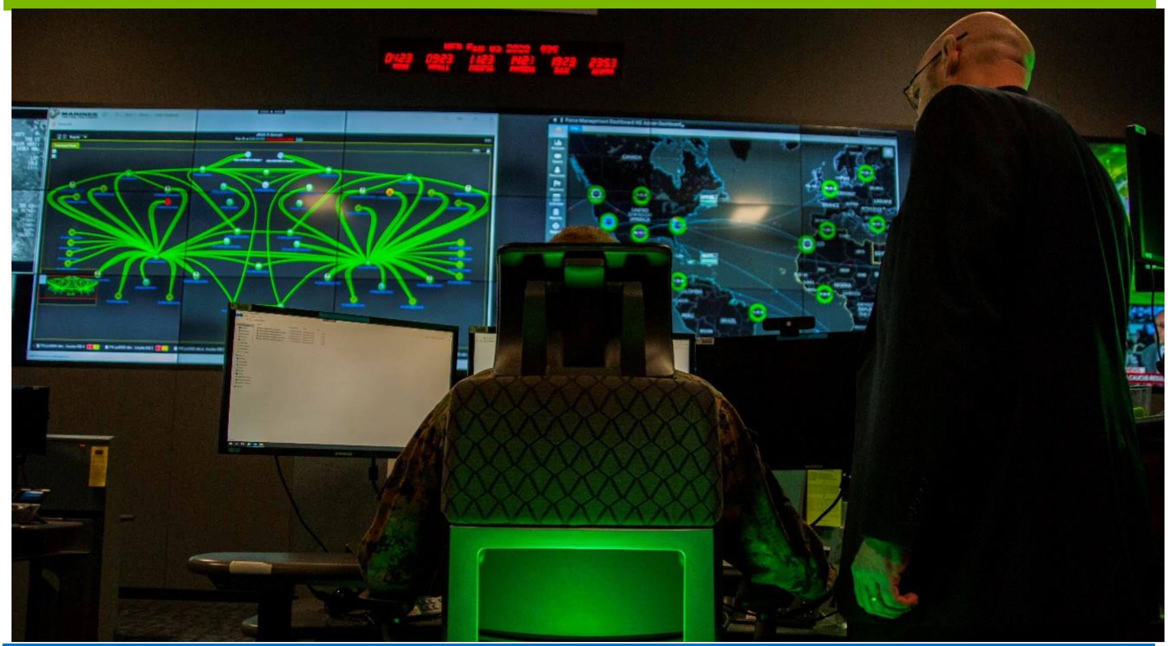
CNA

February 2022



(U) Leveraging AI to Mitigate Civilian Harm

Larry Lewis and Andrew Ilachinski

DOP-2021-U-030953-2Rev

Abstract

Over the last decade, there has been great debate about the introduction of AI to warfare. However, that debate has been primarily about how to make sure that AI applications are not indiscriminate in warfare. There is another important question, in light of international law and the principle of humanity: how can we use AI to protect civilians from harm? And how can we lessen the infliction of suffering, injury, and destruction overall through the use of this emerging technology? This report represents a concrete first step toward meeting this goal. We find that AI can be used to help address patterns of harm and thus reduce the likelihood of harm. We then discuss some potential areas of focus militaries could prioritize in order to reduce risks to civilians overall.

This work was performed under Federal Government Contract No. N00014-16-D-5003.

Cover image credit: Photo By: Marine Corps Staff Sgt. Jacob Osborne

Approved by: **February 2022**

Dr. David Broyles, Director
Special Activities and Intelligence Program
Operational Warfighting Division

Request additional copies of this document through inquiries@cna.org.

Executive Summary

Countries around the world have taken early steps to leverage artificial intelligence (AI) in military capabilities. Although militaries are seeking to leverage the technology of AI for greater effectiveness and efficiency, the idea of adapting AI to military applications has also created considerable controversy. Many concerns have been voiced, including potential bias and lack of fairness, and maintaining human judgment and responsibility in engagement decisions. That said, the chief concern in international discussions is whether military applications of AI could be inherently indiscriminate, unable to differentiate between valid military targets and civilians.

One way to answer this question is to look at specific military applications of AI, including autonomous systems, and examine both technical and operational considerations for how risks to civilians may arise and how they can be mitigated. For example, several presentations during the United Nations Convention on Certain Conventional Weapons meetings on lethal autonomous weapon systems featured examples of autonomous systems that could be used for military warfighting tasks in ways that complied with international law and did not represent an indiscriminate hazard to civilians. Similarly, a previous CNA report (*AI Safety: An Action Plan*) considered some additional military warfighting applications of AI and how risks to civilians from those applications could be minimized through both operational and technical mitigation steps.[1]

Those discussions, however, only address one half of the two-fold responsibilities for civilian protection found in International Humanitarian Law—the negative responsibility that militaries should not direct attacks on civilians. The affirmative responsibility for militaries to take all feasible precautions to protect civilians from harm has been relatively neglected. With regard to AI and autonomy, states should not only be asking how they can meet their negative responsibilities of making sure that AI applications are not indiscriminate in warfare. They should also be asking: ***How can we use AI to protect civilians from harm? And how can AI be used to lessen the infliction of suffering, injury, and destruction of war?***

This report represents a concrete first step toward answering these questions. We begin by framing the problems that lead to civilian harm. If we understand that AI is a tool for solving problems, before we understand how this tool can be used, we need to understand the

[1] Larry Lewis, AI Safety: An Action Plan for the Navy, CNA, DOP-2019-U-021957-1Rev, October 2019.

problems to be solved. What problems need to be solved to better protect civilians or otherwise promote IHL's principle of humanity?. Although the imperative for avoiding civilian harm is universally acknowledged, the specific mechanisms for how such harm occurs have never been characterized in detail. How does civilian harm occur?

After synthesizing our body of work on civilian harm—including analysis of several thousand real-world incidents of civilian harm from military operations—we answer this question, presenting a framework illustrating how civilian harm occurs. We then discuss how civilian harm can be mitigated, including a civilian protection life cycle, which demonstrates a comprehensive approach to mitigating harm. We also discuss some examples of specific mitigation steps that can be taken to reduce civilian harm to show the kinds of actions that are possible for meeting the goal of civilian harm mitigation.

We then present a model approach for identifying opportunities where AI could be used to help address the problem of civilian harm, using the civilian protection life cycle to illustrate potential actions. We find many opportunities for AI applications across the life cycle. This high volume of potential applications ought not surprise us because the problem of civilian harm may be viewed as a microcosm of actions, behaviors, and policies associated with the much larger military operational space overall.

We also discuss specific potential applications of AI that address risk factors we have observed in real-world operations, leveraging techniques that currently exist and in many cases have already been applied to other problems. Although we note that no solution will eliminate the problem of civilian harm—military operations will always have a non-zero risk to civilians—AI can be used to help address patterns of harm we observe and reduce the likelihood of harm. We then discuss some potential areas of focus states could prioritize to reduce risks to civilians overall.

For example, based on our analysis of particularly beneficial mitigation steps for reducing harm to civilians that are amenable to AI applications, we suggest the following functions as promising starting points:

- **Alerting the presence of transient civilians** by using object identification to automatically monitor for additional individuals around the target area and send an alert if they are detected. This application would bring these individuals to the attention of operating forces that may otherwise fixate on the target and miss transient civilian presence.
- **Detecting a change from collateral damage estimate** by finding differences between imagery used to determine the collateral damage estimate and more recent imagery taken in support of an engagement. This application can help identify little details that operating forces might not recognize but that could be cues of unanticipated civilian presence, such as additional vehicles near a building.

- **Alerting a potential miscorrelation** by helping to identify that a miscorrelation has taken place. For example, applications could recognize that a vehicle being tracked is not the same one that was being tracked previously, showing that a swap has occurred between a threat vehicle and a civilian vehicle.
- **Recognizing protected symbols** by using AI/machine learning methods to identify accepted symbols for designating protected objects (e.g., red cross or red crescent) and alerting the operator or the chain of command accordingly. The presence of protected symbols does not mean that the location is, in fact, protected from attack: the location may have lost its protection or an unscrupulous party may be using the symbol to deter attacks, in violation of international law. But this capability would provide a safety net in cases where the protected symbol is present but was missed by operating forces.[2]

Finally, by examining one tragic civilian harm incident in Afghanistan, we find we can draw from potential AI solutions from our AI applications matrix (including several of the abovementioned applications) to help address root causes in that incident and see how such solutions could help to possibly avert civilian harm. Although this is a validating step for our findings, we also note that much more work needs to be done in this area. This report is merely a first step in exploring a vast space of possibilities where details matter greatly. ***Governments, militaries, and academic institutions should be deliberate in developing AI solutions to mitigate harm to civilians, building on this foundation.***

What remains is a matter of will, which we acknowledge is uncertain. Although militaries speak of capabilities that help mitigate civilian harm, such as precision-guided munitions, those capabilities were acquired to engage military targets more effectively. Although militaries may have capabilities that help to mitigate harm in some contexts, militaries have not sought—or even recognized the need—to comprehensively develop capabilities to reduce risk to civilians from all the mechanisms we identify here. Therefore, the set of current capabilities held by militaries is incomplete: much more can be done, and existing risks are not always mitigated by capabilities that do exist. For example, a precision-guided munition has no value in mitigating civilian harm when civilians have been misidentified as a military target and are attacked in that mistaken belief.

In summary, we do not observe militaries around the world seeking to field capabilities based on their value in mitigating civilian harm. We have taken a first step to show how AI-enabled and other applications for reducing the cost of war on civilians are within the realm of the possible. It remains to be seen whether militaries will choose to pursue them.

[2] We note that the Australian Armed Forces have recognized this application as a promising one and have already conducted field experiments showing the utility of this function.

Contents

Introduction

The past decade has seen exponential progress in artificial intelligence (AI), defined as "the ability of a digital computer or computer-controlled robot to perform tasks commonly associated with intelligent beings."[3] AI is having a transformative effect on many areas of life, including commerce, medicine, and banking. Examples include the following:

- Amazon's logistics and delivery system, which uses machine learning (ML) AI to inform and optimize that process. This system includes an ML-driven decision on the best packaging type for each of many millions of products.[4]
- Medical diagnoses using medical imagery can be performed faster and in a repeatable and more economical way using deep learning (DL) AI methods.[5]
- Banks use AI/ML models to detect suspicious behavior and stop fraudulent transactions, leading to a significant reduction in banking fraud, the "biggest challenge for the financial industry."[6]

Countries around the world have taken early steps to leverage AI in military capabilities, including using AI for autonomous systems and functions, decision aids, and optimization in problem solving. For example, the US Department of Defense (DOD) has made serious and public commitments in the form of strategy, policy, organizational changes, and resources to leverage AI. It is also developing systems that leverage modern AI technologies. Examples include the following:

[3] *Encyclopedia Britannica*, s.v. "artificial intelligence," accessed Sept. 20, 2021, https://www.britannica.com/technology/artificial-intelligence.

[4] Amazon, "In the News: How Amazon Is Using Machine Learning to Eliminate 915,000 Tons of Packaging," Jan. 29, 2021, https://aws.amazon.com/blogs/industries/how-amazon-is-using-machine-learning-to-eliminate-915000-tons-of-packaging/.

[5] Ravi Aggarwal et al., "Diagnostic Accuracy of Deep Learning in Medical Imaging: A Systematic Review and Meta-Analysis," *npj Digital Medicine* 4, article 65 (Apr. 7, 2021), https://doi.org/10.1038/s41746-021-00438-z.

[6] Robin Trehan, "How AI Is Transforming Fraud Prevention in Banking and Finance," Deltec, May 11, 2020, https://www.deltecbank.com/2020/05/11/how-ai-is-transforming-risk-in-finance-and-banking/?locale=en.

- In 2017, DOD launched Project Maven, an ongoing effort to tap AI to help DOD analyze what was fast becoming an overwhelming amount of full-motion video being collected in support of counterinsurgency and counterterrorism operations.[7]
- DOD is conducting an initiative to leverage AI to dramatically speed up assessments needed to aid in humanitarian and disaster relief missions, including route analysis, damage assessment, flood water detection, and fire perimeter analysis. This effort has been supported by Johns Hopkins Applied Physics Laboratory, and applications have been tested by state National Guard units.[8]
- DOD is exploring the use of small autonomous, AI-enabled unmanned aerial vehicles (UAVs) for operation by special operations forces in indoor environments.[9]
- In 2020, the US Air Force demonstrated the use of AI on board a US military aircraft for what appears to be the first time. The manned aircraft—a U-2 Dragon Lady with an onboard pilot—used an AI algorithm to control the aircraft's sensor and navigation systems. The AI technology (which did not interact with the aircraft's flight controls) was specifically designed without a manual override.[10]

The US is only one of many countries making AI the center of its strategy. China, another world leader in the development of AI, is doing the same and pursuing AI for a wide range of national applications, including military applications. Examples include the following:

- AI for target identification, including identifying US military aircraft on runways (reporting an accuracy of 92 percent) and naval targets in maritime environments[11]
- AI-enabled capabilities in facial recognition, (Chinese language) textual analysis, and the analysis and synthesis of massive troves of surveillance data, which the country

[7] Cheryl Pellerin, "Project Maven to Deploy Computer Algorithms to War Zone by Year's End," DOD website, July 21, 2017, https://www.defense.gov/Explore/News/Article/Article/1254719/project-maven-to-deploy-computer-algorithms-to-war-zone-by-years-end/.

[8] Luke Strum, "Intel Airmen Sharpen AI Technology for Domestic Response," Air National Guard website, Nov. 25, 2019, https://www.ang.af.mil/Media/Article-Display/Article/2025332/intel-airmen-sharpen-ai-%20technology-for-domestic-response/; Department of Defense, "Mission Initiatives: Threat Reduction and Protection," JAIC website, https://www.ai.mil/mi_threat_reduction_and_protection.html.

[9] Department of Defense, "Mission Initiatives: Joint Warfighting Operations," JAIC website, https://www.ai.mil/mi_joint_warfighting_operations.html.

[10] Aaron Gregg, "In a First, Air Force Uses AI on Military Jet," *Washington Post*, Dec. 16, 2020, https://www.washingtonpost.com/business/2020/12/16/air-force-artificial-intelligence/.

[11] Alex Barker, "Giving Precision Munitions 'Eyes' and a 'Brain': The State of PLA Research on Military Target Recognition," *China Brief*, 21, no. 13 (Jul. 2, 2021), https://jamestown.org/program/giving-precision-munitions-eyes-and-a-brain-the-state-of-pla-research-on-military-target-recognition/.

has been able to develop and refine through extensive surveillance of its domestic population for domestic purposes[12]

- AI-enabled unmanned systems, including patrol boats and swarms of armed drones[13]

Similarly, Russia is making AI a major priority for modernization of its forces, seeing innovation as essential to its defense and status as a great power. Specific efforts include the following:

- Experimentation with unmanned, AI-enabled small tanks and robotic exoskeletons for soldiers that will allow troops to carry more weapons and equipment[14]
- Using AI to automate processes and command and control functions such as collection, processing, storage, and delivery of information necessary to optimize command and control of troops and weapons
- Employing AI on the Su-35S, a heavy long-range fighter, to support pilot decision-making for target acquisition and combat maneuver[15]

These are only a few of the military AI applications being explored by many countries around the world.

Concerns regarding civilian harm risks

Although many militaries are seeking to leverage the technology of AI, the idea of adapting AI to military applications has also created considerable controversy. The most deliberate debate on this issue concerns lethal autonomous weapon systems (LAWS). Based on expressed concerns, in 2014 the United Nations (UN) Convention on Certain Conventional Weapons (CCW) first met for informal discussions on LAWS. This body has now spent almost a decade discussing the ethical, legal, and operational considerations of LAWS, including whether weapon systems operating autonomously (without a human operator) should be allowed to

[12] C. Todd Lopez, "Where It Counts, U.S. Leads in Artificial Intelligence," DOD website, July 9, 2020, https://www.defense.gov/Explore/News/Article/Article/2269200/where-it-counts-us-leads-in-artificial-intelligence/; Yasmin Tadjdeh, "China Threatens U.S. Primacy in Artificial Intelligence (UPDATED)," *National Defense*, Oct. 31, 2020, https://www.nationaldefensemagazine.org/articles/2020/10/30/china-threatens-us-primacy-in-artificial-intelligence.

[13] Yasmin Tadjdeh, "China Threatens U.S. Primacy."

[14] Margarita Konaev and Samuel Bendett, "Russian AI-Enabled Combat: Coming to a City near You?" War on the Rocks, July 31, 2019, https://warontherocks.com/2019/07/russian-ai-enabled-combat-coming-to-a-city-near-you/.

[15] Jeffrey Edmonds et al., *Artificial Intelligence and Autonomy in Russia*, CNA, DRM-2021-U-029303-Final, May 2021.

use lethal force.[16] Although autonomy is the AI application most debated in the international community, there is growing awareness that other applications—decision aids and optimized functions—can also carry risk.

What are the risks involved? Many concerns have been voiced, some valid (e.g., potential bias and lack of fairness, maintaining human judgment and responsibility in engagement decisions) and some less valid (e.g., the fear of robots taking over the world) for the current state of AI technology. However, the chief concern in international discussions is whether military applications of AI could be inherently indiscriminate, unable to differentiate between valid military targets and civilians.

One way to answer this question is to look at specific military applications of AI, including autonomous systems, and examine both technical and operational considerations for how risks to civilians may arise and how they can be mitigated. Several presentations during the UN CCW meetings did indeed feature examples of autonomous systems that could be used for military warfighting tasks in ways that complied with international law and did not represent an indiscriminate hazard to civilians. Similarly, a previous CNA report considered some additional military warfighting applications of AI and how risks to civilians from those applications could be minimized through both operational and technical mitigation steps.[17]

Opportunity for using AI to mitigate civilian harm

Another question should be asked because of the nature of international humanitarian law (IHL). States established IHL to legally obligate them and their armed forces to standards of conduct in armed conflict, with particular emphasis on the protection of civilians. All states took on these obligations willingly because they recognized the moral and strategic importance of doing so. IHL includes both affirmative responsibilities (e.g., militaries should take all feasible precautions to protect civilians from harm) and negative responsibilities (e.g., militaries should not direct attacks on civilians). IHL also affirms the commitment of states to

[16] The CCW is properly referred to as the Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to Be Excessively Injurious or to Have Indiscriminate Effects.

[17] Larry Lewis, AI Safety: An Action Plan for the Navy, CNA, DOP-2019-U-021957-1Rev, October 2019.

the principle of humanity, which "forbids the infliction of all suffering, injury or destruction not necessary for achieving the legitimate purpose of a conflict."[18]

Thus, regarding AI, including autonomous functions, although states should be asking how to meet their negative responsibilities of ensuring that AI applications are not indiscriminate, they should also be asking how to better fulfill their self-determined positive obligations, namely: How can we use AI to protect civilians from harm? And how can we lessen the infliction of suffering, injury, and destruction overall?

We note that CNA is not alone in making this observation. After we began this project, Australia, Canada, Japan, South Korea, the United Kingdom (UK), and the United States submitted a joint paper to the UN CCW that included a similar recommendation to seek answers to these questions. Specifically, the paper recommended the Group of Government Experts on Lethal Autonomous Weapon Systems shift its focus in 2022 to include "identifying examples of ways in which emerging technologies in the area of LAWS could be used to reduce the risks to civilians in military operations."[19] This report represents a concrete first step in meeting this goal.

Approach

If we understand that AI is a tool for solving problems, before we understand how this tool can be used, we need to understand the problems to be solved. What are the problems that need to be solved to better protect civilians or otherwise promote IHL's principle of humanity? Although the imperative for avoiding civilian harm is universally acknowledged, the specific mechanisms for how such harm occurs have never been characterized in detail. How does civilian harm occur? After synthesizing our body of work on civilian harm—including analysis of several thousand real-world incidents of civilian harm from military operations—we produce a framework for how civilian harm occurs. We also discuss some specific risk factors observed in recent US and coalition operations.

We then discuss some examples of specific mitigation steps that can be taken to reduce civilian harm to show what kinds of actions are possible for meeting the goal of civilian harm

[18] International Committee of the Red Cross, *International Humanitarian Law: Answers to Your Questions*, June 9, 2020. This principle of IHL stems in part from the Martens Clause, discussed here: Theodor Meron, "The Martens Clause, Principles of Humanity, and Dictates of Public Conscience," *The American Journal of International Law* 94, no. 1 (Jan. 2000).

[19] Australia, Canada, Japan, South Korea, the United Kingdom, and the United States, *Discussion Paper – Building on Chile's Proposed Four Elements of Further Work for the Convention on Certain Conventional Weapons (CCW) Group of Governmental Experts (GGE) on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems (LAWS)*, submitted to the UN CCW June 2021.

mitigation. We then present a framework for how AI could be used to help address the problem of civilian harm. Finally, we discuss some specific applications of AI that address some risk factors we have observed, illustrating some potential areas of focus states could prioritize to reduce risks to civilians overall.

Understanding Civilian Harm

The tragedy of civilian harm has always been a feature of warfare, and over time states have sought to limit this harm, including the development of IHL. However, over the past several decades, we have come to a better understanding of civilian harm and a more complete view of how to mitigate this harm. CNA has been at the forefront of this effort, applying the scientific method to the problem of civilian harm through analysis of real-world civilian harm incidents. We begin this chapter by describing CNA's body of work on civilian harm, and then we discuss key insights into the mechanisms that cause civilian harm, including a framework for understanding these causes. Finally, based on our identified causes, we discuss specific solutions for mitigating civilian harm, which can serve as a starting point for considering potential AI applications.

CNA's work on civilian harm

Our work on civilian harm began through investigation of a different, but related, operational problem: friendly fire. In Operation Desert Storm in 1991, a significant fraction of US casualties were caused by friendly fire, where US forces were killed or wounded by US engagements. Recognizing friendly fire to be a significant problem, DOD created the Joint Air Defense Operations Joint Engagement Zone Joint Test and Evaluation Activity in an effort to develop capabilities and tactics to help US forces to operate together better and more safely in a common battlespace. This activity was sustained and renamed the All-Service Combat Identification Evaluation Team and later renamed again as the Joint Combat Identification Evaluation Team, with activities including a wider variety of missions and domains.[20]

Developing a methodology for understanding civilian harm

CNA supported these activities by examining challenges in combat identification, including reconstruction and longitudinal analysis of constructive friendly fire incidents occurring in multiple live events. This methodology yielded several findings that had not previously been identified. When the US began Operation Iraqi Freedom in 2003, CNA employed the same methodology for actual friendly fire incidents, using all available information, including investigation reports, operational data (including digital data from combat systems and data links), and media reports to give insights about the root causes of friendly fire. In addition, one

[20] Larry Lewis, *Insights for the Third Offset,* CNA, DRM-2017-U-016281-Final, Sept. 2017.

of the friendly fire incidents resulted in civilian harm because civilians were in the proximity of friendly forces. CNA's detailed reconstruction of this incident resulted in the identification of several causal factors that were not understood before. For example, the incident featured a misassociation: the pilot wrongly associated information regarding the threat with the location of civilians and friendly forces, leading the pilot to engage the wrong location and cause civilian and friendly casualties. These new insights suggested our methodology's value for understanding and addressing the problem of civilian harm.

