REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
05-27-2020

2. REPORT TYPE
FINAL

3. DATES COVERED *(From - To)*
N/A

4. TITLE AND SUBTITLE
Warriors and War Algorithms: Leveraging Artificial Intelligence to Enable Ethical Targeting

5a. CONTRACT NUMBER
N/A

5b. GRANT NUMBER
N/A

5c. PROGRAM ELEMENT NUMBER
N/A

6. AUTHOR(S)
John G. Thorne, Lt Col, USAF

Paper Advisor: Dr Timothy P. Schultz

5d. PROJECT NUMBER
N/A

5e. TASK NUMBER
N/A

5f. WORK UNIT NUMBER
N/A

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Ethics and Emerging Military Technology
Naval War College
686 Cushing Road
Newport, RI 02841-1207

8. PERFORMING ORGANIZATION REPORT NUMBER
N/A

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

10. SPONSOR/MONITOR'S ACRONYM(S)
N/A

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

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 Ethics and Emerging Military Technology Graduate Certificate. The contents of this paper reflect my own personal views and are not necessarily endorsed by the NWC, the Department of the Navy, or the Department of the Air Force.

14. ABSTRACT:
This paper examines how Artificial Intelligence (AI) is likely to enable the US military to execute its warfighting missions in a more ethical way, and thereby better abide by the moral intent of International Humanitarian Law (IHL). It also highlights how human-machine teaming may create risks. It accomplishes this task by examining two key principles of military targeting: distinction and proportionality. AI can greatly assist warfighters in practicing discrimination by enhancing their ability to positively identify their targets. Additionally, the ability of emerging algorithms to ingest and characterize large data sources provides analysts with unprecedented levels of battlespace awareness, which allows them to better discriminate between adversary targets and non-combatants. With regard to proportionality, AI provides an opportunity to more effectively model weapons effects and estimate collateral damage. These models will allow targeting analysts to foresee and avoid previously unknown second-order effects to non-combatants. Fundamentally, partnering a human analyst with an AI that can more quickly identify potential targets and accurately foresee the probable results of a strike is the most ethically preferable course of action at this time. Therefore, it is AI's ability to provide greater foreseeability that enables military professionals to optimally adhere to the principles of distinction and proportionality.

15. SUBJECT TERMS (Key words)
Artificial Intelligence, computer vision, ethics, international humanitarian law, joint targeting

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
					Dr. Thomas E. Creely
a. REPORT	b. ABSTRACT	c. THIS PAGE	N/A	65	19b. TELEPHONE NUMBER *(include area code)*
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED			401-841-6499

Standard Form 298 (Rev. 8-98)

NAVAL WAR COLLEGE
Newport, RI

Warriors and War Algorithms: Leveraging
Artificial Intelligence to Enable Ethical Targeting

John Thorne

Lt Col, USAF

Date Submitted: 28 May 2020

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

Table of Contents

Abstract

This paper examines how Artificial Intelligence (AI) is likely to enable the US military to execute its warfighting missions in a more ethical way, and thereby better abide by the moral intent of International Humanitarian Law (IHL). It also highlights how human-machine teaming may create risks. It accomplishes this task by examining two key principles of military targeting: distinction and proportionality. AI can greatly assist warfighters in practicing discrimination by enhancing their ability to positively identify their targets. Additionally, the ability of emerging algorithms to ingest and characterize large data sources provides analysts with unprecedented levels of battlespace awareness, which allows them to better discriminate between adversary targets and noncombatants. With regard to proportionality, AI provides an opportunity to more effectively model weapons effects and estimate collateral damage. These models will allow targeting analysts to foresee and avoid previously unknown second-order effects to noncombatants. Fundamentally, partnering a human analyst with an AI that can more quickly identify potential targets and accurately foresee the probable results of a strike is the most ethically preferable course of action at this time. Therefore, it is AI's ability to provide greater foreseeability that enables military professionals to optimally adhere to the principles of distinction and proportionality.

Acknowledgments

I would like to thank my wife and children for their support during this academic year. They also unselfishly allowed me to transform our entire dining room into a research and writing center. Dr. Timothy Schultz was invaluable as an advisor and editor. His ability to pass me the right book or article at the right time was critical to the completion of this work as were his unparalleled editing skills. Finally, I would like to thank Dr. Thomas Creely and the entire Ethics and Emerging Military Technology cohort Your collective intellectual and emotional support was much appreciated as our academic year was altered by COVID-19.

INTRODUCTION

This paper examines how Artificial Intelligence (AI) is likely to enable the US military to execute its warfighting missions in a more ethical way and thereby better abide by the moral intent of International Humanitarian Law (IHL). It also highlights challenges introduced by AI. It will accomplish this task by examining two key characteristics of military targeting: distinction and proportionality. Distinction and proportionality are two of the four[1] fundamental principles of ethical warfare acknowledged by IHL, to include the Hague Conventions, the Geneva Conventions and their Additional Protocols, and ICRC-compiled customary law.[2] A closer examination of these two principles reveals how the concept of foreseeability plays in ethical war-fighting. Foreseeability is a key ethical consideration related to protecting noncombatants commonly referenced in Just War Theory. Michael Walzer's foundational work, *Just and Unjust Wars*, argues unforeseen evil may be acceptable in war, but warriors are ethically required to minimize all foreseeable evil during war.[3] While foreseeability is not labeled a fundamental principle, it is an oft-overlooked characteristic that significantly affects the ethical decision-making for the primary four principles. Foreseeability is also the area where the prudent application of AI can produce the greatest benefits.

The examination consists of a brief review of the ethical underpinnings of distinction and proportionality and how both depend on foreseeability. It then evaluates current IHL and identifies potential opportunities where AI may provide increased compliance as well as

---

[1] The other two recognized principles are humanity and military necessity.

[2] "How Does Law Protect in War? Fundamental Principles of IHL," International Committee of the Red Cross accessed January 7, 2020, https://casebook.icrc.org/glossary/fundamental-principles-ihl.

[3] Michael Walzer, *Just and Unjust Wars: A Moral Argument with Historical Illustrations*, 4th ed (New York: Basic Books, 2006), 155.

characterize potential risks. The examination primarily focuses on ethical decisions and actions made during warfare/combat (*jus in bello*) as opposed to strategic decisions that may lead to war (*jus ad bellum*). Before executing an assessment of the potential and pitfalls AI may bring to military targeting, it is crucial to properly define AI and examine its increasing relevance to national security.

Artificial Intelligence is a complex topic both controversial and often misunderstood. Renowned physicist Stephen Hawking warned that AI may be the "worst event in the history of our civilization," unless its development is properly controlled.[4] Even entrepreneur Elon Musk, whose own company is innovating self-driving cars, warned "… AI is far more dangerous than nukes" and must be regulated.[5] Conversely, Rodney Brooks, the founding director of MIT's Computer Science and AI Laboratory,[6] Harvard's professor of Psychology, Steven Pinker, and even Facebook founder Mark Zuckerberg have described these views as alarmist.[7]

The intensity and scrutiny significantly increase when AI is introduced into the context of national security and military operations. Both Musk and Hawking signed an open letter to the United Nations urging for a ban on AI weapons.[8] One high-profile example of public backlash against the perceived 'weaponization' of AI occurred with Project Maven due to perceived ethical concerns. Project Maven was a Department of Defense (DOD) initiative designed to

---

[4] Arjun Kharpal, "Stephen Hawking Says A.I. Could Be 'Worst Event in the History of Our Civilization,'" *CNBC*, November 6, 2017, https://www.cnbc.com/2017/11/06/stephen-hawking-ai-could-be-worst-event-in-civilization.html.
[5] Catherine Clifford, "Elon Musk: 'Mark My Words — A.I. Is Far More Dangerous than Nukes,'" CNBC, March 13, 2018, https://www.cnbc.com/2018/03/13/elon-musk-at-sxsw-a-i-is-more-dangerous-than-nuclear-weapons.html.
[6] Connie Loizos,"This Famous Roboticist Doesn't Think Elon Musk Understands AI," *TechCrunch*, July 19, 2017, http://social.techcrunch.com/2017/07/19/this-famous-roboticist-doesnt-think-elon-musk-understands-ai/.
[7] Catherine Clifford, "Hundreds of A.I. Experts Echo Elon Musk, Stephen Hawking in Call for a Ban on Killer Robots," *CNBC*, November 8, 2017, https://www.cnbc.com/2017/11/08/ai-experts-join-elon-musk-stephen-hawking-call-for-killer-robot-ban.html.
[8] Clifford, "Elon Musk."

leverage AI to assess video from remotely piloted aircraft in an attempt to more quickly and accurately locate Islamic State of Iraq and Syria terrorists. The DOD embarked upon this endeavor with a number of commercial partners, including Google. However, in 2018 approximately 4,000 employees at Google argued the project contradicted Google's ethics which were based on the unofficial motto of "do no evil." The employees then signed a petition demanding "… Google nor its contractors will ever build warfare technology."[9] Shortly afterwards, Google withdrew from the project. These events were applauded by some and criticized by others who viewed the protests as short-sighted or naïve because less-ethical US competitors would still benefit from Google's work in AI.

Peter Thiel, a founder of Palantir, argued that Google's decision to not support Maven, while concurrently working on AI development with Chinese companies in Beijing, was misguided.[10] This is due to the fact that China's declared principle of "civil-military" fusion guarantees commercial research conducted in China will eventually be shared with the Chinese government's internal security forces and military. This sharing has led to a "vibrant dual use economy" which has allowed China to narrow the military power gap between its competitors, including the US.[11] While it was once assumed that economic and technical investment in China and other oppressive regimes would eventually promote democracy, it has had an inverse effect. The regimes have arguably maintained control of advanced technology to prevent the

[9] Daisuke Wakabayashi, and Scott Shane, "Google Will Not Renew Pentagon Contract That Upset Employees," *The New York Times*, June 1, 2018, https://www.nytimes.com/2018/06/01/technology/google-pentagon-project-maven.html.
[10] Peter Thiel, "Good for Google, Bad for America," *The New York Times*, August 1, 2019, https://www.nytimes.com/2019/08/01/opinion/peter-thiel-google.html.
[11] Tai Ming Cheung, *Fortifying China: The Struggle to Build a Modern Defense Economy* (Ithaca: New York:2013), 4.

development of political opposition movements.[12] One example involves China's development of a biometric database, which includes audio monitoring and facial-recognition technology, and is expected to be part of their future "social credit" system that allows the government to reward or punish citizens based on their monitored behavior.[13] This surveillance capability was clearly demonstrated with China's use of AI and facial recognition to identify and monitor protestors during the 2019 Hong Kong protests. Despite these concerns, there is potential for AI to be used responsibly and effectively for military use. For even the invasive Chinese surveillance state has been used for good and has been relatively embraced by the populace for its ability to track the spread of the COVID-19 epidemic in 2020.[14]

A compelling vision for the ethical employment of AI for military purposes was laid out by Lucas Kunce, a Marine who served in Iraq and Afghanistan. He argued that fears of providing advanced technology to the military were unfounded. He provided vignettes describing how AI-enhanced tools may have prevented his team from killing a civilian they believed to be throwing a grenade at them, when he was merely holding a shoe. He also described an engagement where one of his Marines shot and killed a young girl in a car that his unit believed to be a vehicle-borne improvised explosive device.[15] It is this viewpoint which accurately depicts the military as fellow citizens attempting to accomplish a mission given to them by their elected leadership. It also notes that a professional military is not a pack of

[12] Bruce Bueno de Mesquita, and George W. Downs, "Development and Democracy," *Foreign Affairs* 84, no. 5 (2005): 85, https://doi.org/10.2307/20031707.
[13] Elizabeth C. Economy, "China's New Revolution: The Reign of Xi Jinping," *Foreign Affairs* 97, no. 3 (2018): 60–74, http://usnwc.summon.serialssolutions.com/.
[14] Yingzhi Yang, and Julie Zhu, "Coronavirus Brings China's Surveillance State out of the Shadows," *Reuters,* February 7, 2020, https://www.reuters.com/article/us-china-health-surveillance-idUSKBN2011HO.
[15] Lucas Kunce, "Dear Tech Workers, U.S. Service Members Need Your Help," *The New York Times, August 28, 2019*, https://www.nytimes.com/2019/08/28/opinion/military-war-tech-us.html.

indiscriminate killers, such as the pillagers that ravaged Europe during the Thirty Years War, or even the US draft-heavy army of the Vietnam era that fostered the My Lai massacre. Instead, they are professionals responsible for managing violence for the state.[16] It is this characteristic which has been further codified in International Humanitarian Law that should assuage fears that the marriage of AI and the military will result in a future of Skynet, Terminators, and other types of killer robots.

[16] Mick Ryan, "Mastering the Profession of Arms, Part 1: The Enduring Nature," *War on the Rocks*, February 8, 2017, https://warontherocks.com/2017/02/mastering-the-profession-of-arms-part-i-the-enduring-nature/.

SIGNIFICANCE AND BACKGROUND

Defining AI and War Algorithms

The most important concept for understanding AI is the difference between narrow and general AI. Narrow or "weak" AI is a system capable of executing a single task, or a small grouping of related tasks.[17] Narrow AI has made well-publicized progress, such as computers defeating the world's top-rated *Go* or *Chess* players. While those accomplishments are impressive, it is important to remember narrow AI is interacting with a closed system bounded by specific rules. Attempts to have narrow AIs deal with more open systems have met with significant challenges, such as self-driving cars dealing with unpredictable weather and pedestrians. General AI, commonly referred to as artificial general intelligence (AGI), "strong," or "human-level" is considered to be at least fifteen years away.[18] AGI is essentially an AI that is as smart or smarter than the human brain, and has been labeled the most "ambitious scientific quest in human history."[19] The confusion regarding AI goes far beyond these two foundational definitions.

Confusion regarding the exact definition of AI can be tracked back to its origins. In 1955 John McCarthy first utilized the term "artificial intelligence" in a grant proposal with Marvin Minsky to fund a workshop about thinking machines at Dartmouth College. The term was specifically chosen to exclude Norbert Weiner, who had previously coined the term

[17] Melanie Mitchell, *Artificial Intelligence: A Guide for Thinking Humans* (New York: Farrar, Straus and Giroux, 2019), 46.
[18] Mitchell, *Artificial*, 46.
[19] Jeff Clune, "AI-GAs: AI-Generating Algorithms, an Alternate Paradigm for Producing General Artificial Intelligence," *ArXiv:1905.10985*, (January 31, 2020): 1, http://arxiv.org/abs/1905.10985.

"cybernetics" for his own work, but was notoriously difficult to work with.[20] Since that moment, there have been conflicting definitions that have muddied the waters. McCarthy's choice also drew a line between his peers, programmers and practitioners, and others, such as Weiner, who were more concerned about the philosophical and ethical implications of thinking robots.[21] It is this bifurcation that has led to a running concern amongst many experts and the public that AI is being developed in a vacuum outside of ethical considerations. It would not be until 2018 that the US government would attempt to address these issues, as they apply to military and defense applications.

The FY2019 National Defense Authorization Act (NDAA) directed the Department of Defense (DOD) to define AI since no US government definition existed at the time.[22] It is interesting to note the NDAA acknowledged up to five disparate working definitions of AI to use as reference points and potential boundaries. The Act also directed the DOD to stand up a central coordinating authority for oversight and acceleration of the Department's diverse AI and machine learning initiatives.[23] This became the impetus for formalizing the role of the Joint Artificial Intelligence Center (JAIC) in 2018. The JAIC's goal is to implement the DOD's AI strategy of 1) delivering AI-enabled capabilities that address key missions, 2) scaling AI's impact across DOD through a common foundation that enables decentralized development and

---

[20] Kenneth Cukier, "Ready for Robots? How to Think About the Future of AI," *Foreign Affairs* 98, no. 4 (August 2019): 192–98.
[21] Cukier, "Ready," 192
[22] Kelley M Sayler , and Daniel S Hoadley, *Artificial Intelligence and National Security* (Washington, D.C.: U.S. Congressional Research Service, January 30, 2019), 1.
[23] U.S. Congress, House, "John S. McCain National Defense Authorization Act for Fiscal Year 2019," H.R. 5515, 115th Co, January 3, 2018 Pub. L. No. Public Law 115-232 (2018), 1695, https://www.congress.gov/115/bills/hr5515/BILLS-115hr5515enr.pdf.

experimentation, 3) cultivating a leading AI workforce, 4) engaging with commercial, academic, and international allies and partners, and 5) leading in military ethics and AI safety.[24]

Despite this progress, defining AI remains a challenge. The recent Defense Innovation Board (DIB) report on AI ethical principles defines AI as, "a variety of information processing techniques and technologies used to perform a goal-oriented task and the means to reason in the pursuit of that task."[25] This definition, in which the DIB amplified "AI does not equal autonomy," appears to be in harmony with the DOD Strategy's definition of AI: "the ability of machines to perform tasks that normally require human intelligence – for example, recognizing patterns, learning from experience, drawing conclusions, making predictions, or taking action – whether digitally or as the smart software behind autonomous physical systems."[26] This version, pared down from the five original definitions posited in the NDAA, is a step in the right direction, but only at the US government level, as the recent surge of AI-related publications from the commercial sector, academia, intergovernmental organizations (IGOs), non-governmental organizations (NGO), and other countries use a myriad of different definitions.

