

Policing systems are not the only situations where racial bias has emerged: Social Media tracking enabled by AI offers other avenues to inject bias into decisions that were previously the exclusive (and regrettable) purview of humans. Google's search algorithms gained unwanted notoriety in 2015 when its Photos app identified blacks as gorillas.[33] Social media fared no better six years later when Facebook's AI labeled black men as primates. Tony Tran reported,

---

[30] Government Accountability Office, *Artificial Intelligence,* 25.

[31] Julia Angwin et al., "Machine Bias: There's Software Used Across the Country to Predict Future Criminals and It's Biased Against Blacks," *ProPublica,* May 23, 2016, https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

[32] Angwin, "Machine Bias."

[33] Tony Tran, "Facebook Apologizes for Racist AI that Labeled Black Men as 'Primates.'" *The Byte,* September 5, 2021, https://futurism.com/the-byte/facebook-racist-ai.

"Users watching the video saw an automated prompt that asked if they would like to 'Keep seeing videos about primates?"[34] If AI, then, is to be used to make determinations on peoples' intentions, it must overcome these documented cases of racial bias. Considering these Social Media episodes, AI-enabled systems that attempt to classify or categorize persons are possibly biased towards race, reducing their value as they are unlikely to make sufficiently accurate determinations. The Kargu-2 kamikaze drone described earlier has the capability of targeting individual humans, so the general concerns of AI's bias towards race is no small matter.

Misapplication is a second societal risk with ethical applications occurring when the machine performs in a manner not originally intended. Natalie Wolchover documents the extreme case of YouTube's algorithms gone awry. "Two years ago, computer scientists and users began noticing that YouTube's algorithm seemed to achieve its goal by recommending increasingly extreme and conspiratorial content. One researcher reported that after she viewed footage of Donald Trump campaign rallies, YouTube next offered her videos featuring 'white supremacist rants, Holocaust denials and other disturbing content.'"[35] Citing the lines of code that don't fully understand the situation and provide answers that don't align with humans original preferences, Wolchover concedes YouTube's coders didn't set out to radicalize viewers although also cites they can't think of everything.[36] Closely related to misapplication is the spread of misinformation, as New York University researchers "found that from August 2020 to January 2021, misinformation got six times more clicks on Facebook than posts containing factual news."[37] Thus, the possibility of AI to provide an unintentional answer (such as the

[34] Tran, "Facebook Apologizes."

[35] Natalie Wolchover, "AI Will Do What We Ask. That's a Problem," *Quantamagizine,* January 30, 2020, https://www.quantamagazine.org/artificial-intelligence-will-do-what-we-ask-thats-a-problem-20200130/.

[36] Wolchover, "AI Will Do."

[37] Ramishah Maruf, "Researchers Studying Facebook Misinformation Say They Were De-platformed," *CNN Business,* September 5, 2021, https://www.cnn.com/2021/09/05/media/reliable-sources-facebook-researchers-deplatform/index.html.

interrogated being questioned due to factors beyond the original scope of the machine) represents a second ethical risk – and one that will increase. Will militaries seeking to use AI-enabled autonomy be able to assure its coding returns only intentional – and discards unintentional – answers to tactical problems?

A third risk is spoofing, or the fooling of systems by humans for their own ends. It is prohibitively expensive – if not impossible – for machine training phases to cover all possible scenarios or examples AI will be exposed to once fielded.[38] Because systems can be fooled into seeing things humans wouldn't see, nefarious human actors can therefore manipulate AI to derive inaccurate conclusions or make the wrong determination.[39] Spoofing of voices has already affected cybercrime with the fraudulent transfer of $243,000 in 2019 due to an AI-enabled deepfake voice simulating a German CEO convincing a CEO of a UK-based subsidiary to send funds to an apparent Hungarian supplier.[40] [41] The possibility that criminals could create fake digital identities or somehow masquerade for legitimate actors lessens the reliability of data-driven systems for public safety as used by the interrogator. Considering the DoD's emphasis on AI's reliability, spoofing represents a tangible risk towards gaining commanders', users', and strategic leaders' trust in AI.[42]

Spoofing is of particular risk when applying AI in a competitive setting. Modern AI used in the corporate world and in private settings has been developed without concerns towards

[38] Julia Bossman, "Top 9 Ethical Issues in AI," *World Economic Forum,* October 21, 2016, https://www.weforum.org/agenda/2016/10/top-10-ethical-issues-in-artificial-intelligence/

[39] Bossman, "Top Ethical Issues."

[40] Catherine Stupp, "Fraudsters Used AI to Mimic CEO's Voice in Unusual Cybercrime Case," *Wall Street Journal,* August 30, 2019, accessed August 28, 2021, https://www.wsj.com/articles/fraudsters-use-ai-to-mimic-ceos-voice-in-unusual-cybercrime-case-11567157402.

[41] The funds ended up in the criminals' account.

[42] The May 26, 2021, Deputy Secretary of Defense memorandum "Implementing Responsible Artificial Intelligence in the Department of Defense" defines Reliable as "the Department's AI capabilities will have explicit, well-defined used, and *the safety, security, and effectiveness of such capabilities will be subject* to testing and *assurance within those defined uses* across AI capabilities' entire life cycle." [emphasis added]

deception. Considering, though, Sun Tzu's maxim "All war is deception," the use of AI-enabled autonomous weapons in future conflict should surely account for the possibility adversaries will deliberately deceive AI in order to gain military advantage. Sam Tangredi provides, "The DoD must admit that commercial AI largely has been designed not only without any awareness of the character of war, but also without deception as an element of its environment."[43] Spoofing might take the form of insurgents developing "adversarial" graphics to deceive AI systems into the true identities of combatants and non-combatants and then painting "non-enemy" graphics on their vehicles' roofs and "enemy" graphics on poorly-secured public buses. Timing their movements around those of the public buses, insurgents might move undetected while the AI-enabling strikes non-combatants perceived to be enemy forces.[44] Beyond deliberate deception, mere accidents might have similar consequences such as missile defense systems mistaking unusual horizon glare as incoming missile launches, triggering the firing of interceptors and resulting in retaliatory strikes.[45]

Beyond concerns of AI's transferability and adaptability and the general risks associated with AI, there are several technical concerns associated with these AI-enabled unmanned systems. First, accurate recognition of the target must be achieved for AI-enabled autonomous weapons to operate within the bounds of international humanitarian law. Second, their increased reliance on communications complicates their employment. Worse, the possibility of hacking or some other sort of cyber-related intrusion limits the degree to which these sorts of unmanned weapons will be able to operate effectively. AI often implements processes using cloud

[43] Sam Tangredi, "Sun Tzu Versus AI: Why Artificial Intelligence Can Fail in Great Power Competition," *Proceedings 147, no. 5/1,419* (May 2021): 22.

[44] Zachary Arnold and Helen Toner, *AI Accidents: An Emerging Threat What Could Happen and What to Do* (Washington, D.C.: Center for Security and Emerging Technology, July 2021), 10-11, https://cset.georgetown.edu/wp-content/uploads/CSET-AI-Accidents-An-Emerging-Threat.pdf.

[45] Arnold, *AI Accidents,* 7-8.

environments, thus holding at risk their full potential in these types of unmanned systems based on the expectation adversaries will seek to deny access to communications.[46] The quality of data available to AI enabling autonomous unmanned weapons further complicates the situation.[47] The poisoning of training and testing data or model poisoning represent further technical risk. In summation, risks, challenges, and technical problems compound and complicate the operationalization of AI-enabled autonomous weapons.

Concerns over these types of unmanned systems are increasing from a variety of international foreign policy and human security actors.

The larger community has taken notice of AI-enabled autonomy that enhances, or entirely subsumes, lethal weapons' effectiveness and has in many cases identified concerns beyond the comparatively simple technological challenges. Extrapolating views of democratic nations towards general AI to the subject of AI-enabled autonomous weapons, the Center for AI and Digital Policy (CAIDP) shared expected international concern for these variants of unmanned weapons in a March 25, 2021, letter to the U.S. House of Representatives. Recognizing the policy institute had not yet fully evaluated these types of unmanned weapons in its survey of national policies and practices associated with AI, the CAIDP called for their limitation: "Our recent review of country policies strongly indicates support among democratic nations for limits on these systems."[48] The U.S. Congress' Congressional Research Service might concur with the Center's assertion, as its April 19, 2021 report explicitly identifies twice

[46] Department of Defense, *Unmanned Systems Integrated Roadmap 2017-2042*, 18, https://www.defensedaily.com/wp-content/uploads/post_attachment/206477.pdf.

[47] *Unmanned Systems Roadmap*, 19.

[48] Marc Rotenberg and Tuan Nguyen, untitled letter to the Chairman and Ranking Members of the House Armed Services Subcommittees, March 25, 2021, https://www.caidp.org/.

the number of countries supporting a stance prohibiting AI-enabled autonomous weapons as those who support their prospective use.[49]

States are not the only ones to levy concerns with these technologies embedded into unmanned systems, as both inter-governmental and non-governmental organizations share their concerns. A prominent voice among inter-governmental organizations is the United Nations as evidenced by Secretary-General Antonio Guterres tweeting, "Autonomous machines with the power and discretion to select targets and take lives without human involvement are politically unacceptable, morally repugnant and should be prohibited by international law."[50] In a prepared statement subsequently delivered by the United Nations Office at Geneva to the March 25, 2019, convening of the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons, he further emphasized his concern on these weapons types, citing previous Group of Governmental Experts' principles that "'human responsibility for decisions on the use of weapons systems must be retained since accountability cannot be transferred to machines.'"[51] The challenges AI-enabled autonomy creates in lethal weapon systems principally centers around possible violations of international humanitarian law and inferior or inadequate human interaction with their use.[52]

Non-governmental organizations concur. Among these is the Human Rights Watch that has noted challenges AI-enabled autonomous weapons will pose when practically, or attempting

[49] Kelley M. Sayler, "International Discussions Concerning Lethal Autonomous Weapon Systems," Congressional Research Service, April 19, 2021, https://crsreports.congress.gov/product/pdf/IF/IF11294.