Applying the methodology: Iraq and Afghanistan

As insurgencies developed in Iraq and Afghanistan (Operation Enduring Freedom, launched in 2001), the US shifted from major combat operations to counterinsurgency, an approach for which it was largely unprepared. With civilian protection as a central feature of counterinsurgency and the added identification challenges for a military with an irregular threat that does not wear uniforms or other identifying information, civilian harm became a central challenge. CNA analyzed civilian harm in both conflicts, as detailed below.

Iraq: Analysis of escalation of force

We first examined civilian harm in Iraq, where civilian casualties (civilian harm) were primarily caused by escalation of force incidents. These incidents occurred both at military checkpoints and during convoy operations, where military forces engaged vehicles and individuals who appeared to them to be threatening. A sharp rise in civilian casualties from this cause prompted the first military-led tracking of civilian harm: Multinational Forces-Iraq began tracking civilian harm in 2004 to understand the scope of the issue and try to identify solutions. Tracking indicated that US forces caused more than 500 civilian harm from escalation of force incidents in the first half of 2005. In mid-2005, prompted by senior leader direction to mitigate civilian harm, US forces changed their tactics and procedures to reduce risks to civilians, resulting in a significant drop in civilian harm from escalation of force. After this drop, CNA analyzed this civilian harm data and assessed risk factors associated with escalation of force incidents, such as the risk of misidentification and the tendency to engage perceived threats at distances far exceeding those needed for self-defense considerations—risk factors that unfortunately would be seen again in Afghanistan.[21]

Afghanistan: Real-time support

The problem of civilian harm also became a strategic issue in US and international force operations in Afghanistan. Afghan leaders and the international community expressed alarm over escalating numbers of US- and coalition-caused civilian harm in Afghanistan between

[21] Larry Lewis, *Reducing and Mitigating Civilian Casualties: Enduring Lessons*, Joint Staff, Apr. 12, 2013.

2006 and 2009, and US leaders saw the issue of civilian harm becoming a strategic issue in the campaign. For example, the then-commander of US Central Command described civilian harm as becoming a "toxic" issue that threatened the meeting of US and international strategic objectives and strengthened the support of antigovernment elements.

This realization drove a practical approach to the problem of civilian harm in Afghanistan. Like in Iraq, the process began with data: in late 2008, US and international forces began tracking civilian harm. This tracking effort was originally intended to counter external allegations for the purpose of public affairs, but the process of gathering this data became a foundation for a more evidence-based approach to mitigating civilian harm. In 2009, a US military lessons-learned organization effort led by CNA analyzed civilian harm incidents in the midst of operations. We used operational records, combined with open source and civil society information when available, to discover patterns and mechanisms of civilian harm. This analysis yielded new insights that enabled practical, focused steps military forces could take to reduce the risk to civilians in operations in Afghanistan. Changes made based on this analysis included a modified International Security Assistance Force (ISAF) tactical directive governing the conduct of air strikes and new guidance for escalation of force at checkpoints, among many other measures.

Subsequent analysis showed that these steps were effective in reducing civilian harm, with a reduction of 20 percent within the first year and continued reductions thereafter. The rate of civilian harm per operation, representing the relative risk of civilian harm during operations, also decreased when such rates could be measured. Consequently, operations were less likely to cause civilian harm than they were before these changes. This practical approach to civilian harm mitigation included (1) tracking civilian harm, and (2) monitoring risks and adapting to mitigate them. CNA supported US and international forces in both these steps.

Tracking civilian harm

We noted above that the US began tracking incidents of civilian harm in Iraq in response to growing attention to civilians being killed by US forces at checkpoints through escalation of force. However, this tracking was a temporary measure that focused on one particular type of operation: checkpoints. In Afghanistan, ISAF began tracking civilian harm incidents comprehensively for all types of operations. In 2009, the tracking was resourced to make it more robust, and it started being used to promote operational learning. Tracking consisted of a spreadsheet with date, time, location of incident, unit involved, type of operation, number of civilian harm, and other details that could be used for consequence management and trend analysis. This tracking became a best practice replicated in later US conflicts. For example, counterterrorism forces and the US-led counter-ISIS coalition both tracked civilian harm from their operations.

In addition to aiding in consequence management (e.g., determining whether medical care needed to be provided, whether the US should apologize for inadvertent civilian tolls, or whether amends such as compensation were appropriate) and communications (e.g., gathering the latest information regarding civilian tolls from a particular incident), this tracking created a way to factor in contributions from external organizations and individuals with relevant information. Tracking of civilian harm was by no means perfect (as we will discuss later), but it still provided the foundation for learning and adaptation. CNA periodically reviewed the civilian harm tracking cell's database containing details of individual incidents and made corrections and improvements to improve the accuracy of the information that served as a foundation for mitigation efforts.

Monitoring and responding to trends

A second key aspect to addressing the risk to civilians is monitoring metrics that capture that risk and responding to them. In addition to the number of civilian harm, another important measure is the rate of civilian harm—the number of strikes causing civilian harm divided by the total number of strikes. This measure represents the relative risk of civilian harm from military operations. Monitoring the rate over time allows for a better understanding of how the relative risk to civilians is changing, enabling the possibility of early focused interventions in response to emerging and troubling trends.

To support this process, CNA developed metrics to monitor the level of risk to civilians from military operations and how this risk changed over time. Part of our support to ISAF was monitoring trends each month with ISAF-provided data, which enabled an opportunity to respond to risks to civilians in real time. For example, in January 2011, we noticed worsening trends in the civilian harm rates for several types of operations. After alerting ISAF to these trends and the factors behind them, the headquarters rapidly made operational changes that addressed the causes of those trends. As a result, the rates and numbers of civilian harm went back down.[22] This best practice of monitoring and responding to emerging trends was later written into US national policy for civilian harm.[23]

[22] Larry Lewis and Diane Vavrichek, *Rethinking the Drone War*, (Quantico, VA: Marine Corps University Press, 2016).

[23] The White House, July 1, 2016, *Executive Order--United States Policy on Pre- and Post-Strike Measures to Address Civilian Casualties in U.S. Operations Involving the Use of Force*. The commitment to monitor and address trends is contained in Section 4 of the Executive Order. We note that the rate of CIVCAS is dependent on many factors, such as operating environment, adversary tactics that purposely endanger civilians, and type of operation (e.g., air strikes, artillery fire, ground operation), so it is not necessarily a poor reflection on a military force if the rate increases over time or if the average rate for one operation is higher than that for another operation. But at the very least, a military force should be aware of and understand this information.

One example of identifying and responding to trends was our work on identifying and recommending steps to address specific risks of civilian harm during air strikes—a particular area of concern because air strikes tend to cause the most civilian harm per incident compared to other types of force (e.g., small arms fire, artillery fire). We discuss here three areas of risk to civilians we identified and the adaptations that were made over the course of operations in Afghanistan: gatherings of people, self-defense considerations, and double taps.

In Afghanistan, commander's guidance (in the form of tactical directives) emphasized the need to reduce civilian harm, including mitigating harm in light of the risk of unobserved civilians being in buildings. In 2009, that guidance read: "I expect leaders at all levels to scrutinize and limit the use of force like close air support against residential compounds and other locations likely to produce civilian casualties in accordance with this guidance."[24] In our analysis of civilian harm incidents in 2009 to 2010 that fell under that guidance, a pattern emerged: although forces exercised greater care in air strikes of compounds, several air strikes still targeted gatherings of people not inside structures, and the commander's intent was not being applied consistently to those strikes. Subsequent guidance made this case clearer, and as a result, civilian harm from air strikes were reduced because forces adapted to mitigate this risk.

Another common risk contributing to civilian harm occurred when strikes were made in self-defense when military forces were under fire and calling for air support. Analysis of these cases revealed a pattern: although strikes began because of an urgent self-defense situation, attacks sometimes continued after the self-defense situation no longer existed, even though those attacks were not first approved under the rules of engagement governing cases other than self-defense, which require more careful consideration of collateral damage concerns. Guidance was provided stating that self-defense engagements should seek to address only the self-defense situation and should not shift to an offensive mission against identified combatants without prior approval. As a result, civilian harm from air strikes were reduced.

civilian harm were also caused at times as part of the practice of "double taps.". Some news reports alleged that air strikes would first target combatants but then subsequently target civilian first responders rushing to aid the survivors. In actuality, the second air strikes were based on assessments of the first strike not being mission effective, and first responders were harmed because of an inadequate collateral damage estimate associated with the subsequent strike. Additional guidance on these follow-on attacks, accompanied at times by additional surveillance capabilities to aid collateral damage estimates, mitigated the risk to civilians in those cases.

[24] NATO/ISAF, Tactical Directive, July 6, 2009.
https://www.nato.int/isaf/docu/official_texts/Tactical_Directive_090706.pdf.

Other assessments

Soon after we completed our work in Afghanistan supporting operations in real time, we conducted two assessments on civilian harm. The first examined specific risks to civilians from the use of drones, and the second identified overarching lessons from US operations for mitigating civilian harm.[25] These lessons informed CNA's work in drafting two sections of the 2016 Executive Order on civilian harm.[26]

CNA also led the analysis of operational data for several additional assessments, including the 2018 Joint Staff civilian harm Review and more recent assessments examining US- and coalition-caused civilian harm in Mosul, Iraq; Raqqa, Syria; Afghanistan; and Somalia.[27] In addition, we examined civilian harm in Yemen caused by the Saudi-led coalition in Yemen.[28] We have also assessed risks to specific populations or groups, such as children and health care workers.[29] Collectively, we have analyzed about 2,000 real-world civilian harm incidents.

Root cause determinations: Why civilian harm happens

As we consider the concept of using AI to mitigate civilian harm, a key limitation is that AI is a tool for solving specific, well-defined problems. So, the problem of civilian harm must be well characterized to apply AI effectively. This limitation is not unique to AI and also applies to militaries: we have observed that military forces are effective in mitigating civilian harm only if the force understands why civilian harm occurs in the first place. For example, in Afghanistan between 2006 and 2008, the US military and international forces attempted to institute measures to reduce civilian harm, but the number of civilian harm incidents continued to climb. Although mitigation steps were attempted, they were not effective. Our analysis of civilian harm explained why these military measures were not working: they were based on faulty assumptions about how civilian harm happens.

[25] Larry Lewis, *Drone Strikes: Civilian Casualty Considerations*, Joint Staff, June 18, 2013; Lewis, *Reducing and Mitigating Civilian Casualties*.

[26] The White House, July 1, 2016, *Executive Order--United States Policy on Pre- and Post-Strike Measures*.

[27] Department of Defense, *Joint Staff CIVCAS Review*, Apr. 17, 2018, https://www.justsecurity.org/wp-content/uploads/2019/02/Civ-Cas-Study-Redacted-just-security.pdf.

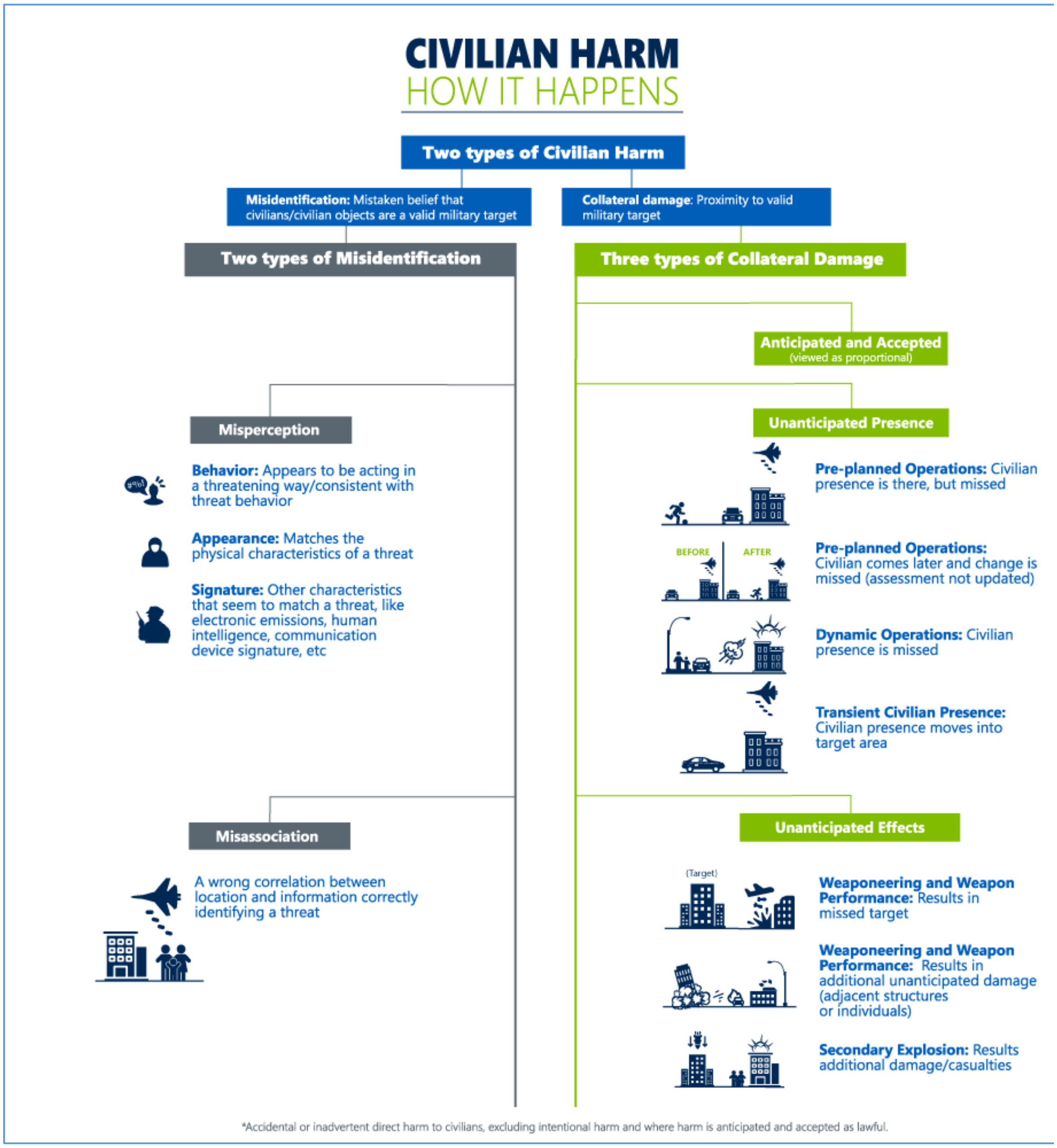
[28] Larry Lewis, *Learning from Yemen*, CNA, IRM-2019-U-019749-Final, May 2019, https://www.cna.org/CNA_files/PDF/IRM-2019-U-019749-Final.pdf.

[29] For example, see Larry Lewis, *Protecting Medical Care in Conflict: A Solvable Problem*, CNA, Oct. 30, 2019, https://humanrightscommission.house.gov/sites/humanrightscommission.house.gov/files/documents/Protecting%20Medical%20Care%20in%20Conflict%20-%20Lewis.pdf.

For example, in Afghanistan, the US military was operating under the assumption that when civilian harm occurred, the military was engaging a valid military target and civilians in the area were harmed in the engagement. We call this a "collateral damage" mechanism. But analysis of real-world incidents showed that about half the time, another mechanism was in play: the military identified what it believed was a valid military target and engaged it in that belief, but the target was, in fact, civilians. This "misidentification" problem is fundamentally different from the "collateral damage" problem and demands different kinds of solutions.

To better define the specific problems that AI applications should seek to solve, we performed a meta-analysis of the thousands of incidents we have examined, defining the specific mechanisms that led to civilian harm over the last decade and a half, including in Afghanistan, Iraq, Syria, Somalia, and Yemen. Based on this meta-analysis, we developed a graphic, shown in Figure 1, illustrating the different mechanisms that can cause civilian harm.

Figure 1. Mechanisms for civilian harm



Source: CNA.

As we stated above, there are two main causes of civilian harm: misidentification and collateral damage.[30] Each main cause has several variations. For example, misidentifications can be either from a *misperception* of civilians as being a threat based on appearance, behavior, or other information, or from a *misassociation*, where information about a valid military target is wrongly ascribed to civilians, resulting in the belief that they are, in fact, the target.

Civilian harm through collateral damage can also be from a variety of causes. First, such harm may have been factored into the proportionality analysis. However, most civilian harm we have observed was from factors that were unanticipated. One kind is *unanticipated presence*, where civilian presence is not observed until after the engagement. This can happen in a few ways. For preplanned operations, civilians may have been present but were missed in the collateral damage estimate. For example, civilians may be present in buildings or vehicles but not observed by military forces or sensors. Alternately, no civilians may have been present at the time the collateral damage estimate was performed, but civilians moved into the area later, unobserved. For dynamic operations, where the engagement process is compressed and often lacks the planning and formalized target approval process of deliberate engagements, it is also possible to miss the presence of civilians. Also, whether the engagement is preplanned or dynamic, civilian harm can occur when civilians move into the target area around the time of the engagement. We call this *transient civilian presence*.

The final type of collateral damage is due to *unanticipated effects*, of which we see three kinds. The first is weaponeering and weapon performance issues that lead to the target being missed, resulting in civilian harm at the affected location. Another is weaponeering and weapon performance issues that affect the intended area but also have unanticipated effects on adjacent areas or structures where civilians are present. Finally, secondary explosions can lead to harm to civilians outside the range of effects of the weapon itself.

A civilian protection life cycle for mitigating risks

Now that we have defined specific mechanisms for how civilian harm can occur, we can start to consider possible steps to mitigate the risk of harm. In our work, we have found that mitigating risk to civilians should consider far more than the "trigger pull" engagement decisions. Rather, we find civilian harm mitigation is strengthened through an adaptive approach to observing patterns and taking steps to mitigate them. Although individual

[30] Here we are referring to direct harm to civilians that is accidental or inadvertent. Although in some cases civilian harm can be accepted as lawful and proportional in an engagement of a valid military target, most cases we have reviewed did not anticipate civilian harm in the engagement decision.

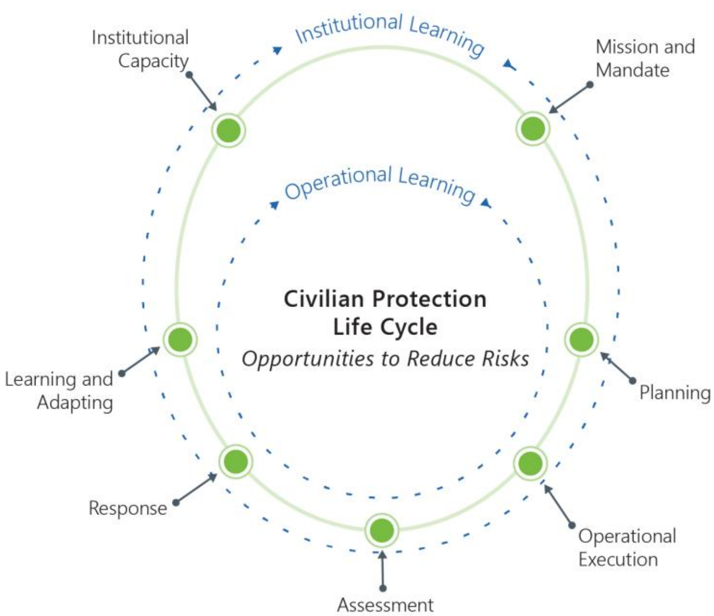
practices and policies can be beneficial for protecting civilians, the most impact is achieved through a comprehensive approach we refer to as the civilian protection life cycle (CPLC). This life cycle reflects attention to civilian protection at all points in the planning and use of military force and includes learning loops so that militaries can adapt and improve to overcome challenges. The CPLC, illustrated in Figure 2, consists of the following elements:

- **Mission and Mandate.** This element involves designating and allocating needed capabilities and authorities to conduct operations in ways that consider the protection of civilians from the beginning. For example, the mission can be shaped to mitigate risks to civilians, capabilities to support civilian harm mitigation (CHM) requirements can be allocated, and authorities and rules of engagement can be tailored to mitigate risks to civilians.
- **Planning.** At strategic down to tactical levels, this element involves conducting planning and developing command guidance that factors in risks to civilians and includes feasible steps and alternatives to help mitigate them. Examples include shaping operations to mitigate specific risks to civilians, developing tactical alternatives to avoid established patterns of harm, tailoring and adapting command guidance to better address patterns of harm, conducting mission rehearsals that emphasize mitigation measures, allocating capabilities for CHM when appropriate, and determining pattern of life for civilians.
- **Operational Execution**. This element involves performing targeting processes that promote accurate identification and delivery of lethal effects while seeking ways to minimize civilian harm and reverberating effects. Examples include exercising tactical patience, deliberate weaponeering to mitigate civilian risks, coordinating details to ensure a common target and avoid misassociation, and screening for transient civilian presence.
- **Assessment.** This element involves considering all available information, with internal and external sources, to determine the best estimate of civilian harm caused by the use of force. Examples include assessing battle damage to include effects on civilians and monitoring for potential civilian harm incidents.
- **Response.** This element involves working to mitigate the tragic consequences of civilian harm to affected individuals and populations, including the provision of urgent medical care, making amends to victims, and acknowledgment and apology when appropriate.
- **Learning and Adapting.** This element involves using assessments, including patterns of harm and trend data, to identify operational refinements to better protect civilians. These assessments also identify institutional requirements that can help address observed challenges. Examples include giving insights to enhance current mitigation

efforts, sharing lessons and data for effective learning, and informing needed institutional improvements.

- **Institutional Capacity.** This element involves designing the force to reduce risks to civilians and addressing observed challenges and requirements across the military institution (e.g., doctrine, training, materiel solutions) to strengthen the ability to mitigate harm to civilians over time.

Figure 2. Civilian protection life cycle



Source: CNA.

The CPLC also includes two learning loops: operational learning, where assessments of causes and trends directly inform the improvement of operational practices and policies within the context of an ongoing operation, and institutional learning, where assessments of challenges and requirements inform needed changes to, for example, doctrine, policy, organization, training, and leadership, together with equipment and facilities.

Examples of practical steps for mitigating civilian harm

Militaries can take various practical steps to mitigate civilian harm more effectively. Some are universal, whereas others are particular to a specific mechanism of civilian harm. For example, three best practices can potentially be applied to all operations:

- **Tactical patience.** Tactical patience involves taking time to verify positive identification and understanding of the operating environment before attacking, when the situation allows.
- **Tactical alternatives**. Tactical alternatives involves considering different options for achieving desired effects in view of potential second-order effects.
- **Shaping.** Shaping involves planning and maneuvering forces to reduce the likelihood of a situation in which significant force in the presence of civilians might be required.

We discuss each of these best practices below.