As with the DOD, other entities have discovered how challenging it can be to define AI. Two of the most recent and compressive assessments of US defense initiatives concede "AI does not have a clear definition"[27] or "overall, devising a good definition of AI is challenging."[28] The ICRC, when reporting on AI, merely defaults to the Oxford Dictionary definition. This trend is

---

[24] "About the JAIC - JAIC," accessed December 31, 2019, https://www.ai.mil/about.html.
[25] "AI Principles: Recommendations on the Ethical Use of Artificial Intelligence by the Department of Defense" (Defense Innovation Board, October 31, 2019), 5, innovation.defense.gov/ai.
[26] "AI Principles: Recommendations on the Ethical Use of Artificial Intelligence by the Department of Defense Supporting Document" (Defense Innovation Board, October 31, 2019), 9–10, innovation.defense.gov/ai.
[27] Tate Nurkin et al., *A Candle in the Dark: US National Security Strategy for Artificial Intelligence*, 2019, 5, https://www.atlanticcouncil.org/wp-content/uploads/2019/12/AC_CandleinDark120419_FINAL.pdf.
[28] *DEPARTMENT OF DEFENSE POSTURE FOR ARTIFICIAL INTELLIGENCE: Assessment.* (S.l.: RAND CORPORATION, 2020), 21–22.

not an indictment of these organizations, but demonstrates how AI is composed of so many specific sub-components that a general definition becomes almost meaningless. For the purpose of this study, the Harvard Law School's Program on International Law and Armed Conflict (HLS PILAC) has successfully demarcated a specific type of AI known as war-algorithms. They define a war-algorithm as "… any algorithm that is expressed in computer code, that is effectuated through a constructed system, and that is capable of operating in relation to armed conflict."[29] This is the definition most apt for this paper as it views computer algorithms, such as narrow-AI, as part of a potentially larger military decision-making system. Additionally, it looks beyond lethal autonomous systems, often feared to be AGI, that monopolize the current AI narrative in the press, and considers other decision-aiding software that may enable combat operations.[30]

Increasing Security Relevance of AI

The 2018 National Defense Strategy states rapid technological advancements are driving changes to the character of warfare, which are having a significant effect on the current and future global security environment.[31] The Department of Defense's list of emerging technologies of concern includes directed energy, biotechnology, robotics, and artificial intelligence.[32] Additionally, the current National Security Strategy acknowledges gene editing, nanotechnology, and artificial intelligence as priority emerging technologies to maintain the

---

[29] Dustin A. Lewis, Gabriella Blum, and Naz K. Modirzadeh, "War-Algorithm Accountability," *SSRN Electronic Journal*, August 2016, vii, https://doi.org/10.2139/ssrn.2832734.
[30] Arthur Michel's explanation of the current fixation about lethal autonomous systems. Arthur Holland Michel, "The Killer Algorithms Nobody's Talking About," *Foreign Policy* (blog), accessed January 21, 2020, http://foreignpolicy.com/2020/01/20/ai-autonomous-weapons-artificial-intelligence-the-killer-algorithms-nobodys-talking-about/.
[31] Jim Mattis, "Summary of the 2018 National Defense Strategy" (Washington, D.C.: Department of Defense, 2018), 3, https://dod.defense.gov/Portals/1/Documents/pubs/2018-National-Defense-Strategy-Summary.pdf.
[32] Mattis, 3.

US's competitive advantage.[33] Unsurprisingly, artificial intelligence (AI) is highlighted by both documents. In fact, the importance of AI is not relegated to the United States. There is a worldwide consensus AI has the potential to modify warfare and the international order.

In 2019 the U.S. Intelligence Community reported the "global race to develop AI… is likely to accelerate the development of highly capable…systems with national security implications."[34] Russia's President, Vladimir Putin, stated "artificial intelligence is the future, not only of Russia, but of all mankind" and "whoever becomes the leader in this sphere will become ruler of the world."[35] Similarly, Chinese President Xi Jinping specified in a 2018 speech to the Politburo that China must "ensure that our country marches in the front ranks where it comes to theoretical research in this important area of AI, and occupies the high ground in critical and AI technologies."[36] The Indian Government's National Institution for Transforming India reports that "AI is poised to disrupt our world" and because India is the "fastest growing economy with the second largest population in the world in the world, has a significant stake in the AI revolution."[37]

These security-focused, realist interpretations of AI are common of many emerging technologies. Key realist thinkers have argued the increasing speed with which technological

---

[33] Executive Office of the President, *National Security Strategy of the United States of America* (United States: White House Office, 2017), 20, https://www.whitehouse.gov/wp-content/uploads/2017/12/NSS-Final-12-18-2017-0905.pdf.
[34] Daniel R. Coats, "Worldwide Threat Assessment of the US Intelligence Community," § Senate Select Committee on Intelligence (2019), 15.
[35] Gregory C. Allen, "Putin and Musk Are Right: Whoever Masters AI Will Run the World," CNN, accessed December 30, 2019, https://www.cnn.com/2017/09/05/opinions/russia-weaponize-ai-opinion-allen/index.html.
[36] Gregory C. Allen, "Understanding China's AI Strategy: Clues to Chinese Strategic Thinking on Artificial Intelligence and National Security" (Washington, D.C.: Center for a New American Security, February 2019), 3, www.cnas.rg.
[37] Anna Roy, "National Strategy for Artificial Intelligence #AIFORALL" (New Delhi, India: National Institution for Transforming India Aayog, June 2018), 5, https://niti.gov.in/writereaddata/files/document_publication/NationalStrategy-for-AI-Discussion-Paper.pdf.

innovations occur will destabilize the future international environment. Throughout history, technological breakthroughs have broken the military power parity between nations, quickly destabilizing their relationship.[38] This uneven growth of power is a key instigator in conflict between states as they try to change the international status quo.[39] In the past, key technologies diffused to other states in a relatively short amount of time to ensure one nation did not have a long-term advantage over potential adversaries.[40] However, the potential revolutionary nature of AI resulted in states prioritizing the advancement of this key technology due to its perceived importance to their national defense. While many nation states are looking towards AI as a source of potential power, a number of international and non-governmental organizations have expressed concerns regarding this emerging technology.

International and non-governmental organizations across the world tend to focus on two themes regarding AI. First, they recommend uses to improve human life. Second, they often express concern regarding the ethical use of AI in warfare, to include lethal autonomous weapon systems (LAWS). The UN's AI for Good Global Summit is an initiative with the goal of constructing a common understanding of AI technologies and capabilities.[41] Currently 36 UN agencies actively engage in leveraging AI into their operations. These efforts range from the UN Institute for Disarmament Research, which is focused on human control and responsibility for AI-enabled weapons, to the UN Offices for Disaster Risk Reduction, which endeavors to

---

[38] Kenneth Neal Waltz, *Theory of International Politics* (Long Grove, Ill.: Waveland Press, 2010), 179.

[39] Robert Gilpin, *War and Change in World Politics* (Cambridge; New York: Cambridge University Press, 1981), 93.

[40] Marshall G. S Hodgson and Edmund Burke, *Rethinking World History: Essays on Europe, Islam, and World History* (Cambridge [England]; New York, NY, USA: Cambridge University Press, 1993), 71.

[41] "United Nations Activities on Artificial Intelligence (AI) 2019" (Geneva, Switzerland: International Telecommunication Union, 2019), v, https://www.itu.int/dms_pub/itu-s/opb/gen/S-GEN-UNACT-2019-1-PDF-E.pdf.

leverage AI to enable better risk assessments for sustainable development.[42] The International Committee of the Red Cross acknowledges it has two primary interests in AI, both of which are focused on military applications: 1) the use of AI in the conduct of war/violence and 2) AI's use to protect victims of armed conflicts.[43] To this end they have advocated for greater international discussions on the potential role of AI in warfare. It is not surprising then, that the list of the top ten most read articles on the Humanitarian Law and Policy Blog in 2019 featured AI in warfare as the topic for the top three spots.[44]

The NGO Human Rights Watch is leading a group of 100 NGOs and 28 countries[45] in the Campaign to Stop Killer Robots with the goal to ban AI-driven weapons that lack meaningful human control.[46] This effort, along with those of the UN and ICRC, provide liberalist and constructivist counternarratives to the realist interpretation by showing how international relations can be shaped by international law and that ideas stemming from collective morals, ethics, and cultures ultimately affect how states interact with each other.[47] However, as this paper will demonstrate, AI-enabled weapons have the potential to increase the capability of the world's militaries to execute their operations more ethically and in accordance with IHL.

---

[42] "United Nations Activities on Artificial Intelligence (AI) 2019," 54 & 58.

[43] "Artificial Intelligence and Machine Learning in Armed Conflict: A Human-Centered Approach" (Geneva, Switzerland: ICRC, June 5, 2019), 1, https://www.icrc.org/en/document/artificial-intelligence-and-machine-learning-armed-conflict-human-centred-approach.

[44] "The Most Read Blog Posts in 2019 Archives," Humanitarian Law & Policy Blog, accessed January 21, 2020, https://blogs.icrc.org/law-and-policy/category/type-of-post/the-most-read-blog-posts-in-2019/.

[45] The U.S. and Russia are not aligned with this movement. China had called for a ban on autonomous weapons, but not their development or production.

[46] "The Campaign to Stop Killer Robots," accessed December 30, 2019, https://www.stopkillerrobots.org/about/.

[47] Jack Snyder, "One World, Rival Theories," *Foreign Policy*, no. 145 (2004): 56, 59, https://doi.org/10.2307/4152944.

ARTIFICIAL INTELLIGENCE AND DISTINCTION

In 2015, a U.S. AC-130 gunship mistakenly engaged a Doctors Without Borders' hospital building in Afghanistan while attempting to engage a number of insurgents. The U.S. Central Command's formal inquiry included:

> *…the personnel involved did not know that they were striking a medical facility. The intended target was an insurgent-controlled site which was approximately 400 meters away from the MSF Trauma Center. The investigation found that an AC-130U Gunship aircrew, in support of a U.S. Special Forces element that was supporting a partnered Afghan ground force, misidentified and struck the MSF Trauma Center. The investigation determined that all members of both the ground force and the AC-130U aircrew were unaware the aircrew was firing on a medical facility throughout the engagement.*[48]

Even more recently, in 2020, Iranian Air Defense operators engaged and destroyed a Ukrainian civilian airliner departing Tehran. While Iran initially blamed the incident on "technical failures," they eventually admitted their personnel incorrectly identified the civilian aircraft for a military missile.[49] These regrettable events in addition to those reported by Lucas Kunce during his deployments in Iraq, while unintentional, were admittedly failures to meet the principle of Distinction.

Article 48 of Additional Protocol I of the Geneva Conventions defines the requirement to distinguish between military and civilian personnel as well as military and civilian objects. It

[48] "April 29: CENTCOM Releases Investigation into Airstrike on Doctors Without Borders," U.S. Central Command, accessed February 8, 2020, https://www.centcom.mil/MEDIA/PRESS-RELEASES/Press-Release-View/Article/904574/april-29-centcom-releases-investigation-into-airstrike-on-doctors-without-borde/.

[49] "UN Aviation Experts to Join Ukraine Airlines Iran Crash Investigation," UN News, January 14, 2020, https://news.un.org/en/story/2020/01/1055312.

clearly states that parties involved in armed conflict must at "all times distinguish between the civilian population and combatants and between civilian objects and military objectives and accordingly shall direct their operations only against military objectives."[50] Additionally, Article 57 states parties are required to do "everything feasible to verify" the military nature of a target and take "constant care" to avoid civilians and their objects.[51] U.S. military legal documents further amplify the importance of Distinction, often referred to as discrimination, by describing its offensive and defensive components. The former assures that civilians and protected objects are not the subject of attack, while the latter requires warring parties to adequately take measures to physically separate as well as mark or identify combatant and noncombatant personnel and objects."[52,53] Despite these clear guidelines, failures in distinction continue to occur due to the nature of war.

Military theorist Carl von Clausewitz's *On War* describes how the nature of war undermines the ability of human beings to complete what are essentially clear tasks. His oft-quoted maxim, "Everything in war is very simple, but the simplest thing is difficult"[54] clearly applies to Distinction. The difficulty stems from war's complexity and is exacerbated by the presence of extreme danger, physical, mental, and moral challenges, as well as a high degree of uncertainty which prevents many military decision-makers from consistently making rational,

---

[50] "Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of International Armed Conflicts (Protocol I), of 8 June 1977" (International Committee of the Red Cross, June 8, 1977), chap. I Art 48, https://ihl-databases.icrc.org/applic/ihl/ihl.nsf/vwTreaties1949.xsp.
[51] "Protocol II," chap. IV Art 57.
[52] Dustin Kouba, ed., *Operational Law Handbook*, 17th ed. (Charlottesville, VA: The Judge Advocate General's Legal Center and School, 2017), 9–10.
[53] *Department of Defense Law of War Manual* (Washington, D.C.: Office of General Counsel, Department of Defense, 2016), 62–63, https://tjaglcspublic.army.mil/dod-low-manual.
[54] Carl von Clausewitz et al., *On War* (Princeton, N.J.: Princeton University Press, 1984), 119.

value-maximizing decisions. These characteristics are commonly referred to as fog and friction.[55] While fog and friction can degrade the ability of military forces to distinguish between combatants and noncombatants, it does not remove the moral obligation to make every effort to increase the level of success in this endeavor. Therefore, before evaluating how AI can mitigate the effects of fog and friction and better meet the intent of Distinction it is important to review the ethical evolutions that led to today's understanding of this principle.

The Ethical Foundations of Distinction

When laws do not exist, it's necessary to leverage ethics to build a foundation.[56] This maxim rings true with moral warfighting. The concepts of a just or moral war date back to ancient and medieval thinking and have evolved from regionally and religiously diverse sources to a broader consensus of what is right and wrong in war. While these initial ideas were originally based on pure definitions of right and wrong, they evolved over time into a more pragmatic set of applicable customs.[57] This evolution was critical to shape the current set of customary laws, to include Distinction, that shape our understanding of just war thinking today.

Fifteenth century French writer Christine de Pizan's work, *The Book of Deeds of Arms and of Chivalry*, summarizes the military code of honor of her era. Her writing was significantly shaped by the late fourth century Roman military theorist Vegetius, as well as the works of fourteenth century Italian jurist John of Leganano and French scholar Honr'e Bonet. Significantly, her work's target audience was not just other scholars but the military practitioners

---

[55] Clausewitz et al., 101.
[56] George R. Lucas, *Ethics and Cyber Warfare: The Quest for Responsible Security in the Age of Digital Warfare* (New York, NY: Oxford University Press, 2017), 40.
[57] David Rodin and Henry Shue, eds., *Just and Unjust Warriors: The Moral and Legal Status of Soldiers* (New York: Oxford University Press - Books Distribution Services, 2010), 19.

of the time, namely the nobility.[58] The concept of Distinction between combatants and noncombatants was a continuing theme she reinforces. She clearly differentiates between members of the profession of arms and "common people, that is, peasants, shepherds, and such…" who deserve protection similar to "priests and churchmen."[59] Her acknowledgement that harm may potentially befall noncombatants is very nuanced and Distinction is an imperfect science. For example, she voices that noncombatants providing material support to the war effort are vulnerable to have their possessions confiscated versus those who did not support the war. She also warns the uncertainty inherent in war will undoubtedly result in injury to some noncombatants, as "for weeds cannon be separated from good plants, because they are so close together that the good ones suffer."[60]

In the sixteenth century, Dominican Priest Francisco di Vitoria's work *Concerning the Laws of War* responded to Spanish Conquistador tactics in the New World. While he agreed the Spanish had the right to colonize and convert the indigenous peoples to Christianity, he protested their repeated actions of killing or capturing the nonmilitary population was immoral.[61] The Principle of Distinction is not just limited to Christian ideology. Historical works from other religions have also espoused, in parallel, the need to practice Distinction.