[50] Antonio Guterres, "Autonomous Machines," Twitter, March 25, 2019, 1:28 PM tweet, https://twitter.com/antonioguterres/status/1110232038081204224?s=20.

[51] Michael Moller, delivering a prepared statement of Antonio Guterres, "Secretary-General's Message to Meeting of the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems," United Nations, March 25, 2019, https://www.un.org/sg/en/content/sg/statement/2019-03-25/secretary-generals-message-meeting-of-the-group-of-governmental-experts-emerging-technologies-the-area-of-lethal-autonomous-weapons-systems.

[52] "Autonomous Weapons that Kill Must Be Banned, Insists UN chief," *United Nations*, March 25, 2019, https://news.un.org/en/story/2019/03/1035381.

to practically, fulfill the intent of international humanitarian law. Human Rights Watch argues that the principles of international humanitarian law require "the humane treatment of others and respect for human life and human dignity" in all circumstances to include lethal combat engagements.[53] This non-governmental organization argues humans inherently feel empathy and compassion thus motivating each other to treat each other humanely, incidents of which are as curiously timeless as warfare and institutionally sanctioned taking of life itself.

Human Rights Watch has further argued that AI-enabled autonomous weapons would base their actions on coding and machine language pre-programmed during their development phases thereby challenging them to adapt to unpredictable scenarios or complexity not envisioned during programming.[54] In cases where humane treatment of combatants – or former combatants – was warranted, this non-governmental organization believes these types of weapons would be unable to adapt in a manner consistent with international humanitarian law.[55] Worse, machine language inherent in AI-enabled autonomous weapons would view humans as objects and remove any dignity to be afforded to either combatants or non-combatants. It concludes: "Even if fully autonomous weapons could adequately protect human life, they would be incapable of respecting human dignity. Unlike humans, these robots would be unable to appreciate fully the value of a human life and the significance of its loss. They would make life-and-death decisions based on algorithms, reducing their human targets to objects."[56]

Neither a non-governmental organization nor an inter-governmental organization, the uniquely situated International Committee of the Red Cross further reinforces these perspectives.

---

[53] Bonnie Docherty, "Heed the Call: A Moral and Legal Imperative to Ban Killer Robots," *Human Rights Watch,* August 21, 2018, https://www.hrw.org/report/2018/08/21/heed-call/moral-and-legal-imperative-ban-killer-robots.

[54] Docherty, "Heed the Call."

[55] Examples include those seeking surrender or wounded to the point armed resistance renders a combatant physically, mentally, and emotionally unable to continue.

[56] Docherty, "Heed the Call."

The Red Cross cites the lack of human dignity as a complicating factor in these types of lethal weapons. Beyond questions of the possibility of persons being killed by the effects of AI-enabled autonomy, the Red Cross further challenges their use by addressing the manner and scope in which persons are killed and the processes by which machines may make these determinations.[57] The transfer of human agency over matters of life and death to machines troubles the Red Cross, as follows: "If human agency is lacking to the extent that machines have effectively, and functionally, been delegated these decisions, then it undermines the human dignity of those combatants targeted, and of civilians that are put at risk as a consequence of legitimate attacks on military targets."[58] It argues for meaningful human control throughout the use of these weapons, identifying that human supervision and the capacity to deactivate AI-enabled autonomous weapons or intervene when necessary must be retained. Oversight of technical development surrounding these unmanned weapons' coding to assure predictability and reliability is important, as are "operational constraints on the task for which the weapon is used, the type of target, the operating environment, the timeframe of operation and the scope of movement over an area."[59]

Multinational corporations are also taking notice of these developments. Many in the C-suite of the large tech firms are supportive of a general trend towards AI-enabled autonomous weapons: "Tech-industry leaders are pushing the Pentagon to adopt commercially developed technologies on a grand scale to counter the rise of China, an initiative that could transform the

[57] International Committee of the Red Cross, *Ethics and Autonomous Weapon Systems: An Ethical Basis for Human Control?* (Geneva, Switzerland: ICRC, April 3, 2018), https://www.icrc.org/en/document/ethics-and-autonomous-weapon-systems-ethical-basis-human-control.

[58] ICRC, *Ethics.*

[59] ICRC.

military and the multibillion-dollar defense-contracting business."[60] But their workforces don't necessarily share these sentiments and have in fact sought to influence the debate. An open letter to Google CEOs noted that tech companies are being implicated in lethal activities simply through their development of AI.[61] Opposition to highly-AI reliant and autonomously-envisioned Project Maven is stated even more plainly: "Building this technology to assist the US Government in military surveillance – and potentially lethal outcomes – is not acceptable."[62]

It is against this backdrop of increasingly probable international competitor development of unmanned systems enabled by AI and capable of autonomous operation and overlaid against international and multinational corporate concerns that the National Security Commission on AI presented its judgments in its 2021 final report. As provided previously, the Final Report judged that, provided *human-authorized* use and *proper* design and testing, these types of unmanned systems can be used in ways consistent with international humanitarian law (emphasis added).[63] It further judged existing DoD procedures are capable of ensuring the US will field safe and reliable systems as evidenced in the November 21, 2012, Department of Defense Directive governing autonomy in weapons systems or the May 26, 2021, Department of Defense

[60] Ryan Tracy, "As Google, Microsoft and Amazon Seek Bigger Defense Role, Some Are Leery," *Wall Street Journal,* September 7, 2021, https://www.wsj.com/amp/articles/tech-industry-seeks-bigger-role-in-defense-not-everyone-is-on-board-11631019600.

[61] Open Letter in Support of Google Employees and Tech Workers, "Researchers in Support of Google Employees: Google should withdraw from Project Maven and commit to not weaponizing its technology," International Committee on Robot Arms Control, accessed September 10, 2021, https://www.icrac.net/open-letter-in-support-of-google-employees-and-tech-workers/.

[62] "Letter to Google C.E.O." *New York Times,* May 30, 2018, https://static01.nyt.com/files/2018/technology/googleletter.pdf.

[63] *Final Report*, 92.

memorandum articulating the department's five principles governing responsible AI.[64] [65] [66] Considering little available evidence that US competitors have adopted procedures with similar regard for safety of non-combatants and rigor in design and testing, the commission implied the United States should continue to develop – and possibly use – AI-enabled autonomous unmanned weapons systems.[67] While not directly addressed in the Final Report, the key words "human-authorized use" and "proper" govern how the United States should conceptually approach these types of unmanned weapons. Clearly establishing "human-authorized use" and "proper" (design and testing) challenge the military ethicist in that the qualities of "proper" and the question of who is "authorized" deserve further exploration. The same dynamic is revealed when considering existing DoD procedures.

[64] *Final Report.*

[65] Department of Defense, *Autonomy in Weapons Systems,* DOD Directive 3000.09 (Washington, DC: Department of Defense, May 8, 2017), https://www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodd/300009p.pdf.

[66] The referenced memorandum is posted to https://media.defense.gov/2021/May/27/2002730593/-1/-1/0/IMPLEMENTING-RESPONSIBLE-ARTIFICIAL-INTELLIGENCE-IN-THE-DEPARTMENT-OF-DEFENSE.PDF.

[67] *Final Report,* 92.

Chapter 2 – The legality and ethics of human-authorized use and proper design and testing.

Having established the current and future security situation, this paper evaluates the guiding keywords of human authorized use and proper design and testing from perspectives of legality and ethics. Legal considerations promote necessary conditions to use them in future war and to guide design and testing. But in questions of authorization or "proper" design and testing, legality falls short; instead, ethics provides the superior conceptual vehicle. Due to their necessity, however, the legal considerations are developed first.

Attempting to effectively define "proper" (design and testing) and refine "human-authorized" use – a legality-based approach is necessary…but not sufficient.

Legality offers options to effectively define "proper" and understand human-authorized use and establishes some necessary elements of these types of weapons' employment. The Department of the Navy's Unmanned Campaign Framework establishes, "respect for the rule of law is not only "right" but also a competitive advantage—it is an enduring position that humans embrace and often demand."[68] Considering the possibility American adversaries will not adopt or adhere to similar legal frameworks or exhibit compatible ethics, the avoidance of unethical practices such as indiscriminate targeting may not advance this advantage; instead, respectful appreciation for the rule of law should be pursued.[69] Both international humanitarian law as previewed previously and Just War Theory bear mention.

Legal considerations surrounding international humanitarian law help shape a working definition of "proper design and testing" and further understand "human-authorized use." International humanitarian law provides two fundamental principles: "persons who are not, or

[68] Department of the Navy, *Unmanned Campaign Framework*, March 16, 2021, 31, https://www.navy.mil/Portals/1/Strategic/20210315%20Unmanned%20Campaign_Final_LowRes.pdf?ver=LtCZ-BPlWki6vCBTdgtDMA%3D%3D.

[69] *Unmanned Campaign Framework*, 31.

are no longer, participating in hostilities must be protected" and "the right of parties to an armed conflict to choose methods and means of warfare is not unlimited."[70] These principles are further refined in the elements of (a) distinction, or the requirement for combatants to distinguish between civilians and combatants; (b) proportionality, stated as the prohibition of attacks that result in excessive civilian life compared to the sought military objective; and (c) accountability, defined as the necessity for parties to conflict to be held accountable for their actions resulting from use of their weapons.[71]

The principle of distinction holds that any combatant must be able to understand what or who is a target and to determine its importance – the defense of "merely following orders" no longer holds when considering the necessity of a combatant to distinguish between legitimate military targets and those beyond the scope governed by international humanitarian law. Considering the application of AI-enabled autonomy in unmanned weapon systems, these systems must therefore be able to determine those targets and objectives of attack authorized under international humanitarian law. They must also possess the inherent capability to determine, without direct human intervention, those targets that would be considered as non-combatants by a human operator. This goes beyond civilians who are not party to the conflict and to extends to those combatants seeking to surrender or with injuries that prevent the mounting of further resistance.