Tactical patience can be applied in a range of environments and missions when it is consistent with meeting mission objectives and following self-defense considerations. Some situations require an immediate use of force: for example, a high-priority target is in danger of being lost or military forces are in a self-defense situation against an immediate threat. In other situations, there is time for further consideration, and forces can use tactical patience as a precautionary measure. Real-world examples of tactical patience include the following:

- **Two pen flares.** A soldier at a checkpoint aims a warning pen flare at a car that has not heeded earlier verbal and visual warnings. The car continues toward the soldier. Noting hazy weather that could hinder visibility, the soldier decides he has time to fire another pen flare instead of resorting immediately to firing at the car as a threat. The car driver sees the second pen flare and stops, averting civilian harm.
- **Children in the road.** An attack helicopter observes two individuals digging in a road and believes they could be laying an improvised explosive device along a road that military forces often travel. Because there is no pressing need to attack immediately, the helicopter repositions to view the scene from a different vantage point. Viewed from this different perspective, the two individuals are clearly children digging in the road. The helicopter does not fire.

Likewise, forces can plan for and employ *tactical alternatives* to mitigate risks to civilians from their operations. In response to command emphasis on being effective and yet sparing civilians in Afghanistan, forces there actively sought to find solutions that presented fewer risks to civilians in their missions while preserving the success of the mission and the safety of the force. For example, one unit switched its approach from conducting raids to catch enemy forces to conducting census operations in partnership with local forces, building better understanding of the local situation while also culling out combatants hiding within the population. Another unit shifted from using air strikes for close air support when under fire to using pre-positioned snipers to reduce the risk of civilian harm that could occur during air strikes on buildings.

In *shaping*, forces consider civilian harm risks in planning, equipment to employ, and the placement and movement of forces to reduce those risks when possible. This practice can be as easy as reconsidering the placement of a checkpoint, for example moving a checkpoint positioned around a curve to a location with more visibility to allow more reaction time and enable tactical patience. This practice can also involve shaping the entire concept of operation to better address risks to civilians, such as bringing in munitions or surveillance assets better suited for the threat and environment or allowing fighters to leave an urban area to reduce the intensity of fighting in highly populated areas and pursuing them later in other locations.

Some mitigation steps are particular to specific mechanisms. For example, for checkpoint operations, civilian harm occurred primarily because of misidentifications based on behavior: civilians could appear to behave in threatening ways or in ways consistent with the behavior of the anticipated threat. Vehicles did not respond to warnings to slow down as they approached military forces, often because typical warning methods (e.g., waving and warning) were not very effective in catching the attention of drivers in dusty, limited-visibility conditions. Human factor failures also contributed, such as when military forces used laser dazzlers intended to warn drivers but the color (green) was interpreted by civilians as a signal to proceed, so they were fired upon when they did not stop.

Finding ways to more effectively communicate with civilian drivers reduces the problem of misidentification based on behavior. This communication could include dedicated signs, rumble strips on the road to alert drivers, and preemptive communication with local populations regarding the procedures and expectations for checkpoints. Understanding this mechanism could also avoid nonoptimal mitigation steps: for example, military forces in Afghanistan discontinued warning shots because of a few instances when those shots led to civilian harm, but analysis showed that removing this step deprived forces of one of their only means of warning civilians and led to a net increase in civilian harm.

Similarly, analysis of artillery fire showed how existing doctrine is geared toward high-end conflict, where considerations of placement of artillery focus on the risk of friendly fire. Thus, artillery rounds are fired to first overshoot the target in the direction opposite friendly forces and to then adjust fires back towards the intended target. Although this approach makes sense in a battlefield absent of civilians, in an urban setting, it can introduce significant risk to civilians. This approach represents civilian harm through collateral damage due to the unnoticed presence of civilians in the area of fire, which is exacerbated by the tactic developed to mitigate friendly fire but not to consider the presence of civilians. A better approach, in the absence of friendly forces in proximity, is to first look for signs of potential civilian presence and then place initial fires in the opposite direction.

Potential civilian harm mitigation measures for AI

We have discussed the potential mechanisms whereby civilian harm occurs and provided a framework for how mitigation steps can be introduced to reduce risks to civilians. But not all these steps are amenable to AI-enabled applications. For example, considering tactical alternatives and providing shaping of the operational environment are both highly complex, open-ended tasks that are not well suited for the powerful but narrow nature of tasks that AI can perform. However, tactical patience—pausing an engagement if certain risk factors are present or if certain requirements are not met—is possibly a function that an AI-enabled application could perform. The next sections describe some other potential mitigation measures that we see as significant contributors to the problem of civilian harm that are defined well enough that an AI application could potentially be successful.

Detecting transient civilian presence

Transient individuals or vehicles moving into the engagement area undetected was one of the leading causes of civilian harm in recent US operations. In these cases, a person or vehicle moved into the field of view and there was insufficient time to abort the attack or steer the weapon into a safer area by the time the operator noticed. For example, when using drones, the operators often zoom in at the last moments before the engagement, increasing the chances of transients. This phenomenon was also a problem in earlier operations in Afghanistan, where forces developed tactics to try to reduce these occurrences: the operator would switch the full-motion video resolution to a wider field of view before the engagement decision to better detect transient civilian presence.

Recognizing protected symbols

Hospitals and other humanitarian groups are protected from attack per international humanitarian law, but they still are attacked all too often, sometimes because militaries fail to recognize them for what they are. These sites often use symbols to visually communicate their protected nature. These symbols can include the red cross, the red crescent, and the blue cross (for historical/cultural sites). Human operators can fail to observe these symbols for multiple reasons: they may miss them in their focus on a perceived threat or they may be viewing through an infrared targeting pod that makes it harder to clearly differentiate colored objects.

Developing a robust civilian pattern of life

For deliberate, preplanned attacks, militaries often conduct a pattern of life assessment and a collateral damage estimate using available imagery. If the attack is not conducted until days, weeks, or longer after the assessment, the conditions on the ground can change. Also, pattern of life assessments are inherently limited. For example, civilian activity may not be evident in the slice of time of the assessment. Also, these assessments tend to factor in threat locations but not civilian or humanitarian locations, tending to sway determinations toward a threat instead of balancing them with other available information regarding civilian and humanitarian entities.

Improving collateral damage estimates

For militaries that have a formal collateral damage estimation process, this process gives an idea of the potential civilian toll that can be expected from a particular attack. But these estimates can suffer from severe shortfalls. One shortfall is that the collateral damage estimate represents a discrete snapshot in time. What if things change between the time of the estimate and the time of the engagement? There is no standard process or tool that evaluates whether a change has occurred that may affect the projected civilian toll. Another shortfall is that these estimates are based on a standard statistical model, whereas the pattern of life of civilians in conflict areas can be significantly disrupted from steady-state conditions. And it is not safe to assume that civilians will simply flee conflict areas. Confirmed civilian harm incidents involving families hiding in buildings in areas of heavy fighting show that civilians may be present even when military forces believe otherwise. Notably, models for collateral damage estimation are not calibrated and refined by actual operational results.

Detecting misassociations

One contributing factor to civilian harm is misassociation, where a surveillance platform follows a correctly identified combatant but over time the target is "swapped" with a civilian entity. The military surveillance does not notice the swap, leading to the engagement of civilians in the mistaken belief that they are a military target. This situation often occurs because there is no critical scrutiny of the nature of the target over time. Is it the same vehicle or group of people? Do they still have the same intelligence signature? Another factor is confusion over the target location, for example when the correctly identified threat is at one location and that target is handed off to someone else who then sees another entity at a different location and mistakes them as the threat.

Deconflicting with critical infrastructure

Although military strikes increasingly include consideration of immediate collateral damage to civilian structures, they often fail to consider the impact on critical infrastructure. Damage to such infrastructure (e.g., electricity, water) can have widespread and lasting effects on civilian populations. Ideally militaries would identify locations of critical infrastructure and then flag potential attacks at or near those locations, but this does not tend to happen in practice.

Leveraging AI Applications for Mitigating Civilian Harm

Having introduced the problem of civilian harm and some specific challenges that increase risk to civilians, we now discuss how AI and ML can be leveraged to mitigate civilian harm. Our discussion has three parts. We first identify specific actions in each of the six basic components of the CPLC with an eye toward actions that are potentially amenable to AI- or ML-enabled approaches. Second, in analogous fashion, we parse the space of AI/ML applications to identify those that are most promising for addressing key elements of the CPLC, prioritized by applications that directly address mechanisms for civilian harm and associated systemic challenges we identified in the previous chapter. Finally, we combine these efforts by explicitly associating a list of potential civilian harm mitigations with a set of specific AI/ML methods and technologies.

Connecting the CPLC with AI/ML applications

Figure 3 shows the CPLC–AI/ML applications (CPLC-AI) matrix, which associates CPLC elements with specific AI/ML applications. The 26 rows of this matrix denote specific CPLC actions (associated with each of the six main CPLC elements, highlighted in **green**), whereas the columns represent a taxonomy (described below) of 33 basic classes of AI and ML applications (organized into 11 "top level" domains, highlighted in **blue**, such as autonomous unmanned aerial systems (UASs)/unmanned surface vessels (USVs) cyberspace, and computer vision). If a given AI/ML application (e.g., column index *A*) is of value to and can be reasonably leveraged to help support a given CPLC action (e.g., row index *C*), then the (*C*, *A*) matrix entry contains the symbol "■"; otherwise the matrix element is empty. The numbers in parentheses highlighted in **green** and **blue** denote the number of "■" entries that appear in the corresponding row and column, respectively.

Neither the AI/ML applications taxonomy nor the veracity of matrix elements is definitive (the international community of AI researchers has not yet reached a consensus on how AI ought to even be defined, much less settled on a universally agreed-upon taxonomy of methods and applications).[31] Nonetheless, the matrix arguably captures the core gestalt of the two main "concept spaces" being woven together in this paper, namely actions that can be taken to help reduce civilian harm and AI/ML technologies that can be leveraged to facilitate their efficacy. Note that although the AI/ML taxonomy is by no means sacrosanct (our adherence to three

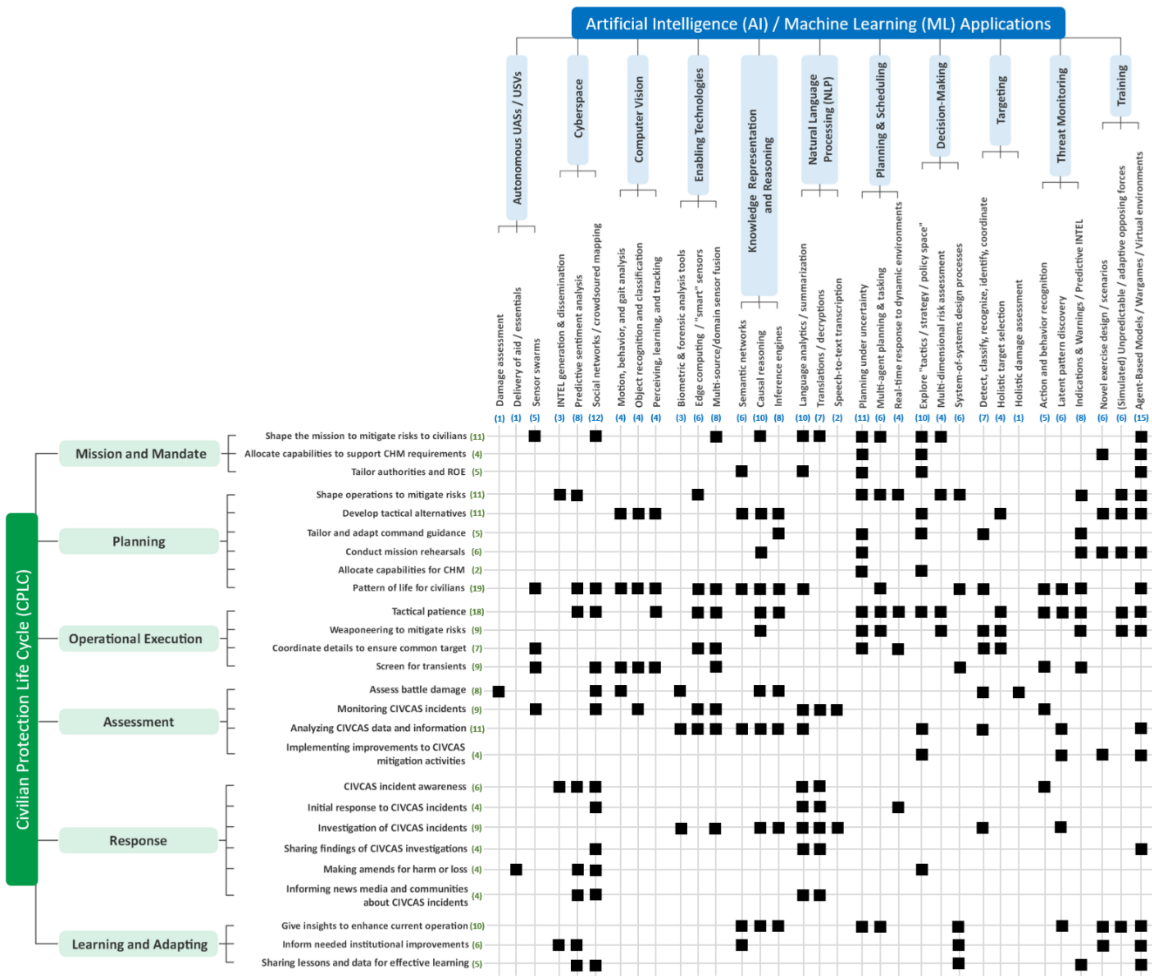
[31] Andrew Ilachinski, *AI, Robots, and Swarms*, CNA, DRM-2017-U-014796-Final, Jan. 2017.

applications per "top level" domain is arbitrary and was chosen mostly for expedience and to save space), the top level of the taxonomy is based partly on a recent survey of state-of-the-art methods[32] and partly on how AI and ML are characterized in the National Security Commission on Artificial Intelligence's (NSCAI's) final report.[33] The values of the matrix elements themselves were adjudicated entirely by the authors of this paper. Although the presence (or absence) of individual "■" entries in Figure 3 may be questioned, what matters most is the degree to which this matrix (or plausibly credible variants of this matrix) reveals the overall connective tissue that binds elements of the CPLC and extant AI/ML applications.

[32] See Appendix E ("AI/ML Approaches, Methods, and Algorithms Taxonomy") in Andrew Ilachinski, *Artificial Intelligence: Emerging Themes, Issues, and Narratives*, CNA, DOP-2020-U-028073-Final, Oct. 2020.

[33] NSCAI, *Final Report*, Mar. 2021, https://www.nscai.gov/wp-content/uploads/2021/03/Full-Report-Digital-1.pdf.

Figure 3. CPLC-AI matrix



Source: CNA.

To better understand the CPLC-AI matrix, the next sections take a closer look at the AI/ML applications taxonomy, starting with the domains spread across the top row (highlighted in **blue**). Each domain depicts an area for which AI or ML applications already exist, and can be brought to bear on problems germane to the CPLC in the near- to mid-term.

Autonomous UASs/UAVs

This domain includes autonomous UASs, USVs, and drones in general. Unmanned systems are among the most publicly visible AI-enabled technologies, spanning the spectrum from Defense

Advanced Research Projects Agency (DARPA)-level challenges and prototypes[34] to deployed military systems[35] (e.g., systems used to identify targets of interest on the battlefield)[36] to driverless cars.[37] Among the myriad possible ways autonomous robots can be used to help support elements of the CPLC are damage assessment (in concert with other technologies),[38] delivery of aid and essentials,[39] and as parts of larger networked sensor swarms (that can be used to vastly enhance situational awareness (SA) over wide areas).[40]

In early 2020, shortly after the COVID-19 pandemic took hold of the world, the US-based drone maker Draganfly partnered with the Australian Department of Defense and University of South Australia to deploy special "pandemic drones" that can detect coughing, sneezing, respiratory rate, and even fever from a distance.[41]

Two recent research programs based in part on robot and swarm technologies are (1) DARPA's OFFensive Swarm-Enabled Tactics program, which seeks to develop swarms of collaborative autonomous systems to surveil operational areas, buildings, and objects to provide real-time actionable intelligence to troops in urban environments,[42] and (2) a large-scale autonomous surveillance system called the Roborder project,[43] which consists of unmanned mobile robots

[34] "DARPA Subterranean Challenge Announces Systems Competition Teams for Final Event," DARPA website, May 3, 2021, https://www.darpa.mil/news-events/2021-05-03.

[35] Dan Gettinger, *The Drone Databook*, The Center for the Study of the Drone at Bard College, Sept. 2019, https://dronecenter.bard.edu/files/2019/10/CSD-Drone-Databook-Web.pdf.

[36] Arthur Holland Michel, *Unarmed and Dangerous: The Lethal Applications of Non-Weaponized Drones*, The Center for the Study of the Drone at Bard College, Mar. 2020, https://dronecenter.bard.edu/files/2020/03/CSD-Unarmed-and-Dangerous-Web.pdf.

[37] Lawrence D. Burns, *Autonomy: The Quest to Build the Driverless Car—And How It Will Reshape Our World*, (New York: Ecco Press, 2018).

[38] Xiaoyu Zhu, Junwei Liang, and Alexander Hauptmann, "MSNet: A Multilevel Instance Segmentation Network for Natural Disaster Damage Assessment in Aerial Videos," Dec. 30, 2020, arXiv:2006.16479 [cs.CV].

[39] A notable recent example is the US company Zipline's lightweight fixed-wing drones delivering blood to 25 hospitals and clinics across Rwanda in 2019. Evan Ackerman and Michael Koziol, "In the Air with Zipline's Medical Delivery Drones," *IEEE Spectrum* (Apr. 30, 2019), https://spectrum.ieee.org/in-the-air-with-ziplines-medical-delivery-drones/particle-1.

[40] Dan Popescu et al, "A Survey of Collaborative UAV–WSN Systems for Efficient Monitoring," *Sensors* 19, no. 21 (Oct. 28, 2019), https://www.mdpi.com/1424-8220/19/21/4690/htm.

[41] "Can a Pandemic Drone Help Stop the Spread of COVID-19?" Draganfly website, July 3, 2020, https://draganfly.com/news/can-a-pandemic-drone-help-stop-the-spread-of-covid-19/.

[42] "OFFensive Swarm-Enabled Tactics (OFFSET)," DARPA website, https://www.darpa.mil/program/offensive-swarm-enabled-tactics. Technical details appear in Timothy H. Chung, "OFFensive Swarm-Enabled Tactics (OFFSET)," Mar. 2021, https://apps.dtic.mil/sti/pdfs/AD1125864.pdf.

[43] "Roborder: Autonomous Swarm of Heterogeneous Robots for Border Surveillance," Roborder website, accessed Sept 20, 2021, https://roborder.eu/wp-content/uploads/2021/03/ROBORDER_General_v2.0.pdf.

(including aerial, ground, water surface, and underwater vehicles) capable of functioning either as standalone systems or in swarms that are deployed with multimodal adaptive sensors. Pilot cases include early identification and tracking of illegal activities and communications, and detection of "accidents" that occur at borders.[44] A recent paper provides a comprehensive survey of the use of mini-UAVs for remote sensing.[45]

Cyberspace

In the broadest sense, this domain denotes a set of applications derived principally from (or that otherwise involve or directly leverage) data that are communicated over computer networks.[46] civilian harm-related applications include the generation and dissemination of intelligence,[47] crowdsourced mapping technologies[48] (which also involve *Computer Vision*),[49] and predictive sentiment analysis[50] (which combines social networks[51] and *Natural Language Processing*). The Global Database of Events, Language, and Tone provides a vast set of data that describes human behavior at the societal level over time; it is designed to monitor "…the world's broadcast, print, and web news from nearly every corner of every country in over 100 languages and identifies the people, locations, organizations, themes, sources, emotions, counts, quotes, images and events driving our global society every second of every day."[52]

---

[44] "Roborder: Aims & Objectives," Roborder website, https://roborder.eu/the-project/aims-objectives/.

[45] Tian-Zhu Xiang, Gui-Song Xia, and Liangpei Zhang, "Mini-Unmanned Aerial Vehicle-Based Remote Sensing: Techniques, Applications, and Prospects," *IEEE Geoscience and Remote Sensing Magazine* 7, no. 3 (Sept. 2019).

[46] Note that our use of the term "cyberspace" deliberately focuses on its data communication–centric meaning and not the more typical focus of issues dealing with cybersecurity. B. Geluvaraj, P. Satwik, and T. Kumar, "The Future of Cybersecurity: Major Role of Artificial Intelligence, Machine Learning, and Deep Learning in Cyberspace," S. Smys, R. Bestak, J.Z. Chen, and I. Kotuliak, eds, in International Conference on Computer Networks and Communication Technologies, Singapore: Springer, 2019.

[47] Daniel Ish, Jared Ettinger, and Christopher Ferris, *Evaluating the Effectiveness of Artificial Intelligence Systems in Intelligence Analysis*, RAND Corporation, RR-A464-1, 2021.

[48] Andrew Ilachinski, *Applications of Social Media to Military Operations: Overview and Assessment*, CNA, DME-2013-U-005368-Final, July 2013.

[49] Kotaro Hara and Jon E. Froehlich, "Characterizing and Visualizing the Physical World Accessibility at Scale Using Crowdsourcing, Computer Vision, and Machine Learning," *ACM SIGACCESS Accessibility and Computing* 113 (Oct. 2015).

[50] Yousef Mourabit et al., "A New Sentiment Analysis System of Tweets Based on Machine Learning Approach," *International Journal of Scientific and Technology Research* 9, no. 12 (Dec. 2020).

[51] Qiaoyu Tan, Ninghao Liu, and Xia Hu, "Deep Representation Learning for Social Network Analysis," Apr. 18, 2019, arXiv:1904.08547 [cs.SI].

[52] GDELT homepage, accessed Sept. 20, 2021, https://www.gdeltproject.org/.

The military's interest in using social media as a new form of intelligence (social media intelligence or SOCMINT)[53] in general, and crowdsourcing[54] in particular, dates to DARPA's 2009 Red Balloon Contest (also known as the Network Challenge).[55] The contest was designed to probe how the internet and social networking may be used to solve a distributed, time-critical geolocation problem. Specifically, the challenge was to find 10 red weather balloons that were deployed at undisclosed locations across the continental United States. An MIT Media Lab team located all of them within nine hours by using social media. Apart from demonstrating the utility of social media in solving a real-world "spatial search" problem (all the more impressive when recalling that in 2009 social media was only nascent), the result also showed that a "solution" was possible via crowdsourcing despite efforts to provide false information on the location of the balloons.[56]

A widely used system that demonstrates the power of crowdsourcing is Ushahidi (which means "testimony" or "witness" in Swahili), a freely available open-source platform for data collection, visualization, and interactive mapping. It was introduced in 2007 as a one-stop source of data about violence in Kenya following Kenya's disputed 2007 presidential election. The idea was to visualize the physical distribution of locations where specific violent events occurred using eyewitness reports of violence (sent in by email or text message) and Google Maps. Since then, Ushahidi has been in continual development and has been used to track the evolution of a variety of regional events, including the disaster relief efforts following the earthquake in Haiti in January 2010, the Syrian revolution in 2011, and the earthquake in Nepal in 2015.[57]

As another illustrative example of the power of crowdsourcing, a recent paper in *Armor: Mounted Maneuver Journal* cited an event from 2017 at which 1st Brigade Combat Team, 4th Infantry Division was conducting a reconnaissance-in-force at the National Training Center. The brigade successfully redirected its lead battalion to avoid an ambush at a chokepoint based

53 Bruce Forrester and Kees den Hollander, "The Role of Social Media in the Intelligence Cycle," in Next-Generation Analyst IV, Proceedings Vol. 9851, May 12, 2016.