The great Hindu epic the *Mahabharata,* written around 200 BCE, provides a glimpse into their early beliefs regarding Distinction as a law of war. For example, warriors must not kill aged men, women of all ages, nor children. Also, a "…peaceful citizen walking along the road,"

---

58 Gregory M. Reichberg, Henrik Syse, and Endre Begby, eds., *The Ethics of War: Classic and Contemporary Readings* (Malden, MA; Oxford: Blackwell Pub, 2006), 211.
59 Reichberg, Syse, and Begby, 221–22.
60 Reichberg, Syse, and Begby, 222.
61 George R. Lucas, *Military Ethics: What Everyone Needs to Know* (New York, NY: Oxford University Press, 2016), 87–88.

nor "spectators should not be killed." Additionally, it acknowledges the existence of some protected personnel amongst the adversary army, such as "camp followers, war musicians, and gate guards."[62] The *Mahabharata* also identifies that non-military structures, such as "gardens, temples and other places of public worship should be left unmolested." [63]

Distinction in Islam has been primarily divided between how to treat Muslims and non-Muslims. More specifically, noncombat immunity is not a universal rule, but a means to protect those deemed "innocent" under Sharia law.[64] Ninth century Sunni scholar Muhammad ibn al-Hasan al Shaybani reiterates Muhammad's directives regarding warfare. His armies were ordered to not kill women and males who have not reached puberty should not be killed, though adult males, even noncombatants, could be slain. This mandate to not kill does not always equate to no injury. Twelfth century philosopher Ibn Rushd suggested three forms of damage that an enemy may suffer: personal, property, and liberty. Therefore, women and children who were taken as booty can be considered the victims of a form of injury.[65] Using religion as a discriminator is not a purely Islamic ideal as medieval scholars considered it morally allowable to treat Christians and Muslims (i.e. Saracens) differently. Dutch scholar Hugo Grotius, however, was able to lay out an argument that disaggregated the reason for a war from the conduct of war and contended differences in religion should not affect how noncombatants are treated.[66]

---

[62] Surya P. Subedi, "The Concept in Hinduism of 'Just War,'" *Journal of Conflict & Security Law* 8, no. 2 (2003): 355, www.jstor.org/stable/26294280.
[63] Subedi, 356.
[64] Hodgson and Burke, *Rethinking World History*, 59.
[65] John Kelsay, *Islam and War: A Study in Comparative Ethics*, 1st ed (Louisville, Ky: Westminster/John Knox Press, 1993), 62–63.
[66] Rodin and Shue, *Just and Unjust Warriors: The Moral and Legal Status of Soldiers*, 179–80.

Over time, these various interpretations of Distinction have evolved into what has become today's customary law. Distinction's ethical foundation continues to undergo debate among scholars. These debates are framed within the context of just war thinking which coalesced around Michael Walzer's 1977 *Just and Unjust Wars,* which responded to the multiple wars of the twentieth century that led to the deaths of millions of noncombatants.

Distinction and Current Just War Thinking

Walzer describes a standing war convention that describes moral and legal reasons for going to war (*jus ad bellum*) and how combatants can morally act in war (*jus in bello*). This convention is designed for practical application that acknowledges warriors engaged in combat lack the capability to do a thorough ethical analysis of each target they engage. The two major principles of the war convention deal specifically with Distinction. The first is "…that once war has begun, soldiers are subject to attack at any time (unless they are wounded or captured page)."[67] The second states "…that noncombatants cannot be attacked at any time."[68] He does indicate the line between combatants and noncombatants is less clear than the line that divides uniformed military members from all civilians.

Walzer concedes the difficulty of labeling civilians as combatants, but explains how current militaries are dependent on logistics and technologies supplied by civilians.[69] While individual civilians who contribute to the war effort may be difficult to distinguish, buildings are not. Destroying facilities that provide a military advantage to the enemy are acknowledged as

[67] Walzer, *Just and Unjust Wars*, 138.
[68] Walzer, 151.
[69] Walzer, 145.

morally allowable targets.[70] Walzer does distinguish between facilities whose sole function is military related as opposed to those that support the military and the general population. A tank factory would be targetable while a food processing plant may not.[71] While Walzer's writings generally agree with current customary law, some scholars argue the convention, as he describes it, does not support the most ethically allowable forms of warfare.

Just war revisionist scholar Jeff McMahon has led the charge that the current laws of war and Walzer's war convention are not the most moral.[72] His argument focuses on two key points related to the moral equivalency of combatants. He maintains it is morally preferable to target combatants belonging to a state that began a war for unjust reasons, for example aggression, than it is to target those defending against the invaders, who are relatively morally innocent.[73] In a similar vein, he argues it may be morally correct to injure civilians if they are more culpable for the cause of an unjust war or if they significantly contribute to the warfighting capability of their side.[74] His viewpoint, while logically sound, does not appear to be practical for implementation during the fog and friction of actual combat. However, AI has the potential to mitigate this fog and friction and allow for more ethical targeting.

Distinction in Practice with AI

The two primary ways AI can better enable military targeting operations to meet the ethical and legal standards of Distinction are through assisting in positive identification (PID) and expanding overall battlespace awareness. The earlier vignettes of the MSF hospital and the

---

[70] Helen Frowe, *The Ethics of War and Peace: An Introduction*, Second edition, The Ethics Of (London: Routledge/Taylor & Francis Group, 2016), 109.
[71] Walzer, *Just and Unjust Wars*, 146.
[72] Rodin and Shue, *Just and Unjust Warriors: The Moral and Legal Status of Soldiers*, 97.
[73] Rodin and Shue, 162.
[74] Rodin and Shue, 162.

Ukrainian airliner both share similar root causes, human error. The investigation of the hospital engagement revealed that both the ground force commander requesting fire and the AC-130 crew confused the hospital with another compound with "an outer perimeter wall, with multiple buildings inside of it."[75] Similarly, the Iranian military initially blamed the airliner shootdown on a technical failure before admitting their human operator mistook the plane for an incoming missile[76] because it appeared to have "the flying posture and altitude of an enemy target" [77]

It is important to note these errors were exacerbated by the perceived need to act quickly under the stress of a lethal threat that did not exist at the time. Algorithms can mitigate these types of stress-initiated mistakes in warfare. At this time, the best systems to achieve this goal are known as *human in the loop.* These systems allow an algorithm (narrow AI) to sense and define the environment and possibly recommend actions but require a human to make the definite decision to act. This is in contrast to *human on the loop* systems, which allow algorithms to make decisions while a human being monitors, and can override, its processes. Even more reliant on AI are *human out of the loop systems*, which are fully autonomous with humans only monitoring the actions/outputs of the system.[78] Since life-and-death decisions require the utmost care the best way to address the intent of the principle of Distinction through improved PID and battlespace awareness is via *human in the loop* systems. Other systems do not have future value as they evolve, but there is a high level of potential ethical risk inherent in

---

75 "U.S. Department of Defense Releases Report of Investigation Finding That October 2015 Air Strike on Doctors Without Borders Hospital in Kunduz, Afghanistan, Was Not a War Crime," *American Journal of International Law* 110, no. 3 (2016): 582, https://doi.org/10.1017/S0002930000054038.
76 "UN Aviation Experts to Join Ukraine Airlines Iran Crash Investigation."
77 Farnaz Fassihi, "Iran Says It Unintentionally Shot Down Ukrainian Airliner," *The New York Times*, January 10, 2020, sec. World, https://www.nytimes.com/2020/01/10/world/middleeast/missile-iran-plane-crash.html.
78 Paul Scharre, *Army of None: Autonomous Weapons and the Future of War*, First edition (New York; London: W. W. Norton & Company, 2018), 29.

these systems, to include technical safety, algorithmic bias, and malicious/unintended use concerns.[79] Therefore human-machine teaming, or more precisely warriors and war-algorithm teaming, is the optimal course.

Positive Identification (PID) and AI

The Department of Defense's (DoD) primary form of Distinction is through PID. PID is defined as "the reasonable certainty that a functionally and geospatially defined object of attack is a legitimate military target in accordance with the Law of War and applicable ROE."[80] More simply, it answers the Who?, What?, and Where? questions regarding an entity. PID is also acknowledged as the foundational consideration in the DoD's collateral damage methodology, which its Proportionality assessments are based on. It states that an assessment begins with the question, "Can I PID the object I want to affect?"[81] While PID can technically be achieved through a combination of different intelligence sources, invariably a visual component is required in the final analysis. Therefore, the most commonly known algorithms associated with PID are related to imagery interpretation. Project MAVEN represents this type of warrior/war-algorithm partnership and the level of scrutiny applied to them.

MAVEN and other programs like it leverage Computer Vision to detect and identify objects in imagery. Computer Vision is type of AI that allows computers to "see" the world and relies on machine learning to refine its accuracy over time. For example, one system for Google's self-driving cars decreased the number of human corrections it required from every 700

---

[79] Brian T. Molloy, "Project Governance for Defense Applications of AI: An Ethics-Based Approach" (Newport, RI, Naval War College, 2020), 13.
[80] "Chairman of the Joint Chiefs of Staff Instruction: No-Strike and the Collateral Damage Estimation Methodology" (Department of Defense, February 13, 2009), A-6, https://www.aclu.org/files/dronefoia/dod/drone_dod_3160_01.pdf.
[81] "CJCSI 3160.01," D-A-7.

miles to every 5,000 miles over a period of 12 months. Similarly, another system was able to detect malignant skin lesions at the same level as 21 dermatologists.[82] Object recognition accuracy has also increased as processor speeds and machine learning networks increase in capability. The winning algorithm of the 2010 ImageNet Large Scale Visual Recognition Challenge had an accuracy of 72% while in 2017 the winning algorithm had an accuracy of 98%.[83]

Air Force Lieutenant General Jack Shanahan, Director of the Joint Artificial Intelligence Center (JAIC), acknowledged that military uses of Computer Vision would have a learning curve when deployed in different theaters. He stated, "once you deploy it to a real location, it's flying against a different environment than it was trained on….Still works of course … but it's just different enough in this location, say that there's more scrub brush or there's fewer buildings or there's animals running around that we hadn't seen in certain videos. That is why it's so important in the first five days of a real-world deployment to optimize or refine the algorithm."[84] This also highlights the importance of pairing the algorithm with an experienced analyst.

In 2017 imagery analysts leveraged their expertise to refine target identification algorithms in the Middle East which led to an increase of accuracy from 60% to 80%.[85] Far from deterring deployed units from employing the system, the 80% success rate created an

---

[82] Serena Yeung et al., "Bedside Computer Vision — Moving Artificial Intelligence from Driver Assistance to Patient Safety," *The New England Journal of Medicine* 378, no. 14 (2018): 1271, https://doi.org/10.1056/NEJMp1716891.
[83] Mitchell, *Artificial Intelligence*, 86–90.
[84] "The Pentagon's New Artificial Intelligence Is Already Hunting Terrorists - Defense One," accessed April 21, 2020, https://www.defenseone.com/technology/2017/12/pentagons-new-artificial-intelligence-already-hunting-terrorists/144742/.
[85] "Lt. Gen. Jack Shanahan Media Briefing on A.I.-Related Initiatives With," U.S. DEPARTMENT OF DEFENSE, accessed April 21, 2020, https://www.defense.gov/Newsroom/Transcripts/Transcript/Article/1949362/lt-gen-jack-shanahan-media-briefing-on-ai-related-initiatives-within-the-depart/.

increased demand, especially from the US Special Operations Command and the elite Joint Special Operations Command, both of which employ some of the most experienced imagery analysts in the intelligence community.[86] This highlights the military's understanding of the importance of algorithms augmenting human PID skills, but also that their efforts will improve future algorithms. This is especially true when algorithms can be used to study full-motion video streams on a 24/7 basis.

Leveraging Computer Vision to monitor video is often interpreted as a means to relieve imagery analysts from a duty that is often perceived as dull or repetitive. However, leveraging AI to monitor video has a significant advantage. In the book *Thinking, Fast and Slow*, Daniel Kahneman defines intelligence as "not only ability to reason, it is also the ability to find relevant material in memory and to deploy attention when needed."[87] It is no surprise a narrow AI would struggle to perfectly identify potential targets in a complex environment while a human analyst has the better capability to reason. A computer algorithm, however, does have a significant ability with regard to memory, to include accuracy and speed, as well as an increased capability to maintain focus over time.

At this point the advantages of pairing warrior and war-algorithms will increase the ability to adhere to the intent of Distinction through more accurate and timelier PID. This capability also applies to other potentially life-saving situations under IHL. The same algorithms that detect targets can also support the Geneva Convention's requirement to "…take all possible measures to search for and collect the shipwrecked, wounded and sick" after every naval

[86] "The Pentagon's New Artificial Intelligence Is Already Hunting Terrorists - Defense One."
[87] Daniel Kahneman, *Thinking, Fast and Slow*, 1st pbk. ed (New York: Farrar, Straus and Giroux, 2013), 46.

engagement.[88] It may also aid in the ability "for teams to search for, identify and recover the dead from battlefield areas."[89] These same identification algorithms have already been leveraged to support humanitarian efforts. In 2019, the JAIC partnered with the Indiana National Guard to test AI systems that would identify objects in support of floods and forest fires.[90] The AI reportedly provides analysts all the information to make assessments within a couple of minutes compared to the many man-hours, days, or weeks it would normally take.[91]

Some allowances in IHL may cause confusion with algorithms that are too narrowly focused, for example one that was designed to identify individuals carrying weapons. Article 22 of the Geneva Convention I allows for the protection of medical units. This includes armed orderlies or the creation of a defensive picket with sentries.[92] Article 19 of Geneva Convention IV also warns that a protected medical facility does not lose its protection due to the presence of small arms or ammunition stocks introduced by recently admitted patients.[93] It is possible that these physical signatures, which are also common to legitimate military targets, could result in a mis-identification of a medical facility. However, human teaming with increased battlespace awareness, to be discussed in the next section, should mitigate this possibility.

Another possible consequence that must be addressed is negative effects of human-machine teaming. Previous studies of *human out of the loop* systems have identified the potential

---

88 "Geneva Convention Relative to the Protection of Civilian Persons in Time of War of 12 August 1949" (International Committee of the Red Cross, August 12, 1949), secs. II-Article22, https://ihl-databases.icrc.org/applic/ihl/ihl.nsf/vwTreaties1949.xsp.
89 "Protocol II," secs. 1-Article 33.
90 "Intel Airmen Sharpen AI Technology for Domestic Response," Air National Guard, accessed April 21, 2020, https://www.ang.af.mil/Media/Article-Display/Article/2025332/intel-airmen-sharpen-ai-technology-for-domestic-response/.
91 "Intel Airmen Sharpen AI Technology for Domestic Response."
92 "Geneva Conventions," I-Article 22.
93 "Geneva Conventions," IV-Article 19.

for the human operators to suffer from a declining sense of agency to include "decrements in vigilance such as reduced sensitivity to important signals, complacent or excessive trust in system ability."[94] Possibly, this loss of agency could affect operators in a *human in the loop* system as well. After building a level of confidence with a specific algorithm, an analyst could fall prey to confirmation bias instead of critically assessing the recommendations of the system. Human-machine teaming could also atrophy the original skills of the analysts as they become more accustomed to automated processes. This could result in degraded performance if the automated systems are unavailable in a contested environment. One of the additional factors that led to the Kunduz hospital incident was the fact the e-mail system of the AC-130 was non-operational, forcing the crew to rely on verbal communications with the ground force.[95] Acknowledging the risks inherent in human-machine teaming and appropriately training analysts and their leadership on proper techniques to prevent these issues are critical steps in assuring decision-makers can continue to make ethical decisions, even when their technology is degraded.