Secondarily, proportionality must also be factored into the development of unmanned systems enabled with AI and operating autonomously. These systems must be able to ascertain collateral damage both in pre-planned situations and dynamic environments where the presence

[70] International Justice Resource Center, "International Humanitarian Law," accessed September 4, 2021, https://ijrcenter.org/international-humanitarian-law/.

[71] This paper adopts the principles used in the *Final Report.* See page 92.

of non-combatants was neither expected nor anticipated. Unmanned systems in future war must gain the wherewithal to determine when excessive force is possible and, more importantly, not authorized under the purview of international humanitarian law. Beyond the weapons themselves, AI-enabled autonomous weapons have the potential of increasing the use of violence instead of minimizing it or restoring the peace, meaning discussions of proportionality must extend beyond the simple use of these types of unmanned weapons.[72]

Last, AI-enabled autonomous weapons systems must adhere to the element of accountability to operate within the confines of international humanitarian law. Manned weapons systems used by human operators offer the assurance of moral agency, or the ability to determine appropriate outcomes and to be held accountable by higher authorities for their actions.[73] Moral agency and the capacity for moral judgment (that is, choosing right from wrong) holds that responsibility for action governs the element of accountability. Mapping this element to these types of unmanned weapons systems and considering that general AI may emerge within years, Wendell Wallach and Colin Allen argue it is necessary to start building functional morality in machines and to imbue basic ethical sensitivity for humans' own safety. [74] The programming of morality into AI, or the transference of moral agency to machines, and the policing of machines' behavior by other machines is both a necessity and a possibility.

Using unclassified versions of rules of engagement to establish bounds on ethical conduct, AI researcher Ron Arkin "concluded it is possible to program some kind of ethical

---

[72] Heather Roff, "Lethal Autonomous Weapons and Jus Ad Bellum Proportionality," *Case Western Reserve Journal of International Law,* 47 (2015): 50, https://scholarlycommons.law.case.edu/cgi/viewcontent.cgi?referer=&httpsredir=1&article=1006&context=jil.

[73] "Moral Agent," Ethics Unwrapped, University of Texas McCombs School of Business, accessed September 4, 2021, https://ethicsunwrapped.utexas.edu/glossary/moral-agent.

[74] Wendell Wallach and Colin Allen, *Moral Machines: Teaching Robots Right from Wrong* (Oxford: Oxford University Press, 2009), 4.

standard - a conscience if you will - for machines to follow in lethal actions."[75] AI entrepreneur Amir Husain comments, "If one robot begins acting badly, whether through faulty programming or enemy hacking, the rest of the network needs to be able to disable and if necessary destroy it…just as society takes action when an individual human violates generally accepted ethics."[76] Accepting that programming solutions might be achieved, the international humanitarian law's element of accountability demands machines exhibit a degree of programmed moral agency if artificial-intelligence enabled autonomous weapons are to be used in future war.

In addition to international humanitarian law, and in manners like unmanned weapons systems under the direct control of a human operator, Just War theory also bears on these AI-enabled autonomous weapons systems. In *ante bellum* situations, or prior to war, these types of unmanned systems must be engineered for retention of human control over their effects throughout their operational use and lifecycle.[77] In war, or *in bello*, Just War Theory holds that commanders must retain decision-making authority over objectives and recognizes that the dynamic nature of conflict since the dawn of warfare creates situations where objectives mutate and evolve as conflict progresses.[78] Using this perspective, unmanned systems enabled with AI and operating autonomously must remain responsive to human commanders in order to satisfy Just War Theory's *in bello* considerations and assure international humanitarian law's principles of discrimination and proportionality. Finally, following the termination of hostilities in a *post bellum* environment, the possible commitment of war crimes or other violent acts exceeding

[75] M. Shane Riza, *Killing Without Heart: Limits on Robotic Warfare in an Age of Persistent Conflict,* (Washington, DC: Potomac Books, May 1, 2013), 133.

[76] Sydney Freedberg, "Should Pentagon Let Robots Kill Humans? Maybe," *Breaking Defense,* July 10, 2017, https://breakingdefense.com/2017/07/should-pentagon-let-robots-kill-humans-maybe/.

[77] Heather Roff and Richard Moyes, "Meaningful Human Control, AI and Autonomous Weapons," (briefing paper prepared for the Informal Meeting of Experts on Lethal Autonomous Weapons Systems, UN Convention on Certain Conventional Weapons, April 2016), 3, accessed September 6, 2021, https://article36.org/wp-content/uploads/2016/04/MHC-AI-and-AWS-FINAL.pdf.

[78] Roff, "Meaningful Human Control," 4.

warfare's governing rules means that someone must remain responsible and accountable for those crimes.[79] The evolution towards greater autonomy in unmanned weapons and, more subtly, the application of AI that determines its own conclusions challenges *post bellum* notions of responsibility and accountability in that machines might inherently arrive at actions resulting in the commitment of war crimes not originally envisioned by the human commander authorizing their use, the developer designing the autonomous unmanned weapon system, or the coders of the AI.

Curiously, both legal frameworks make arguments for their development and use in future war. The elements of international humanitarian law speaking to distinction and the *ante bellum* consideration for human commanders to retain control of manned or unmanned weapons assigned to or operating in support of their commands might challenge developers and testers, but AI's inherent ability to detect patterns at speeds unattainable by humans means that distinguishing combatants and noncombatants is both possible and legally permissible. Further, communications with commanders' headquarters might be enhanced in situations where unmanned machines enabled with AI are employed because of their superior ability to execute routine tasks and speeds far exceeding human capabilities. Collateral damage estimators already rely on AI to identify potential occurrences and to make recommendations spanning selected weapons systems to the so-called weaponeering options of the selected weapon system.[80] [81] Engineering machines to seek guidance from human commanders when the situation differs from the anticipated might satisfy *in bello* considerations. Projections of possible moral agency

[79] Roff, "Meaningful Human Control," 5.

[80] Douglas Stansbury, "AI – Conquering a Relative Disadvantage," (MMAS diss., U.S. Army Command and General Staff College, 2019), 11.

[81] Weaponeering is refers to establishing the types of weapons needed to achieve required levels of damage on a target and considers factors such as the reliability of the selected weapons, the vulnerability of the target to attack, and the accuracy of munitions.

resident in machines as provided above speaks to adherence to the element of accountability and might assuage *post bellum* concerns.

Beyond those legal considerations, though, the ethics of AI-enabled autonomous systems requires examination: to wit, the legality-based question of "can" these systems be properly authorized and used and the ethics-based question of "should" reveal a piercing difference. Considering the ethical risks inherent in AI in a general sense, the prospect of these unmanned systems being used in future war, and indeed in current conflicts such as Libya and Nagorno-Karabakh as provided earlier, portends a challenging future inherent with risks and challenges. While necessary, simply relying on legal arguments is insufficient to understand the bounds of human-authorized use and "proper" [design and testing].

Exploring the ethical nature of AI-enabled autonomous weapons: the value of ethics to define "proper."

As opposed to strictly legal arguments as so-called lagging frameworks, ethics provides a superior vehicle to practically establish which humans should authorize the use of these types of unmanned systems and to operationally define "proper." Recalling the common definition of "proper" and its associations to society and morality, ethics enables humans to make judgments in opaque cognitive environments and in unanticipated circumstances. As opposed to legality that relies on the principle of past precedent, ethics is an inherently forward-looking, futurizing field in that it seeks to recommend how humans should behave in novel situations or faced with new technologies.

Ethics will remain valuable as AI-enabled autonomous weapons become more prevalent and as the character of the profession of arms and its interactions with humans extend past weapons under the direct control of a human operator to these types of unmanned weapons

systems. The profession of arms often requires evaluation of the greater good or the more advantageous utility.[82] Joseph Chapa establishes the distinction of retributive justice and distributive justice with the former necessitated in circumstances where the actions of guilty parties justify their punishment. Conversely, distributive justice – and particularly in a military setting – is the distribution of justice to the combatants to spare the noncombatants. Chapa states, "Distributive justice leaves us with the uncomfortable proposition that just warfighters must do ugly things to prevent uglier ones."[83] Considering autonomous weapons intended for operation independent of direct human involvement, then, the capability of evaluating the so-called uglier proposition must be expressly designed into the machine. AI-enabled weapons must be able to recognize disadvantageous outcomes.

In addition to the evaluation of greater good, the profession of arms and conflict between humans necessitates personal moral judgment.[84] Chapa continues: "If just warriors are eventually to leave the profession whole, they must come to recognize this moral remainder they will bear; and the earlier they recognize it the better."[85] Although AI-enabled autonomous weapons might lack inherent morality, their use can and should reflect the inherent morality of not only those employing them but also those designing and testing them and even those developing their coding. Particularly relevant due to AI's "intelligent" aspect, the entire community of users, code writers, companies, etc. bear ethical accountability given the possibility of uncontrollability – and because the entire community must recognize they *should* bear a moral remainder.

---

[82] This is best encapsulated in consequence-based ethics.

[83] Joseph Chapa, "Moral Philosophy as a Force Protection Measure," *Strategy Bridge,* July 19, 2019, https://thestrategybridge.org/the-bridge/2019/7/16/moralphilosophy-as-a-force-protection-measure.

[84] This speaks to the value of virtues-based ethics.

[85] Chapa, "Moral Philosophy."