54 Kathryn B. Laskey, "Crowdsourced Decision Support for Emergency Responders," in 18th International Command and Control Research and Technology Symposium, Alexandria, VA, June 19-21, 2013, https://apps.dtic.mil/sti/pdfs/ADA588344.pdf.

55 John C. Tang et al., "Reflecting on the DARPA Red Balloon Challenge," *Communications of the ACM* 54, no. 4 (April 2011), https://cacm.acm.org/magazines/2011/4/106587-reflecting-on-the-darpa-red-balloon-challenge/fulltext.

56 Alex Rutherford et al., "Impossible by Conventional Means: Ten Years on from the DARPA Red Balloon Challenge," Aug. 13, 2020, arXiv:2008.05940v1,.

57 Sergio De Simone, "Ushahidi and the Power of Crowdsourcing," InfoQ, June 27, 2015; Ushahidi homepage, https://www.ushahidi.com/.

on a "cyber-recon team" report containing two critical pieces of enemy information obtained entirely from open-source information on Facebook, Snapchat, and Tinder.[58]

DARPA's Air Space Total Awareness for Rapid Tactical Execution program seeks to develop new advanced low-cost sensors, AI algorithms, and virtual testing environments to deconflict airspace activities of friendly forces and create a better common operating picture.[59]

Computer Vision

Computer Vision (CV), apart from representing one of the earliest[60] and still most intensely researched areas in AI and ML,[61]serves as a methodological backbone to several powerful derivative technologies that can be leveraged to mitigate civilian harm, including basic object recognition and classification;[62] motion, behavior, and gait analysis;[63] and general perceiving, learning, and tracking in dynamic environments (in concert with *Robotics,* see above).[64]

---

[58] Christopher Lowman and Gerald Prater, "Expansion of the Reconnaissance and Security BCT into the Cyber Domain: Lessons Learned from NTC Rotation 17-07.05," July 2017, unpublished white paper. During the National Training Center exercise, the best estimate of opposing-force (OPFOR) locations based "entirely on social-media trolling … was surprisingly consistent with templated OPFOR locations derived from other sources," see Curt Taylor, "It's Time for Cavalry to Get Serious about Cyber Reconnaissance," *Armor: Mounted Maneuver Journal* (Fall 2018), https://mcoe.azurewebsites.us/Armor/eARMOR/content/issues/2018/Fall/4Taylor18.pdf.

[59] DARPA states, "ASTARTE will not only provide a continuously updating, real-time, four-dimensional (space and time) moving picture of the battlespace for friendly forces but will also use its sensor network to detect and map adversary locations, increasing situational awareness." See "Real-Time Airspace Awareness and De-Confliction for Future Battles," DARPA website, Apr. 7, 2020, https://www.darpa.mil/news-events/2020-04-07.

[60] Although CV has been studied for decades (dating back to the early 1960s), a significant milestone occurred in February 2015, when Microsoft Research announced a system that for the first time surpassed human-level performance on ImageNet, a widely used dataset. See Kaiming He et al., "Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification," Feb. 6, 2015, arXiv:1502.01852v1 [cs.CV].

[61] Joel Janai et al., "Computer Vision for Autonomous Vehicles: Problems, Datasets and State of the Art," Mar. 17, 2021, arXiv:1704.05519v3 [cs.CV].

[62] L.E. Carvalho and A. von Wangenheim, "3D Object Recognition and Classification: A Systematic Literature Review," *Pattern Analysis and Applications* 22(Feb. 2019).

[63] Salisu Ibrahim Yusuf, Steve Adeshina, and Moussa Mahamat Boukar, "Parameters for Human Gait Analysis: A Review," in 2019 15th International Conference on Electronics, Computer and Computation (ICECCO), Dec. 2019.

[64] Muhammad Saputra, Andrew Markham, and Niki Trigoni, "Visual SLAM (Simultaneous Localization and Mapping) and Structure from Motion in Dynamic Environments: A Survey," *ACM Computing Surveys*51, no. 2 (June 2018).

A well-known recent example of this kind of technology is Project Maven,[65] which consists of tools for identifying potential targets (e.g., vehicles, buildings, and people) from UAV imagery. Also, DARPA's Target Recognition and Adaption in Contested Environments program uses ML to locate and identify targets.[66]

Other examples of AI/ML-enabled CV systems or research programs applicable to civilian harm mitigation include the following:

- The Army's Next-Generation Squad Weapon, which is slated to replace the M249 squad automatic weapon and the M4/M4A1 carbine and will purportedly be equipped with automatic target recognition, target tracking, wireless communication able to transmit fire control data to others, and facial recognition technology[67]
- Synthetic aperture radar (SAR) target recognition with DL[68]
- Drone-based thermal imaging and recognition[69] and
- Using Wi-Fi[70] and shadows[71] to "see" behind walls and around corners, respectively, and using a single laser shot[72] (fired through a keyhole) to track moving objects

[65] Project Maven (formally known as the Algorithmic Warfare Cross-Functional Team) was established in 2017 to "accelerate the DOD's integration of big data and machine learning." The ownership of Project Maven (as of this writing) is being transferred from the Washington Headquarters Service to the Under Secretary of Defense for Intelligence & Security, as stipulated in the original memo by then-Deputy Defense Secretary Robert Work, whose office established the program. See Robert Work, Memorandum, Subject: Establishment of an Algorithmic Warfare Cross-Functional Team (Project Maven), Apr. 26, 2017, https://dodcio.defense.gov/Portals/0/Documents/Project%20Maven%20DSD%20Memo%2020170425.pdf.

[66] Pat Host, "Deep Learning Analytics Develops DARPA Deep Machine Learning Prototype," Defense Daily, May 11, 2016, https://www.defensedaily.com/deep-learning-analytics-develops-darpa-deep-machine-learning-prototype/advanced-transformational-technology/.

[67] The award date is scheduled for November 2021, with delivery of the first rifle/automatic rifle systems to the Army in May 2022. See Mathew Cox, "Army's Next Infantry Weapon Could Have Facial-Recognition Technology," Real Clear Defense, June 3, 2019.

[68] Ryan J. Soldin, "SAR Target Recognition with Deep Learning," in IEEE Applied Imagery Pattern Recognition Workshop (AIPR), Washington, DC, Oct. 2018.

[69] David C. Schedl, Indrajit Kurmi, and Oliver Bimber, "Search and Rescue with Airborne Optical Sectioning," *Nature Machine Intelligence* 2 (Nov. 23, 2020).

[70] Mingmin Zhao et al., "Through-Wall Human Pose Estimation Using Radio Signals," *Computer Vision and Pattern Recognition (CVPR)* (2018), http://openaccess.thecvf.com/content_cvpr_2018/papers/Zhao_Through-Wall_Human_Pose_CVPR_2018_paper.pdf.

[71] Charles Saunders, John Murray-Bruce, and Vivek Goyal, "Computational Periscopy with an Ordinary Digital Camera," *Nature* 365 (Jan. 23, 2019), https://www.nature.com/articles/s41586-018-0868-6.

[72] Christopher A. Metzler, David B. Lindell, and Gordon Wetzstein, "Keyhole Imaging: Non-Line-of-Sight Imaging and Tracking of Moving Objects Along a Single Optical Path," *IEEE Transactions on Computational Imaging* 7 (Dec. 22, 2020).

Enabling Technologies

Enabling Technologies are a set of miscellaneous AI/ML-enabled technologies (all of which depend on other applications) that includes biometric and forensic analysis tools,[73] edge computing (and "smart" sensors),[74] and multi-sensor and multi-domain sensor fusion.[75] Enabling technologies run the gamut from (1) conceptual—as in the "AI stack" framework, which consists of a set of interdependent technology layers to help visualize, organize, plan, and prioritize strategic AI/ML capabilities and investments[76]—to (2) cross-disciplinary—as in an Army Research Office–funded project that applied insights from cognitive neuroscience to develop new "brain training" methods to help soldiers avoid civilian harm by friendly fire[77]—to (3) visionary—as in collaborative AI at the tactical edge (CATE), a recent effort to develop a prototype AI/ML architecture and framework that enables simple, rapid integration of collaborative, multiagent AI technology into the processing, exploitation, and dissemination chain at the edge, sensor, and tactical level. CATE "imagines a soldier on patrol … [in] the background, collaborative AI agents scan city cameras, review patterns of life, providing an AI enabled over watch. The AI agents determine there is a threat and alert the soldier, who never [has] to look down at a screen and take his eyes off his immediate surroundings."[78]

Knowledge Representation and Reasoning

Knowledge Representation and Reasoning (KR&R) constitutes a critical set of nascent technologies that are being developed to represent contextually rich information about the world.[79] Although many more specific techniques and applications can be subsumed into this

[73] Mayank Vatsa, Richal Singh, and Angshul Majumdar, eds., *Deep Learning in Biometrics*, (Boca Raton, FL: CRC Press, 2018).

[74] Mario Molinara, Alessandro Bria, Saverio De Vito, and Claudio Marrocco, eds., "Artificial Intelligence for Distributed Smart Sensing," special issue, *Pattern Recognition Letters* (Jan. 2021).

[75] Erik Blasch et al., "Machine Learning/Artificial Intelligence for Sensor Data Fusion–Opportunities and Challenges," *IEEE Aerospace and Electronic Systems Magazine* 36, no. 7 (July 2021).

[76] Andrew Moore, Martial Hebert, and Shane Shaneman, "The AI Stack: A Blueprint for Developing and Deploying Artificial Intelligence," in Ground/Air Multisensor Interoperability, Integration, and Networking for Persistent ISR IX, Proceedings Vol. 10635, May 4, 2018. The lead author of this paper served as dean of the Carnegie Mellon University School of Computer Science from 2014 to 2018 and was a member of the NSCAI.

[77] Adam Biggs, Matthew Cain, and Stephen Mitroff, "Cognitive Training Can Reduce Civilian Casualties in a Simulated Shooting Environment," *Psychological Science* (July 13, 2015).

[78] Susan Toth and William Hughes, "The Journey to Collaborative AI at the Tactical Edge (CATE)," in Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications III, Proceedings Vol. 11746, Apr. 12, 2021.

[79] Pierre Marquis, Odile Papini, and Henri Prade, eds., *A Guided Tour of Artificial Intelligence Research: Volume I: Knowledge Representation, Reasoning and Learning*, (Springer-Verlag, May 2020).

one category, Figure 3 identifies three areas that can most readily be leveraged to help mitigate civilian harm: semantic networks and ontologies,[80] causal reasoning and discovery,[81] and inferential reasoning.[82]

An example of a large-scale KR&R system (albeit one that is still at the basic research front) is DARPA's Collection and Monitoring via Planning for Active Situational Scenarios (COMPASS) program. COMPASS is designed to leverage AI, game theory, and modeling and estimation to both identify stimuli that yield the most information about an adversary's intentions and provide decision-makers high-fidelity intelligence on how to respond (with positive and negative tradeoffs for each course of action).[83]

AI/ML-enabled knowledge representation tools can also be used to mitigate common shortfalls of "mission handoff," defined as the "process of passing an ongoing mission from one unit to another with no discernible loss of continuity" (e.g., SA, adversary composition, allies, host nation forces, civilian populace).[84] Typical problems include short handoff timeframes, the dynamic nature of the operational environment, mismatches between data offered by an outgoing unit and data required by an incoming unit, and formatting mismatches (e.g., in intelligence briefs). AI/ML knowledge-based query tools (combined with methods to "understand" unstructured textual data, see *Natural Language Processing* applications, below) can be used to gain insight into all relevant data.[85]

Natural Language Processing

Natural Language Processing (NLP), like CV, is a catchall phrase that consists of a vast set of interrelated methodologies and applications centered on the automatic manipulation of text-

---

[80] Ji Han et al., "Semantic Networks for Engineering Design: State of the Art and Future Directions," *Journal of Mechanical Design* 144, no. 2 (Sept. 9, 2021).

[81] Judea Pearl, "The Seven Tools of Causal Inference, with Reflections on Machine Learning," *Communications of the ACM* 62, no. 3 (Mar. 2019).

[82] Sagir Yusuf and Chris Baber, "Inferential Reasoning for Heterogeneous Multi-Agent Mission," *International Scholarly and Scientific Research & Innovation* 14, no. 10 (2020).

[83] Fotis Barlos et al., *Collection and Monitoring via Planning for Active Situational Scenarios (COMPASS): Strategic Multi-Layer Assessment (SMA) Report*, Sandia National Laboratories, SAND2020-0136R, Jan. 2020, https://www.osti.gov/servlets/purl/1592839.

[84] Department of Defense, *Foreign Internal Defense,* Joint Publication 3-22, Aug. 17, 2018, https://www.jcs.mil/Portals/36/Documents/Doctrine/pubs/jp3_22.pdf.

[85] See pages 89–90 in Stephan De Spiegeleire, Matthijs Maas, and Tim Sweijs, *Artificial Intelligence and the Future of Defense: Strategic Implications for Small and Medium-Sized Force Providers*, The Hague Centre for Strategic Studies, 2017, https://www.jstor.org/stable/resrep12564.

and speech-based language.[86] Specific NLP tasks (for which there are numerous extant ML applications) include machine translation,[87] text classification and summarization,[88] semantic parsing,[89] sentiment analysis,[90] speech-to-text transcription,[91] and natural language inference and knowledge extraction.[92] Additionally, NLP methods can be used to develop training tools to foster cultural awareness.[93] As for almost all other cases, the entries in Figure 3 represent only a small illustrative subset of a vastly larger set of NLP applications that may be leveraged to help mitigate civilian harm.

Planning and Scheduling

Planning and Scheduling (P&S) represents a core set of "classic AI" methods that consist of deciding on a course of action and steps to take in complex dynamic and uncertain

---

[86] Amirsina Torfi et al., "Natural Language Processing Advancements by Deep Learning: A Survey," Feb. 27, 202, arXiv:2003.01200v4 [cs.CL].

[87] Darminder Ghataoura and Sam Ogbannaya, "Application of Image Captioning and Retrieval to Support Military Decision Making," in International Conference on Military Communication and Information Systems (ICMCIS), the Hague, Netherlands, May 4-5, 2021, IEEE.

[88] Shenguluan Hou and Ruqian Lu, "Knowledge-Guided Unsupervised Rhetorical Parsing for Text Summarization," *Information Systems* 94 (Dec. 2020).

[89] Hossam Elzayady, Khaled Badran, and Gouda Salama, "Arabic Opinion Mining Using Combined CNN - LSTM Models," *International Journal of Intelligent Systems and Applications* 12, no. 4 (Aug. 2020.)

[90] Recent work on applying ML and DL to sentiment analysis (with relevance to military applications) includes (1) Liang-Chu Chen, Chia-Meng Lee, and Mu-Yen Chen, "Exploration of Social Media for Sentiment Analysis Using Deep Learning," *Soft Computing* 24 (2020); (2) Mohammed El-Jawad, Rania Hodlod, and Yasser Omar, "Sentiment Analysis of Social Media Networks Using Machine Learning," in 14th International Computer Engineering Conference, Cairo, Egypt, Dec. 29-30, 2018, IEEE; and (3) Yogesh Chandra and Antoreep Jana, "Sentiment Analysis Using Machine Learning and Deep Learning," in 7th International Conference on Computing for Sustainable Global Development, New Delhi, India, Mar. 12-14, 2020, IEEE.

[91] Abdelaziz Abdelhamid et al., "End-to-End Arabic Speech Recognition: A Review," in 19th Conference on Language Engineering (ESOLEC'19), Sept. 2020.

[92] Robert E. Wray, James Kirk, and John Laird, "Language Models as a Knowledge Source for Cognitive Agents," Sept. 20, 2021, arXiv:2109.08270v2 [cs.AI].

[93] Sodiq Adewole et al., "Dialogue-Based Simulation for Cultural Awareness Training," Feb. 1, 2020, arXiv:2002.00223v1 [cs.CY].

environments.[94] Basic civilian harm-centric methods include planning under uncertainty,[95] multi-agent planning and tasking,[96] and real-time response to dynamic environments.[97]

The use of AI for dynamic planning (at least pertaining to optimizing and scheduling military logistics problems) goes back several decades to DARPA's Dynamic Analysis and Replanning Tool program, introduced in 1991 and used for planning logistics during Operation Desert Storm.[98] Other examples include AFRANCI (developed in 2006), which combines neural network modeling with symbolic algorithms and has been used as a P&S tool to support a civilian rescue scenario,[99] and a coordinated continuous Monte Carlo tree-search algorithm, which has been applied to planning search and rescue missions for UAV teams.[100] (See *Training* for several wargaming-related course of action examples.)

Decision-Making

Decision-Making overlaps with both KR&R and P&S (to the extent that all three application areas entail methods designed to adjudicate a course of action) and includes applications at the "decision" end of the broader knowledge representation/planning/decision-making process.[101] civilian harm-centric methods include those that can be used to explore novel

---

94 See Chapter 11 in Stuart Russell and Peter Norvig, *Artificial Intelligence: A Modern Approach*, 4th Edition, (Pearson, May 2021).

95 Yang Zhen, Zhang Wanpeng, and Liu Hongfu, "Real-Time Strategy Game Tactical Recommendation Based on Bayesian Network," *Journal of Physics: Conference Series* 1168, no. 3 (2019).

96 Alejandro Torreno et al., "Cooperative Multi-Agent Planning: A Survey," *ACM Computing Surveys* 50, no. 6 (Nov. 2017).

97 Chao Chen et al., "NECTAR-An Agent-Based Dynamic Task Allocation Algorithm in the UAV Swarm," *Complexity* 2020 (Sept. 2020).

98 Sara Reese Hedberg, "DART: Revolutionizing Logistics Planning," *IEEE Intelligent Systems* 17 (May 2002).

99 Specifically, the testbed problem was to use the system to "decide" whether an ambulance or firefighter should rescue a civilian. The civilian is assumed to be somewhere in a burning building, and the decision is based on an "agent's" location and the civilian's life condition. Dynamic variables may include the positions of the ambulance, firefighter, civilian, burning building, fire brigade, and nearest refuge; metrics describing the "life condition" of the firefighter and civilian; the building's degree of volatility; and the innate difficulty of performing the civilian rescue. The analogy to similar CIVCAS-related "problems" ought to be obvious. See Francisco Reinaldo et al., "A Tool for Multi-Strategy Learning," *Advances in AI Research in Computing Science* 26 (2006).

100 Chris Baker et al., "Planning Search and Rescue Missions for UAV Teams," in Proceedings of the Twenty-second European Conference on Artificial Intelligence, Aug. 2016.

101 Yash Shrestha, Shiko Ben-Mehahem, and Georg von Krogh, "Organizational Decision-Making Structures in the Age of Artificial Intelligence," *California Management Review* 61, no. 4 (Aug. 2019).

tactics and strategies in abstract military operational mission and policy spaces[102] and perform multi-dimensional risk assessments[103] and those that are rigorously grounded in a system-of-systems approach to design processes in general.[104]

The most notable recent successes of AI/ML-enabled decision-making (outside of DOD) have come from research into game-playing algorithms. Examples include the following:

- AlphaZero, which, starting from random play and using no domain knowledge except for game rules, required only 24 hours to achieve a superhuman level of play in chess, shogi (a Japanese variant of chess), and Go, and defeated a world champion program in each105 (a later version, MuZero, matched AlphaZero's superhuman performance without any knowledge of game rules)[106]
- An AI that learned to play all 57 Atari video games[107]
- AlphaStar, which defeated 99.8 percent of human Starcraft II gamers[108]
- Pluribus, the first AI to defeat human professional players in a multiplayer game[109]

An important implicit additional dimension that can be used to characterize all AI-enabled decision-making processes is the aggregated human-AI decision-making relationship.[110] For

[102] Bonnie Johnson and William Treadway, “Artificial Intelligence—An Enabler of Naval Tactical Decision Superiority,” *AI Magazine* 40, no. 1 (Mar. 2019).

[103] Nicola Paltrinieri, Louise Comfort, and Genserik Reniers, “Learning About Risk: Machine Learning for Risk Assessment,” *Safety Science* 118 (Oct. 2019).

[104] Kanstantsin Miatliuk, Adam Wolniakowski, and Pawel Kolosowski, “Engineering System of Systems Conceptual Design in Theoretical Basis of Hierarchical Systems,” in 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Toronto, Ontario, Canada, Oct. 11–14, 2020, IEEE.

[105] David Silver et al., “Mastering Chess and Shogi by Self-Play with a General Reinforcement Learning Algorithm,” Dec. 5, 2017, arXiv:1712.01815v1.

[106] Julian Schrittwieser et al., “Mastering Atari, Go, Chess and Shogi by Planning with a Learned Model,” Nov. 19, 2019, https://arxiv.org/abs/1911.08265.

[107] Adria Badia et al., “Agent57: Outperforming the Atari Human Benchmark,” Mar. 30, 2020, arXiv:2003.13350.

[108] Compared to classic board games like chess, StarCraft II entails much greater real-world-like complexity. For example, the game includes hundreds of “pieces” (soldiers in the factions' armies) that move simultaneously in real time, not in an orderly turn-based fashion. See Oriol Vinyals et al., “Grandmaster Level in StarCraft II Using Multi-Agent Reinforcement Learning,” *Nature* 575 (Oct. 30, 2019), https://www.nature.com/articles/s41586-019-1724-z.

[109] Unlike chess, in which two players always have perfect access to all game-relevant information, poker includes multiple simultaneous players, and decisions must be based on imperfect (i.e., hidden) information and human-psychology-centric actions that include bluffing. See Noam Brown and Tuomas Sandholm, “Superhuman AI for Multiplayer Poker,” *Science* 365 (Aug. 30, 2019).

[110] Dominik Dellerman et al., “The Future of Human-AI Collaboration: A Taxonomy of Design Knowledge for Hybrid Intelligence Systems,” May 7, 2021, arXiv:2105.03354 [cs.AI].

example, the Sensing for Asset Protection with Integrated Electronic Networked Technology system, developed by the UK's Defence Science and Technology Laboratory, is designed as an AI-enabled decision-support tool that combines autonomous sensing with fusion and sensor management to provide both SA over large areas and context-driven decision-making.[111]

Targeting

Targeting denotes a class of AI/ML-enabled applications that combine AI's innate ability to discover patterns in extremely high-dimensional abstract data spaces (see *Threat Monitoring*, below) with methods of adjudicating multiple conflicting criteria "action selection" problems developed by the complex system theory (CST)[112] and agent-based modeling (ABM)[113] communities. These methods can be applied to the entire F2T2EA targeting cycle (find, fix, track, target, engage, and assess),[114] but the most civilian harm-centric ones are those that include CST/ABM-enabled approaches to detect, classify, recognize, and identify potential targets. Holistic target selection and damage assessment (i.e., the two other applications in Figure 3 that appear under the main heading *Targeting*) refer to AI/ML-enabled methods that respect the interdependent dynamic relationships that define a given socio-cultural-political-religious and physical system being targeted for attack or assessed for damage (after an attack).