The final advantage that PID algorithms provide is they allow and enable highly trained human analysts to apply their judgment and experience on more complex problems beyond PID. Algorithms excel at sorting through high-volumes of data (i.e. images) and can highlight areas of unusual activity that analysts can provide context. The best examples of this are coherent change detection algorithms that compare images of the same location taken at different times. The algorithms then automatically flag areas with the most activity for a human analyst to review and assess the potential causes and effects of the activity. The US Air Force's most recent

[94] B. Berberian, "Man-Machine Teaming: A Problem of Agency," *IFAC PapersOnLine* 51, no. 34 (2019): 119, https://doi.org/10.1016/j.ifacol.2019.01.049.
[95] "U.S. Department of Defense Releases Report of Investigation Finding That October 2015 Air Strike on Doctors Without Borders Hospital in Kunduz, Afghanistan, Was Not a War Crime," 581.

Intelligence Surveillance, and Reconnaissance Flight Plan explains how AI algorithms can free analysts from linear, industrial-age processes, commonly known as Processing Exploitation and Dissemination (PED) to a "…sensing, identifying, attributing, and sharing (SIAS)-based, multi-domain, multi-intelligence, prototyping, rapidly-experimenting culture."[96] This shift will improve the military's ability to better foresee adversary and neutral party actions. This may lead to greater battlespace awareness, which allows for greater Distinction as the military will not only be able to PID a specific entity but will also better highlight additional targets, protected facilities, and personnel during a conflict.

Battlespace Awareness and AI

As mentioned previously, Distinction, especially in the context of PID, is often associated with visual signatures. There are other algorithms that can leverage other sources of intelligence to create a better understanding of the greater environment instead of a single target. This concept is battlespace awareness. New war algorithm-based systems, such as DIA's Machine-Assisted Analytic Rapid-Repository System (MARS) are constantly improving the military's ability to process bulk data from disparate sources and build a foundational understanding of adversary, friendly, and neutral entities to allow for more lawful and moral battlefield engagements. Simply put, successful systems have evolved from a single "kill chain" focused against a single target into a greater "kill web" that defines targets and their relationships with other objects in the battlespace.[97]

---

[96] Lt Gen Dash Jamieson, "Next Generation ISR Dominance Flight Plan 2018-2028" (Headquarters Air Force, 2018), https://www.airuniversity.af.edu/ISR/#isr-flight-plan.

[97] The "Kill Web" concept was first introduced to me in discussions with Dr. Timothy Schultz, Naval War College, Newport, RI

One common advantage that is attributed to *human out of the loop* systems, particularly lethal autonomous weapons, is they provide a tactical advantage by allowing one side to see their target and shoot first. However, this is a narrow, if not sensationalized, view of the true advantage war algorithms can provide. While decision-making speed is important to tactical success, other significant variables lead to success at the operational and strategic levels of warfare. The fixation that rapid decisions are the primary driver for successful military operations is a common oversimplification of John Boyd's Orient-Observe-Decide-Act (OODA) model. What is often overlooked is his argument that information that enables truly meaningful actions is just as important, as is the need for accurate foundational information that enables sound orientation to measure against new observations.[98] Algorithms that create foundational intelligence and allow forces to engage military targets of significant strategic value are an oft-overlooked but critical component to more effective and moral warfighting.

Early databases designed to provide battlespace awareness, such as DoD's Modernized Integrated Database (MIDB), were driven by manual inputs from analysts primarily using a limited number of vetted and classified sources. This produced information that was usually of high accuracy, but only for a relatively limited number of targets. Frustrations with the timeliness of the data due to the ponderous data input process and stovepipe nature of the quality control and review system eventually precipitated a counter revolution in 2016. The result: Terminator.

Terminator was a program created jointly between DoD programmers and USAF analysts to turn the massive amounts of intelligence collected in Iraq and Syria into a useful decision-

[98] Frans P. B Osinga, *Science, Strategy and War: The Strategic Theory of John Boyd* (London; New York: Routledge, 2007), 234–36.

making aid in the fight against the Islamic State of Iraq and Syria (ISIS). Like MIDB, this system relied primarily on human input, but its user-friendly interface allowed for potential targets to be crowd-sourced as did its ability to ingest potential targets from other static databases. This led to the synthesis of over 3,000 points of interest in the initial 96 hours of Terminator's existence.[99] Crowd-sourcing the data from hundreds of analysts across the world identified more entities across the battlespace than found in MIDB, and in a timelier manner. The confidence-level for each entity could vary significantly from "confirmed" to "possible" depending on the number and quality of the sources used, which required additional manual analysis before an analyst could develop the point of interest into a target.[100] However, even concentrations of many low-confidence datapoints proved useful in highlighting areas that required more attention from intelligence collection platforms. Despite its short-term success against (ISIS) and later in Afghanistan, Terminator's manually-intensive processes were not optimized for dealing with larger, more complex datasets across the world and turning them into the foundational intelligence decision-makers require. This mission is the purview of MARS.

DIA's Machine-Assisted Analytic Rapid-Repository System (MARS) is designed to support rigorous analytical standards while maintaining flexibility by being "a system capable of ingesting and managing large volumes of it [data], and making it available to both humans and machines."[101] This flexibility will include its ability to ingest data that historically was not utilized at a large scale by the military, such as publicly available information (PAI), such as

---

[99] "Self-Developed Intelligence Program Strengthens Coalition Fight Against," U.S. Air Force, accessed April 23, 2020, https://www.af.mil/News/Article-Display/Article/1078078/self-developed-intelligence-program-strengthens-coalition-fight-against-isil/.

[100] "Defense Intelligence Agency Strategic Approach" (Defense Intelligence Agency, September 2018), 11, https://www.dia.mil/Portals/27/Documents/About/DIA_Strategic_Approach.pdf.

[101] "DIA's Vision of MARS: Decision Advantage for the 21st Century," Defense Intelligence Agency, accessed August 5, 2019, https://www.dia.mil/News/Articles/Article-View/Article/1855910/dias-vision-of-mars-decision-advantage-for-the-21st-century/.

social media. It will convert the data into foundational intelligence that creates a virtual model of the world that will signifantly increase "warfighter's ability to mitigate risks and defeat adversaries." Most importantly, MARS is designed to use algorithms to increase the military's ability to foresee adversary actions. This will be accomplished by enabling "the simulation of courses of action, allowing operators to quickly and fully grasp the likely effects of proposed activities or movements."[102]

The use of algorithms to ingest PAI then organize and correlate it into useable intelligence is seen as the key to the future of systems like MARS. The sheer volume of PAI compared to classical and exquisite intelligence sources is staggering (Figure 1). For example, DATAMINR, a private company, reports its own algorithms integrate over 10,000 different data sources, to include text, graphics, videos, and audio in multiple languages.[103] PAI has become so important to military intelligence processes that Lieutenant General VeraLinn "Dash" Jamieson, Deputy Chief of Staff for ISR, recently stated that PAI "will constitute between 60 – 80 percent of our intelligence assessments at the speed and scale of modern warfare."[104]

---

[102] "DIA's Vision of MARS."

[103] Jason Wilcox, "The Multi-Dimensional Value of Public Twitter Data for Real-Time Event Detection," accessed April 21, 2020, https://www.dataminr.com/blog/the-multi-dimensional-value-of-public-twitter-data-for-real-time-event-detection.

[104] Over The Horizon, "Publicly Available Information: The Secret to Unclassified Data, Part I," OTH, April 8, 2019, https://othjournal.com/2019/04/08/publicly-available-information-the-secret-to-unclassified-data-part-i/.

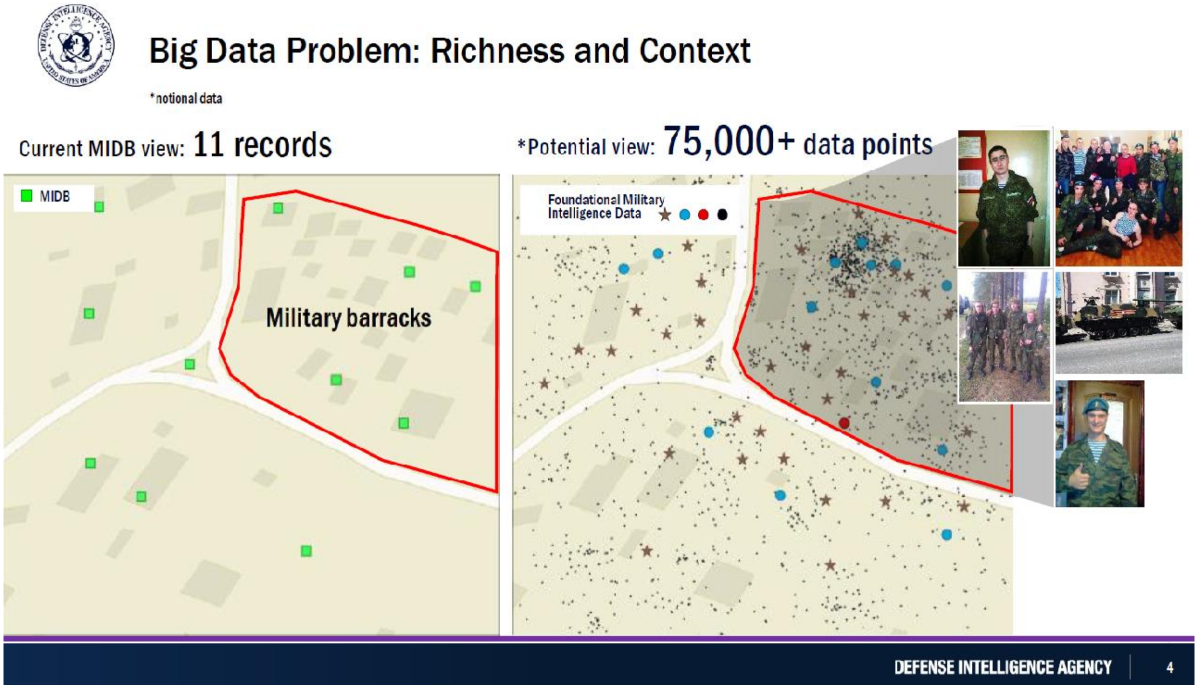

Figure 1. Notional Representation of MARS Battlespace Awareness Capability with PAI[105]

Dangers of Leveraging Big Data for Battlespace Awareness

Lt Gen Shanahan, director of the JAIC, has gone on the record to state one of his biggest concerns regarding AI and warfare is the accuracy of the data. "I treat it as mineral ore: there's a lot of crap. You have to filter out the impurities from the raw material to get the gold nuggets."[106] Therefore it is crucial that intelligence agencies maintain tradecraft standards and methodologies to grade the reliability and accessibility of these new sources of data just as they do with classical clandestine sources.

105 "Machine Assisted Analytic Rapid Repository System (MARS)" (Defense Intelligence Agency, June 21, 2019), https://nova.afceachapters.org/file/31587/download?token=43UDQrZf.

106 Theresa Hitchens, "EXCLUSIVE Pentagon's AI Problem Is 'Dirty' Data: Lt. Gen. Shanahan," *Breaking Defense* (blog), accessed April 21, 2020, https://breakingdefense.com/2019/11/exclusive-pentagons-ai-problem-is-dirty-data-lt-gen-shanahan/.

There is also the potential for data to create biases. If there are a higher number of data sources in one specific region compared to another, it may appear that the area with more reporting has a higher concentration of adversary activity. In reality, the adversary may be practicing better operational security and signature management in the areas of lesser activity. Additionally, data sources, especially the media, can have reporting biases. Clearly Iraqi, Syrian, Iranian, Russian, and Israeli news sources have all described the war against ISIS in radically different ways.

Finally issues of privacy will likely arise with the use of PAI. Social media apps do not have terms and conditions that state one's data may be used to kill a combatant. The ethical debates over whether the benefit of potentially saving lives in combat are equal to people's privacy are expected to continue similarly to current debates over government surveillance.

Distinction and Foreseeability – the Key to Proportionality

In summary, the true promise of AI is that it can increase ethical warfighting by providing ability to PID targets tactically and operate proactively at the strategic and operational levels through battlespace awareness. During the 2014-2017 air campaign against ISIS, the US-led Coalition delivered over 110,000 munitions and over 90% of strikes were against targets that were not completely pre-planned.[107] The other 10% were against targets that could have been under evaluation for several weeks.[108] Most of the dynamic strikes were in defense of local ground forces fighting ISIS. Reactive strikes in self-defense have the potential to cause injury to

[107] "United States' Airstrike Rules of Engagement Reviewed," *AOAV* (blog), April 29, 2019, https://aoav.org.uk/2019/an-investigation-into-the-united-states-rules-of-engagement-governing-the-use-of-airstrikes/.
[108] "Department of Defense Press Briefing by General Harrigian via Teleconf," U.S. DEPARTMENT OF DEFENSE, accessed April 21, 2020, https://www.defense.gov/Newsroom/Transcripts/Transcript/Article/1193060/department-of-defense-press-briefing-by-general-harrigian-via-teleconference/.

noncombatants and their property because the Law of War "does not limit a commander's inherent right of self-defense."[109]

By using systems such as MARS to gain a more comprehensive understanding of an adversary's military warfighting capabilities there is a greater chance to significantly minimize the amount of injury noncombatants are exposed to. This would be accomplished by identifying key adversary centers of gravity, whose destruction would provide a significant military advantage, and engaging them at a specific time and place that would limit collateral damage. Additionally, a large living database of adversary combatant and noncombatant entities would also allow for greater care to be taken when dynamic, short-notice, strikes are required. How this can occur is the topic of discussion for the next section.

[109] "CJCSI 3160.01," D-3.

ARTIFICIAL INTELLIGENCE AND PROPORTIONALITY

The formidable challenges of meeting the Principle of Distinction in combat pale in comparison to assessing whether an action meets the Principle of Proportionality. Whereas Distinction relies on an individual to correctly recognize the combatant status of a person or object, Proportionality relies on the synthesis of less concrete concepts, such as human judgment, values, and foresight. Unlike Distinction, Proportionality also exists as both a *jus ad bellum* and a *jus in bello* consideration. This analysis primarily concerns itself with *jus in bello* applications.

From a *jus ad bellum* perspective, Proportionality is often considered the most difficult of the seven conditions required to commence a just war. This is due to the difficulty in predicting the potential cumulative damage a war would cause against a less tangible moral good, such as self-determination.[110] Therefore it should not be surprising that international law often avoids these calculations by promoting the defense of sovereignty as a good that surpasses any potential harm.[111]

Brown University's *Cost of War* project estimates direct violence during the post-9/11 U.S.-led wars in Afghanistan, Iraq, and Syria resulted in over 800,000 deaths. They assess over 355,000 (41%) of those deaths were civilians. These numbers are dwarfed by the even greater number of people wounded or who died due to second-order effects resulting from displacement, malnutrition, or sickness.[112] Even with this data, from a *jus ad bellum* perspective, it is extremely difficulty to weigh the costs of the Afghanistan and Iraq campaigns against the

---

[110] Frowe, *The Ethics of War and Peace*, 56–57.
[111] Frowe, 58.
[112] "Costs of War," accessed April 13, 2020, https://watson.brown.edu/costsofwar/.

potential good these wars accomplished. However, it is relatively easier to digest each individual military engagement from a *jus in bello* perspective.

Article 51 of Additional Protocol I of the Geneva Conventions codifies Proportionality as forbidding "an attack which may be expected to cause incidental loss of civilian life, injury to civilians, damage to civilian objects, or a combination thereof, which would be excessive in relation to the concrete and direct military, advantage anticipated."[113] Article 57 further guides the military to avoid any "attack which may be expected to cause" collateral damage and "take all feasible precautions in the choice of means and methods of attack with a view to avoiding, and in any event to minimizing, incidental loss of civilian life, injury to civilians and damage to civilian objects."[114] IHL are further refined for military professionals, to include the requirement for predictive analysis.

The DoD Law of War Manual states, "incidental damage to the civilian population and civilian objects is unfortunate and tragic, but inevitable," but a Proportionality assessment can weigh the need for engaging a military target against "the expected harms to determine whether the latter are disproportionate in comparison to the former."[115] Interestingly, the DoD Operational Law Handbook takes a less permissive approach: "proportionality requires commanders to refrain from attacks in which the expected harm incidental to such attacks would be excessive in relation to the concrete and direct military advantage anticipated to be gained."[116] From a compliance standpoint, it appears two critical components can be measured if an attack is proportional. They are the intent of the attacker and the result of the attack, to include the

[113] "Protocol II," chaps. II-51-5b.
[114] "Protocol II," IV-57-2a-ii/iii.
[115] *Department of Defense Law of War Manual*, 61.
[116] Kouba, *Operational Law Handbook*, 11.

military advantage gained and the amount of collateral damage that occurred. Out of these components, the amount of collateral damage is usually easier to assess and intent is more difficult to assess. The debate between results and intent drives the continuing ethical debate on Proportionality.