Although the character of the profession of arms should remain constant, the character of war and conflict are changing. Decisions will need to be made at machine speed, and the migration towards cyberwarfare and the incorporation of hyper-sonics means that decision space for combatants will be compressed in time, further necessitating ethical agility in the decision-making process. The combatant that determines how to leverage autonomous weapons systems stands to obtain a competitive advantage over its adversaries.

Beyond the character of conflict, though, Christian Brose argues its basic nature will not change because the nature of war is inextricably linked to the unchanging nature of humans. Whether it involves longbows or source code, war will always be violent, politically motivated, and composed of the same three elemental functions that new recruits learn in basic training: move, shoot, and communicate.[86] AI-enabled autonomous systems therefore afford an opportunity to master those three elemental functions due to inherent advantages in operating speed.

Curiously, these systems might even promote more ethical behavior in conflict. Brose provides, "greater autonomy can not only enhance military effectiveness; it can also allow more humans to pay more attention to the ethics of war than ever before."[87] There are ethical questions related to AI in general and in lethal autonomous weapons systems, but their use might also ironically assist in the resolution of ethical questions. "Autonomous systems will enable humans to spend less time on menial problems and more time on moral ones."[88] Unmanned lethal systems embedded with AI in either teammate or wingman roles can assist with tasks and activities such as identification of legitimate targets, tracking multiple targets and prioritizing

[86] Christian Brose, "The New Revolution in Military Affairs: War's Sci-Fi Future," *Foreign Affairs,* May/June 2019, https://www.foreignaffairs.com/articles/2019-04-16/new-revolution-military-affairs.
[87] Brose, "The New Revolution."
[88] Brose.

those according to pre-established guidelines, and assessment of tactical effectiveness.[89] Tracking friendly positions and promoting their safety represent a second possibility to be realized by these technologies. The U.S. Department of Defense's Unmanned Systems Roadmap asserts this level of automation will alleviate the human operator of task level activities associated with the engagement of a target, allowing the operator to focus on the identified threat and the decision to engage.[90] There are clear advantages to be realized through the convergence of the technologies of AI, automation, and unmanned systems.

Responsibility in autonomous systems is possible, similarly advantageous, and might even be increased in situations expressly designed with responsible use in mind.[91] AI systems that are explicitly programmed to seek human certainty in uncertain events - or those they haven't previously recognized or dealt with before - could promote responsible behavior. Kenneth Cukier provides that AI systems might be developed to learn human preferences. Developers should "explicitly program them to be uncertain about people's objectives and to possess the ability to learn more about them by observing human behavior."[92] Autonomous machines programmed in this manner will then strive to maximize their interactions with humans, especially in new or unfamiliar situations.

[89] *Unmanned Systems Roadmap 2017-2042*, 23.

[90] *Unmanned Systems Roadmap 2017-2042.*

[91] Deputy Secretary of Defense memorandum dated May 26, 2021, defines the ethical AI principle of Responsibility as "DoD personnel will exercise appropriate levels of human judgment and care, while remaining responsible for the development, deployment, and use of AI capabilities." On top of the two conditions identified in the Final Report's judgment, this professional paper aims to provide a working model qualifying, defining, and operationalizing "appropriate." See also https://media.defense.gov/2021/May/27/2002730593/-1/-1/0/IMPLEMENTING-RESPONSIBLE-ARTIFICIAL-INTELLIGENCE-IN-THE-DEPARTMENT-OF-DEFENSE.PDF.

[92] Kenneth Cukier, "Ready for Robots? How to Think About the Future of AI," *Foreign Affairs,* July/August 2019, https://www.foreignaffairs.com/reviews/review-essay/2019-06-11/ready-robots.

The relationship between AI and ethical decision making – it boils down to quality.

AI will play a more important role in ethical decision making as machine learning is refined. It is a value-laden technology that derives its power from a constellation of actors: the companies that commission it, the code-writers that train and test it, and the users who employ it. When applied in a military setting, that constellation expands to the states and governments that authorize its use and commanders that make life and death decisions with it. Ethically operationalizing this technology in a military setting requires a mind-bending perspective around what responsibility and quality mean. Christian Brose notes, "The old belief that software merely supports hardware must be inverted: future militaries will be distinguished by the quality of their software, especially their AI."[93]

When considering AI-enabled autonomous weapons, Ian Barbour's middle ground of Technology as an Instrument of Power correlating with Kranzberg's First Law is instructive: "A third basic position holds that technology is neither inherently good nor inherently evil but is an ambiguous instrument of power whose consequences depend on social context."[94] AI-enabled weapons systems are not necessarily good nor evil. Rather, it is the ethics engineered into its development and use, or the lack thereof, that will have consequential effect on societies and either result in their positive use or negative outcomes. An appreciation for the ethics of technology takes Brose's assertion to a new level: the quality of AI-enabled weapons systems extends beyond the actual operation of the software to the values engineered in the systems.

The values of accountability, consent, identity, and, to some degree, privacy each speak to a dimension of quality. First, autonomous weapons and the reliance on neural networks and inherent hidden layers create challenges with accountability and understanding of decision-

[93] Brose, "The New Revolution."
[94] Ian Barbour, *Ethics in an Age of Technology* (San Francisco: HarperCollins, 1993), 15.

making, and so-called black box AI further challenges these types of systems. This element of AI must be overcome if AI-enabled autonomous weapons are to be employed, as they directly compete with the international humanitarian law principle of accountability. Rather, developers should strive for so-called explainable AI and "develop machine-learning systems that provide an explanation for their decisions and recommendations and allow users to know when and why the system will succeed or fail."[95]

Autonomous weapons development therefore must provide for an explanation of how the machine arrived at Conclusion X over Conclusion Y. Cukier cites Judea Pearl's contribution: "Deep learning has its own dynamics, it does its own repair and its own optimization, and it gives you the right results most of the time. But when it doesn't, you don't have a clue about what went wrong and what should be fixed."[96] AI will operate both using rules programmed into the machine and the rules it develops to solve problems, so engineering accountability into machines is necessary to assure quality.

Consent represents a second dimension of ethical quality to be incorporated into responsible AI-enabled weapon systems, a dimension of increasing relevance as artificial intelligence improves. In general terms, AI raises questions if individuals have been provided with the opportunity to offer informed consent when engaging in interactions with machines.[97] The same would apply in a military context when using AI-enabled autonomous weapons in that commanders might need to be provided with the opportunity to consent to their use in support of their mission. To what degree should a senior commander gain the consent of a subordinate

---

[95] Government Accountability Office, *Artificial Intelligence,* 79.

[96] Cukier, "Ready for Robots?"

[97] Bernd Carsten Stahl, Job Timmermans, and Catherine Flick, "Ethics of Emerging Information and Communication Technologies: On the Implementation of Responsible Research and Innovation," *Science and Public Policy,* Volume 44, Issue 3, June 2017, https://www.researchgate.net/publication/308348361_Ethics_of_Emerging_Information_and_Communication_Technologies_On_the_implementation_of_responsible_research_and_innovation.

commander if an AI-enabled autonomous weapon is employed in either direct or some enabling capacity in his or her operational environment? What about an adjacent commander whose operational environment will be affected by an AI-enabled autonomous weapon employed from an adjacent joint area of operations? On top of these layers, multinational operations – the expected norm of future war – are complicated in that for every nation that nominally supports development and use of these types of unmanned weapons, there are two who do not support them. This has implications for how militaries structure forces and capabilities in that AI-enabled autonomous weapons might be affecting a commander's operational environment without that commander even realizing it. Development of these systems should account for this dimension of quality.

Crucially, identity represents a third dimension of ethical quality that must be addressed to implement responsible lethal autonomous weapons systems. AI systems often take on traditional functions fulfilled by humans, meaning that it "may alter our view on what it means to be a human or individual."[98] Adoption of these technologies in military capabilities may alter how individuals view the profession of arms. Harkening back to the ideally unchanging character of the profession of arms, just war theory, and Chapa's distributive justice, AI-enabled lethal weapons systems must not alter how militaries and their service members view themselves as members of the profession of arms. They do, however, increase those with moral remainder, meaning enduring and sustainable ethics education of the entire community and the constellation of actors from the user to the commander is even more important, an aspect explored later in this professional paper.

[98] Stahl, "Ethics of Emerging Information."

Last, AI – to include that enabling autonomous weapons – should respect persons' privacy. Privacy is individually and collectively viewed as a most-sacred right: as per Samuel Warren and Supreme Court Justice Louis Brandeis, "The principle which protects [privacy…] is in reality not the principle of private property, but that of an inviolate personality."[99] For all of this importance, though, AI has an inherent dilemma in that the most effective AI systems will be those benefitting from the most data, and, similarly, lesser availability of data will yield comparatively less-effective machines. Kenneth Cukier cites privacy regulations that limit or otherwise discourage the aggregation of data of a personal nature as an inherent drawback. He states, "Minimizing the quantity and type of data that can be used in AI systems may seem wise in an era when companies and countries are vacuuming up all the personal data they can and paying little attention to the risks of misuse."[100] When prospective regulations responding for calls to collect on data deemed absolutely necessary result in reduced volume within the training and testing data sets, less effective machine performance should be expected. The volume of data expected to train and test AI-enabled autonomous unmapped weapons and to employ them in an ethical manner, though, competes with this dynamic. This dilemma coming to the fore based on AI's ambiguous nature is particularly relevant in an autonomous lethal weapon system designed to operate against humans – there are some elements of inherent good and inherent bad.

As an emergent technology, AI embodies Barbour's argument that it is neither inherently good nor inherently bad, and, for better or worse, AI-enabled and autonomous weapons systems will be used as instruments of power. The DoD must recognize that the quality of software in these systems extends beyond the operation of the systems towards how they behave when

[99] Samuel Warren and Louis Brandeis, "The Right to Privacy," *Harvard Law Review* Vol. 4, No. 5. (Dec. 15, 1890): 205.