To illustrate what we mean, at least intuitively, think of the complex food webs in natural ecologies (nature's own best exemplars of "complex adaptive systems"). What are the "most important" species in a given ecology? One approach to answering this question is to apply Google's PageRank algorithm,[115] which ranks the "importance" of webpages, to determine which species are critical for sustaining biological niches in the ecology.[116] The method effectively maps websites to "species" (thought of as "nodes" of a food web) and uses a generalized form of PageRank to identify the key species whose individual loss (to the food

---

111 UK Ministry of Defence, *SAPIENT Middleware Interface Control Document*," May 11, 2020, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/895246/SAPIENT_Interface_Control_Document_v5.0.pdf.

112 George Mobus and Michael Kalton, *Principles of Systems Science (Understanding Complex Systems),* (Springer, 2016).

113 Uri Wilensky and William Rand, *An Introduction to Agent-Based Modeling*, (Cambridge, MA: MIT Press, 2015).

114 Merel A. C. Ekelhof, "Lifting the Fog of Targeting: 'Autonomous Weapons' and Human Control Through the Lens of Military Targeting," *Naval War College Review* 71, no. 3 (2018).

115 Amy Langville and Carl Meyer, *Google's PageRank and Beyond*, (Princeton, NJ: Princeton University Press, 2006).

116 Stefano Allesina and Mercedes Pascual, "Googling Food Webs: Can an Eigenvector Measure Species' Importance for Coextinctions?" *PLoS Computational Biology* 5, no. 9 (2009).

web) would result in the maximal number of other extinctions (i.e., the "most important" species in an ecosystem). Although it has long been known that the collapse of ecosystems may be triggered by the extinction of critical species[117] and that even a "small" species loss may lead to multiple cascading coextinctions,[118] the typically vast web of mutual interactions among species (and their environment) makes it difficult to predict which components of an ecosystem are the most important. The key takeaway from this simple example is that there are AI/ML-enabled tools (considerably more sophisticated than PageRank) that allow us to analogously determine which targets, if attacked, are most likely to entail civilian harm.

DARPA's AI-enabled Mosaic Warfare program offers an explicitly "complex systems of systems"-based approach to warfighting designed "around 'tiles' of capabilities (i.e., functions: sensors and shooters), rather than uniquely shaped 'puzzle pieces' (i.e., platforms) that must be fitted into specific slots in battle plan in order for it to work."[119] Although Mosaic Warfare may still be years away from deployment, even in notional form, its core "kill web, not kill chain" vision makes it particularly replete with opportunities for civilian harm mitigation if those opportunities are intentionally built into the design.

Finally, just as for *Decision-Making*, an important implicit additional class of AI/ML-enabled targeting applications is applications that leverage a hybrid human-AI collaborative effort.[120] The US Army's Aided Threat Recognition from Mobile Cooperative and Autonomous Sensors (ATR-MCAS) system provides a glimpse of the "art of the possible" in the near to medium term.[121] The ATR-MCAS prototype is an AI-enabled system that consists of networked state-of-the-art air and ground vehicles that leverage sensors and edge computing. Being developed explicitly as a "battlefield teammate" to soldiers, ATR-MCAS includes sensors enabling vehicles to navigate within areas of interest to identify, classify, and geolocate entities, obstacles, and potential threats. It is capable of aggregating and distributing target data, which can be used to make recommendations and predictions based on the combined threat picture, and its "AI-enabled decision support agent" recommends responses (e.g., which threats to prioritize).

Although only recently announced (in January 2020), DARPA's Artificial Social Intelligence for Successful Teams program seeks to develop AI/ML-enabled systems built from agents with the

[117] Jose Montoya, Stuart Pimm, and Ricard Sole, "Ecological Networks and Their Fragility," *Nature* 442 (2006).

[118] Jose Montoya and Ricard Sole, "Small World Patterns in Food Webs," *Journal of Theoretical Biology* 214 (2002).

[119] "DARPA Tiles Together a Vision of Mosaic Warfare," DARPA website, Sept. 24, 2018, https://www.darpa.mil/work-with-us/darpa-tiles-together-a-vision-of-mosiac-warfare.

[120] Jason Cody, Karina Roundtree, and Julie Adams. "Human-Collective Collaborative Target Selection," *ACM Transactions on Human-Robot Interaction* 10, no. 2 (Mar. 2021).

[121] Patrick Ferraris, "Aided Detection on the Future Battlefield," Defense Visual Information Distribution Service, Jan. 24, 2020, https://www.dvidshub.net/news/360225/aided-detection-future-battlefield.

ability to create shared mental models that "…demonstrate the basic machine social skills needed to infer the goals and situational knowledge of human partners, predict what they will need, and offer context-aware actions in order to perform as adaptable and resilient AI teammates."[122]

Threat Monitoring

Threat Monitoring leverages a broad set of AI/ML-enabled pattern recognition methods and applications. Compared to humans, AI can assimilate and find patterns in vastly larger data spaces on much shorter time scales. CPLC-specific applications include behavior recognition,[123] latent pattern discovery,[124] and transforming multi-domain datastreams (see *Enabling Technologies*) into actionable intelligence.[125]

Training

Training, in the context of mitigating civilian harm, subsumes and conflates myriad other AI/ML-related modeling and simulation applications and includes novel exercise design and

122 "ASIST agents must operate in increasingly complex and specialized environments; be adaptable to sudden perturbations in the mission or team, like the loss of communication with a key teammate; and use noisy multi-channel observations to represent the world and do complex inference and prediction." See Joshua Elliot, "Artificial Social Intelligence for Successful Teams (ASIST)," DARPA website, accessed Sept. 20, 2021, https://www.darpa.mil/program/artificial-social-intelligence-for-successful-teams.

123 Nagesh Jadhav and Rekha Sugandhi, "Survey on Human Behavior Recognition Using Affective Computing," in 2018 IEEE Global Conference on Wireless Computing and Networking (GCWCN), Lonavala, India, Nov. 23-24, 2018, IEEE.

124 Samira Ranaei and Arho Suominen, "Using Machine Learning Approaches to Identify Emergence: Case of Vehicle Related Patent Data," in 2017 Portland International Conference on Management of Engineering and Technology (PICMET), Portland, OR, July 9-13, 2017, IEEE.

125 James L. Regens, "Augmenting Human Cognition to Enhance Strategic, Operational, and Tactical Intelligence," *Intelligence and National Security* 34, no. 5 (2019).

scenario generation,[126] adaptive and unpredictable opposing forces,[127] and various forms of virtual environments,[128] including those generated in part using agent-based models.[129]

AI/ML and military wargaming are a particularly potent combination that can be applied in two (partly overlapping) ways: (1) using traditional (non–AI/ML-enabled) wargames as exploratory teaching aids to help identify the holistic effects that AI may have on existing military operational concepts (such as those that appear in the CPLC),[130] and (2) harnessing AI/ML technologies by directly embedding them within wargames. In the latter case, AI/ML can be used to develop an "all-knowing" *Alexa*-like front-end interface that facilitates conventional wargaming practices[131] or as an innate decision-support tool to "discover" alternative courses of action.[132] A recent survey summarizes the state-of-the-art AI/ML-enabled adaptive learning systems.[133]

Before making some general remarks about the gestalt of the CPLC-AI matrix in Figure 3, we note that implicit in all 11 of the domains that appear as columns in this matrix is AI's inherent

---

126 Robert Sottilare, "A Hybrid Machine Learning Approach to Automated Scenario Generation (ASG) to Support Adaptive Instruction in Virtual Simulations and Games," in I3M Defense & Homeland Security Simulation Workshop, Budapest, Hungary, 2018.

127 Jeremy Ludwig and Bart Presnell, "Developing an Adaptive Opponent for Tactical Training," in First International Conference on Adaptive Instructional Systems (AIS), Orlando, FL, July 26–31, 2019.

128 Christina Cook, "Designing a Virtual Embedded Scenario-Based Military Simulation Training Program Using Educational and Design Instructional Strategies," (Doctoral thesis, University of Central Florida, 2018).

129 Wenhui Fan et al., "Multi-Agent Modeling and Simulation in the AI Age," *Tsinghua Science and Technology*, 26, no. 5 (Oct. 2021).

130 ED McGrady and Justin Peachy, eds., *Representing Artificial Intelligence in Wargames*, *Connections Conference Working Group 2 Proceedings*, Dec. 2020, https://paxsims.files.wordpress.com/2020/12/connections_wg2_2020_final-4a.pdf.

131 Benjamin Jensen, Scott Cuomo, and Chris Whyte, "Wargaming with Athena: How to Make Militaries Smarter, Faster, and More Efficient with Artificial Intelligence," War on the Rocks, June 5, 2018, https://warontherocks.com/2018/06/wargaming-with-athena-how-to-make-militaries-smarter-faster-and-more-efficient-with-artificial-intelligence/.

132 AlphaZero has recently been applied to wargaming. See Glenn Moy and Slava Shekh, "The Application of AlphaZero to Wargaming," in Advances in Artificial Intelligence, Proceedings of 32nd Australasian Joint Conference, Adelaide, SA, Australia, Dec. 2019. More focused discussions of how AI can be used for course-of-action analysis in military wargames are given by (1) Peter J. Schwartz et al., "AI-Enabled Wargaming in the Military Decision Making Process," In Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications II, Proceedings Vol. 11413, 2020, and (2) William DeBerry et al., "The Wargame Commodity Course of Action Automated Analysis Method," *The Journal of Defense Modeling and Simulation: Applications, Methodology, Technology* (July 19, 2021).

133 Tumaini Kabudi, Ilias Pappas, and Dag Olsen, "AI-Enabled Adaptive Learning Systems: A Systematic Mapping of the Literature," *Computers and Education: Artificial Intelligence* 2 (2021).

ability to **discover novelty in unimaginably vast "possibility spaces"** that no human is capable of parsing as quickly or efficiently.[134]

Commentary about the CPLC-AI matrix in Figure 3

What does the CPLC-AI matrix in Figure 3 actually show us? The most obvious takeaway is that, without exception, each of the 26 CPLC elements has **significantly more than a single AI/ML application associated with it**: the minimum, maximum, and average values are 2, 19, and 7.3, respectively. Similarly, the minimum, maximum, and average number of CPLC elements with which a given AI/ML application is associated are 1, 14, and 5.6, respectively. Although the presence (or absence) of any given entry may be questioned, the unassailably rich network of potential associations testifies to the enormous latent general applicability of AI/ML technologies.

Remember that AI/ML-enabled technologies that support each of the 11 domains that appear in Figure 3 already exist (albeit to varying degrees of efficacy and subject to the vagaries of ongoing basic research), which means that a broad arsenal of AI/ML-enabled tools can be harnessed to mitigate civilian harm. To be sure, the devil is in the details: there are no panacea "solutions" to the general civilian harm problem (whether AI/ML enabled or not; some specific options are discussed in the next section). But the mere fact that Figure 3 contains so many plausibly assigned "■" entries associating elements of the CPLC with potential AI/ML applications strongly suggests that a vast heretofore largely untapped reservoir of things can immediately be done to enhance essentially all components of the CPLC.

Error! Reference source not found. shows the number of links in the CPLC-AI matrix within each of the main (i.e., top-level) CPLC and AI/ML application elements.

Table 1. Number of " ·" entries in each of the top-level categories in the CPLC-AI matrix in Figure 3

		Artificial Intelligence (AI) / Machine Learning (ML) Applications											
		USVs	Cyber	CV	ET	KR/R	NLP	P&S	DM	Tgt	TM	Trn	Total
Civilian Protection Life	M&M	1	1	0	1	2	3	4	4	0	0	4	20
	Planning	1	4	5	2	8	1	6	6	2	6	9	50
	Op Exec	2	1	3	5	3	0	5	3	5	6	2	35

[134] See Joel Lehman et al., "The Surprising Creativity of Digital Evolution: A Collection of Anecdotes from the Evolutionary Computation and Artificial Life Research Communities," Nov. 21, 2019, arXiv:1803.03453,; and Giorgio Franceshelli and Mirco Mosolesi, "Creativity and Machine Learning: A Survey," Apr. 20, 2021, arXiv:2104.02726v2.

		Artificial Intelligence (AI) / Machine Learning (ML) Applications											
Cycle (CPLC)	**Assess**	**2**	**2**	**2**	**6**	**5**	**4**	**0**	**2**	**2**	**3**	**3**	**31**
	Response	**1**	**9**	**0**	**2**	**2**	**11**	**1**	**1**	**1**	**2**	**1**	**32**
	Lrn&Adapt	**0**	**4**	**0**	**0**	**4**	**0**	**2**	**3**	**0**	**2**	**6**	**21**
	Total	**7**	**21**	**10**	**16**	**24**	**19**	**18**	**19**	**10**	**19**	**25**	

Source: CNA.

The plethora of potential applications ought not surprise us because the problem of civilian harm may be viewed as a microcosm of actions, behaviors, and policies associated with the much larger military operational space. And, as numerous studies (and real-world operations) demonstrate, there is virtually no militarily relevant domain of activity to which AI/ML-enabled technologies cannot be applied, even if only in principle.[135]

As argued earlier, the CPLC-AI matrix in Figure 3 (and its derivative tallies in **Error! Reference source not found.**) ought not be imbued with any deeper meaning apart from its overall general credibility and the informed plausibility of its individual entries. However, we point out a few salient features that may not be immediately obvious:

- **A few entries are effectively "one offs,"** meaning that the range of their associability (or applicability) is natively limited to no more than a few cases. For example, "damage assessment" via *Autonomous UASs/USVs* in the first column of Figure 3 is an innately focused application limited to supporting the "assess battle damage" secondary component of CPLC's *Assessment* element, and "allocate capabilities for CHM" in the eighth row under the CPLC's *Planning* element entails an inherently small set of relevant AI/ML-enabled capabilities.
- **Some CPLC elements and AI/ML applications are only marginally associable**. Examples include those with zero entries in **Error! Reference source not found.**, such as the intersections of *Learning and Adapting* and *Autonomous UASs/USVs*, *Response* and *CV*, and *Operational Execution* and *NLP*.

[135] An extensive taxonomy of AI applications to military operations (woven partly around elements of the OODA loop: observe, orient, decide, and act) appears in Appendix I, "Mindmap of Possible Military Applications of AI," in Ilachinski, *Artificial Intelligence: Emerging Themes, Issues, and Narratives*. For a basic taxonomy, see Stoney Trent and Scott Lathrop, "A Primer on Artificial Intelligence for Military Leaders," *Small Wars Journal* (Aug. 2018).

- **Certain specific AI/ML capabilities that fall under a given broad application area may be applied to more than one CPLC element, including those that do not ostensibly belong to the same area.** For example, although *Training* may not intuitively be expected to play a role for CPLC elements other than, for example, *Planning* or *Learning and Adapting*, the basic underlying methodologies are applicable to all parts of the CPLC.
- **The main CPLC elements with the largest average number of associated AI/ML applications are *Operational Execution* and *Planning*.** They have averages of 43/4 ≈ 10.8 and 54/6 ≈ 9.0, respectively (see **Error! Reference source not found.**). Except for the lack of any obvious applicability of robotic technologies to planning, what these two elements have most in common is a need to adjudicate among a potentially massive abstract "space of possibilities." This is something ML is particularly adept at, as well as combining multiple simultaneous capabilities, such as causal reasoning, multi-dimensional risk assessment, and AI/ML-enabled wargaming.
- **The secondary-level CPLC element with the largest number of associated AI/ML applications (i.e., the element with the greatest innate potential for leveraging AI/ML-enabled technology) is "pattern of life for civilians" under *Planning*.** Its 19 associations include at least one AI/ML application from each of the top-level application domains. The element with the next largest number of associations is "tactical patience" under *Operational Execution* (with 18 associations), followed by "shape the mission to mitigate risks to civilians" under *Mission and Mandate*; "shape operations to mitigate risks" and "develop tactical alternatives," both under *Planning*; and "analyzing civilian harm data" under *Assessment* (all with 11 associated AI/ML applications each). Heuristically, we expect such multifaceted "problems" to require a rich and complementary assembly of methods and technologies (e.g., sensors, behavior recognition, inference engines).
- **The AI/ML applications domains associated with the largest number of CPLC elements are *Training* (26) and *KR&R* (24).** *Cyberspace* is a close third with 23 associations. The methods underlying these three areas fall under much broader and well-studied classes of complex systems–based multiagent models, simulations, and knowledge graphs (that codify entire semantic and causal ontologies), respectively. It is thus not surprising that they are also the most generally associable with the gamut of CPLC elements.
- **The specific AI/ML applications with the largest number of associated CPLC elements are "agent-based models/wargames/virtual environments" under *Training* (15), "social networks/crowdsourced mapping" under *Cyberspace* (12), and "planning under uncertainty" under *P&S* (11)**, although "causal reasoning" (under KR&R) and "language analytics/summarization" (under NLP), with

10 associations each, are not far behind. These are all methodologically and technologically mature applications ready to be harnessed.

What Figure 3 does *not* show, at least explicitly, is that multiple simultaneous algorithms, applications, and technologies can be combined synergistically to develop powerful standalone multipurpose AI/ML-enabled systems.

There is perhaps no better example of this than GAIA (Generating Alternative Interpretations for Analysis), funded in part by DARPA and developed by the US Army Research Laboratory, the University of Illinois at Urbana-Champaign, and Columbia University. GAIA won the "Best Demo" award at the 58th Annual Meeting of the Association for Computational Linguistics in July 2020.[136] It is described by its developers as a "fine-grained multimedia knowledge extraction system" and is the "first comprehensive, open source multimedia knowledge extraction system that takes a massive stream of unstructured, heterogeneous multimedia data from various sources and languages as input, and creates a coherent, structured knowledge base, indexing entities, relations, and events, following a rich, fine-grained ontology."[137] GAIA combines multiple simultaneous ML methods, including text and visual knowledge extraction and correlation, object detection and recognition of public figures, cross-media knowledge fusion (text, images, videos, speech, and optical character recognition), and behavioral and evidential pattern recognition. It provides real-time tracking of ongoing events, issues contextually and dynamically relevant alerts, and includes a nominal ability to "predict" changes to an environment or operations as well anticipate other topics and data that might be related to ongoing incidents.[138]

The Command and Control Incident Management Emergency Response Application (C2IMERA) system is an example of a large-scale military-grade tool (albeit one that uses AI and ML in a rudimentary fashion) that is already deployed and provides planning, force generation, emergency management, real-time SA, and command and control (C2) monitoring and execution functions.[139] It was developed by the Air Force Life Cycle Management Center's

[136] Manling Li et al., "GAIA: A Fine-Grained Multimedia Knowledge Extraction System," in Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations, July 2020, https://aclanthology.org/2020.acl-demos.11v2.pdf.

[137] Li et al., "GAIA." p. 1.

[138] *GAIA: A Fine-Grained Multimedia Knowledge Extraction System* (Blender Lab, UIUC), Video, http://blender.cs.illinois.edu/software/gaia-ie/gaia.mp4.

[139] US Air Force, *C2IMERA Overview* (Kessel Run, Mar. 26, 2020), Video, https://www.youtube.com/watch?v=8meuTt1RQbE. C2IMERA is mandated for use across all Air Combat Command installations and is currently used at more than 40 US Air Force bases across the continental United States and Europe.

Detachment 12 (also known as "Kessel Run"), a software development and acquisitions unit.[140] C2IMERA combines two core C2 capabilities, real-time SA and coordination, by making all pertinent data about specific installations, environments, assets, and personnel accessible in one centralized location. In 2019, it was used by C2 specialists at Moody Air Force Base, Georgia, to track Hurricane Dorian, helping prepare the base for the approaching storm and identifying fly-away options outside the path of the hurricane.[141] More recently, in Afghanistan, C2IMERA was used as a civilian harm-mitigation tool to support the recent noncombatant evacuation operation.[142] As the evacuations from Afghanistan took place, C2IMERA was used to keep abreast of a rapidly changing environment and ensure the safe transit of more than 124,000 US civilian and military personnel, allies and partners, and Afghans from Kabul. The Combined Air and Space Operations Center also used C2IMERA to receive automated alerts of the incident and response, which provided an additional level of coordination in near real time.[143] Although C2IMERA currently leverages AI or ML only minimally, virtually all elements of this already proven and deployed system can only be enhanced by integrating various AI/ML-enabled technologies.

Now that we have gotten a bird's eye view of the landscape of CPLC and AI/ML (and discussed a litany of possible associations), we next examine some specific ways in which various potential mitigations for civilian harm may benefit from AI/ML-enabled technologies. In the previous chapter, we discussed the different mechanisms that cause civilian harm and some specific problems that need to be addressed to reduce risks to civilians. We discuss some of these risks again but in a different way, highlighting the limitations of the environment in which AI/ML-enabled technologies will be operating.

Reducing but not eliminating risks to civilians

Figure 4 shows a notional schematic illustrating key elements of the environment many AI applications seeking to mitigate civilian harm will be dealing with. It highlights the critical agents and features in an operating area, including the targeting area (indicated in **purple**),

---

[140] US Air Force, Kessel Run homepage, https://kesselrun.af.mil/.

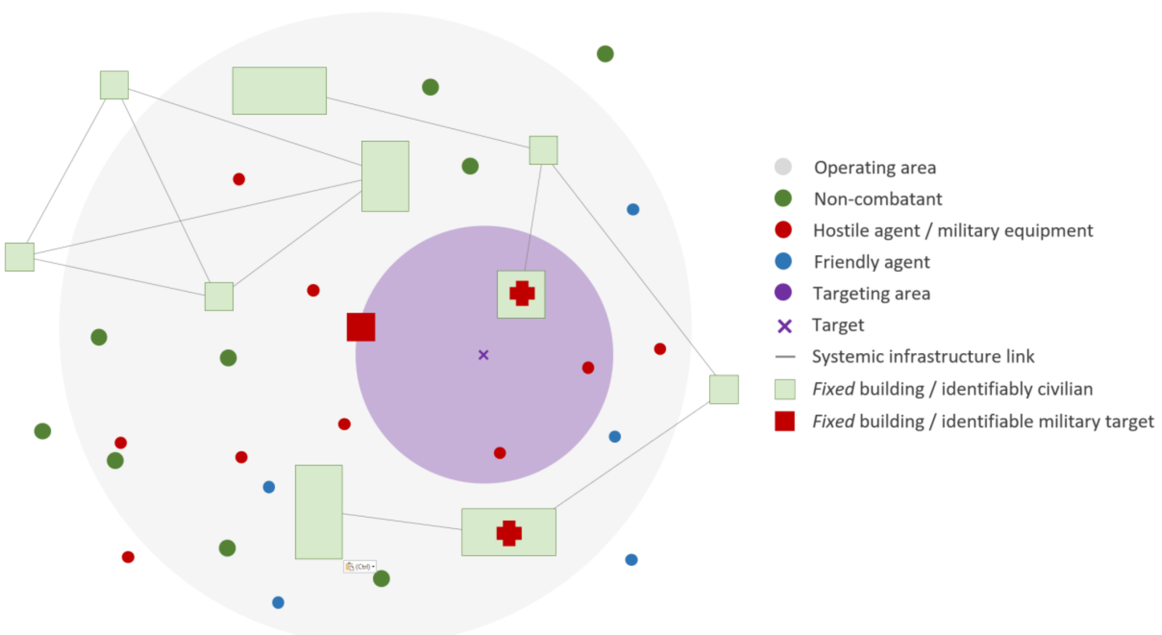
[141] "Adaptable Command and Control System Allowed Georgia Air Force Base to Track and Prepare for Hurricane Dorian," Leidos, Oct. 1, 2019, https://www.leidos.com/insights/adaptable-command-and-control-system-allowed-georgia-air-force-base-track-and-prepare.