The Ethical Foundations of Proportionality

While the ethical and legal definitions of Distinction are in harmony this is not the case with regard to Proportionality. The simplified legal dichotomy between combatants and noncombatants is at odds with the complex nature of Proportionality that relies on value judgments, risk assessments, and other factors debated by ethicists.[117] The ethical debates regarding Proportionality increased in intensity over the 20th century for key reasons. The first is the second half of the 20th century saw an increasing emphasis on the significance of individual human rights, which is a foundational assumption of Proportionality.[118]

Second, the increasing destructive power of weapons, specifically air-delivered munitions, had the potential to cause significant collateral damage. This fear was illustrated by H.G. Well's 1907 book, *The War in the Air*, which described a fictional world-wide war involving giant airships with great but indiscriminate destructive power. By the end of the conflict, the attacks were more destructive than decisive and led to second-order effects, such as economic disaster and famine that killed many civilians.[119] Despite these fears, international attempts to limit the damage from airpower were limited. The 1923 Hague Rules of Air

---

[117] Seth Lazar, "War," in *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta, Spring 2020 (Metaphysics Research Lab, Stanford University, 2020), https://plato.stanford.edu/archives/spr2020/entries/war/.
[118] Rodin and Shue, *Just and Unjust Warriors: The Moral and Legal Status of Soldiers*, 5.
[119] H. G Wells and Patrick Parrinder, *The War in the Air* (London; New York: Penguin, 2005), 176–78, 250, 274.

Warfare, acknowledged, "…the bombardment of cities, towns, villages, dwellings or buildings is legitimate provided that there exists a reasonable presumption that the military concentration is sufficiently important to justify such bombardment, having regard to the danger thus caused to the civilian population."[120] The U.S. use of atomic bombs against Hiroshima and Nagasaki in 1945 has been debated and framed as either, "among the worst acts in the dismal history of warfare" or a rational way to end the war with Japan.[121] The rational argument stems from a purely utilitarian calculation that indicated the amount of damage caused by the weapons would be less than the foreseeable damage an invasion of the home islands would cause, which Winston Churchill described as "…vast , indefinite butchery…".[122]

Ethical debates regarding Proportionality further increased when the era of precision-guided weapons began. The ability to accurately strike targets led to the increased public expectation that noncombatants and civilian facilities ought not be injured or damaged.[123] This expectation led to further debate that precision-guided weapons could actually increase collateral damage due to a resulting moral hazard that results in a lower threshold to employ weapons in general. Finally, the advent of cyberwarfare has increased ethical debates regarding Proportionally, especially with regard to measuring secondary effects. For example, the Stuxnet attack against Iran's nuclear weapons programs has been hailed as "the first truly ethical weapon

---

[120] Jean-Marie Henckaerts et al., eds., *Customary International Humanitarian Law*, vol. 1 & 2 (Cambridge, UK: Cambridge University Press, 2005), 298, ht.
[121] John Forge and John Forge, "Proportionality, Just War Theory and Weapons Innovation," *Science and Engineering Ethics* 15, no. 1 (2009): 32, https://doi.org/10.1007/s11948-008-9088-z.
[122] Walzer, *Just and Unjust Wars*, 266–67.
[123] James Igoe Walsh, "Precision Weapons, Civilian Casualties, and Support for the Use of Force," *Political Psychology* 36, no. 5 (2015): 508, https://doi.org/10.1111/pops.12175.

ever created" because no human was harmed; conversely, it was criticized because the malware eventually escaped and infected numerous civilian systems.[124]

The third reason the 20th century saw an increase in ethical debates regarding Proportionality was the growing normalization of just war thinking, highlighted by Walzer's foundational work. Over the last fifty years, his work and its underlying moral principles have been debated by generations of revisionists, especially in the context of the overarching generalizations inherent in IHL. Walzer acknowledged that noncombatant immunity does not prevent injury to noncombatants because their proximity to combat puts them at risk. This risk creates a responsibility for combatants to take a degree of care to protect the rights of noncombatants to not be injured.[125] The first phase of this care is based on the moral doctrine of double effect (DDE).

13th century Catholic philosopher Thomas Aquinas introduced the concept of DDE in *Summa Theologica* while debating if killing in self-defense was morally allowable. His logic centered round the argument that a single act may have two different effects, even though one effect may be intended.[126] Therefore, the act of saving one's own life is morally permissible even if it has the unintended consequence of killing the attacker. However, DDE is not a blank check. Four conditions must be met.

1. The original act cannot be evil
2. The good effect cannot directly result from the bad effect
3. The actor must only intend the good result, the bad result may be foreseeable but cannot be intended

---

[124] Lucas, *Military Ethics*, 212.
[125] Walzer, *Just and Unjust Wars*, 152.
[126] Alison McIntyre, "Doctrine of Double Effect," in *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta, Spring 2019 (Metaphysics Research Lab, Stanford University, 2019), https://plato.stanford.edu/archives/spr2019/entries/double-effect/.

4. The good result must be morally equal to or greater than the evil result[127]

Here Proportionality obtains its inherent requirement to have the military necessity of an attack outweigh any collateral damage, and it creates the requirement for some level of foresight to anticipate the potential positive and negative results.

Walzer then adds additional constraints because even though DDE can provide "blanket justification" for large numbers of unintended, but foreseeable, noncombatant deaths, the protection of human rights requires a greater commitment.[128] He names this commitment double intent. It requires the "foreseeable evil be reduced as far as possible" even to the extent that it may impose costs to the combatants executing the attack.[129] It is this change that makes Proportionally much more morally complicated than Distinction, because it possesses deontological as well as utilitarian aspects. Deontology, associated with ethical theorist Immanuel Kant, judges an action to be morally correct as long as the actor's intentions were good. He acknowledged life was filled with too many variables that may affect the final results of an action; therefore, only the original intent should be judged.[130] Conversely, Utilitarianism, also called consequentialism, is an ethical framework that judges the morality of an action based on the consequences of an action, not the intent of the actor. More specifically, it argues the action that produces the most positive results for the most people is the morally preferable

[127] Lawrence M. Hinman, *Ethics: A Pluralistic Approach to Moral Theory*, 5th ed (Boston, MA: Wadsworth Pub Co, 2013), 163.
[128] Walzer, *Just and Unjust Wars*, 153–54.
[129] Walzer, 155.
[130] Hinman, *Ethics*, 6.

choice.[131] Walzer's version of the War Conventions was published in 1977, and his definition of Proportionality has been debated by multiple waves of revisionist ethicists.[132]

Some revisionists, such as David Rodin, argue against Walzers's belief that combatants on both sides of a conflict are morally equivalent. Rodin states individual actions in a conflict are independently amoral apart from how they support achieving the greater war aim. Therefore, combatants engaged in an unjust war may perceive their actions to win the war as positive, however in a larger sense they are immoral.[133] A number of revisionists also believe that DDE and the Geneva Conventions do not go far enough to protect noncombatants, and they put additional requirements on the military to preserve civilian life.

Tony Coady generally agrees with Walzers' positive commitment to preserve noncombatant lives, but he highlights minimizing evil results may require choosing a completely different action that results in less collateral damage.[134] Others take an even more protective view of noncombatants. For example, Colm McKeogh advocates replacing DDE for the foreseeable harm principle. It morally allows damage to civilian property but not civilian deaths unless the killing was both "unforeseen and reasonably unforeseeable."[135] Stephen Nathanson rejects this principle as too close to practical pacifism and advocates for an almost-as-stringent precautionary principle. He argues the precautionary principle requires a higher level of care than the current Geneva Conventions because it demands combatants take every effort possible

---

131 Hinman, 124.

132 For a detailed description of revisionist argument see: Seth Lazar, "Just War Theory: Revisionists Versus Traditionalists," *Annual Review of Political Science* 20, no. 1 (May 11, 2017): 37–54, https://doi.org/10.1146/annurev-polisci-060314-112706.

133 Rodin and Shue, *Just and Unjust Warriors: The Moral and Legal Status of Soldiers*, 4–5.

134 Frowe, *The Ethics of War and Peace*, 149.

135 Frowe, 151.

to predict collateral damage and minimize it as opposed to the Conventions' requirement of "constant care."[136] Therefore harm to noncombatants is morally permissible if, and only if, extensive analysis decreases the risk to noncombatants. In a similar vein, Seth Lazar ties the morality of an engagement to the probability the identity of combatant or noncombatant can be verified. He concludes in his study that a noncombatant's right to live is violated when the attacker is less sure of its targets or noncombatant identity.[137] With these concepts in mind, it is now time to evaluate how AI allows military forces to better adhere to not only IHL but possibly even more stringent moral requirements illuminated by ethicists.

Proportionality in Practice with AI

Military, legal, and ethical practitioners agree Proportionality requires a comparison between two complex concepts. The first, military necessity, depends on the principle of Distinction to identify combatant personnel and related facilities and their potential value to achieving the war's ultimate objective. In the DoD lexicon, this is often referred to as basic target development and intermediate target development. Basic target development is essentially analogous to PID (answering who? what? where? when?) while intermediate target development requires battlespace awareness to define a target's overall significance to the adversary's warfighting capacity (essentially the why?). The second concept of the Proportionality equation is the foreseeable harm to noncombatants and protected facilities. The DoD refers to the practice of assessing and minimizing this harm as advanced target development. Advanced target development consists of two key components, predicting weapons effects and estimating collateral damage. Narrow AI algorithms currently play a significant role in these processes and

---

[136] Frowe, 154.
[137] Seth Lazar, "Risky Killing and the Ethics of War," *Ethics* 126, no. 1 (2015): 97, https://doi.org/10.1086/682191.

have the potential to increase the military's ability to better foresee and mitigate potential harm to noncombatants, especially with regard to collateral damage estimates.

Predicting Weapons Effects

Current weapons technology enables aircraft to engage targets with near-perfect accuracy. This puts greater pressure on targeting analysts to identify the highest value targets and the best weapons to strike the target while minimizing collateral damage. The recent campaign against ISIS has been labeled by military leaders as the most precise air campaign in history.[138] Despite this precision, coalition airstrikes are responsible for over 1,300 confirmed noncombatant deaths, almost 8,300 estimated deaths, and over 5,700 estimated injuries.[139] These numbers demonstrate the increasing pressure on targeting analysts to select not only the correct target, but the exact weapon and aimpoint combination (a process commonly referred to as weaponeering) to minimize injuries to noncombatants while achieving their directed objective.

Weaponeering is both a science and an art as it relies on automated systems and physics-based computer modelling plus the experience of human analysts to navigate a chaotic and ever-changing environment.[140] The DoD's Joint Technical Coordinating Group for Munitions Effectiveness (JTCG/ME) provides software that models the accuracy and explosive yield of warhead, guidance, and fusing combinations found in the Joint Munitions Effectiveness Manuals (JMEM). It also provides vulnerability data for potential targets, based on their size and

---

138 "General: Airpower Key to ISIL Fight; Strikes to Continue," U.S. Air Force, accessed October 23, 2019, https://www.af.mil/News/Article-Display/Article/658205/general-airpower-key-to-isil-fight-strikes-to-continue/. and again in 2017,

139 "US-Led Coalition in Iraq & Syria," accessed April 25, 2020, https://airwars.org/conflict/coalition-in-iraq-and-syria/.

140 "CJCSI 3160.01," D-1.

construction.[141] The JMEM software allows the human analyst to create a specific engagement scenario by manually selecting a warhead, guidance system, fuse combination, and target characteristics. The JMEM algorithm then analyzes hundreds to thousands of iterations of that scenario, usually leveraging a Monte Carlo method, to determine a probable level of damage that scenario would create. Without a Monte Carlo simulation, which randomly assigns realistic values to each variable in a single scenario then combines the results of numerous scenarios to forecast probable results,[142] a human analyst would be forced to rely on less-accurate, manual mathematical models.

In the event the algorithm reports a low probability of achieving a specific effect against a target, the human analyst has the ability to change the desired aimpoints for the weapons or replace the original weapon combination with a different warhead, guidance system, or fusing solution and then run this new scenario through the software. Once the system arrives at an acceptable weaponeering solution, the data is then analyzed by different software to measure the probability of causing collateral damage to adjacent noncombatant facilities and personnel. If the projected collateral damage is considered disproportionate by the military commander, the analyst can return to the original JMEM software to offset the original aimpoints to mitigate damage to the collateral concerns.[143] Otherwise the analysts must create and execute a new scenario and repeat the entire process. Essentially, this process results in satisficing, or

---

[141] "RDT&E Budget Item Justification: PB 2021 Operational Test and Evaluation, Defense" (Department of Defense, February 2020), 7–8, https://apps.dtic.mil/descriptivesum/Y2021/Other/OTE/stamped/U_0605131OTE_6_PB_2021.pdf.

[142] "Explained: Monte Carlo Simulations," MIT News, accessed May 17, 2020, http://news.mit.edu/2010/exp-monte-carlo-0517.

[143] "Joint Publication 3-09 Joint Fires Support" (Washington, D.C.: Department of Defense, April 10, 2019), IV–6, https://www.jcs.mil/Portals/36/Documents/Doctrine/pubs/jp3_09.pdf.

accepting the first result that is considered "good enough."[144] In the case of target development, satisficing means an analyst will accept the first solution they arrive at that achieves the required damage to a target while still staying below the required collateral damage threshold even if this solution is not the most optimal. Advances in AI, however, provide the ability to achieve optimal solutions by more accurately characterizing a target and potential second-order effects that may result from engaging complex targets, even those that house chemical weapons or homemade explosives.

The National Geospatial Intelligence Agency (NGA) is leveraging AI algorithms to support battlespace awareness initiatives. While MARS has the potential to quickly identify large quantities of combatant or noncombatant objects in the battlespace, NGA seeks to provide high-fidelity analysis on objects in a timely manner. NGA and the Intelligence Advanced Research Project Activity (IARPA) have acknowledged, "manually constructed models are accurate and reliable, but making them is time consuming. Of particular concern, the manual process cannot accommodate the provision of models to support rapid-response military efforts or humanitarian crises."[145]

In response they are developing algorithms that can swiftly analyze geospatial information and create accurate 3D object models with real physical properties, "from multiple data sources including commercial satellite panchromatic and multi-spectral imagery for global coverage, and airborne imagery.[146] These models will allow target analysts to create more

---

[144] David A. Cooper, Jessica D. Blankshain, and Nikolas K. Gvosdev, eds., "Foreign Policy Analysis," in *Decision-Making in American Foreign Policy: Translating Theory into Practice* (Cambridge: Cambridge University Press, 2019), 110–11, https://doi.org/10.1017/9781108566742.003.

[145] Dr. HakJae Kim, "New Solution Models 3-D Object Data," *Pathfinder Magazine* 14, no. 2 (2016): 27, https://issuu.com/NGA_GEOINT.

[146] "CORE3D," accessed April 21, 2020, https://www.iarpa.gov/index.php/research-programs/core3d.

accurate estimates to ensure their weaponeering solutions achieve the necessary military effect on the first try, which will then avoid the need to execute reattacks that may levy additional risk on local noncombatants.

In 2015 a coalition airstrike against an ISIS Vehicle-borne Improvised Explosive Device (VBIED) factory in Iraq led to a reported 70 civilian casualties. An investigation revealed the casualties were caused by the secondary explosions resulting from the large quantity of explosives ISIS stored at the facility.[147] The ethical debate regarding who was morally responsible for the civilian casualties revolved around the perception that the coalition should have foreseen these second-order effects or if ISIS was culpable for placing an unmarked weapons facility in a residential neighborhood. Additional Protocol II of the Geneva Conventions prohibits attacks on installations containing dangerous forces, "if such attack may cause the release of dangerous forces and consequent severe losses among the population."[148] The ability to quantify and predict secondary explosions with any certainty is difficult, at best; yet Michael Walzer argues it is a worthy question to determine who put the noncombatants at risk in the first place by allowing them in a battle zone.[149]

DoD targeting instructions direct forces to consider "potential release and dispersal hazards" when attacking sites containing weapons of mass destruction (WMD) or toxic industrial materials (TIM).[150] The DoD is constantly attempting to refine methods to measure second-order effects from these types of targets. The main office charged with this analysis is the

---

147 "Letter to His Excellency Dr. Mark T. Esper, Secretary of Defene" (The Hague: Ministry of Defence of the Kingdom of the Netherlands, January 13, 2020), 5, https://zoek.officielebekendmakingen.nl/blg-928164.pdf.
148 "Protocol II," Article 15.
149 Walzer, *Just and Unjust Wars*, 159.