[100] Cukier, "Ready for Robots?"

presented with novel environments or situations. Therefore, engineering the values of accountability, consent, identity, and privacy from the start is additionally necessary to assure quality and to promote productive uses of national power.

The NSCAI's Final Report addresses emergent ethical concerns associated with general AI.[101] It might be used in manners inconsistent with our constitutional principles protecting the rights and liberties of individuals. Repression of individual freedoms becomes a greater possibility in an environment characterized by AI-enabled systems to maintain public order. Rights to stand trial in a jury of peers and the accused' innocence until proven guilty in a court of law is at risk with AI due to its ability to generate, at machine speed, a false narrative that looks completely genuine and approaching an advanced "deepfake" capability. Equal protections under the law are similarly at risk because errors or biases coded into AI might be magnified by machine learning. AI threatens the right to reasonable privacy and the protection of unwarranted search and seizure: simply due to the volume of personal information residing in the cloud or with companies, AI affords less and less protection against unauthorized or undesired disclosures.[102] Quality AI will assuage these general concerns, and to retain human-authorized use and maintain "proper design and testing" of these consequential weapons systems, it is up to the humans to assure the requisite quality. Applying ethics is the means to do so.

[101] See also the *Final Report's* "Key Considerations for the Development of Responsible Development and Fielding of AI (Abridged)" in its Appendix C starting on 633.

[102] In a sense, and even in open, liberal societies, humans are now being digitally stopped and frisked, or at least frisked, by digital technologies every day. The situation is far worse in authoritarian regimes with explicit intention to do so.

Chapter 3 – Ethically operationalizing human-authorized use and proper design and testing

Having demonstrated in the previous chapter how ethics provides the means for the Department of Defense to fulfill the Final Report's first judgment and its inherent conditions, this chapter pivots towards the relationship between ethics and the operationalization of AI-enabled autonomous weapons. The chapter starts with those ethical considerations to be applied to different elements of the DoD's workforce both in the AI field and those in non-technical roles. It then concludes with some common thematic aspects.

Applying ethics to the Department of Defense's development and use of AI-enabled autonomous weapons: a seven-archetype model.

The Department of Defense's Unmanned Systems Roadmap finds "situational ethical reasoning is currently not coherently implementable by AI in the range of scenarios that military forces may encounter."[103] Because lethal weapons systems operating autonomously and enabled by AI have already arrived on the battlefield, it is up to humans to imbue ethics into these systems. Ethics provides the means to determine the bounds of "human-authorized use" and "proper" design and testing. Understanding the ethical implications of these unmanned systems for the humans responsible for them is in order.

In view of its Final Report assertion that the United States Government is a long way from being AI-ready, the NSCAI in coordination with the Defense Innovation Board and the Joint AI Center proposed a seven-layer model to inspire the community towards AI-readiness.[104] AI worker archetypes span both technical and enabling roles as follows with the full model and summarized recommendations included as an appendix.

- AI Experts lead the ethical design of AI-driven technologies.

[103] *Unmanned Systems Roadmap 2017-2042*, 23.
[104] *Campaign for an AI Ready Force,* 8.

- AI Developers are data-focused and responsible for model training and tuning.
- Deployment Specialists install and maintain hardware and software.
- End Users use AI-enabled systems during normal operations.
- Non-technical Tactical Leaders provide domain knowledge and create the tactical requirements.
- Non-technical Strategic Leaders create and oversee enterprise objectives, and
- Support Roles provide those functions necessary to support AI development and employment.[105]

Ethical considerations also apply to each of these archetypes when projecting the introduction of AI-enabled autonomous weapons into future war as nominally supported by the Final Report.

The technical roles fulfilled by AI experts, developers, and deployment specialists all must apply ethics when approaching autonomous unmanned systems. First, experts are expected to oversee testing and evaluation and should exhibit mastery of the algorithms governing these systems' operation.[106] Because these types of weapon systems may face novel situations and/or adaptive foes as described previously, these experts must be able to project those ethical risks inherent in these situations. The possibility of a communications-denied environment, hacking, or spoofing means that experts must challenge the community to identify how they will uniquely affect AI-enabled autonomous unmanned systems. Not only must these systems be "properly" designed and tested, but the AI experts themselves must understand the ethics that may apply. And experts must develop and recommend concepts of machines' moral agency in order to promote these systems' legal and ethical use in future war.

[105] *Campaign for an AI Ready Force,* 10-12.
[106] *Campaign for an AI Ready Force,* 10.

Developers must also understand the ethical implications of AI-enabled autonomy in order to fully exercise their roles and responsibilities associated with data curation; model training, testing, and tuning; and exploration of new opportunities.[107] AI might be able to make faster collateral damage estimates and weaponeering decisions, but only if it is programmed to do so. Documented bias in AI, misapplication of data, and breaches of privacy represent the most-prominent ethical risks that developers must account for in these unmanned systems' development. Developers ignorant of these ethical risks challenge the legality of these systems' employment vis-à-vis international humanitarian law and Just War Theory. The possibility that drones might directly target humans – as advertised in the Turkish autonomous unmanned systems used by the Government of National Accord to target the Libyan National Army – only exacerbates the risks associated with bias inherent in AI.[108] They must also recognize when so-called black box AI might be grafted into these systems – and subsequently avoided – in order that accountability is retained.

Third, deployment specialists must also understand the ethics of AI to promote "proper" design and testing. Deployment specialists armed with technical expertise represent the common point of contact for end users and exercise responsibility to install and maintain hardware and software.[109] Recalling Brose's argument that the quality of AI-enabled weapons systems extends beyond the actual operation of the software to the values engineered in the systems, deployment specialists must be able to maintain quality control and quality assurance of the AI enabling autonomous unmanned weapons. Quality's dimensions of privacy, accountability, identity, and consent must each be clearly understood and recognized by deployment specialists as they field

[107] *Campaign for an AI Ready Force.*

[108] Additional information on the capabilities and technical features of the Kargu Rotary Wing Attack UAV is available at https://www.stm.com.tr/en/kargu-autonomous-tactical-multi-rotor-attack-uav.

[109] *Campaign for an AI Ready Force*, 11.

and support users' employment of AI-enabled autonomous weapons in order that users' may be properly sensitized to, colloquially, what right looks like. Deployment specialists' assurance of quality represents a third facet of "proper" design and testing.

In the parlance of the NSCAI's model for the federal government's AI workforce, so-called enablers provide the means to define "human authorized use." The enablers of end users, non-technical tactical leaders, non-technical strategic leaders, and support roles each provide important elements of "human authorized use." The Department of Defense already recognizes the general impact AI will have on those outside the AI technical community; to apply AI ethically, its recommendations for ethical use argue for role-relevant training and education programs. The department's AI principles establish, "It is imperative that junior officers, enlisted service members, and civilians are exposed to AI in their training and education early in their careers, and that DoD provides opportunities for continued learning throughout their careers."[110] Based on the proliferation of AI-enabled autonomous weapons plus the well-placed international concerns surrounding it, the ethics of this technology to be considered by non-technical tactical and strategic leaders and by support roles becomes even more important.

The Campaign for an AI Ready Force (hereafter referred to as the Campaign) establishes non-technical tactical leaders as those experts with domain-particular subject matter expertise who will create the tactical requirements for these types of systems.[111] It specifies "they should be trained to understand the basics of data collection and management, AI decision making, and AI specific ethics" in order that they might effectively contribute to the future AI workforce.[112]

[110] Defense Innovation Board, *AI Principles: Recommendations on the Ethical Use of AI by the Department of Defense,* 9, accessed September 7, 2021, https://media.defense.gov/2019/Oct/31/2002204458/-1/-1/0/DIB_AI_PRINCIPLES_PRIMARY_DOCUMENT.PDF.

[111] *Campaign for an AI Ready Force*, 11.

[112] *Campaign for an AI Ready Force*.

These tactical leaders are the local authorizers of these weapons systems and therefore must appreciate the ethical risks developed previously and, during system testing, advise on those domain-specific novel situations not necessarily apparent to the technical community. If, as postulated, AI-autonomy is to create additional capacity for tactical leaders to consider issues of ethics, then those tactical leaders must demonstrate the competency and capacity to do so.

Non-technical tactical leaders are also likely best postured to understand how adversaries may adapt to these unmanned weapons systems. They bring experience in practically applying the rules of engagement and using lethal force via manned or even unmanned systems (albeit under the control of a human operator). They can practically define both the intended effects of these future weapons systems and the range of environments they may operate in. For all this domain knowledge, though, they must also appreciate the ethics concerns inherent in AI lest they overlook aspects of privacy intrusions or inherent bias. This appreciation for ethical concerns can promote an effective dialogue with developers, thus providing for "proper design and testing." It will also rightly weigh on their determination of questions surrounding authority and human-authorized use. They must understand when to trust the AI to arrive at the appropriate answer and to question, and subsequently control, the machine when it does not. Given the increasing computational power in machines, non-technical tactical leaders authorizing AI-enabled autonomous weapons must also demonstrate correspondingly increasing capability to make decisions of an ethical nature.

Guiding the non-technical tactical leaders are the non-technical strategic leaders who "oversee the creation of strategic and enterprise objectives, the deployment and scaling of new systems, and manage the careers of developers and experts."[113] A forward-looking

[113] *Campaign for an AI Ready Force*.

understanding of where these systems and their technological successors are trending is necessary to assure adherence to the principles of international humanitarian law and Just War theory. They must recognize who should consent to human-machine interactions and identify liabilities for AI that misinforms or underperforms. Directly addressing the ethics of AI, the strategic leaders must evaluate where data must be obtained to support "proper design and testing" but also how it can be ethically obtained to protect the rights of the individual as enshrined in the U.S. Constitution. They must assure consent is established when using the data necessary to train and test these types of unmanned systems. Based on their management of developers' and experts' careers, strategic leaders must also assure their ethical development.