[142] Richard Blumenstein, "Kessel Run's C2IMERA Used During Afghan Evacuation," Wright-Patterson AFB, Sept. 23, 2021, https://www.wpafb.af.mil/News/Article-Display/Article/2786182/kessel-runs-c2imera-used-during-afghan-evacuation/.

[143] Mark Pomerleau, "Air Force Software Tool Helped Coordinate Afghanistan Evacuation of Civilians," C4ISRNet, Sept. 23, 2021, https://www.c4isrnet.com/battlefield-tech/c2-comms/2021/09/23/air-force-software-tool-helped-coordinate-afghanistan-evacuation-of-civilians/.

combatants (**red**), friendly military forces (**blue**), civilians (**dark green**), and surrounding area, including buildings (**light green** rectangles, with a superimposed red "**+**" denoting a hospital) and infrastructure (designated by links between buildings). Leaving aside for the moment certain irreducible real-world uncertainties, such as the exact numbers and locations of combatants and civilians (detailed in the examples that follow), these are the basic elements that any AI/ML-enabled technology must be able to sense, understand, draw inferences from, and otherwise make decisions about in the context of mitigating civilian harm. The next few figures show that even a perfect world—one with few or no uncertainties, with clear demarcations between "hostile" and "nonhostile," and in which targeting areas (and concomitant weapon blast zones) that preclude any reasonable likelihood of collateral damage are all easily identifiable—will have a non-zero risk to civilians.[144]

Figure 4. Notional schematic illustrating key elements of the civilian harm AI problem



Source: CNA.

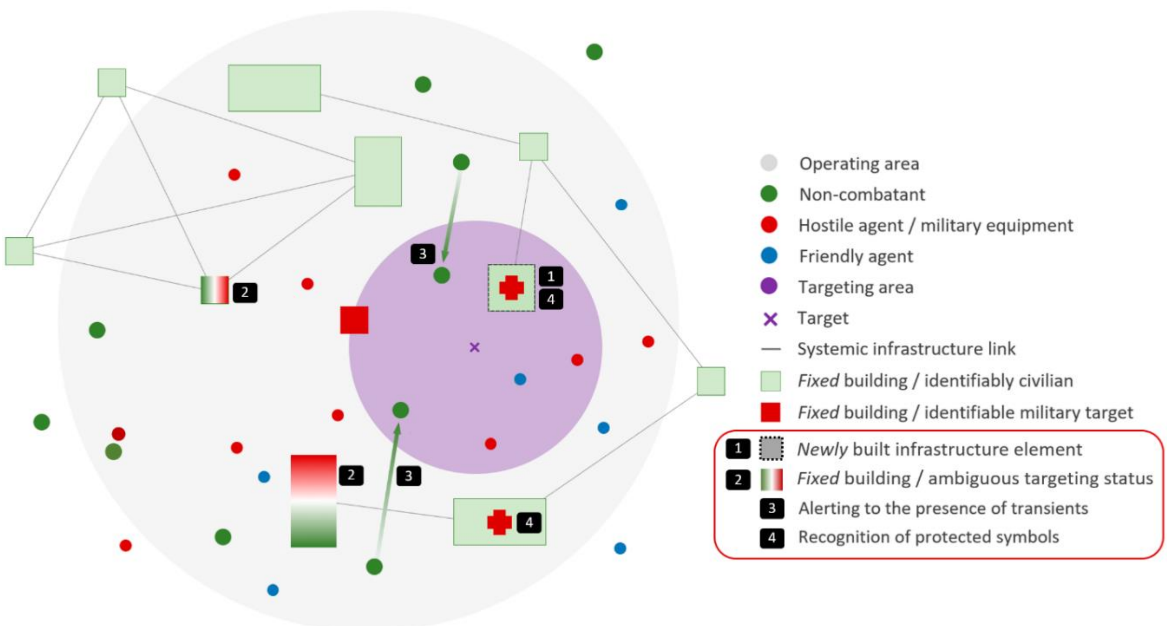
Why is a zero-risk scenario unattainable? Consider Figure 5, which adds a few changes to Figure 4 to reflect the nature of real-world ambiguities:

1. Buildings may exist that do not appear in out-of-date or incomplete targeting datasets.

[144] This is an important point because some commanders in the field have stressed a goal of no CIVCAS. Over the long term, even with the use of AI to mitigate civilian harm more effectively, this laudable goal is not possible. However, we can still use technology to reduce these risks.

2. The "targeting status" of other known buildings may be changed (e.g., a warehouse may have been transformed into a religious center).
3. Civilian transients may have entered an otherwise properly designated targeting area immediately before an "attack" order was issued.
4. An old, weathered (maybe already combat-torn) protected symbol designating, for example, a hospital, may go unrecognized and be incorrectly targeted.

Figure 5. Schematic adding some inherent uncertainty to the notionally "perfect" setting in Figure 4



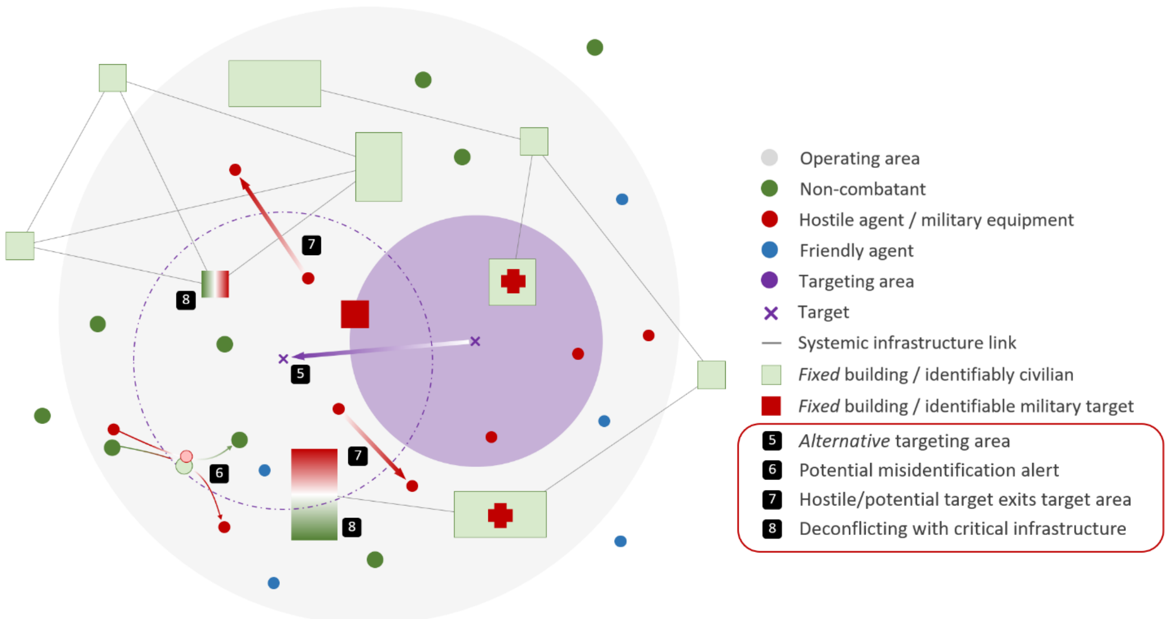
Source: CNA.

Figure 6 adds four more layers of ambiguity and uncertainty:

5. The targeting area may be moved to another location (do military units have sufficient and up-to-the-minute SA?).
6. Correctly identified hostiles that were previously outside the original targeting area may be erroneously misidentified with civilians that inadvertently stray into the new targeting area.
7. Hostiles that were previously in the newly designated targeting area maneuver outside of it.
8. Otherwise-legitimate military targets within the new target area are parts of a larger infrastructure that includes elements which, if attacked, would harm civilians. Examples include both those that are explicit, like power grids (which may be easier

to detect visually and incorporate into targeting strategies), and those that are implicit, like a sewer system (which may be hidden from visual inspection alone).

Figure 6. Schematic adding a second layer of ambiguity or uncertainty to Figure 4



Source: CNA.

Although far from complete, Figure 5 and Figure 6 highlight some of the ways in which harm to civilians (including civilian harm) can arise. These are consistent with the mechanisms provided in Figure 1 (but using real-world examples gives us more detail on those mechanisms), with the addition of other considerations such as the protection of critical infrastructure.

The potential for harm to civilians and critical infrastructure can be mitigated in practice by the following:

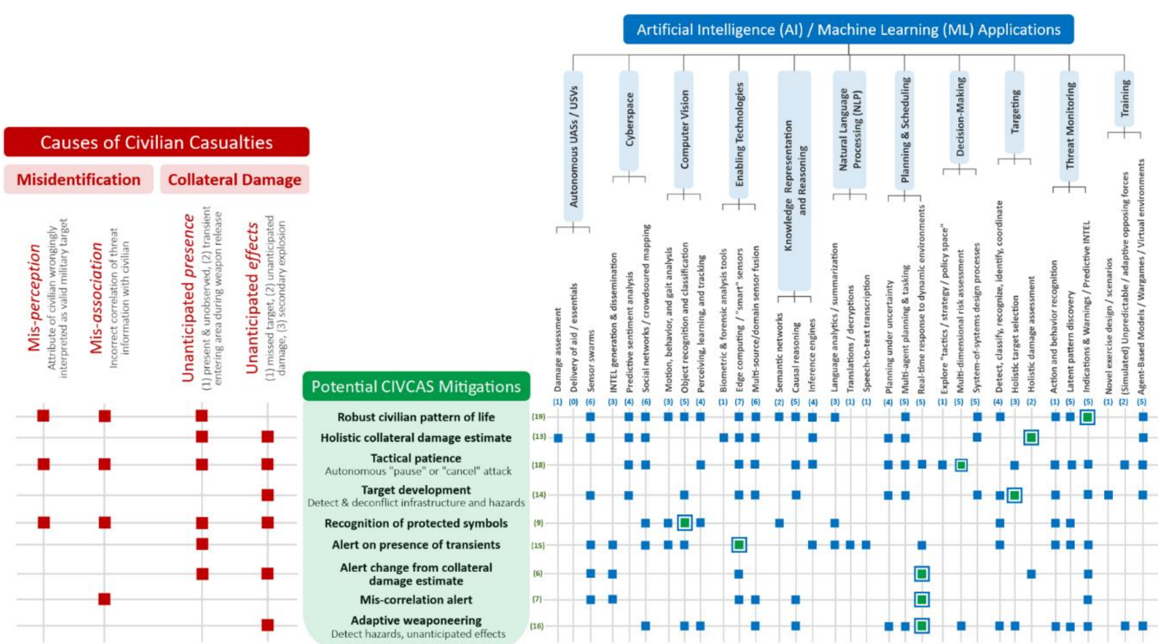
- Continually monitoring all militarily significant and civilian elements within an operating area with the greatest possible fidelity
- Continually assessing, and drawing predictive actionable inferences from, the overall activity within the operating area (i.e., its "pattern of life")
- Recognizing, on a fundamental "system of systems" level of analysis, that **all elements within a military operational area are organically and dynamically entwined** (along with attendant implications for holistic targeting and damage assessment)

We will now leverage the general relationship between the CPLC and AI/ML applications introduced in the last section to show how specific AI/ML-enabled technologies can be brought to bear on this problem.

CIVCAS Cause-Mitigations-AI-applications (CIVCAS-AI) matrix

Figure 7 shows the **CIVCAS Cause-Mitigations-AI-applications (CIVCAS-AI)** matrix. The figure actually contains two linked matrices: one that associates potential civilian harm mitigations with each of four leading causes of civilian harm (and whose entries are highlighted in **red** on the left) and another that associates potential civilian harm mitigations with specific AI/ML applications (the same ones that appear in Figure 3) civilian harm(and whose entries are highlighted in **blue** on the right). The numbers in parentheses in the second matrix (highlighted in **green** and **blue**), denote the number of "■" entries that appear in the corresponding row and column, respectively. Entries with the symbol "▣" identify the most intuitively harnessable AI/ML applications for a given civilian harm mitigation.

Figure 7. CIVCAS Cause-Mitigations-AI-applications (CIVCAS-AI) Matrix



Source: CNA.

The potential civilian harm mitigations listed in the middle column of Figure 7 mainly address the *Planning* and *Operational Execution* stages (simply because we expect AI/ML-enabled applications to have the maximum leverageable impact within these two phases). Mitigations either enact or extend a specific CPLC element (e.g., "pattern of life" under *Planning* and "tactical patience" under *Operational Execution* so that associated applications directly mirror those that appear in Figure 3) or implicitly combine one or more elements (e.g., "adaptive weaponeering," which simultaneously addresses "shape operations to mitigate risks" under *Planning* and "weaponeering to mitigate risks" under *Operational Execution*).

The right-most matrix in Figure 7 shows that **each of the nine potential civilian harm mitigations stands to benefit from multiple kinds of AI/ML applications**. Indeed, it was with one eye focused on achieving such a potential synergy that we were motivated to construct this list. We discuss each of the nine mitigation measures in turn.

Robust civilian pattern of life. Tools that can help human decision-makers better understand civilian patterns of life similarly derive from multiple overlapping AI/ML technology domains, starting with AI/ML-enabled tools for general predictive intelligence. Other applications include using sensor swarms, object recognition and classification, behavior pattern analysis, and language analytics to develop robust datasets for further analysis; applying crowdsourced mapping, causal reasoning, and latent pattern recognition techniques to those datasets to infer (or "discover") otherwise latent behaviors; and using ABM techniques to develop data-informed simulations to help anticipate target requirements and develop general civilian risk–based policy options.

Holistic collateral damage estimate, target development, and **adaptive weaponeering.** Although each of these three potential civilian harm mitigations has unique requirements (as evidenced by their slightly differing sets of associated applications), they all depend critically on the degree to which they can leverage one key capability, namely the capacity to perceive, integrate, and anticipate systemwide dynamical elements. Of course, as for other civilian harm mitigations, the utility of these mitigation tools depends on having other AI/ML applications to support them (e.g., the use of complex network analysis and pattern recognition tools to study—and discover—connectivity patterns of materiel and information flow through a system). Of importance, each of these three applications would enhance the quality of decision-making by better informing the actual magnitude of risks to civilians and associated infrastructure during the targeting process.

Tactical patience. Tactical patience is a term originally coined in 2009 during the Afghanistan campaign. A key element of command guidance in that operation, it involves three of the four elements of the OODA loop (namely, observe, orient, and decide). Exercising tactical patience means that forces take additional time as available to acquire (and deliberate about) more accurate information to make better decisions. Doing so also affords an opportunity to revisit earlier or existing assumptions and incomplete or imprecise data, consider alternative courses of action, and explore options to mitigate ubiquitous "fog of war" uncertainties, particularly those that may inadvertently lead to civilian harm.

Although the class of AI/ML-enabled tools designed to facilitate "multi-dimensional risk assessment" under *Decision-Making* may be expected to provide the most obvious immediate benefit (as highlighted by the symbol "■" in Figure 7), many other technologies can also be harnessed. Some possibilities include (1) leveraging smart sensor grids, predictive sentiment analysis of social networks, and crowdsourced mapping techniques to enhance SA; (2) applying edge-computing technologies and threat monitoring tools to help reveal latent patterns of behavior that are possibly "invisible" to human analysts alone; and (3) using knowledge representation, reasoning, and ML algorithms to explore large multidimensional "decision spaces" to help develop informed and rapid "on the fly" alternative courses of action.

Recognition of protected symbols. This is probably the most straightforward of the applications posited here, with AI/ML methods being used to identify accepted symbols for designating protected objects and alerting the operator or the chain of command accordingly. The presence of protected symbols does not mean that the location is, in fact, protected from attack: the location may have lost its protection or an unscrupulous party may be using the symbol to deter attacks, in violation of international law. But this capability would provide a safety net in cases when the protected symbol is present but was missed by operating forces.[145]

Alert on presence of transient civilians. One of the most frequent mechanisms for civilian harm in recent US operations was the movement of civilians into the target area right around the time of the engagement. An AI/ML object-identification functionality akin to that used in Project Maven could be used to automatically monitor for additional individuals around the target area and send an alert when they are detected. This process would bring such individuals to the attention of operating forces that may otherwise fixate on the target and miss transient civilian presence.

Alert change from collateral damage estimate. AI/ML methods could be used to find differences between imagery used to determine the collateral damage estimate and more recent imagery taken in support of an engagement. This process can help identify little details that operating forces might not recognize but that could be cues of unanticipated civilian presence, such as additional vehicles near a building.

Misassociation alert. Steps can also be taken to help identify that a misassociation has taken place. For example, applications could recognize that a vehicle being tracked is not the same

[145] We note that the Australian Armed Forces have recognized this application as a promising one and have already conducted field experiments showing the utility of this function.

one that was being tracked previously, showing that a swap has occurred between a threat vehicle and a civilian vehicle.

A drone strike in Kabul: Could AI have helped?

On August 29, 2021, the US announced that it had used a drone to strike the vehicle of an ISIS-K suicide bomber in Kabul, Afghanistan, averting an imminent threat to the airport during an airborne evacuation of the country. Soon, however, news reports suggested that the bomber was not the only casualty: a family was also present in the courtyard and was tragically killed in the blast. Subsequent reporting showed that there was another fundamental error: the driver of the vehicle was not in fact a suicide bomber but rather an aid worker with a humanitarian organization registered with the US Agency for International Development.[146]

Although the event is being investigated by the US military, the causes of the incident (from our framework of causes in Figure 7) appear to include both misidentification (based on appearance and behavior) and collateral damage from the presence of transient civilians. In the aftermath of this tragedy, it is worthwhile to ask: could AI-enabled applications have helped? And if so, how?

We can see several possible mitigation steps from Figure 7. For example, alerting the presence of transient civilians may have given decision-makers additional time to scrutinize the engagement area. This alert could have been combined with functions to enable tactical patience, perhaps an automated pause to an engagement in light of detected factors that could include transient civilians. Another AI application could be a functionality to develop a more robust civilian pattern of life to provide additional context to the assumptions made about the vehicle that described it as a threat. For example, reportedly the vehicle drove to the humanitarian organization's location as part of its stops during the day before it was attacked. Although the US military considers visits to suspected threat locations as part of its target development and pattern of life, there is no accompanying civilian pattern of life process to try to identify locations that could show an individual or vehicle is not a threat.

[146] Phil Stewart and Idrees Ali, "U.S. Says Kabul Drone Strike Killed 10 civilians, Including Children, in 'Tragic Mistake,'" Reuters, Sept. 18, 2021, https://www.reuters.com/world/asia-pacific/us-military-says-10-civilians-killed-kabul-drone-strike-last-month-2021-09-17/.

Conclusions

Although militaries speak of capabilities that help mitigate civilian harm, such as precision-guided munitions, those capabilities were acquired to engage military targets more effectively. They may have some benefit in mitigating harm in some circumstances but not all. For example, a precision-guided munition has no value in mitigating civilian harm when civilians have been misidentified as a military target and the munition is engaged in that mistaken belief. We do not see militaries around the world seeking to field capabilities based on their value in mitigating civilian harm.

The emerging technology of AI presents an opportunity for militaries to pursue this goal: AI-enabled and other applications for reducing the cost of war on civilians are within the realm of the possible. Both in light of legal commitments and out of interest in doing everything possible to spare civilians from harm in the waging of war, states should be asking themselves: How can we use AI to protect civilians from harm? And how can we lessen the infliction of suffering, injury, and destruction overall during armed conflict?

In this report we have provided an initial framework and methodology for identifying AI applications to help mitigate civilian harm. For example, based on our analysis of particularly beneficial mitigation steps for reducing harm to civilians that are amenable to AI applications, we suggest the following functions as promising starting points:

- **Alerting the presence of transient civilians** by using object identification to automatically monitor for additional individuals around the target area and send an alert if they are detected. This application would bring these individuals to the attention of operating forces that may otherwise fixate on the target and miss transient civilian presence.
- **Detecting a change from collateral damage estimate** by finding differences between imagery used to determine the collateral damage estimate and more recent imagery taken in support of an engagement. This application can help identify little details that operating forces might not recognize but that could be cues of unanticipated civilian presence, such as additional vehicles near a building.
- **Alerting a potential miscorrelation** by helping to identify that a miscorrelation has taken place. For example, applications could recognize that a vehicle being tracked is not the same one that was being tracked previously, showing that a swap has occurred between a threat vehicle and a civilian vehicle.
- **Recognizing protected symbols** by using AI/ML methods to identify accepted symbols for designating protected objects (e.g., red cross or red crescent) and alerting

the operator or the chain of command accordingly. The presence of protected symbols does not mean that the location is, in fact, protected from attack: the location may have lost its protection or an unscrupulous party may be using the symbol to deter attacks, in violation of international law. But this capability would provide a safety net in cases where the protected symbol is present but was missed by operating forces.[147]

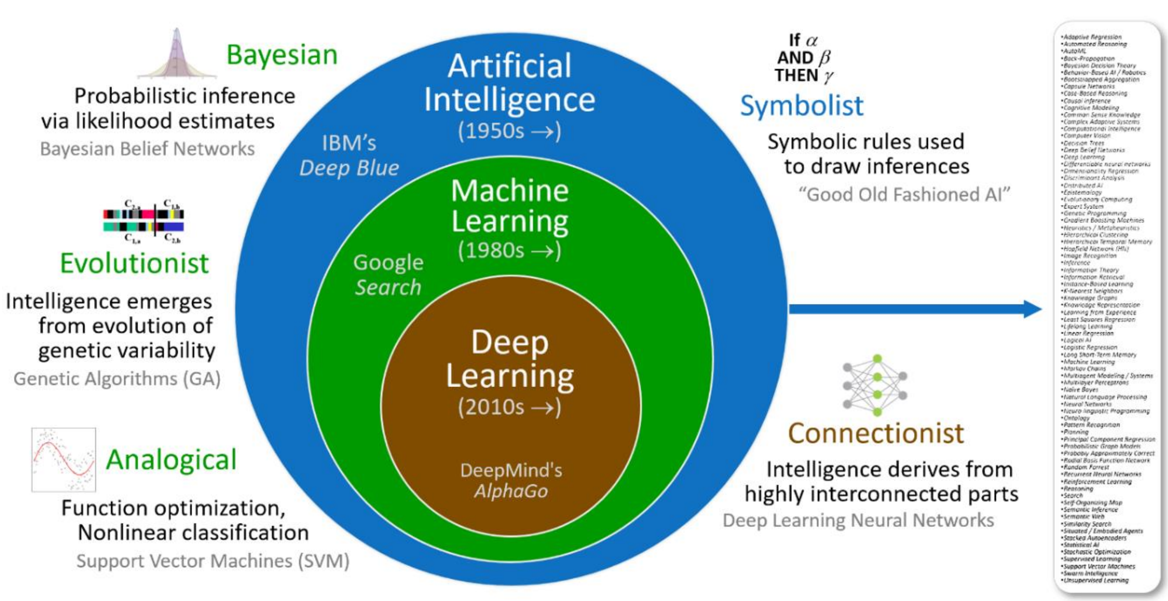
Finally, by examining a recent tragic civilian harm incident in Afghanistan, we found we could identify mechanisms that led to civilian harm from our root cause framework and thus were able to see potential AI solutions from our AI applications matrix that could have helped to avert civilian harm in that incident. The potential utility of such applications for this real world incident illustrates the possibility that AI applications can be applied in specific and practical ways to help reduce, if not eliminate, civilian harm in armed conflict.