150 "Joint Publication 3-09 Joint Fires Support," IV–6.

Defense Threat Reduction Agency's (DTRA) Hazard Prediction and Assessment Capability (HPAC).

DTRA-HPAC algorithms are capable of assessing plume hazards from targets such as hydroelectric dams, chemical plants, and CBR storage. They also provide graphic and quantitative assessments of the significant danger of releasing chemical, biological, or radiological clouds into the atmosphere, producing widespread, long-term, and potentially lethal effects on civilians and noncombatants.[151] Much like other intelligence assessments, HPAC's effectiveness is limited by of the amount of detailed knowledge of the type and quantity of dangerous material stored at a facility as well as local weather patterns. Increased battlespace awareness algorithms provide an opportunity to supplement the data on these types of targets. Nevertheless, algorithms that have the potential to mitigate harm to noncombatants who may have been put at risk by their own government is a level of care that approaches ethical principles that advocate for greater foreseeability in combat. The larger challenge is estimating the number of noncombatants residing in a hazard plume for an accurate collateral damage estimate.

Estimating Collateral Damage

The Department of Defense's current collateral damage methodology is designed to "provide logical and repeatable" methods to ensure due diligence in limiting civilian suffering while enabling the commander to assess risk in the accomplishment of military objectives" and

[151] "CJCSI 3160.01," D-A-33.

"mitigate, to the best of our ability, the unintended consequences of that military action."[152] War algorithms have played a critical role in ensuring the process remains both logical and repeatable. While the current methodology clearly meets the Geneva Convention's IHL baseline of "constant care," emerging Computer Vision algorithms can better predict noncombatant population density and activity that could allow the military to meet more rigorous standards of avoiding harm to noncombatants. This would include Nathanson's precautionary principle of taking every effort to predict collateral damage. It would also potentially meet the intent of McKeough's foreseeable harm principle to avoid foreseeable and reasonably unforeseeable harm by increasing the threshold of what is reasonably foreseeable.

As previously discussed, current algorithms are increasingly capable of modeling the amount of damage a weapon can cause to buildings, both targeted and protected. Measuring potential noncombatant casualties is more complicated. While Distinction can provide significantly detailed data on an individual target, measuring the presence and vulnerability of noncombatants across the battlespace can be a significant challenge. Estimating the numbers of noncombatants in an area adjacent to a target often relies on a pre-determined population density table for a specific region. The density is often tailored to provide estimates for day, night, and episodic events.[153] A common source for this data is Oak Ridge National Laboratory's LandScan Global program. The dataset currently has a 1km resolution and is updated on an

[152] "Annex 3-60 Targeting" (Maxwell Air Force Base, AL: Curtis E. Lemay Center, March 15, 2019), 70, https://www.doctrine.af.mil/Portals/61/documents/Annex_3-60/3-60-Annex-TARGETING.pdf.
[153] "Joint Publication 3-60 Joint Targeting" (Washington, D.C.: Department of Defense, January 31, 2013), D-A-34, https://www.cfr.org/blog/jp-3-60-joint-targeting-and-us-targeted-killings.

annual basis. Its algorithm combines census, imagery analysis, and geographic data for specified regions.[154]

The primary automated tool that leverages this data is the Digital Precision Strike Suite Collateral Damage Estimation (DCiDE) program.[155] Its algorithm essentially produces a casualty estimate for decision makers by merging the output of JMEM weaponizing programs against collateral concerns, primarily buildings, refined by population density data. The collateral concerns are regularly imported from foundational intelligence databases. As mentioned previously, current intelligence databases are primarily populated by military facilities not potential noncombatant facilities. Therefore, human imagery analysts must use their training and experience to identify and characterize many noncombatant facilities, usually in a time-compressed environment. Greater battlespace awareness tools, such as MARS, will signifantly decrease the number of collateral concerns required to be identified during the advanced target development process.

Because the population density data is only updated annually and is based on peacetime social norms, their accuracy is degraded during conflicts. Additionally, the current system admittedly does not account for "unknown transient civilian or noncombatant personnel and/or equipment in the vicinity of a target area. This includes cars passing on roads, people walking down the street, or other noncombatant entities whose presence in the target area cannot be predicted to reasonable certainty within the capabilities and limitations of intelligence collection means."[156] Recent improvements in Computer Vision algorithms have the potential to both

[154] "Home | LandScan™," accessed April 26, 2020, https://landscan.ornl.gov/.
[155] "Exhibit R-2," 9.
[156] "CJCSI 3160.01," D 4-5.

improve the timeliness and accuracy of population models in addition to predicting the potential presence and actions of transient noncombatants.

Stanford University recently reported how a novel computer vision algorithm is able to predict demographic characteristics by analyzing the types of vehicles present in the area.[157] Their analysis of over 50 million Google Street View images from across 200 cities was able predict "income, segregation levels, per capita carbon emission and crime rates."[158] They assess their approach coupled with growing satellite imagery data has potential to predict real-time census data..[159] While the current study is limited to datasets in the U.S. it has the potential to provide data to collateral damage estimates that is more timely than current LandScan data and more refined than its current 1 km resolution.

An even more challenging problem is attempting to predict human traffic patterns in a potential combat environment. Two studies show promising algorithms that have the potential to provide near real-time predictive assessments of human traffic patterns. One study leveraged a large network of cameras in crowded train terminals over two years to refine an algorithm that can track the trajectories of individual pedestrians in highly crowded areas.[160] Over two years they were able to create over 100 million individual trajectories, which in itself is an achievement because Computer Vision had only previously been reliably successful at detecting humans in isolation. But they were able to analyze how the movement of one entity can affect the movement of another, and this breakthrough provides the potential to predict movements.[161]

---

157 "Fine-Grained Car Detection for Visual Census Estimation," 1, accessed April 14, 2020, http://vision.stanford.edu/pdf/gebru2017aaai.pdf.
158 "Fine-Grained Car Detection for Visual Census Estimation," 6.
159 "Fine-Grained Car Detection for Visual Census Estimation," 6–7.
160 Alexandre Alahi, Vignesh Ramanathan, and Li Fei-Fei, "Tracking Millions of Humans in Crowded Space in Crowded Spaces," n.d., 1.
161 Alahi, Ramanathan, and Fei-Fei, 1, 19.

A second study in support of allowing autonomous cars to better predict pedestrian actions also provides significant value. The system attempts to replicate the human ability to recognize pedestrian patterns, which are complex social interactions, and avoid collisions with others with the intent to predict all potential trajectories.[162] The study identified trends that led to individuals to pool together, follow each other, as well as change pacing to avoid collisions.[163] This type of predictive analysis would be crucial in executing time-critical target engagements against high-value targets in urban environments. Nevertheless, as ground truth changes during a war the data would need to be updated as the conflict progressed. It must be acknowledged that during a conflict the demand for surveillance and reconnaissance platforms often outstrips the available supply. Therefore, systems with the capability to better predict human traffic patterns during an engagement would likely need to be prioritized for time-sensitive missions at the cost of collecting data in support of greater battlespace awareness.

The Danger and Promise of AI and Proportionality

Proportionality calculations currently depend on a number of narrow but complementary AIs guided by human inputs and judgment. It's clear the disparate algorithms are not seamlessly integrated. University of Michigan's Dr. Melanie Mitchell contended, "A pile of narrow intelligences will never add up to a general intelligence. General intelligence isn't about the number of abilities, but about the integration between those abilities."[164] It is apparent that current trends concerning Distinction and Proportionality indicate these algorithms will continue

[162] Agrim Gupta et al., "Social GAN: Socially Acceptable Trajectories with Generative Adversarial Networks," *ArXiv:1803.10892 [Cs]*, March 28, 2018, 1, http://arxiv.org/abs/1803.10892.
[163] Gupta et al., 7.
[164] Mitchell, *Artificial Intelligence*, 46.

to improve. It's been said that Moore's law no longer merely applies to the doubling of the number of transistors that can be fit on a single silicon chip, but that human technological abilities double every 12-18 months.[165] This increase in capability and the likely increases in algorithm accuracy and speed pose a serious challenge for the moral use of AI in targeting: overconfidence.

Militaries continue to gain access to more and more sensors, to include micro-satellites, remotely-piloted and unmanned collection platforms, and publicly-available information. The ability to synthesize it could lead to actions that are considered unethical. First, decision-makers may begin to believe, incorrectly, that the use of algorithms has moved target analysis from a probabilistic assessment to a certainty. This belief that target intelligence and collateral damage assessments are 100% accurate could lead to the belief that strikes are significantly less risky, leading some decision-makers to more easily authorize strikes. This idea of riskless warfare becoming a moral hazard has already been considered in relation to remotely piloted aircraft (RPA) since wars could become more frequent if one's own pilots are not at risk.[166]

However, there are safeguards against this slippery slope. From an ethical standpoint, there is a growing normalized understanding that the greater the technology one possesses, the greater the expectation to use it responsibly.[167] From a legal standpoint, the "Rendulic Rule" should prevent decision-makers from confusing probability with certainty. The Rendulic case concerns liability for battlefield acts and confirms "commanders and personnel should be evaluated based on information reasonably available at the time of decision."[168] Therefore, a

---

[165] Lucas, *Military Ethics*, 165.
[166] Frowe, *The Ethics of War and Peace*, 236.
[167] Rodin and Shue, *Just and Unjust Warriors: The Moral and Legal Status of Soldiers*, 73.
[168] Kouba, *Operational Law Handbook*, 10.

Marine who shoots at a noncombatant car he believes to be an imminent threat based on limited information does not break the Law of War. Similarly, it is conceivable a leader who makes a decision to begin a war based on the misperception of their own information could be held liable. Despite this threat of moral hazard, there is an even greater positive effect the growing use of AI can have on probability assessments.

DoD and Intelligence Community initiatives to standardize data formats and share information seamlessly may further enhance foreseeability in target development. This evolution would require an increased emphasis on further integrating targeting algorithms and incorporating machine learning. Even though combat inherently exists in an open-system, the variables common to target development for kinetic engagement are well known (target characteristics, weapon characteristics, and noncombatant characteristics). It may, therefore, be extremely beneficial to incorporate the models and data used to execute weaponeering and collateral damage estimates into a consolidated closed system and construct a single AI that could produce and analyze multiple targeting scenarios for the same target concurrently. This would allow the targeting process to evolve from producing single scenarios that result in mere satisficing and allow for the synthesis of a targeting solution that is both operationally and ethically optimal (see Figures 2 and 3). This system could then rank-order the scenarios for decision-makers, identifying the strike parameters with the best probability of achieving the military object and the lowest probability of civilian casualties.[169] This idea is not too far removed from building a virtual model of an aircraft and projecting its maintenance needs based on its performance after

[169] In this sense the targeting scenarios would be analogous to using Drosophila to evaluate genetic theories. See Nathan Ensmenger, "Is Chess the Drosophila of Artificial Intelligence? A Social History of an Algorithm," *Social Studies of Science* 42, no. 1 (2012): 5–30, https://doi.org/10.1177/0306312711424596.

flying virtual missions in a virtual world nor from having an AI teach itself chess or go. [170]

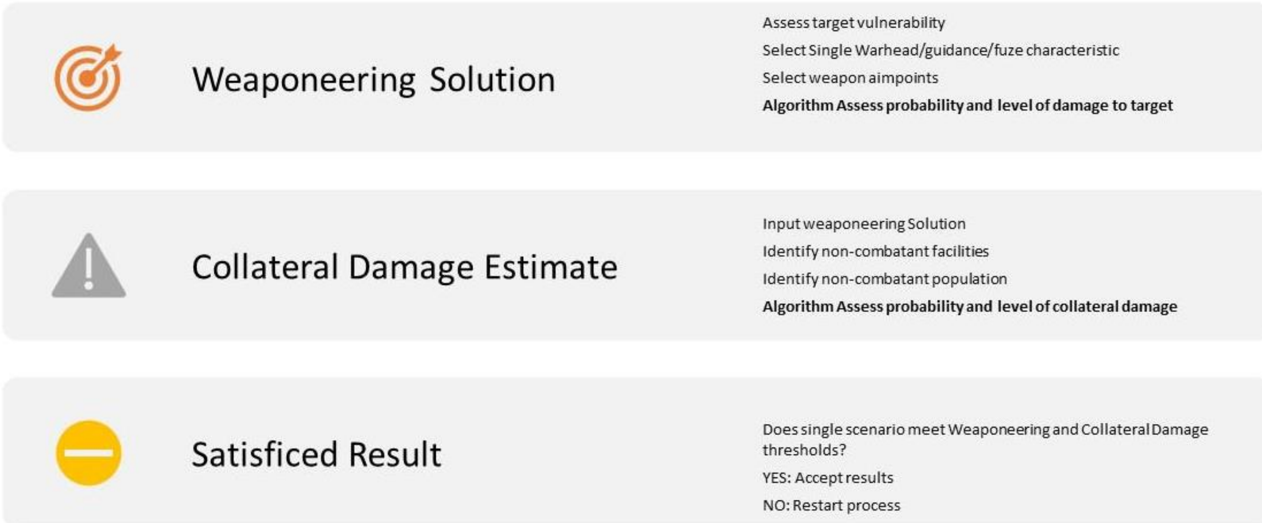

Figure 2: Satisficing in Advanced Target Development

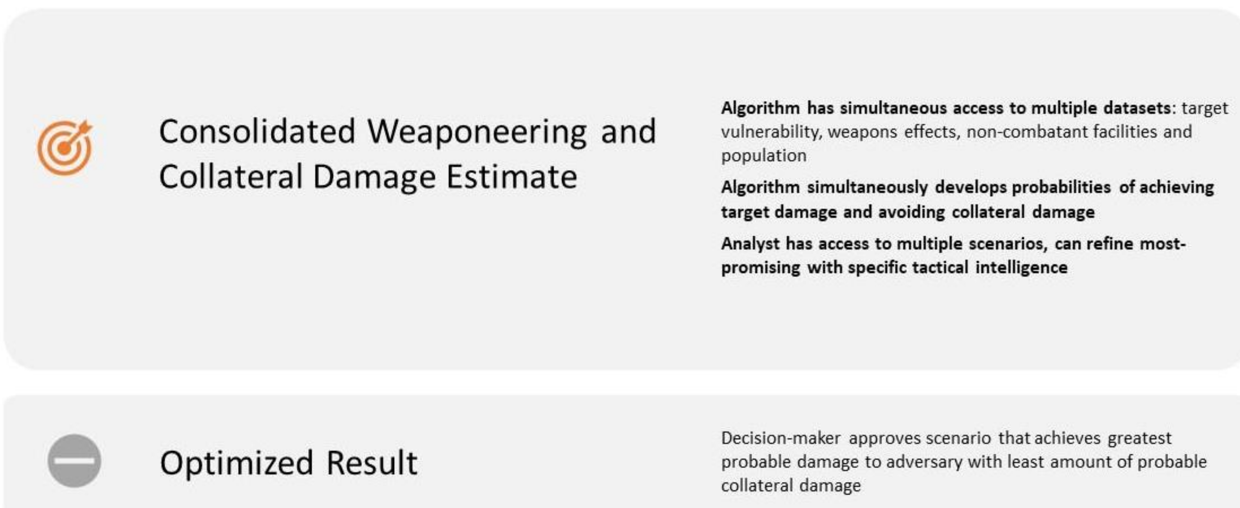

Figure 3: Optimized Advanced Target Development Processes

---

170 "Air Force Partners to Create B-1B 'Digital Twin,'" Air Force Materiel Command, accessed April 27, 2020, https://www.afmc.af.mil/News/Article-Display/Article/2163089/air-force-partners-to-create-b-1b-digital-twin/.