Recalling the international relations aspect of the topic, non-technical strategic leaders must also define and describe how these types of systems contribute either to deterrence by denial, or the making of an adversary's attack so costly they will not even attempt it, or deterrence by punishment, or the capability to effectively penalize an adversary should it take unwanted action. Should AI-enabled autonomous weapons systems be developed and employed to, for instance, create a web of machine-speed capable ballistic missile defense that effectively prevents adversaries from threatening the homeland? Or should they be developed with an offensively-oriented portfolio to respond – at machine speed – to an adversary attack? The speed of warfare enabled by autonomy challenges conventional deterrence concepts in that escalating responses can be triggered faster than ever before, and it is up to the non-technical strategic leaders to reframe how the DoD contributes to deterrence in an age of technology.[114]

Another international relations aspect is the so-called "intensity of interests" striating national interests into three general layers: vital interests (i.e., those national security actors are

[114] Lee, "The Third Revolution."

willing to die for), important interests (…willing to kill for), and peripheral interests (… willing to fund or pay for).[115] The use of unmanned systems as a means of reducing risk to friendly casualties is thoroughly documented. But in situations where the nation is merely willing to kill to achieve its objectives (that is, achievement of an important – but not vital – national interest), a challenge emerges when the adversary is willing to die for their defense of the targeted (vital) interest. Given the possibility of an escalatory adversarial response, the question surrounding use of AI-enabled autonomous weapon systems in these situations will therefore no longer be confined to one of legality (i.e., "can" they be used?) but one of ethics (i.e., "should" they be used?). Both concepts speak to the obligation of strategic leaders to appropriately guide the community both from a necessity of assuring "proper design and testing" and to (a) provide the framework in which these weapons are authorized by use either at the level of the strategic leader or delegated to a tactical leader and (b) align the use of AI-enabled autonomous unmanned weapons to the appropriate intensity of national interests.

Those in support roles also contribute to ethical application of these technologies. Including acquisition officers, contracting officers, human resources specialists, legislative affairs personnel, and legal professionals, those in support roles must also enable human authorization of these systems and assure "proper design and testing."[116] The Unmanned Systems Roadmap provides "[the Department of Defense] should strengthen connections to the private sector so that as AI/[machine learning] solutions mature, [the department] is able to procure the most promising solutions and use them in unmanned systems."[117] Those supporting

[115] Derek Reveron and Jim Cook, "From National to Theater: Developing Theater Strategy," *Joint Forces Quarterly* 70 (3rd Quarter 2013): 114, https://ndupress.ndu.edu/Portals/68/Documents/jfq/jfq-70/JFQ-70_113-120_Reveron-Cook.pdf.

[116] *Campaign for an AI Ready Force*, 12.

[117] *Unmanned Systems Roadmap 2017-2042*, 19.

acquisitions and contracts must recognize not only the inherent public-private nature of these systems' development but the inherent ethical risks – and the opportunities well-placed and ethical private AI companies provide the Department of Defense. Unmanned surveillance systems like the U.S. Navy's Scan Eagle are increasingly contracted in that the Department of Defense funds the obtained services but contracts out both the ownership and the operation of the surveillance systems.[118] Because technology develops rapidly, contracted solutions will probably increase. While the offensive use of AI-enabled autonomous weapons should remain the purview of uniformed military, situations involving ballistic missile defense, countering hyper-sonic weapons, or the defeat of swarming attack drones might involve elements of a contracted response. Therefore, those DoD contracting officers must also have a firm grasp of the ethics of these technologies to assure "proper design and testing" and maintain humans' authorization of their use.

Human resource specialists also play a role in operationalizing the ethics of these technologies. They must recognize those candidates for AI technical roles with a firm grasp of the ethical risk and challenges and those without, and they must structure human resources systems that hire, promote, and develop those technical candidates valuing the ethics of AI-enabled autonomy. Considering the wide range of viewpoints at the international level and the role of the U.S. Congress in shaping foreign policy (either though committee oversight or legislatively authorized funding), legislative specialists must also remain cognizant of the ethics of AI and where it intersects with these types of unmanned weapons. Legal professionals ground the community in the basics and principles of international humanitarian law and Just War

[118] "ScanEagle – Mini-UAV (Unmanned Aerial Vehicle)," *Naval Technology*, accessed September 7, 2021, https://www.naval-technology.com/projects/scaneagle-uav/#:~:text=Under%20a%20contract%20placed%20in%20April%202005%2C%20US,a%20contract%20for%20ScanEagle%2C%20also%20used%20in%20Iraq.

theory, two necessities to fully realize ethical use of these systems, "proper design and testing," and human-authorized use.

The Campaign establishes end users as those using AI-enabled systems within their normal scope of operations.[119] They might likely possess little or no expertise on AI. These users could be the F-35 Joint Strike Fighter pilot supported by an AI-enabled loyal wingman or missile operators providing ballistic missile defense. Or they could be those charged with overcoming adversary anti-access/area denial environments or fulfilling on-the-ground counter-terrorism roles and precision targeting of high-value individuals. They might be pursuing retreating enemy forces like the Government of National Accord's reported pursuit of the Libyan National Army or attacking adversary capabilities in depth using joint fires. In all situations, though, end users will bear accountability for the weapon systems they employ as per international humanitarian law.

The advent of AI-enabled autonomy potentially challenges notions of end users. If these systems can perform tasks autonomously, and if they can use AI to increase sensitivity to their environments, then don't these machines become the end users in and of themselves? Recalling future war likely featuring humans' machines versus humans' machines, end users will remain human in that only humans will bear Chapa's moral remainder. Machines might aid humans' capabilities to wage war, but war will remain an act of humans seeking gain, political or otherwise. Therefore, basic understanding of ethics and the ethics of AI must permeate end users as these types of unmanned systems prospectively proliferate.

[119] *Campaign for an AI Ready Force,* 11.

Analysis of ethical considerations across the community: three common themes emerge.

Evaluation of the ethical considerations of the three AI technical worker archetypes and the four AI enabling working archetypes reveals some common themes germane to societies and morality. Investment in people resonates when operationalizing ethical AI-enabled autonomous weapons. More importantly than the machines that are working their way to the modern battlefield, humans will retain their predominate stamp on future warfare. Ethics education across the scope of risks and challenges emerging in AI therefore underpins these investments and provides a common cognitive, conceptual, and definitional framework to harmonize the technical and non-technical archetypes. The entire community must demonstrate role-relevant competence in the principles of ethical AI-enabled autonomy, its consequences, and its virtues.[120] They must also gain the experience and exposure to determine how and when each of these ethical lens might uniquely apply vis-à-vis Jack Kem's Ethical Triangle.[121] Particularly with the introduction of AI-enabled autonomy, ethics education is no longer confined to legal specialists and the clergy – it has become the purview and obligation of all. Rather, ethics – and the education that supports it – is akin to literacy that underpins every other field.[122] Chapa's moral remainder means ethics education will become even more important both in individual terms across the AI-worker archetypes and thematically across the entire community.

[120] Principles to be applied include addressing the moral obligations to assure accountability and the values of identify, consent, and privacy. A consequential question might be what is the greater good to be obtained by using AI-enabled autonomy as per Chapa's concept of distributive justice? Will these types of unmanned weapons prevent worse outcomes, or will they cause worse outcomes? The so-called "Front Page of the Washington Post" check is relevant to these sorts of unmanned weapons in that it is up to the humans to assure the actions of their unmanned AI-enabled autonomous weapons don't inadvertently compromise their basic values and demonstrates application of virtue ethics.

[121] Jack Kem, "Ethical Decision Making: Using the "Ethical Triangle," U.S. Army Command and General Staff College, 4, accessed September 10, 2021, https://www.cgscfoundation.org/wp-content/uploads/2016/04/Kem-UseoftheEthicalTriangle.pdf.

[122] "Ask an Ethicist: What is Ethics Education?" *Penn State News,* accessed September 10, 2021, https://news.psu.edu/story/377841/2015/10/30/impact/ask-ethicist-what-%E2%80%98ethics-education%E2%80%99.

These types of weapons systems also increase the importance of public-private partnerships. As provided in the Unmanned Systems Roadmap, "many AI/[machine learning] solutions rely on large, integrated cloud technologies for data storage, processing, and dissemination."[123] Defense innovation is predominately a facet of the private sector and one that the Department of Defense benefits from. Considering the quality of software as a defining element of future war's prospective reliance on AI-enabled autonomy, rapidly accessing private software is a critically important aspect to ethically apply these weapons systems. Further, solutions to challenges, ethical or otherwise, and unforeseen opportunities will likely emerge from the private sector given its comparative advantages in innovation and timely research and development. U.S. Government must therefore recognize that "proper design and testing" is only possible through effective public-private partnerships. The result of these partnerships, provided thorough application of the ethics explored in this professional paper, will also assure human authorization for use both overcomes AI's ethical risk and challenges and accounts for the necessary legal frameworks.

An additional aspect associated with AI-enabled autonomous weapons is the so-called security dilemma presented in international relations theory. Robert Jervis establishes that policies and developments which increase one state's security tend to decrease that of others meaning that an unintended spiral may result as these types of systems are further developed. [124] While beyond the scope of this analysis to fully explore, it is important to recognize that these systems exist in both military spheres and political circles. In simpler terms, though, operationalizing ethics in an environment of increasing AI-enabled autonomy serves to mitigate and control the unintended effects associated with the security dilemma.

---

[123] *Unmanned Systems Roadmap 2017-2042,* 19.
[124] Robert Jervis, "Cooperation under the Security Dilemma," *World Politics* 30*,* no. 2 (1978): 169.