[147] We note that the Australian Armed Forces have recognized this application as a promising one and have already conducted field experiments showing the utility of this function.

Appendix: AI/ML Approaches, Methods, and Algorithms Taxonomy

In this report (e.g., Figure 7) we discuss 33 basic classes of AI and ML applications organized into 11 “top-level” domains. Although the lists contain most of the key technologies that may be used to protect civilians and reduce civilian harm and are based partly on how AI and ML are characterized in the NSCAI’s final report,[148] the list provided in this report is neither definitive nor complete. Ideally, Figure 3 and Figure 7 would be accompanied by a third matrix that associates applications with specific methods. For expediency’s sake (because it would take us too far afield from the focus of this paper to provide the proper context for forging such an explicit association), we limit our discussion to summarizing the rudiments of what an AI/ML methodological taxonomy looks like. We draw from material previously published in a 2020 survey of state-of-the-art methods.[149]

Figure 8. The Five Tribes of AI Functions



Source: CNA.

[148] NSCAI, *Final Report*.

[149] Ilachinski, *Artificial Intelligence: Emerging Themes, Issues, and Narratives*.

Figure 8 is a schematic illustration of the "Five Tribes" of AI:[150] (1) Bayesian approaches, which rely on probabilistic inferences via likelihood estimates; (2) symbolist approaches, which are throwbacks to the 1980s' and 1990s' "Good Old-Fashioned AI" and rely on logic and symbolic rules; (3) analogical approaches, which consist largely of classical function optimization and nonlinear classification methods; (4) evolutionist approaches, which are inspired by biological evolution (wherein "solutions" to a problem—even intelligence itself—are "grown" or evolved using basic natural evolutionary processes such as recombination, crossover, and mutation); and (5) connectionist approaches, which is another label for what is arguably today's most popular class of deep neural network learning techniques. Drilling down even a single level from AI's topmost "Five Tribes" decomposition reveals a (deliberately) too-small-to-comfortably-read litany of specific methods, functions, and algorithms (the list that appears on the right-hand side of Figure 8). The takeaway point of this alphabetized list is that, despite its apparent length, it contains only a small subset of extant AI methods!

AI, having started in the 1950s, includes (but is more general than) the ML focus that ensued in the 1980s, which in turn is more general than the DL techniques currently in fashion. "Machine learning" is a catchall phrase that refers to a wide variety of techniques designed to detect patterns in and learn and make predictions from data. Specific techniques include the following:[151]

- Bayesian belief networks, which are graph models whose nodes represent some objects or states of a system and whose links denote probabilistic relationships among those nodes
- Deep learning, which refers to a class of ML algorithms designed to find multiple high levels of abstract representations of patterns in data
- Genetic algorithms and other evolutionary programming techniques that mimic the dynamics of natural selection[152]
- Inductive logic programming, designed to infer a hypothesis from a knowledge base and a set of positive and negative examples[153]

[150] The "Five Tribes" decomposition is borrowed from Pedro Domingos, *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World,* (New York: Basic Books, 2015).

[151] Russell and Norvig, *Artificial Intelligence: A Modern Approach*.

[152] Zbigniew Michalewicz and David B. Fogel, *How to Solve It: Modern Heuristics,* (New York: Springer-Verlag, 2005).

[153] Stephen H. Muggleton and Hiroaki Watanabe, eds., *Latest Advances in Inductive Logic Programming*, (London: Imperial College Press, 2014).

- Neural networks, which are inspired by the structure and function of biological neural networks[154]
- Reinforcement learning, which is inspired by behavioral psychology and refers to a technique whereby learning proceeds by adaptively constructing a sequence of actions that collectively maximize some long-term reward[155]
- Support vector machines, which are essentially multidimensional binary classification algorithms[156]

Although all ML techniques require a dataset (or multiple datasets) as a source of training data, the learning can proceed in one of three ways: supervised, semi-supervised, or unsupervised. In supervised learning, each training data element is explicitly labeled as an input-output pair, where the output is the "correct" desired value that one wishes the system to learn to associate with a given input (thereby learning the general rules by which to associate input-output pairs not in the original training set) and the "output" represents a "supervisory signal." In unsupervised learning, the system attempts to discover hidden structure in data on its own—that is, no reward signals are given to "nudge" the system as it processes the training data. Semi-supervised learning refers to a class of supervised learning techniques that also use unlabeled training data. Reinforcement learning may be considered a form of semi-supervised learning in that it neither uses input-output pairs for training nor is completely unsupervised; instead, the type of feedback the system receives depends on its response. For correct responses, it receives the same type of response as any supervised learning system does (e.g., response is "correct"); for incorrect responses, it is told that an "incorrect response" was given but is not informed of what the correct response was.

Figure 9 shows a top-level view of a taxonomy of ML organized by functional similarity, wherein each major cluster of algorithms is accompanied by a visual schematic that serves as a mnemonic reference for what a given category of algorithms is designed to do.

The mind map in Figure 10 contains the full high-resolution version of this taxonomy and includes 100-plus specific algorithms and embedded hot-link references to primary and secondary reference sources.

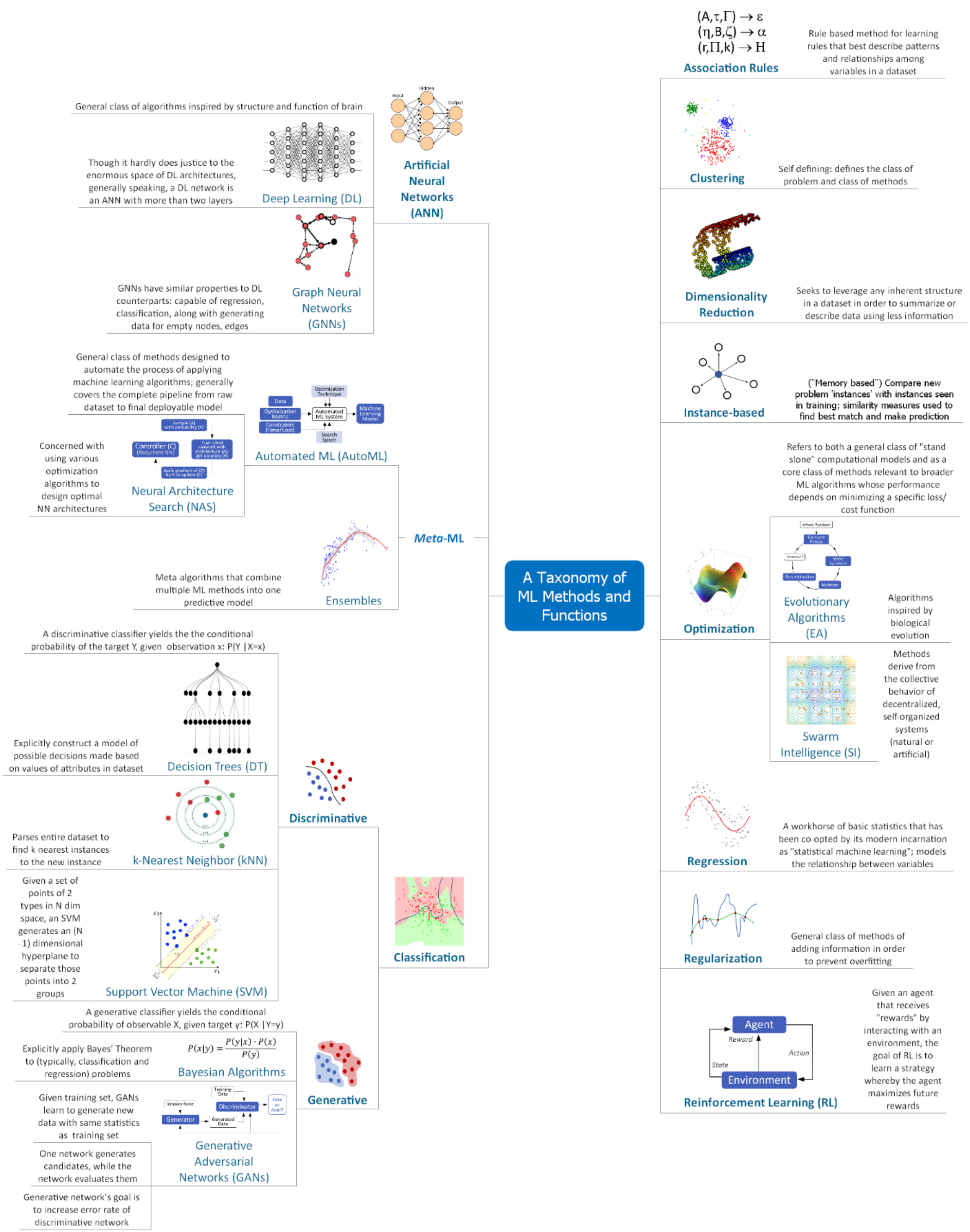
---

[154] Mohamad Hassoun, *Fundamentals of Artificial Neural Networks*, (Cambridge, MA: MIT Press, 2003).

[155] Richard S. Sutton and Andrew G. Barto, *Reinforcement Learning: An Introduction*, (Cambridge, MA: MIT Press, 1998).

[156] Nello Cristianini, *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*, (New York: Cambridge University Press, 2000).

Figure 9. High level taxonomy of ML methods



Source: CNA.

Figure 10. Full taxonomy of ML methods

[Embed the file = “High res mindmap of AI-ML approaches, methods, and algorithms taxonomy”]

Source: CNA.

Figures

Abbreviations

ABM	agent-based modeling
AI	artificial intelligence
ATR-MCAS	Aided Threat Recognition from Mobile Cooperation and Autonomous Sensors
C2	command and control
C2IMERA	Command and Control Incident Management Emergency Response Application
CATE	collaborative AI at the tactical edge
CCW	Convention for Certain Conventional Weapons
COMPASS	Collection and Monitoring via Planning for Active Situational Scenarios
CPLC	civilian protection life cycle
CST	complex system theory
CV	Computer Vision
DARPA	Defense Advanced Research Projects Agency
DL	Deep learning
DOD	Department of Defense
IHL	international humanitarian law
ISAF	International Security Assistance Force
KR&R	Knowledge Representation and Reasoning
LAWS	lethal autonomous weapon systems
ML	machine learning
NLP	Natural Language Processing
NSCAI	National Security Commission on Artificial Intelligence
OODA	observe, orient, decide, act
P&S	Planning and Scheduling
SA	situational awareness
UAS	unmanned aerial system
UAV	unmanned aerial vehicle
UK	United Kingdom
UN	United Nations
USV	unmanned surface vessel

References

Abdelhamid, Abdelaziz, Hamzah Alsayadi, Islam Hegazy, and Zaki Taha Fayed. "End-to-End Arabic Speech Recognition: A Review." In 19th Conference on Language Engineering (ESOLEC'19), Sept. 2020.

Ackerman Evan, and Michael Koziol. "In the Air with Zipline's Medical Delivery Drones." *IEEE Spectrum* (Apr. 30, 2019). https://spectrum.ieee.org/in-the-air-with-ziplines-medical-delivery-drones/particle-1.

"Adaptable Command and Control System Allowed Georgia Air Force Base to Track and Prepare for Hurricane Dorian." Leidos. Oct. 1, 2019. https://www.leidos.com/insights/adaptable-command-and-control-system-allowed-georgia-air-force-base-track-and-prepare.

Adewole, Sodiq, Erfaneh Gharavi, Benjamin Shpringer, Martin Bolger, Vaibhav Sharma, Sund Ming Yang, and Donald E. Brown. "Dialogue-Based Simulation for Cultural Awareness Training." Feb. 1, 2020. arXiv:2002.00223v1 [cs.CY].

Aggarwal, Ravi, Viknesh Sounderajah, Guy Martin, Daniel S. W. Ting, Alan Karthikesalingam, Dominic King, Hutan Ashrafian, et al. "Diagnostic Accuracy of Deep Learning in Medical Imaging: A Systematic Review and Meta-Analysis." *npj Digital Medicine* 4, article 65 (Apr. 7, 2021). https://doi.org/10.1038/s41746-021-00438-z.

Allesina, Stefano, and Mercedes Pascual. "Googling Food Webs: Can an Eigenvector Measure Species' Importance for Coextinctions?" *PLoS Computational Biology* 5, no. 9 (2009).

Amazon. "In the News: How Amazon Is Using Machine Learning to Eliminate 915,000 Tons of Packaging." Jan. 29, 2021. https://aws.amazon.com/blogs/industries/how-amazon-is-using-machine-learning-to-eliminate-915000-tons-of-packaging/.

Australia, Canada, Japan, South Korea, the United Kingdom, and the United States. *Discussion Paper – Building on Chile's Proposed Four Elements of Further Work for the Convention on Certain Conventional Weapons (CCW) Group of Governmental Experts (GGE) on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems (LAWS)*. Submitted to the UN CCW June 2021.

Badia, Adria, Bilal Piot, Steven Kapturowski, Pablo Sprechmann, Alex Vitvitskyi, Daniel Guo, and Charles Blundell. "Agent57: Outperforming the Atari Human Benchmark." Mar. 30, 2020. arXiv:2003.13350.

Baker, Chris, Sarvapali Ramchurn, W.T. Luke Teacy, and Nicholas R. Jennings. "Planning Search and Rescue Missions for UAV Teams." In Proceedings of the Twenty-second European Conference on Artificial Intelligence, Aug. 2016.

Barker, Alex. "Giving Precision Munitions 'Eyes' and a 'Brain': The State of PLA Research on Military Target Recognition." *China Brief* 21, no. 13 (July 2, 2021). https://jamestown.org/program/giving-precision-munitions-eyes-and-a-brain-the-state-of-pla-research-on-military-target-recognition/.

Barlos, Fotis, Anna Skinner, Richard Peeke, Robert Mohan, Andjelka Kelic, Walter Beyeler, Rossitza Homan, et al. *Collection and Monitoring via Planning for Active Situational Scenarios (COMPASS): Strategic Multi-Layer Assessment Report*. Sandia National Laboratories. SAND2020-0136R. Jan. 2020. https://www.osti.gov/servlets/purl/1592839.

Biggs, Adam, Matthew Cain, and Stephen Mitroff. "Cognitive Training Can Reduce Civilian Casualties in a Simulated Shooting Environment." *Psychological Science* (July 13, 2015).

Blasch, Erik, Tien Pham, Chee-Yee Chong, Wolfgang Koch, Henry Leung, Dave Braines, and Tarek Abdelzaher. "Machine Learning/Artificial Intelligence for Sensor Data Fusion–Opportunities and Challenges." *IEEE Aerospace and Electronic Systems Magazine* 36, no. 7 (July 2021).

Blumenstein, Richard. "Kessel Run's C2IMERA Used During Afghan Evacuation." Wright-Patterson AFB. Sept. 23, 2021. https://www.wpafb.af.mil/News/Article-Display/Article/2786182/kessel-runs-c2imera-used-during-afghan-evacuation/.

Brown, Noam, and Tuomas Sandholm. "Superhuman AI for Multiplayer Poker." *Science* 365 (Aug. 30, 2019).

Burns, Lawrence D. *Autonomy: The Quest to Build the Driverless Car—And How It Will Reshape Our World*. New York: Ecco Press, 2018.

"Can a Pandemic Drone Help Stop the Spread of COVID-19?" Draganfly website. July 3, 2020. https://draganfly.com/news/can-a-pandemic-drone-help-stop-the-spread-of-covid-19/.

Carvalho, L.E., and A. von Wangenheim. "3D Object Recognition and Classification: A Systematic Literature Review." *Pattern Analysis and Applications* 22 (Feb. 2019).

Chandra Yogesh, and Antoreep Jana. "Sentiment Analysis Using Machine Learning and Deep Learning." In 7th International Conference on Computing for Sustainable Global Development, New Delhi, India, Mar. 12-14, 2020, IEEE.

Chen, Chao, Weidong Bao, Tong Men, Xiaomin Zhu, Ji Wang, and Rui Wang. "NECTAR-An Agent-Based Dynamic Task Allocation Algorithm in the UAV Swarm." *Complexity* 2020 (Sept. 2020).

Chen, Liang-Chu, Chia-Meng Lee, and Mu-Yen Chen. "Exploration of Social Media for Sentiment Analysis Using Deep Learning." *Soft Computing* 24 (2020).

Chung, Timothy H. "OFFensive Swarm-Enabled Tactics (OFFSET)." Mar. 2021. https://apps.dtic.mil/sti/pdfs/AD1125864.pdf.

Cody, Jason, Karina Roundtree, and Julie Adams. "Human-Collective Collaborative Target Selection." *ACM Transactions on Human-Robot Interaction* 10, no. 2 (Mar. 2021).

Cook, Christina. "Designing a Virtual Embedded Scenario-Based Military Simulation Training Program Using Educational and Design Instructional Strategies." Doctoral thesis, University of Central Florida, 2018.

Cox, Mathew. "Army's Next Infantry Weapon Could Have Facial-Recognition Technology." Real Clear Defense. June 3, 2019.

Cristianini, Nello. *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*. New York: Cambridge University Press, 2000.

"DARPA Subterranean Challenge Announces Systems Competition Teams for Final Event." DARPA website. May 3, 2021. https://www.darpa.mil/news-events/2021-05-03.

"DARPA Tiles Together a Vision of Mosaic Warfare." DARPA website. Sept. 24, 2018. https://www.darpa.mil/work-with-us/darpa-tiles-together-a-vision-of-mosiac-warfare.

DeBerry, William, Richard Dill, Kenneth Hopkinson, Douglas Hodson, and Michael Gramaila. "The Wargame Commodity Course of Action Automated Analysis Method." *The Journal of Defense Modeling and Simulation: Applications, Methodology, Technology* (July 19, 2021).

Dellerman, Dominik, Adrian Calma, Nikolaus Lipusch, Thorsten Weber, Sascha Weigel, and Philipp Ebel. "The Future of Human-AI Collaboration: A Taxonomy of Design Knowledge for Hybrid Intelligence Systems." May 7, 2021. arXiv:2105.03354 [cs.AI].

Department of Defense. *Foreign Internal Defense*. Joint Publication 3-22. Aug. 17, 2018. https://www.jcs.mil/Portals/36/Documents/Doctrine/pubs/jp3_22.pdf.

Department of Defense. *Joint Staff civilian harm Review*. Apr. 17, 2018. https://www.justsecurity.org/wp-content/uploads/2019/02/Civ-Cas-Study-Redacted-just-security.pdf.

Department of Defense. "Mission Initiatives: Joint Warfighting Operations." JAIC website. https://www.ai.mil/mi_joint_warfighting_operations.html.

Department of Defense. "Mission Initiatives: Threat Reduction and Protection." JAIC website. https://www.ai.mil/mi_threat_reduction_and_protection.html.

De Simone, Sergio. "Ushahidi and the Power of Crowdsourcing." InfoQ. June 27, 2015.

De Spiegeleire, Stephan, Matthijs Maas, and Tim Sweijs. *Artificial Intelligence and the Future of Defense: Strategic Implications for Small and Medium-Sized Force Providers*. The Hague Centre for Strategic Studies. 2017. https://www.jstor.org/stable/resrep12564.

Domingos, Pedro. *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World*. New York: Basic Books, 2015.

Edmonds, Jeffrey, Samuel Bendett, Anya Fink, Mary Chestnut, Dmitry Gorenburg, Michael Kofman, Kasey Stricklin, et al. *Artificial Intelligence and Autonomy in Russia*. CNA. DRM-2021-U-029303-Final. May 2021.

Ekelhof, Merel A. C. "Lifting the Fog of Targeting: 'Autonomous Weapons' and Human Control Through the Lens of Military Targeting." *Naval War College Review* 71, no. 3 (2018).

El-Jawad, Mohammed, Rania Hodlod, and Yasser Omar. "Sentiment Analysis of Social Media Networks Using Machine Learning." In 14th International Computer Engineering Conference, Cairo, Egypt, Dec. 29-30, 2018, IEEE.

Elliot, Joshua. "Artificial Social Intelligence for Successful Teams (ASIST)." DARPA website. Accessed Sept. 20, 2021. https://www.darpa.mil/program/artificial-social-intelligence-for-successful-teams.

Elzayady, Hossam, Khaled Badran, and Gouda Salama. "Arabic Opinion Mining Using Combined CNN - LSTM Models." *International Journal of Intelligent Systems and Applications* 12, no. 4 (Aug. 2020).

Encyclopedia Britannica. s.v. "artificial intelligence." Accessed Sept. 20, 2021. https://www.britannica.com/technology/artificial-intelligence.

Fan, Wenhui, Peiyu Chen, Daiming Shi, Xudong Guo, and Li Kou. "Multi-Agent Modeling and Simulation in the AI Age." *Tsinghua Science and Technology* 26, no. 5 (Oct. 2021).

Ferraris, Patrick. "Aided Detection on the Future Battlefield." Defense Visual Information Distribution Service. Jan. 24, 2020. https://www.dvidshub.net/news/360225/aided-detection-future-battlefield.

Forrester, Bruce, and Kees den Hollander. "The Role of Social Media in the Intelligence Cycle." In Next-Generation Analyst IV, Proceedings Vol. 9851. May 12, 2016.

Franceshelli, Giorgio, and Mirco Mosolesi, "Creativity and Machine Learning: A Survey." Apr. 20, 2021. arXiv:2104.02726v2.

GAIA: A Fine-Grained Multimedia Knowledge Extraction System. Blender Lab, UIUC. Video. http://blender.cs.illinois.edu/software/gaia-ie/gaia.mp4.

GDELT homepage. Accessed Sept. 20, 2021. https://www.gdeltproject.org/.

Geluvaraj, B., P. Satwik, and T. Kumar. "The Future of Cybersecurity: Major Role of Artificial Intelligence, Machine Learning, and Deep Learning in Cyberspace." S. Smys, R. Bestak, J.Z. Chen, and I. Kotuliak, eds. In International Conference on Computer Networks and Communication Technologies. Singapore: Springer, 2019.

Gettinger, Dan. *The Drone Databook*. The Center for the Study of the Drone at Bard College. Sept. 2019. https://dronecenter.bard.edu/files/2019/10/CSD-Drone-Databook-Web.pdf.