While board games are less subject to external factors than targeting, experience has shown that forms of narrow AI can analyze the "rules of the game" and achieve the ultimate goal of "winning." In this respect, an AI has the advantage over a human in the amount of experience it can accumulate in a significantly short amount of time. For example, a chess-learning algorithm can be become an expert: it "starts out knowing only the rules of chess, with no embedded human strategies. In just a few hours, it plays more games against itself than have been recorded in human chess history."[171] Chess Grandmaster Gary Kasparov has ascertained that the rise of AI chess experts does not render human chess players obsolete. He has acknowledged that sometimes the algorithms, while learning, produce some truly novel approaches to the game that have not been attempted by humans and other approaches that are less successful. However, he is convinced there is much humans and machines can learn from each other and the best way ahead is for both sides to work together. This human-machine understanding is even more important for ethical and legal targeting.

Proportionality is a human decision that can be enhanced by recommendations from algorithms. In most cases, senior leaders do not have the time or expertise to digest the nuances of DoD's targeting and collateral damage methodologies, which number at over 500 pages. In fact, former President George W. Bush's statement, "I don't do nuance" encapsulates this concept well.[172] This can create a risk of having the data being skewed by cognitive biases, which could lead to the most ethical-targeting solution not being selected. These cognitive biases may have been on display in June 2019 during planning for an airstrike against Iranian targets in

[171] Garry Kasparov, "Chess, a Drosophila of Reasoning," *Science (American Association for the Advancement of Science)* 362, no. 6419 (December 7, 2018): 1087, http://usnwc.summon.serialssolutions.com/2.0.0/link/0/.
[172] Cooper, Blankshain, and Gvosdev, "Foreign Policy Analysis," 97.

response to their shootdown of a U.S. RQ-4 Global Hawk RPA. The decision to not strike was based on the President's belief that the expected 150 Iranian deaths were disproportional.

The cognitive issue was that the model used to explain the potential costs to him was oversimplified. During the briefings the President was provided with an estimate that the strike would result in "150 dead people."[173] However, it appears he took the body count as a certainty and not a probability. The estimate of 150 potential casualties was oversimplified into guaranteed deaths. He also did not assess the nuance that the estimate was based on very specific attack parameters that could be adjusted. It was reported that the high estimate he used for his decision was based on the number of individuals at the target during the day, a worst-case scenario, vice the number of personnel that would be there at night when the strike would actually occur.[174] While the action of providing an oversimplified assessment of casualties allowed POTUS's mental framework to better process this data among the many other variables involved in the decision, it is clear it had a decisive impact on the final decision to not strike and may have warranted more detail.

The second cognitive influence to his decision may have been anchoring. Anchoring occurs when an individual prematurely considers a specific value for an unknown amount before estimating that amount.[175] It is possible the President may have found any number of potential casualties that exceeded the low number he experienced during the two previous Syria strikes as unpalatable. This is due to anchoring. It is clear the numbers from the estimate caused some

[173] Peter Baker, Maggie Haberman, and Thomas Gibbons-Neff, "Urged to Launch an Attack, Trump Listened to the Skeptics Who Said It Would Be a Costly Mistake," *New York Times*, June 21, 2019, ProQuest Central, https://search.proquest.com/docview/2244404959?accountid=322.
[174] Baker, Haberman, and Gibbons-Neff.
[175] Kahneman, *Thinking, Fast and Slow*, 119.

cognitive dissonance. "They gave me very odd numbers," Mr. Trump said about his national-security team. "I wanted an accurate count." That estimate came later on Thursday: 150 potential casualties, or about 40 to 50 at each strike, Mr. Trump explained on Saturday.[176] As the two previous strikes reportedly caused significant physical destruction with a low body count, his baseline of acceptable casualties may have already started artificially low based on earlier experiences. If the 150 was put in the context of the 7,000 civilians reportedly killed by Russian airstrikes in Syria, he may have decided 150 was proportional.

A single AI with the potential to quickly evaluate and prioritize hundreds of potential strike scenarios against a single target can mitigate anchoring. Instead of running the risk a proportionality decision for "Strike C" is being subconsciously biased from data from previous "Strike A" or "Strike B," the individual decision can be properly contextualized.[177] In short, the decision-maker is comparing ethical apples-to-apples vice apples-to-oranges. It is this ability for algorithms to clearly and quickly characterize complex data for human beings that makes AI a key enabler for ethical decision-making in future conflicts.

---

[176] Michael Bender and Gordon Lubold, "Trump Bucked National-Security Aides on Proposed Iran Attack; 'These People Want to Push Us into a War... It's so Disgusting,' President Told Confidant," *Wall Street Journal (Online)*, June 23, 2019, ProQuest Central, https://search.proquest.com/docview/2244847710?accountid=322.

[177] This does not imply that a comprehensive count of civilian casualties has no value, for an in-depth assessment of the ethics of proportionality over time please see, Jeff McMahan, "Proportionality and Time," *Ethics* 125, no. 3 (2015): 696–719, https://doi.org/10.1086/679557.

CONCLUSION AND TOPICS FOR FUTURE RESEARCH

Lethal Autonomous Weapons and Artificial General Intelligence are both compelling topics that continue to dominate the headlines. These topics are worthy of the continuing debate regarding their ethical employment. This debate can be informed by acknowledging how other forms of AI, most notably narrow-AI war algorithms, may lead to greater *jus in bello* ethical warfighting. AI can clearly improve warfighter's ability to positively identify targets through greater battlespace awareness. It can also help prevent injuries to noncombatants by allowing targeting analysts to more accurately predict the amount of collateral damage a strike may cause by creating a process that yields optimal results vice merely dabbling in satisficing.

Lucas Kunce's vision of human-machine teaming in combat is a goal worthy of pursuing and refining. Military professions are responsible for managing violence on behalf of their state, but they are not perfect. Carl von Clausewitz's words regarding decision-making during combat are still true today: "During an operation decisions have usually to be made at once; there may be no time to review the situation or even to think it through."[178] The use of AI, however, can provide the time and space to aid decisions made during combat and have the potential to mitigate the actions Kune's units experienced in combat that resulted in the death of noncombatants. In fact, a greater understanding of our own capabilities and those of potential adversaries may provide strategic decision-makers with the ability to decide which wars are worth fighting in the first place, an inherently ethical responsibility. In the words of Sun Tzu, "now if the estimates made in the temple before hostilities indicate victory it is because calculations show one's strength to be superior to that of his enemy; if they indicate defeat, it is

[178] Clausewitz et al., *On War*, 102.

because calculations show that one is inferior. With many calculations, one can win; with few one cannot. How much less chance of victory has one who makes none at all!"[179] Putting warfighters in the best position to not only win a war, but win it justly is a cause all citizens in a state should embrace. Therefore, it is morally preferable to explore the potential of human-machine teaming beyond the principles of Discrimination and Proportionality.

This paper primarily focused on how narrow AI can enable more ethical targeting processes, specifically by aiding analysts during target development. There are other components of the overall joint targeting process that would benefit from a similar evaluation of the potential of human-machine teaming. The most promising subject matter is the combat assessment and reattack recommendation process. Combat assessment is a process measuring the damage to the target (battle damage assessment), the damage to noncombatants (collateral damage assessment), and how well the weapons performed.[180] All of these assessments could conceivably lead to greater ethical warfighting if they were further aided by narrow-AIs.

More accurate battle damage assessments could prevent unnecessary restrikes, thereby decreasing the risk to noncombatants and friendly forces. Measuring collateral damage and weapon performance faster and more accurately fulfills a moral obligation to prevent harm to future noncombatants. Chris Woods of Air Wars, an organization reporting the number of noncombatant casualties during ongoing conflicts, stated the "military learn from their mistakes by looking at how civilians died."[181] The Department of Defense continues to invest in research to analyze the effects of airstrikes in Iraq, Syria, and Afghanistan with the goal to "improve the

---

[179] Sunzi, *The Illustrated Art of War* (New York: Oxford University Press, 2005), chaps. I–28.
[180] "Chairman of the Joint Chiefs of Staff Instruction: Methodology for Combat Assessment" (Department of Defense, March 8, 2019), 1, https://www.jcs.mil/Portals/36/Documents/Doctrine/training/jts/cjcsi_3162_02.pdf?ver=2019-03-13-092459-350.
[181] "France's War without Accountability," accessed May 15, 2020, https://airwars.org/news-and-investigations/french-non-accountability-for-civilian-harm/.

warfighter's ability to get the right weapon on the right target, achieve the desired effect, and minimize collateral damage while optimizing scarce resources."[182] It is likely the emerging technologies that enable more effective target development could be leveraged to better assess the results of the process.

[182] "Exhibit R-2," 15.

Bibliography

"About the JAIC - JAIC." Accessed December 31, 2019. https://www.ai.mil/about.html.

"AI Principles: Recommendations on the Ethical Use of Artificial Intelligence by the Department of Defense." Defense Innovation Board, October 31, 2019. innovation.defense.gov/ai.

"AI Principles: Recommendations on the Ethical Use of Artificial Intelligence by the Department of Defense Supporting Document." Defense Innovation Board, October 31, 2019. innovation.defense.gov/ai.

Air Force Materiel Command. "Air Force Partners to Create B-1B 'Digital Twin.'" Accessed April 27, 2020. https://www.afmc.af.mil/News/Article-Display/Article/2163089/air-force-partners-to-create-b-1b-digital-twin/.

Alahi, Alexandre, Vignesh Ramanathan, and Li Fei-Fei. "Tracking Millions of Humans in Crowded Space in Crowded Spaces," n.d., 20.

Allen, Gregory C. "Putin and Musk Are Right: Whoever Masters AI Will Run the World." CNN. Accessed December 30, 2019. https://www.cnn.com/2017/09/05/opinions/russia-weaponize-ai-opinion-allen/index.html.

———. "Understanding China's AI Strategy: Clues to Chinese Strategic Thinking on Artificial Intelligence and National Security." Washington, D.C.: Center for a New American Security, February 2019. www.cnas.rg.

"Annex 3-60 Targeting." Maxwell Air Force Base, AL: Curtis E. Lemay Center, March 15, 2019. https://www.doctrine.af.mil/Portals/61/documents/Annex_3-60/3-60-Annex-TARGETING.pdf.

U.S. Central Command. "April 29: CENTCOM Releases Investigation into Airstrike on Doctors Wit." Accessed February 8, 2020. https://www.centcom.mil/MEDIA/PRESS-RELEASES/Press-Release-View/Article/904574/april-29-centcom-releases-investigation-into-airstrike-on-doctors-without-borde/.

"Artificial Intelligence and Machine Learning in Armed Conflict: A Human-Centered Approach." Geneva, Switzerland: ICRC, June 5, 2019. https://www.icrc.org/en/document/artificial-intelligence-and-machine-learning-armed-conflict-human-centred-approach.

Baker, Peter, Maggie Haberman, and Thomas Gibbons-Neff. "Urged to Launch an Attack, Trump Listened to the Skeptics Who Said It Would Be a Costly Mistake." *New York Times*. June 21, 2019. ProQuest Central. https://search.proquest.com/docview/2244404959?accountid=322.

Bender, Michael, and Gordon Lubold. "Trump Bucked National-Security Aides on Proposed Iran Attack; 'These People Want to Push Us into a War... It's so Disgusting,' President Told Confidant." *Wall Street Journal (Online)*. June 23, 2019. ProQuest Central. https://search.proquest.com/docview/2244847710?accountid=322.

Berberian, B. "Man-Machine Teaming: A Problem of Agency." *IFAC PapersOnLine* 51, no. 34 (2019): 118–23. https://doi.org/10.1016/j.ifacol.2019.01.049.

"Chairman of the Joint Chiefs of Staff Instruction: Methodology for Combat Assessment." Department of Defense, March 8, 2019. https://www.jcs.mil/Portals/36/Documents/Doctrine/training/jts/cjcsi_3162_02.pdf?ver=2019-03-13-092459-350.

"Chairman of the Joint Chiefs of Staff Instruction: No-Strike and the Collateral Damage Estimation Methodology." Department of Defense, February 13, 2009. https://www.aclu.org/files/dronefoia/dod/drone_dod_3160_01.pdf.

Cheung, Tai Ming. *Fortifying China: The Struggle to Build a Modern Defense Economy*, 2013.

Clausewitz, Carl von, Michael Howard, Peter Paret, and Bernard Brodie. *On War*. Princeton, N.J.: Princeton University Press, 1984.

Clifford, Catherine. "Elon Musk: 'Mark My Words — A.I. Is Far More Dangerous than Nukes.'" CNBC, March 13, 2018. https://www.cnbc.com/2018/03/13/elon-musk-at-sxsw-a-i-is-more-dangerous-than-nuclear-weapons.html.

———. "Hundreds of A.I. Experts Echo Elon Musk, Stephen Hawking in Call for a Ban on Killer Robots." CNBC, November 8, 2017. https://www.cnbc.com/2017/11/08/ai-experts-join-elon-musk-stephen-hawking-call-for-killer-robot-ban.html.

Clune, Jeff. "AI-GAs: AI-Generating Algorithms, an Alternate Paradigm for Producing General Artificial Intelligence." *ArXiv:1905.10985 [Cs]*, January 31, 2020. http://arxiv.org/abs/1905.10985.

Coats, Daniel R. Worldwide Threat Assessment of the US Intelligence Community, § Senate Select Committee on Intelligence (2019).

Cooper, David A., Jessica D. Blankshain, and Nikolas K. Gvosdev, eds. "Foreign Policy Analysis." In *Decision-Making in American Foreign Policy: Translating Theory into Practice*, 14–51. Cambridge: Cambridge University Press, 2019. https://doi.org/10.1017/9781108566742.003.

"CORE3D." Accessed April 21, 2020. https://www.iarpa.gov/index.php/research-programs/core3d.

"Coronavirus Brings China's Surveillance State out of the Shadows." *Reuters*, February 7, 2020. https://www.reuters.com/article/us-china-health-surveillance-idUSKBN2011HO.

"Costs of War." Accessed April 13, 2020. https://watson.brown.edu/costsofwar/.

Cukier, Kenneth. "Ready for Robots? How to Think About the Future of AI." *Foreign Affairs* 98, no. 4 (August 2019): 192–98.

"Defense Intelligence Agency Strategic Approach." Defense Intelligence Agency, September 2018. https://www.dia.mil/Portals/27/Documents/About/DIA_Strategic_Approach.pdf.

Department of Defense Law of War Manual. Washington, D.C.: Office of General Counsel, Department of Defense, 2016. https://tjaglcspublic.army.mil/dod-low-manual.

DEPARTMENT OF DEFENSE POSTURE FOR ARTIFICIAL INTELLIGENCE: Assessment. S.l.: RAND CORPORATION, 2020.

U.S. DEPARTMENT OF DEFENSE. "Department of Defense Press Briefing by General Harrigian via Teleconf." Accessed April 21, 2020. https://www.defense.gov/Newsroom/Transcripts/Transcript/Article/1193060/department-of-defense-press-briefing-by-general-harrigian-via-teleconference/.

Defense Intelligence Agency. "DIA's Vision of MARS: Decision Advantage for the 21st Century." Accessed August 5, 2019. https://www.dia.mil/News/Articles/Article-View/Article/1855910/dias-vision-of-mars-decision-advantage-for-the-21st-century/.

Ensmenger, Nathan. "Is Chess the Drosophila of Artificial Intelligence? A Social History of an Algorithm." *Social Studies of Science* 42, no. 1 (2012): 5–30. https://doi.org/10.1177/0306312711424596.

Executive Office of the President. *National Security Strategy of the United States of America*. United States: White House Office, 2017. https://www.whitehouse.gov/wp-content/uploads/2017/12/NSS-Final-12-18-2017-0905.pdf.

MIT News. “Explained: Monte Carlo Simulations.” Accessed May 17, 2020. http://news.mit.edu/2010/exp-monte-carlo-0517.

Fassihi, Farnaz. “Iran Says It Unintentionally Shot Down Ukrainian Airliner.” *The New York Times*, January 10, 2020, sec. World. https://www.nytimes.com/2020/01/10/world/middleeast/missile-iran-plane-crash.html.

“Fine-Grained Car Detection for Visual Census Estimation.” Accessed April 14, 2020. http://vision.stanford.edu/pdf/gebru2017aaai.pdf.