Conclusion – Ethics is integral to AI-enabled autonomous weapons

The modern battlefield will be awash with AI-enabled unmanned weapons operating autonomously. Beyond current conflicts, future war will possibly be fought at machine speed, and the race to retain or gain dominance in this field cannot be ignored. While cognizant of, and perhaps because of, international views on these types of weapons, the United States must effectively develop and use these systems and do so in a manner that promotes ethics and legality.

This paper adopts "human authorized use" and "proper design and testing" as its guiding elements not only due to their citation in the Final Report without future substantive explanation but because their exploration and examination drives the necessary ethical considerations. Their analysis coupled with an understanding of the ethical risks and challenges revealed in the emerging technologies of AI-enabled autonomous weapons can and should impact the entire workforce ranging from those in technical roles to those using or enabling these systems. In addition to investing in the weapons systems themselves, investment in people via ethics education and similar forms of professional development is just – if not more – important as the character of war adapts and mutates.

The so-called killer robots have arrived. It is up to the community of those in AI technical roles and those enabling and enabled by AI to realize and fulfill their ethical responsibilities in order that these machines retain and support the humanity necessary to control the monopoly on violence. Appreciation of AI's ethical risks and challenges and how they intersect with autonomous unmanned weapons and affect societies and morality is the place to start.

Appendix - AI Workforce Model for Federal Government extended to summarize recommendations that promote human-authorized use and assure "proper" design and testing

The model developed by the NSCAI, the Defense Innovation Board, and the Joint Artificial Intelligence Center provides general ethical considerations for the adoption of AI.[125] The per-archetype recommendations put forth in this paper aim to further expand these considerations on the probability the Department of Defense will continue to develop and use AI-enabled autonomous unmanned weapons as judged in the Final Report. This appendix extends the model by identifying the ethical considerations necessary to promote human-authorized use in an ethical manner and assure "proper" design and testing.

Build Consensus			Questions for Departments		Ethical Considerations Relevant to AI-Enabled Autonomous Weapons
Worker Archetypes	Output	Capabilities (Ethics Throughout)	Training, Education, and Recruitment	Organizational Needs and Composition	What is necessary to promote human-authorized use and assure proper design and testing?
AI Expert	Leads the ethical design, development, and deployment of AI-driven technologies; oversees test and evaluation (verification and validation) to determine technology readiness; helps maintain and leverage supporting data architecture; translates requirements into capabilities; translates technical topics for senior leaders	Expert in data science, machine learning (e.g., deep learning), AI lifecycle, applied ethics and one or more of the following: natural language processing, computer vision, robotics, human-computer interfaces; human centered systems engineering; algorithmic and computational theory	How will the national security community train or recruit and integrate AI experts?	How many AI experts does the national security workforce need? Where should they be? Should they be uniformed, civilian, or contractors?	• Determination of ethical risk when facing novel situations and/or adaptive foes. • Mitigation of the communications-denied environment, hacking, or spoofing. • Development of concepts of moral agency within machines.
AI Developer	Data selection and preprocessing; model selection, training, and validation; partnership with domain knowledge experts and end users; discovery of local opportunities	Computational statistics and data science; programming (e.g. Python or R); model development using an ML library.	Who trains developers for the national security workforce? When will they be identified and trained?	How many developers does the national security workforce need? Should they be uniformed, civilian, contractors, and/or contracted companies?	• Continued improvement of automated collateral damage estimators and weaponeering. • Addressing documented bias particularly when paired with facial recognition software. • Reliance on explainable AI.
Deployment Specialist	Infrastructure installation and maintenance, review input/output sent by end-users, additions to training data sets, rough examination of training data sets, training/testing existing models, deployment	Hardware/software installation and maintenance, training data management, model verification/validation, algorithm deployment, data cleansing	Education equivalent to a technical certification offered by a military program or vocational training.	How many AI technicians does the national security workforce need? Where should they be?	• Quality control and quality assurance of AI enabling autonomous weapons. • Sensitizing users to situations where privacy, accountability, identity, and/or consent might be at risk.

[125] The Table is provided in the *Campaign for an AI Ready Force* starting on page 8.

Build Consensus			Questions for Departments		Ethical Considerations Relevant to AI-Enabled Autonomous Weapons
Worker Archetypes	Output	Capabilities (Ethics Throughout)	Training, Education, and Recruitment	Organizational Needs and Composition	What is necessary to promote human-authorized use and assure proper design and testing?
End User	Daily business augmented/enabled by AI	Use of systems and apps	Normal systems training	Ubiquitous	• Basic understanding of ethics and the ethics of AI. • Recognition of their enduring accountability as per IHL.
Non-technical Tactical Leader	Gathers tactical requirements to guide the development of new AI-enabled capabilities, oversees deployment to ensure tactical requirements are met; partners with technicians, data engineers, and AI experts; leads normal operations.	Tactical domain implementation expert, basic data collection and management, basic understanding of AI decision making within the context of use and the sources of failures and errors, ethics applied to tactical use	How will the national security community train and educate tactical leaders? How much do they need to know?	How many tactical leaders should the national security enterprise have?	• Competence in AI's ethical concerns. • Familiarity with issues surrounding trust in AI and practical experience (through training, education, etc) making trust/don't trust decisions. • Mastery in ethical decision making.
Non-technical Strategic Leader	Oversees the creation of strategic and enterprise objectives, considers the ethics of new capabilities, oversees deployment and scaling, partnership with experts, developers, and tactical leaders; career management.	Basics and ethics of AI lifecycle, strategic and enterprise expertise, tactical domain management, software development processes	When and where will leaders learn about AI? How much do they need to know?	How will leaders incentivize AI competence? How many leaders need to be competent, and at what point in their careers?	• Retaining control over consent decisions. • Identifying liability and accountability for underperforming AI. • Evaluation of data sources' reliability and veracity. • Promoting the ethical development of the other six worker archetypes. • Integration of AI-enabled autonomous weapons into deterrence concepts. • Determination of when AI-enabled autonomous weapons might intensify or inadvertently escalate national interests.
Support Roles	Acquisition and contracting of AI hardware and software, services, and identification of commercial opportunities; legal support; legislative affairs, human resources, etc.	Understanding of software purchasing, data boundaries/limitations and rights; funding requirements; and compute purchases, identification of skill and qualifications of AI practitioners; legal and ethical aspects of development and deployment	When and where will support experts learn about AI? How much do they need to know?	What parts of the support workforce needs to learn about AI demands?	• Prioritization of ethical private AI companies and contractor providers. • Selection of candidates with firm understandings of ethics. • Cognizance of the ethics of AI.

Bibliography

Angwin, Julia, Jeff Larson, Surya Mattu and Lauren Kirchner. “Machine Bias: There’s Software Used Across the Country to Predict Future Criminals. And It’s Biased Against Blacks.” *ProPublica,* May 23, 2016. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

Arnold, Zachary, and Helen Toner. “AI Accidents: An Emerging Threat What Could Happen and What to Do.” *Center for Security and Emerging Technology*, July 2021. https://cset.georgetown.edu/wp-content/uploads/CSET-AI-Accidents-An-Emerging-Threat.pdf.

Barbour, Ian. *Ethics in an Age of Technology*. San Francisco: HarperCollins, 1993.

Bossman, Julia. “Top 9 Ethical Issues in Artificial Intelligence.” *World Economic Forum*, October 21, 2016. https://www.weforum.org/agenda/2016/10/top-10-ethical-issues-in-artificial-intelligence/.

Brose, Christian. "The New Revolution in Military Affairs: War’s Sci-Fi Future." *Foreign Affairs,* May/June 2019. https://www.foreignaffairs.com/articles/2019-04-16/new-revolution-military-affairs.

Chapa, Joseph. “Moral Philosophy as a Force Protection Measure.” *Strategy Bridge,* July 19, 2019. https://thestrategybridge.org/the-bridge/2019/7/16/moralphilosophy-as-a-force-protection-measure.

Chuter, Andrew. “British Shell Out Seed Funding For ‘Loyal Wingman’ Combat Drone.” *Defense News,* January 25, 2021. https://www.defensenews.com/global/europe/2021/01/25/british-shell-out-seed-funding-for-loyal-wingman-combat-drone/.

Cox, Matthew. “To Defeat Enemy Drone Swarms, Troops May Have to Take a Back Seat to Machines, General Says.” *Military.com,* January 25, 2021. https://www.military.com/daily-news/2021/01/25/defeat-enemy-drone-swarms-troops-may-have-take-back-seat-machines-general-says.html#:~:text=To%20Defeat%20Enemy%20Drone%20Swarms%2C%20Troops%20May%20Have,Multinational%20Simulation%20Center%2C%20Grafenwoehr%2C%20Germany%2C%20in%20March%2C%202018.

Cukier, Kenneth. "Ready for Robots? How to Think About the Future of AI." *Foreign Affairs*, July/August 2019. https://www.foreignaffairs.com/reviews/review-essay/2019-06-11/ready-robots.

Defense Innovation Board. *AI Principles: Recommendations on the Ethical Use of Artificial Intelligence by the Department of Defense.* Accessed September 7, 2021.

https://media.defense.gov/2019/Oct/31/2002204458/-1/-1/0/DIB_AI_PRINCIPLES_PRIMARY_DOCUMENT.PDF.

_____. *Campaign for an AI Ready Force.* Accessed September 5, 2021. https://media.defense.gov/2019/Oct/31/2002204191/-1/-1/0/CAMPAIGN_FOR_AN_AI_READY_FORCE.PDF.

Department of Defense. *Autonomy in Weapons Systems.* DOD Directive 3000.09. Washington, DC: Department of Defense, May 8, 2017. https://www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodd/300009p.pdf.