Ghataoura, Darminder, and Sam Ogbannaya. "Application of Image Captioning and Retrieval to Support Military Decision Making." In International Conference on Military Communication and Information Systems (ICMCIS), the Hague, Netherlands, May 4-5, 2021, IEEE.

Gregg, Aaron. "In a First, Air Force Uses AI on Military Jet*."* *Washington Post*. Dec. 16, 2020. https://www.washingtonpost.com/business/2020/12/16/air-force-artificial-intelligence/.

Han, Ji, Serhad Sarica, Feng Shi, and Jianxi Luo. "Semantic Networks for Engineering Design: State of the Art and Future Directions." *Journal of Mechanical Design* 144, no. 2 (Sept. 9, 2021).

Hara, Kotaro, and Jon E. Froehlich, "Characterizing and Visualizing the Physical World Accessibility at Scale Using Crowdsourcing, Computer Vision, and Machine Learning." *ACM SIGACCESS Accessibility and Computing* 113 (Oct. 2015).

Hassoun, Mohamad. *Fundamentals of Artificial Neural Networks*. Cambridge, MA: MIT Press, 2003.

He, Kaiming, Xiangu Zhang, Shaoqing Ren, and Jian Sun. "Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification." Feb. 6, 2015. arXiv:1502.01852v1 [cs.CV].

Hedberg, Sara Reese. "DART: Revolutionizing Logistics Planning." *IEEE Intelligent Systems* 17 (May 2002).

Host, Pat. "Deep Learning Analytics Develops DARPA Deep Machine Learning Prototype." Defense Daily. May 11, 2016. https://www.defensedaily.com/deep-learning-analytics-develops-darpa-deep-machine-learning-prototype/advanced-transformational-technology/.

Hou, Shenguluan, and Ruqian Lu. "Knowledge-Guided Unsupervised Rhetorical Parsing for Text Summarization." *Information Systems* 94 (Dec. 2020).

IARPA. *Request for Information*. IARPA-RFI-19-06. Mar. 26, 2019. https://www.iarpa.gov/engage-with-us/rfi?scroll_position=448.9639892578125&rfi_filter=past.

Ilachinski, Andrew. *AI, Robots, and Swarms*. CNA. DRM-2017-U-014796-Final. Jan. 2017.

Ilachinski, Andrew. *Applications of Social Media to Military Operations: Overview and Assessment*. CNA. DME-2013-U-005368-Final. July 2013.

Ilachinski, Andrew. *Artificial Intelligence: Emerging Themes, Issues, and Narratives*. CNA. DOP-2020-U-028073-Final. Oct. 2020.

International Committee of the Red Cross. *International Humanitarian Law: Answers to Your Questions*. June 9, 2020.

Ish, Daniel, Jared Ettinger, and Christoper Ferris. *Evaluating the Effectiveness of Artificial Intelligence Systems in Intelligence Analysis*. RAND Corporation. RR-A464-1. 2021.

Jadhav, Nagesh, and Rekha Sugandhi. "Survey on Human Behavior Recognition Using Affective Computing." In 2018 IEEE Global Conference on Wireless Computing and Networking (GCWCN), Lonavala, India, Nov. 23-24, 2018, IEEE.

Janai, Joel, Fatma Guney, Aseem Behl, and Andreas Geiger. "Computer Vision for Autonomous Vehicles: Problems, Datasets and State of the Art." Mar. 17, 2021. arXiv:1704.05519v3 [cs.CV].

Jensen, Benjamin, Scott Cuomo, and Chris Whyte. "Wargaming with Athena: How to Make Militaries Smarter, Faster, and More Efficient with Artificial Intelligence." War on the Rocks. June 5, 2018. https://warontherocks.com/2018/06/wargaming-with-athena-how-to-make-militaries-smarter-faster-and-more-efficient-with-artificial-intelligence/.

Johnson, Bonnie, and William Treadway. "Artificial Intelligence—An Enabler of Naval Tactical Decision Superiority." *AI Magazine* 40, no. 1 (Mar. 2019).

Kabudi, Tumaini, Ilias Pappas, and Dag Olsen. "AI-Enabled Adaptive Learning Systems: A Systematic Mapping of the Literature." *Computers and Education: Artificial Intelligence* 2 (2021).

Konaev, Margarita and Samuel Bendett. "Russian AI-Enabled Combat: Coming to a City near You?" War on the Rocks. July 31, 2019. https://warontherocks.com/2019/07/russian-ai-enabled-combat-coming-to-a-city-near-you/.

Langville, Amy, and Carl Meyer. *Google's PageRank and Beyond*. Princeton, NJ: Princeton University Press, 2006.

Laskey, Kathryn B. "Crowdsourced Decision Support for Emergency Responders." In 18th International Command and Control Research and Technology Symposium, Alexandria, VA, June 19-21, 2013. https://apps.dtic.mil/sti/pdfs/ADA588344.pdf.

Lehman, Joel, Jeff Clune, Dusan Misevic, Christoph Adami, Lee Altenberg, Julie Beaulieu, Peter Bentley, et al. "The Surprising Creativity of Digital Evolution: A Collection of Anecdotes from the Evolutionary Computation and Artificial Life Research Communities." Nov. 21, 2019. arXiv:1803.03453.

Lewis, Larry, and Diane Vavrichek. *Rethinking the Drone War*. Quantico, VA: Marine Corps University Press, 2016.

Larry Lewis, *AI Safety: An Action Plan for the Navy.* CNA, DOP-2019-U-021957-1Rev, October 2019.

Lewis, Larry. *Drone Strikes: Civilian Casualty Considerations*. Joint Staff. June 18, 2013.

Lewis, Larry. *Insights for the Third Offset*. CNA. DRM-2017-U-016281-Final. Sept. 2017.

Lewis, Larry. *Learning from Yemen*. CNA. IRM-2019-U-019749-Final. May 2019. https://www.cna.org/CNA_files/PDF/IRM-2019-U-019749-Final.pdf.

Lewis, Larry. *Protecting Medical Care in Conflict: A Solvable Problem*. CNA. Oct. 30, 2019. https://humanrightscommission.house.gov/sites/humanrightscommission.house.gov/files/documents/Protecting%20Medical%20Care%20in%20Conflict%20-%20Lewis.pdf.

Lewis, Larry. *Reducing and Mitigating Civilian Casualties: Enduring Lessons*. Joint Staff. Apr. 12, 2013.

Li, Manling, Alireza Zareian, Ying Lin, Xiaoman Pan, Spencer Whitehead, Brian Chen, Bo Wu, et al. "GAIA: A Fine-Grained Multimedia Knowledge Extraction System." In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations, July 2020. https://aclanthology.org/2020.acl-demos.11v2.pdf.

Lopez, C. Todd. "Where It Counts, U.S. Leads in Artificial Intelligence." DOD website. July 9, 2020. https://www.defense.gov/Explore/News/Article/Article/2269200/where-it-counts-us-leads-in-artificial-intelligence/.

Lowman, Christopher, and Gerald Prater. "Expansion of the Reconnaissance and Security BCT into the Cyber Domain: Lessons Learned from NTC Rotation 17-07.05." July 2017. Unpublished white paper.

Ludwig, Jeremy, and Bart Presnell. "Developing an Adaptive Opponent for Tactical Training." In First International Conference on Adaptive Instructional Systems (AIS), Orlando, FL, July 26–31, 2019.

Marquis, Pierre, Odile Papini, and Henri Prade, eds. *A Guided Tour of Artificial Intelligence Research: Volume I: Knowledge Representation, Reasoning and Learning*. Springer-Verlag, May 2020.

McGrady, ED, and Justin Peachy, eds. *Representing Artificial Intelligence in Wargames, Connections Conference Working Group 2 Proceedings*, Dec. 2020. https://paxsims.files.wordpress.com/2020/12/connections_wg2_2020_final-4a.pdf.

Meron, Theodor. "The Martens Clause, Principles of Humanity, and Dictates of Public Conscience." *The American Journal of International Law* 94, no. 1 (Jan. 2000).

Metzler, Christopher A., David B. Lindell, and Gordon Wetzstein. "Keyhole Imaging: Non-Line-of-Sight Imaging and Tracking of Moving Objects Along a Single Optical Path." *IEEE Transactions on Computational Imaging* 7 (Dec. 22, 2020).

Miatliuk, Kanstantsin, Adam Wolniakowski, and Pawel Kolosowski. "Engineering System of Systems Conceptual Design in Theoretical Basis of Hierarchical Systems." In 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Toronto, Ontario, Canada, Oct. 11–14, 2020, IEEE.

Michalewicz, Zbigniew, and David B. Fogel. *How to Solve It: Modern Heuristics*. New York: Springer-Verlag, 2005.

Michel, Arthur Holland. *Unarmed and Dangerous: The Lethal Applications of Non-Weaponized Drones*. The Center for the Study of the Drone at Bard College. Mar. 2020. https://dronecenter.bard.edu/files/2020/03/CSD-Unarmed-and-Dangerous-Web.pdf.

Mobus, George, and Michael Kalton. *Principles of Systems Science (Understanding Complex Systems)*. Springer, 2016.

Molinara, Mario, Alessandro Bria, Saverio De Vito, and Claudio Marrocco, eds. "Artificial Intelligence for Distributed Smart Sensing." Special issue, *Pattern Recognition Letters* (Jan. 2021).

Montoya, Jose, and Ricard Sole, "Small World Patterns in Food Webs." *Journal of Theoretical Biology* 214 (2002).

Montoya, Jose, Stuart Pimm, and Ricard Sole. "Ecological Networks and Their Fragility." *Nature* 442 (2006).

Moore, Andrew, Martial Hebert, and Shane Shaneman. "The AI Stack: A Blueprint for Developing and Deploying Artificial Intelligence." In Ground/Air Multisensor Interoperability, Integration, and Networking for Persistent ISR IX, Proceedings Vol. 10635, May 4, 2018.

Mourabit, Yousef, Youssef El Habouz, Mustapha Lydiri, and Hicham Zougagh. "A New Sentiment Analysis System of Tweets Based on Machine Learning Approach." *International Journal of Scientific and Technology Research* 9, no. 12 (Dec. 2020).

Moy, Glenn, and Slava Shekh. "The Application of AlphaZero to Wargaming." In Advances in Artificial Intelligence, Proceedings of 32nd Australasian Joint Conference, Adelaide, SA, Australia, Dec. 2019.

Muggleton, Stephen H., and Hiroaki Watanabe, eds. *Latest Advances in Inductive Logic Programming*. London: Imperial College Press, 2014.

NATO/ISAF, Tactical Directive, July 6, 2009. https://www.nato.int/isaf/docu/official_texts/Tactical_Directive_090706.pdf.

NSCAI. *Final Report*. Mar. 2021. https://www.nscai.gov/wp-content/uploads/2021/03/Full-Report-Digital-1.pdf.

"OFFensive Swarm-Enabled Tactics (OFFSET)." DARPA website. https://www.darpa.mil/program/offensive-swarm-enabled-tactics.

Paltrinieri, Nicola, Louise Comfort, and Genserik Reniers. "Learning About Risk: Machine Learning for Risk Assessment." *Safety Science* 118 (Oct. 2019).

Pearl, Judea. "The Seven Tools of Causal Inference, with Reflections on Machine Learning." *Communications of the ACM* 62, no. 3 (Mar. 2019).

Pellerin, Cheryl. "Project Maven to Deploy Computer Algorithms to War Zone by Year's End." DOD website. July 21, 2017. https://www.defense.gov/Explore/News/Article/Article/1254719/project-maven-to-deploy-computer-algorithms-to-war-zone-by-years-end/.

Pomerleau, Mark. "Air Force Software Tool Helped Coordinate Afghanistan Evacuation of Civilians." C4ISRNet. Sept. 23, 2021. https://www.c4isrnet.com/battlefield-tech/c2-comms/2021/09/23/air-force-software-tool-helped-coordinate-afghanistan-evacuation-of-civilians/.

Popescu, Dan, Florian Stoican, Grigore Stamatescu, Oana Chenaru, and Loretta Ichim. "A Survey of Collaborative UAV–WSN Systems for Efficient Monitoring." *Sensors* 19, no. 21 (Oct. 28, 2019). https://www.mdpi.com/1424-8220/19/21/4690/htm.

Ranaei, Samira, and Arho Suominen. "Using Machine Learning Approaches to Identify Emergence: Case of Vehicle Related Patent Data." In 2017 Portland International Conference on Management of Engineering and Technology (PICMET), Portland, OR, July 9-13, 2017, IEEE.

"Real-Time Airspace Awareness and De-Confliction for Future Battles." DARPA website. Apr. 7, 2020. https://www.darpa.mil/news-events/2020-04-07.

Regens, James L. "Augmenting Human Cognition to Enhance Strategic, Operational, and Tactical Intelligence." *Intelligence and National Security* 34, no. 5 (2019).

Reinaldo, Francisco, Marcus Siqueira, Rui Camacho, and Luis Paulo Reis. "A Tool for Multi-Strategy Learning." *Advances in AI Research in Computing Science* 26 (2006).

"Roborder: Aims & Objectives." Roborder website. https://roborder.eu/the-project/aims-objectives/.

"Roborder: Autonomous Swarm of Heterogeneous Robots for Border Surveillance." Roborder website. Accessed Sept. 20, 2021. https://roborder.eu/wp-content/uploads/2021/03/ROBORDER_General_v2.0.pdf.

Russell, Stuart, and Peter Norvig. *Artificial Intelligence: A Modern Approach*, 4th Edition. Pearson, May 2021.

Rutherford, Alex, Manuel Cebrian, Inho Hong, and Iyad Rahwan. "Impossible by Conventional Means: Ten Years on from the DARPA Red Balloon Challenge." Aug. 13, 2020. arXiv:2008.05940v1.

Saputra, Muhammad, Andrew Markham, and Niki Trigoni. "Visual SLAM (Simultaneous Localization and Mapping) and Structure from Motion in Dynamic Environments: A Survey." *ACM Computing Surveys* 51, no. 2 (June 2018).

Saunders, Charles, John Murray-Bruce, and Vivek Goyal. "Computational Periscopy with an Ordinary Digital Camera." *Nature* 365 (Jan. 23, 2019). https://www.nature.com/articles/s41586-018-0868-6.

Schedl, David C., Indrajit Kurmi, and Oliver Bimber. "Search and Rescue with Airborne Optical Sectioning." *Nature Machine Intelligence* 2 (Nov. 23, 2020).

Schrittwieser, Julian, Ioannis Antonoglou, Thomas Hubert, Karen Simonyan, Laurent Sifre, Simon Schmitt, Arthur Guez, et al. "Mastering Atari, Go, Chess and Shogi by Planning with a Learned Model." Nov. 19, 2019. https://arxiv.org/abs/1911.08265.

Schwartz, Peter J., Daniel O'Neill, Meghan Bentz, Adam Brown, Brian Doyle, Olivia Liepa, Robert Lawrence, et al. "AI-Enabled Wargaming in the Military Decision Making Process." In Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications II, Proceedings Vol. 11413, 2020.

Shrestha, Yash, Shiko Ben-Mehahem, and Georg von Krogh. "Organizational Decision-Making Structures in the Age of Artificial Intelligence." *California Management Review* 61, no. 4 (Aug. 2019).

Silver, David, Thomas Hubert, Julian Schrittwieser, Ioannis Antonoglou, Matthew Lai, Arthur Guez, Marc Lanctot, et al. "Mastering Chess and Shogi by Self-Play with a General Reinforcement Learning Algorithm." Dec. 5, 2017. arXiv:1712.01815v1.

Soldin, Ryan J. "SAR Target Recognition with Deep Learning." In IEEE Applied Imagery Pattern Recognition Workshop (AIPR), Washington, DC, Oct. 2018.

Sottilare, Robert. "A Hybrid Machine Learning Approach to Automated Scenario Generation (ASG) to Support Adaptive Instruction in Virtual Simulations and Games." In I3M Defense & Homeland Security Simulation Workshop, Budapest, Hungary, 2018.

Stewart, Phil, and Idrees Ali. "U.S. Says Kabul Drone Strike Killed 10 Civilians, Including Children, in 'Tragic Mistake.'" Reuters. Sept. 18, 2021. https://www.reuters.com/world/asia-pacific/us-military-says-10-civilians-killed-kabul-drone-strike-last-month-2021-09-17/.

Strum, Luke. "Intel Airmen Sharpen AI Technology for Domestic Response." Air National Guard website. Nov. 25, 2019. https://www.ang.af.mil/Media/Article-Display/Article/2025332/intel-airmen-sharpen-ai-%20technology-for-domestic-response/.

Sutton, Richard S., and Andrew G. Barto. *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press, 1998.

Tadjdeh, Yasmin. "China Threatens U.S. Primacy in Artificial Intelligence (UPDATED)." *National Defense*. Oct. 31, 2020. https://www.nationaldefensemagazine.org/articles/2020/10/30/china-threatens-us-primacy-in-artificial-intelligence.

Tan, Qiaoyu, Ninghao Liu, and Xia Hu. "Deep Representation Learning for Social Network Analysis." Apr. 18, 2019. arXiv:1904.08547 [cs.SI].

Tang, John C., Manuel Cebrian, Nicklaus A. Giacobe, Hyun-Woo Kim, Taemie Kim, and Douglas Wickert. "Reflecting on the DARPA Red Balloon Challenge." *Communications of the ACM* 54, no. 4 (Apr. 2011). https://cacm.acm.org/magazines/2011/4/106587-reflecting-on-the-darpa-red-balloon-challenge/fulltext.

Taylor, Curt. "It's Time for Cavalry to Get Serious about Cyber Reconnaissance." *Armor: Mounted Maneuver Journal* (Fall 2018). https://mcoe.azurewebsites.us/Armor/eARMOR/content/issues/2018/Fall/4Taylor18.pdf.

The White House. July 1, 2016. *Exective Order--United States Policy on Pre- and Post-Strike Measures to Address Civilian Casualties in U.S. Operations Involving the Use of Force*.

Torfi, Amirsina, Rouzbeh Shirvani, Yaser Keneshloo, Nader Tavaf, and Edward Fox. "Natural Language Processing Advancements by Deep Learning: A Survey." Feb. 27, 2021. arXiv:2003.01200v4 [cs.CL].

Torreno, Alejandro, Eva Onaindia, Antonin Komenda, and Michal Stolba. "Cooperative Multi-Agent Planning: A Survey." *ACM Computing Surveys* 50, no. 6 (Nov. 2017).

Toth, Susan, and William Hughes. "The Journey to Collaborative AI at the Tactical Edge (CATE)." In Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications III, Proceedings Vol. 11746, Apr. 12, 2021.

Trehan, Robin. "How AI Is Transforming Fraud Prevention in Banking and Finance." Deltec. May 11, 2020. https://www.deltecbank.com/2020/05/11/how-ai-is-transforming-risk-in-finance-and-banking/?locale=en.

Trent, Stoney, and Scott Lathrop. "A Primer on Artificial Intelligence for Military Leaders." *Small Wars Journal* (Aug. 2018).

UK Ministry of Defence. *SAPIENT Middleware Interface Control Document*. May 11, 2020. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/895246/SAPIENT_Interface_Control_Document_v5.0.pdf.

Ushahidi homepage. https://www.ushahidi.com/.

US Air Force. *C2IMERA Overview*. Kessel Run. Mar. 26, 2020. Video. https://www.youtube.com/watch?v=8meuTt1RQbE.

US Air Force. Kessel Run homepage. https://kesselrun.af.mil/.

Vatsa, Mayank, Richal Singh, and Angshul Majumdar, eds. *Deep Learning in Biometrics*. Boca Raton, FL: CRC Press, 2018.

Vinyals, Oriol, Igor Babuschkin, Wojciech Czarnecki, Michael Mathieu, Andrew Didzik, Junyoung Chung, David Choi, et al. "Grandmaster Level in StarCraft II Using Multi-Agent Reinforcement Learning." *Nature* 575 (Oct. 30, 2019). https://www.nature.com/articles/s41586-019-1724-z.

Wilensky, Uri, and William Rand. *An Introduction to Agent-Based Modeling*. Cambridge, MA: MIT Press, 2015.

Work, Robert. Memorandum. Subject: Establishment of an Algorithmic Warfare Cross-Functional Team (Project Maven). Apr. 26, 2017. https://dodcio.defense.gov/Portals/0/Documents/Project%20Maven%20DSD%20Memo%2020170425.pdf.

Wray, Robert E., James Kirk, and John Laird. "Language Models as a Knowledge Source for Cognitive Agents." Sept. 20, 2021. arXiv:2109.08270v2 [cs.AI].

Xiang, Tian-Zhu, Gui-Song Xia, and Liangpei Zhang. "Mini-Unmanned Aerial Vehicle-Based Remote Sensing: Techniques, Applications, and Prospects." *IEEE Geoscience and Remote Sensing Magazine* 7, no. 3 (Sept. 2019).

Yusuf, Sagir, and Chris Baber. "Inferential Reasoning for Heterogeneous Multi-Agent Mission." *International Scholarly and Scientific Research & Innovation* 14, no. 10 (2020).

Yusuf, Salisu Ibrahim, Steve Adeshina, and Moussa Mahamat Boukar. "Parameters for Human Gait Analysis: A Review." In 2019 15th International Conference on Electronics, Computer and Computation (ICECCO), Dec. 2019.

Zhao, Mingmin, Tianhong Li, Mohammad Abu Alsheikh, Yonglong Tian, Hang Zhao, Antonio Torralba, and Dina Katabi. "Through-Wall Human Pose Estimation Using Radio Signals." *Computer Vision and Pattern Recognition (CVPR)* (2018). http://openaccess.thecvf.com/content_cvpr_2018/papers/Zhao_Through-Wall_Human_Pose_CVPR_2018_paper.pdf.

Zhen, Yang, Zhang Wanpeng, and Liu Hongfu. "Real-Time Strategy Game Tactical Recommendation Based on Bayesian Network." *Journal of Physics: Conference Series* 1168, no. 3 (2019).

Zhu, Xiaoyu, Junwei Liang, and Alexander Hauptmann. "MSNet: A Multilevel Instance Segmentation Network for Natural Disaster Damage Assessment in Aerial Videos." Dec. 30, 2020. arXiv:2006.16479 [cs.CV].

This report was written by CNA's Operational Warfighting Division (OPS).

OPS focuses on ensuring that US military forces are able to compete and win against the nation's most capable adversaries. The major functional components of OPS work include activities associated with generating and then employing the force. *Force generation* addresses how forces and commands are organized, trained, scheduled, and deployed. *Force employment* encompasses concepts for how capabilities are arrayed, protected, and sustained at the operational level in peacetime and conflict, in all domains, against different types of adversaries, and under varied geographic and environmental conditions.

CNA is a not-for-profit research organization that serves the public interest by providing in-depth analysis and result-oriented solutions to help government leaders choose the best course of action in setting policy and managing operations.

3003 Washington Boulevard, Arlington, VA 22201
www.cna.org ● 703-824-2000

Dedicated to the Safety and Security of the Nation

DOP-2021-U-030953-2Rev

3003 Washington Boulevard, Arlington, VA 22201
www.cna.org ● 703-824-2000