Forge, John, and John Forge. “Proportionality, Just War Theory and Weapons Innovation.” *Science and Engineering Ethics* 15, no. 1 (2009): 25–38. https://doi.org/10.1007/s11948-008-9088-z.

“France’s War without Accountability.” Accessed May 15, 2020. https://airwars.org/news-and-investigations/french-non-accountability-for-civilian-harm/.

Frowe, Helen. *The Ethics of War and Peace: An Introduction*. Second edition. The Ethics Of. London: Routledge/Taylor & Francis Group, 2016.

“Fundamental Principles of IHL | How Does Law Protect in War? - Online Casebook.” Accessed January 7, 2020. https://casebook.icrc.org/glossary/fundamental-principles-ihl.

U.S. Air Force. “General: Airpower Key to ISIL Fight; Strikes to Continue.” Accessed October 23, 2019. https://www.af.mil/News/Article-Display/Article/658205/general-airpower-key-to-isil-fight-strikes-to-continue/.

“Geneva Convention Relative to the Protection of Civilian Persons in Time of War of 12 August 1949.” International Committee of the Red Cross, August 12, 1949. https://ihl-databases.icrc.org/applic/ihl/ihl.nsf/vwTreaties1949.xsp.

Gilpin, Robert. *War and Change in World Politics*. Cambridge; New York: Cambridge University Press, 1981.

Gupta, Agrim, Justin Johnson, Li Fei-Fei, Silvio Savarese, and Alexandre Alahi. “Social GAN: Socially Acceptable Trajectories with Generative Adversarial Networks.” *ArXiv:1803.10892 [Cs]*, March 28, 2018. http://arxiv.org/abs/1803.10892.

Henckaerts, Jean-Marie, Louise Doswald-Beck, Carolin Alvermann, and International Committee of the Red Cross, eds. *Customary International Humanitarian Law*. Vol. 1 & 2. 2 vols. Cambridge, UK: Cambridge University Press, 2005. ht.

Hinman, Lawrence M. *Ethics: A Pluralistic Approach to Moral Theory*. 5th ed. Boston, MA: Wadsworth Pub Co, 2013.

Hitchens, Theresa. “EXCLUSIVE Pentagon’s AI Problem Is ‘Dirty’ Data: Lt. Gen. Shanahan.” *Breaking Defense* (blog). Accessed April 21, 2020. https://breakingdefense.com/2019/11/exclusive-pentagons-ai-problem-is-dirty-data-lt-gen-shanahan/.

Hodgson, Marshall G. S, and Edmund Burke. *Rethinking World History: Essays on Europe, Islam, and World History*. Cambridge [England]; New York, NY, USA: Cambridge University Press, 1993.

“Home | LandScan™.” Accessed April 26, 2020. https://landscan.ornl.gov/.

Horizon, Over The. “Publicly Available Information: The Secret to Unclassified Data, Part I.” OTH, April 8, 2019. https://othjournal.com/2019/04/08/publicly-available-information-the-secret-to-unclassified-data-part-i/.

Air National Guard. “Intel Airmen Sharpen AI Technology for Domestic Response.” Accessed April 21, 2020. https://www.ang.af.mil/Media/Article-Display/Article/2025332/intel-airmen-sharpen-ai-technology-for-domestic-response/.

Jamieson, Lt Gen Dash. “Next Generation ISR Dominance Flight Plan 2018-2028.” Headquarters Air Force, 2018. https://www.airuniversity.af.edu/ISR/#isr-flight-plan.

“Joint Publication 3-09 Joint Fires Support.” Washington, D.C.: Department of Defense, April 10, 2019. https://www.jcs.mil/Portals/36/Documents/Doctrine/pubs/jp3_09.pdf.

“Joint Publication 3-60 Joint Targeting.” Washington, D.C.: Department of Defense, January 31, 2013. https://www.cfr.org/blog/jp-3-60-joint-targeting-and-us-targeted-killings.

Kahneman, Daniel. *Thinking, Fast and Slow*. 1st pbk. ed. New York: Farrar, Straus and Giroux, 2013.

Kasparov, Garry. “Chess, a Drosophila of Reasoning.” *Science (American Association for the Advancement of Science)* 362, no. 6419 (December 7, 2018): 1087. http://usnwc.summon.serialssolutions.com/2.0.0/link/0/eLvHCXMwtV1LS8NAEF6sRelFbH2_yFHRymY32SQHD1IVwVO1RXsqk31gwaal9YH_3n2kbSwqehDCEnZJIDOT2W8mk_kQouQU1-d8ggxAskQC8IiYb32BL9NExZRzHVCk3MSNjXv60IybneR2xsg5m_tXxes5rXrzI-0flD-9qZ7Q59oE9KiNQI-_MoPGo3S06HB8MbKcBb0nMAjRFKPbVGwRnk7e9C8oegopgxvQbmg0eLU5dRjlhcR56sCPbRmG296kc3fYMDUSTIv-kDJSUDwLcpfmHJyP3f5YEN2wb2VHbYLZ8e_8vDrX4Hq6ZJqd90WPP5_JrN6-K6FSlJgqvUa786mPcgUtT66ZCwYsKGitopUczXvnTu5VtCCzGlpy_J7vNVTN5Tn2DvP23kdryLMqOfHAmynEGyhvqpB11L66bDWu6zlPRd0AAB2Cx3EaEEX8FAc-MKVRtVAcgxKYM8pjEnEV-kKD3YAyABkymQb6SADTSKUx2UCL2SCTW8gLORORAJGAiAKcSo0OSRgnAJQLygTZRpvuebtD14ykO5HEzrcru6gyU_4eKitt2nIflV_G2Rs_sOL9AD0OHVw.

Kelsay, John. *Islam and War: A Study in Comparative Ethics*. 1st ed. Louisville, Ky: Westminster/John Knox Press, 1993.

Kharpal, Arjun. “Stephen Hawking Says A.I. Could Be ‘Worst Event in the History of Our Civilization.’” CNBC, November 6, 2017. https://www.cnbc.com/2017/11/06/stephen-hawking-ai-could-be-worst-event-in-civilization.html.

Kim, Dr. HakJae. “New Solution Models 3-D Object Data.” *Pathfinder Magazine* 14, no. 2 (2016): 27. https://issuu.com/NGA_GEOINT.

Kouba, Dustin, ed. *Operational Law Handbook*. 17th ed. Charlottesville, VA: The Judge Advocate General’s Legal Center and School, 2017.

Lazar, Seth. “Just War Theory: Revisionists Versus Traditionalists.” *Annual Review of Political Science* 20, no. 1 (May 11, 2017): 37–54. https://doi.org/10.1146/annurev-polisci-060314-112706.

———. “Risky Killing and the Ethics of War.” *Ethics* 126, no. 1 (2015): 91–117. https://doi.org/10.1086/682191.

———. “War.” In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Spring 2020. Metaphysics Research Lab, Stanford University, 2020. https://plato.stanford.edu/archives/spr2020/entries/war/.

“Letter to His Excellency Dr. Mark T. Esper, Secretary of Defene.” The Hague: Ministry of Defence of the Kingdom of the Netherlands, January 13, 2020. https://zoek.officielebekendmakingen.nl/blg-928164.pdf.

Lewis, Dustin A., Gabriella Blum, and Naz K. Modirzadeh. "War-Algorithm Accountability." *SSRN Electronic Journal*, August 2016. https://doi.org/10.2139/ssrn.2832734.
U.S. DEPARTMENT OF DEFENSE. "Lt. Gen. Jack Shanahan Media Briefing on A.I.-Related Initiatives With." Accessed April 21, 2020. https://www.defense.gov/Newsroom/Transcripts/Transcript/Article/1949362/lt-gen-jack-shanahan-media-briefing-on-ai-related-initiatives-within-the-depart/.
Lucas, George R. *Ethics and Cyber Warfare: The Quest for Responsible Security in the Age of Digital Warfare*. New York, NY: Oxford University Press, 2017.
———. *Military Ethics: What Everyone Needs to Know*. New York, NY: Oxford University Press, 2016.
"Machine Assisted Analytic Rapid Repository System (MARS)." Defense Intelligence Agency, June 21, 2019. https://nova.afceachapters.org/file/31587/download?token=43UDQrZf.
Mattis, Jim. "Summary of the 2018 National Defense Strategy." Washington, D.C.: Department of Defense, 2018. https://dod.defense.gov/Portals/1/Documents/pubs/2018-National-Defense-Strategy-Summary.pdf.
McIntyre, Alison. "Doctrine of Double Effect." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Spring 2019. Metaphysics Research Lab, Stanford University, 2019. https://plato.stanford.edu/archives/spr2019/entries/double-effect/.
McMahan, Jeff. "Proportionality and Time." *Ethics* 125, no. 3 (2015): 696–719. https://doi.org/10.1086/679557.
Mesquita, Bruce Bueno de, and George W. Downs. "Development and Democracy." *Foreign Affairs* 84, no. 5 (2005): 77–86. https://doi.org/10.2307/20031707.
Michel, Arthur Holland. "The Killer Algorithms Nobody's Talking About." *Foreign Policy* (blog). Accessed January 21, 2020. http://foreignpolicy.com/2020/01/20/ai-autonomous-weapons-artificial-intelligence-the-killer-algorithms-nobodys-talking-about/.
Mitchell, Melanie. *Artificial Intelligence: A Guide for Thinking Humans*. New York: Farrar, Straus and Giroux, 2019.
Molloy, Brian T. "Project Governance for Defense Applications of AI: An Ethics-Based Approach." Naval War College, 2020.
Nurkin, Tate, Stephen Rodriguez, Atlantic Council of the United States, and Brent Scowcroft Center on International Security. *A Candle in the Dark: US National Security Strategy for Artificial Intelligence*, 2019. https://www.atlanticcouncil.org/wp-content/uploads/2019/12/AC_CandleinDark120419_FINAL.pdf.
"Opinion | Dear Tech Workers, U.S. Service Members Need Your Help - The New York Times." Accessed October 1, 2019. https://www.nytimes.com/2019/08/28/opinion/military-war-tech-us.html.
Osinga, Frans P. B. *Science, Strategy and War: The Strategic Theory of John Boyd*. London; New York: Routledge, 2007.
"Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of International Armed Conflicts (Protocol II), of 8 June 1977." International Committee of the Red Cross, June 8, 1977. https://ihl-databases.icrc.org/applic/ihl/ihl.nsf/vwTreaties1949.xsp.
"RDT&E Budget Item Justification: PB 2021 Operational Test and Evaluation, Defense." Department of Defense, February 2020. https://apps.dtic.mil/descriptivesum/Y2021/Other/OTE/stamped/U_0605131OTE_6_PB_2021.pdf.

Reichberg, Gregory M., Henrik Syse, and Endre Begby, eds. *The Ethics of War: Classic and Contemporary Readings*. Malden, MA ; Oxford: Blackwell Pub, 2006.

Rodin, David, and Henry Shue, eds. *Just and Unjust Warriors: The Moral and Legal Status of Soldiers*. New York: Oxford University Press - Books Distribution Services, 2010.

Roy, Anna. "National Strategy for Artificial Intelligence #AIFORALL." New Delhi, India: National Institution for Transforming India Aayog, June 2018. https://niti.gov.in/writereaddata/files/document_publication/NationalStrategy-for-AI-Discussion-Paper.pdf.

Ryan, Mick. "Mastering the Profession of Arms, Part 1: The Enduring Nature." *War on the Rocks*, February 8, 2017. https://warontherocks.com/2017/02/mastering-the-profession-of-arms-part-i-the-enduring-nature/.

Sayler, Kelley M, and Daniel S Hoadley. "Artificial Intelligence and National Security." Washington, D.C.: U.S. Congressional Research Service, January 30, 2019. https://crsreports.congress.gov.

Scharre, Paul. *Army of None: Autonomous Weapons and the Future of War*. First edition. New York ; London: W. W. Norton & Company, 2018.

U.S. Air Force. "Self-Developed Intelligence Program Strengthens Coalition Fight Agains." Accessed April 23, 2020. https://www.af.mil/News/Article-Display/Article/1078078/self-developed-intelligence-program-strengthens-coalition-fight-against-isil/.

Snyder, Jack. "One World, Rival Theories." *Foreign Policy*, no. 145 (2004): 53–62. https://doi.org/10.2307/4152944.

Subedi, Surya P. "The Concept in Hinduism of 'Just War.'" *Journal of Conflict & Security Law* 8, no. 2 (2003): 339–61. www.jstor.org/stable/26294280.

Sunzi. *The Illustrated Art of War*. New York: Oxford University Press, 2005.

"The Campaign to Stop Killer Robots." Accessed December 30, 2019. https://www.stopkillerrobots.org/about/.

Humanitarian Law & Policy Blog. "The Most Read Blog Posts in 2019 Archives." Accessed January 21, 2020. https://blogs.icrc.org/law-and-policy/category/type-of-post/the-most-read-blog-posts-in-2019/.

"The Pentagon's New Artificial Intelligence Is Already Hunting Terrorists - Defense One." Accessed April 21, 2020. https://www.defenseone.com/technology/2017/12/pentagons-new-artificial-intelligence-already-hunting-terrorists/144742/.

Thiel, Peter. "Opinion | Good for Google, Bad for America." *The New York Times*, August 1, 2019, sec. Opinion. https://www.nytimes.com/2019/08/01/opinion/peter-thiel-google.html.

TechCrunch. "This Famous Roboticist Doesn't Think Elon Musk Understands AI." Accessed January 1, 2020. http://social.techcrunch.com/2017/07/19/this-famous-roboticist-doesnt-think-elon-musk-understands-ai/.

Thornberry, Mac. John S. McCain National Defense Authorization Act for Fiscal Year 2019, Pub. L. No. Public Law 115-232 (2018). https://www.congress.gov/115/bills/hr5515/BILLS-115hr5515enr.pdf.

UN News. "UN Aviation Experts to Join Ukraine Airlines Iran Crash Investigation," January 14, 2020. https://news.un.org/en/story/2020/01/1055312.

"United Nations Activities on Artificial Intelligence (AI) 2019." Geneva, Switzerland: International Telecommunication Union, 2019. https://www.itu.int/dms_pub/itu-s/opb/gen/S-GEN-UNACT-2019-1-PDF-E.pdf.

AOAV. "United States' Airstrike Rules of Engagement Reviewed," April 29, 2019. https://aoav.org.uk/2019/an-investigation-into-the-united-states-rules-of-engagement-governing-the-use-of-airstrikes/.

"U.S. Department of Defense Releases Report of Investigation Finding That October 2015 Air Strike on Doctors Without Borders Hospital in Kunduz, Afghanistan, Was Not a War Crime." *American Journal of International Law* 110, no. 3 (2016): 579–87. https://doi.org/10.1017/S0002930000054038.

"US-Led Coalition in Iraq & Syria." Accessed April 25, 2020. https://airwars.org/conflict/coalition-in-iraq-and-syria/.

Wakabayashi, Daisuke, and Scott Shane. "Google Will Not Renew Pentagon Contract That Upset Employees." *The New York Times*, June 1, 2018, sec. Technology. https://www.nytimes.com/2018/06/01/technology/google-pentagon-project-maven.html.

Walsh, James Igoe. "Precision Weapons, Civilian Casualties, and Support for the Use of Force." *Political Psychology* 36, no. 5 (2015): 507–23. https://doi.org/10.1111/pops.12175.

Waltz, Kenneth Neal. *Theory of International Politics*. Long Grove, Ill.: Waveland Press, 2010.

Walzer, Michael. *Just and Unjust Wars: A Moral Argument with Historical Illustrations*. 4th ed. New York: Basic Books, 2006.

Wells, H. G, and Patrick Parrinder. *The War in the Air*. London; New York: Penguin, 2005.

Wilcox, Jason. "The Multi-Dimensional Value of Public Twitter Data for Real-Time Event Detection." Accessed April 21, 2020. https://www.dataminr.com/blog/the-multi-dimensional-value-of-public-twitter-data-for-real-time-event-detection.

Yeung, Serena, N. Lance Downing, Li Fei-Fei, and Arnold Milstein. "Bedside Computer Vision — Moving Artificial Intelligence from Driver Assistance to Patient Safety." *The New England Journal of Medicine* 378, no. 14 (2018): 1271–73. https://doi.org/10.1056/NEJMp1716891.