_____. *Implementing Responsible Artificial Intelligence in the Department of Defense.* Accessed September 4, 2021. https://media.defense.gov/2021/May/27/2002730593/-1/-1/0/IMPLEMENTING-RESPONSIBLE-ARTIFICIAL-INTELLIGENCE-IN-THE-DEPARTMENT-OF-DEFENSE.PDF.

_____. *Unmanned Systems Roadmap 2017-2042.* Accessed September 5, 2021. https://www.defensedaily.com/wp-content/uploads/post_attachment/206477.pdf.

Department of the Navy. *Unmanned Campaign Framework.* March 16, 2021. https://www.navy.mil/Portals/1/Strategic/20210315%20Unmanned%20Campaign_Final_LowRes.pdf?ver=LtCZ-BPlWki6vCBTdgtDMA%3D%3D.

Freedberg, Sydney. "Should Pentagon Let Robots Kill Humans? Maybe." *Breaking Defense*, July 10 2017. https://breakingdefense.com/2017/07/should-pentagon-let-robots-kill-humans-maybe/.

Freedman, David, "US Is Only Nation with Ethical Standards for AI Weapons. Should We Be Afraid?" *Newsweek Magazine,* September 15, 2021. https://www.newsweek.com/2021/09/24/us-only-nation-ethical-standards-ai-weapons-should-we-afraid-1628986.html.

Government Accountability Office. *Artificial Intelligence: An Accountability Framework for Federal Agencies and Other Entities.* GAO-21-519SP (Washington, DC: GAO, June 2021). https://www.gao.gov/assets/gao-21-519sp.pdf.

Gurcan, Metin. Turkish Drone Sets Off International Buzz Over 'Killer Robots'. *Al-Monitor,* June 8, 2021. https://www.al-monitor.com/originals/2021/06/turkish-drone-sets-international-buzz-over-killer-robots.

Human Rights Watch. "Heed the Call: A Moral and Legal Imperative to Ban Killer Robots." August 21, 2018. https://www.hrw.org/report/2018/08/21/heed-call/moral-and-legal-imperative-ban-killer-robots.

International Committee of the Red Cross. *Ethics and Autonomous Weapon Systems: An Ethical Basis for Human Control?* Geneva, Switzerland: ICRC, April 3, 2018. https://www.icrc.org/en/document/ethics-and-autonomous-weapon-systems-ethical-basis-human-control.

International Justice Resource Center. "International Humanitarian Law." Accessed September 4, 2021. https://ijrcenter.org/international-humanitarian-law/.

Jervis, Robert. "Cooperation Under the Security Dilemma." World Politics 30, no. 2 (1978): 167-214.

Kem, Jack. "Ethical Decision Making: Using the "Ethical Triangle." *U.S. Army Command and General Staff College*. Accessed September 10, 2021. https://www.cgscfoundation.org/wp-content/uploads/2016/04/Kem-UseoftheEthicalTriangle.pdf.

Kranzberg, Melvin. "Technology and History: 'Kranzberg's Laws." *Technology and Culture,* no 3 (July 1986): 544-560.

Lee, Kai-Fu. "The Third Revolution in Warfare." *The Atlantic*, September 11, 2021. https://www.theatlantic.com/technology/archive/2021/09/i-weapons-are-third-revolution-warfare/620013/.

"Letter to Google C.E.O." *New York Times*, May 30, 2018. https://static01.nyt.com/files/2018/technology/googleletter.pdf.

Maruf, Ramishah. "Researchers Studying Facebook Misinformation Say They Were De-platformed." *CNN Business,* September 5, 2021. https://www.cnn.com/2021/09/05/media/reliable-sources-facebook-researchers-deplatform/index.html.

Moller, Michael, delivering a prepared statement of Antonio Guterres. "Secretary-General's message to Meeting of the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems." *United Nations,* March 25, 2019. https://www.un.org/sg/en/content/sg/statement/2019-03-25/secretary-generals-message-meeting-of-the-group-of-governmental-experts-emerging-technologies-the-area-of-lethal-autonomous-weapons-systems.

National Security Commission on Artificial Intelligence. *Final Report.* Arlington, VA: National Security Commission on Artificial Intelligence, 2021. Accessed September 13, 2021. https://www.nscai.gov/wp-content/uploads/2021/03/Full-Report-Digital-1.pdf.

O'Conner, Tom. "Russia's Military Challenges U.S. and China by Building a Missile That Makes Its Own Decisions." *Newsweek,* 20 July 20, 2017. https://www.newsweek.com/russia-military-challenge-us-china-missile-own-decisions-639926.

Penn State News. "Ask an Ethicist: What is Ethics Education?" Accessed September 10, 2021. https://news.psu.edu/story/377841/2015/10/30/impact/ask-ethicist-what-%E2%80%98ethics-education%E2%80%99.

Qazizai, Fazelminallah. "The Drone Unit That Helped the Taliban Win the War." *Newslines Magazine,* September 15, 2021. https://newlinesmag.com/reportage/the-drone-unit-that-helped-the-taliban-win-the-war/.

Reuters Staff. "China Eyes Artificial Intelligence for New Cruise Missiles." *Reuters,* August 19, 2016. https://www.reuters.com/article/us-china-defence-missiles-idUSKCN10U0EM.

Reveron, Derek, and Jim Cook. "From National to Theater: Developing Theater Strategy." *Joint Forces Quarterly*, 70 (3rd Quarter 2013): 113-120. https://ndupress.ndu.edu/Portals/68/Documents/jfq/jfq-70/JFQ-70_113-120_Reveron-Cook.pdf.

Riza, M. Shane. *Killing Without Heart: Limits on Robotic Warfare in an Age of Persistent Conflict.* Washington, DC: Potomac Books, May 1, 2013.

Roff, Heather. "Lethal Autonomous Weapons and Jus Ad Bellum Proportionality." *Case Western Reserve Journal of International Law*, 47 (2015): 37-52. Accessed September 13, 2021. https://scholarlycommons.law.case.edu/cgi/viewcontent.cgi?referer=&httpsredir=1&article=1006&context=jil.

Roff, Heather, and Richard Moyes. "Meaningful Human Control, Artificial Intelligence and Autonomous Weapons." Briefing paper prepared for the Informal Meeting of Experts on Lethal Autonomous Weapons Systems, UN Convention on Certain Conventional Weapons, April, 2016. https://article36.org/wp-content/uploads/2016/04/MHC-AI-and-AWS-FINAL.pdf.

Rotenberg, Marc, and Tuan Nguyen, "Untitled letter to the Chairman and Ranking Members of the House Armed Services Subcommittees," 25 March 2021. https://www.caidp.org/.

Sayler, Kelley M. "International Discussions Concerning Lethal Autonomous Weapon Systems." *Congressional Research Service*. April 19, 2021. https://crsreports.congress.gov/product/pdf/IF/IF11294.

Sharp, Rachel. "Revealed: How Iran's Top Nuclear Scientist Was Assassinated by a Killer AI Machine Gun That Allowed Sniper Based 1,000 Miles Away to Fire 15 Bullets After Disguised Spy Car Had Pinpointed His Location." *DailyMail.com,* September 20, 2021. https://www.dailymail.co.uk/news/article-10004269/How-Irans-nuclear-scientist-assassinated-killer-robot-machine-gun-kitted-AI.html.

Stahl, Bernd Carsten, Job Timmermans, and Catherine Flick. "Ethics of Emerging Information and Communication Technologies: On the Implementation of Responsible Research and Innovation." *Science and Public Policy* 44, no. 3 (June 2017).

https://www.researchgate.net/publication/308348361_Ethics_of_Emerging_Information_and_Communication_Technologies_On_the_implementation_of_responsible_research_and_innovation.

Stansbury, Douglas Stansbury. "Artificial Intelligence – Conquering a Relative Disadvantage." MMAS diss., U.S. Army Command and General Staff College, 2019.

Stupp, Catherine. "Fraudsters Used AI to Mimic CEO's Voice in Unusual Cybercrime Case." *Wall Street Journal,* August 30, 2019. https://www.wsj.com/articles/fraudsters-use-ai-to-mimic-ceos-voice-in-unusual-cybercrime-case-11567157402.

Tangredi, Sam. "Sun Tzu Versus AI: Why Artificial Intelligence Can Fail in Great Power Competition." *Proceedings* 147, no. 5/1,419 (May 2021).

Tracy, Ryan. "As Google, Microsoft and Amazon Seek Bigger Defense Role, Some Are Leery." *Wall Street Journal,* September 7, 2021. https://www.wsj.com/amp/articles/tech-industry-seeks-bigger-role-in-defense-not-everyone-is-on-board-11631019600.

Tran, Tony. "Facebook Apologizes for Racist AI that Labeled Black Men as 'Primates'." *The Byte,* September 5, 2021. https://futurism.com/the-byte/facebook-racist-ai.

United Nations. "Autonomous Weapons that Kill Must Be Banned, Insists UN Chief." March 25, 2019. https://news.un.org/en/story/2019/03/1035381.

University of Texas McCombs School of Business: Ethics Unwrapped. "Moral Agent." Accessed September 4, 2021. https://ethicsunwrapped.utexas.edu/glossary/moral-agent.

Wallach, Wendell, and Colin Allen. *Moral Machines: Teaching Robots Right from Wrong.* Oxford: Oxford University Press, 2009.

Warren, Samuel, and Louis Brandeis. "The Right to Privacy." *Harvard Law Review* 4, no. 5. (Dec. 15, 1890): 193-220.

Wolchover, Natalie. "Artificial Intelligence Will Do What We Ask. That's a Problem." *Quantamagizine,* January 30, 2020. https://www.quantamagazine.org/artificial-intelligence-will-do-what-we-ask-thats-a-problem-20200130/.

Wolfgang, Ben. "Drones Have Outsized Impact on Future of War." *Washington Times,* January 12, 2021. https://www.washingtontimes.com/news/2021/jan/12/drones-have-outsized-impact-future-war/.