LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

LLNL-TR-791947

Artificial Intelligence, the Final Piece to the Counterforce Puzzle?

R. Loss

September 30, 2019

Disclaimer

This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

Artificial Intelligence, the Final Piece to the Counterforce Puzzle?

Rafael Loss[1]

Introduction

It is widely expected that artificial intelligence (AI) will disrupt and fundamentally transform many aspects of social, economic, and political life over the coming decades.[2] As we are in the early stages of AI adoption in most fields, many discussions of its potential necessarily impact remain vague and superficial, including where national security is concerned.[3] Nevertheless, some scholars have productively explored AI's impact on the character of warfare,[4] the balance of power among states,[5] and human society itself.[6] This article contributes to this growing body of literature by examining how this new technology might interact with one of the most prominent features of the post-World War II strategic environment: nuclear weapons. Specifically, it asks whether AI-driven improvements to intelligence, surveillance, and reconnaissance (ISR) capabilities, i.e.,

[1] Rafael Loss was a 2019 nuclear scholar with the Project on Nuclear Issues at the Center for Strategic and International Studies. He is grateful to Leah Matchett, Brad Roberts, Lindsey Sheppard, Wes Spain, and Simone Williams for their helpful comments and suggestions. He is also indebted to the 2019 Nuclear Scholars Initiative class for countless enlightening discussions and to the PONI team for facilitating a great program. Part of this work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. The views and opinions expressed herein do not necessarily reflect those of the United States government or Lawrence Livermore National Security, LLC. LLNL-TR-776479-DRAFT

[2] For the purposes of this study, I use "artificial intelligence"/"AI" to refer to the suite of machine-learning technologies that enable computer vision/automated image recognition, i.e., the automated identification and classification of objects from imagery data.

[3] See Paige Gasser, Rafael Loss, and Andrew Reddie, *Assessing the Strategic Effects of Artificial Intelligence*, Workshop Summary, Lawrence Livermore National Laboratory, September 2018, https://cgsr.llnl.gov/content/assets/docs/Final_AI_Workshop_Summary.pdf. For a comprehensive yet concise overview of the national security implications of AI, see Kelley M. Sayler, *Artificial Intelligence and National Security* (Washington, D.C.: Congressional Research Service, 2019), https://crsreports.congress.gov/product/pdf/R/R45178. Other relevant reports include Greg Allen and Taniel Chan, *Artificial Intelligence and National Security* (Cambridge, MA: Belfer Center for Science and International Affairs, 2017); JASON, *Perspectives on Research in Artificial Intelligence and Artificial General Intelligence Relevant to DoD* (McLean, VA: The MITRE Corporation, 2017); Edward Geist and Andrew J. Lohn, *How Might AI Affect the Risk of Nuclear War?* (Santa Monica, CA: RAND Corporation, 2018); Nicholas D. Wright, ed., *AI, China, Russia, and the Global Order: Technological, Political, Global, and Creative Perspectives* (Washington, D.C.: Department of Defense, 2018); and Vincent Boulanin, ed., *The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk, Volume I: Euro-Atlantic Perspectives* (Solna, Sweden: SIPRI, 2019).

[4] Mark Gilchrist, "Emergent Technology, Military Advantage, and the Character of Future War," *The Strategy Bridge*, July 26, 2018, https://thestrategybridge.org/the-bridge/2018/7/26/emergenttechnology-military-advantage-and-the-character-of-future-war.

[5] Michael C. Horowitz, "Artificial Intelligence, International Competition, and the Balance of Power," *Texas National*
Security Review 1, no. 3 (2018), 37-57.

[6] Kenneth Payne, "Artificial Intelligence: A Revolution in Strategic Affairs?" *Survival* 60, no. 5 (2018), 7-32.

automated image recognition/computer vision, could enable an effective counterforce capability and thereby imperil nuclear deterrence and weaken first-strike stability.[7] Since the advent of the nuclear age, analysts have been divided over the meaning of nuclear weapons for international politics. Proponents of the "nuclear revolution" argue that nuclear weapons, particularly a secure second-strike capability, which would allow a state to retaliate with nuclear force after suffering a first strike, satisfy a state's security needs because they present a potential for unacceptable damage that would deter any potential challenger from even considering aggression.[8] In their view, international politics in the shadow of nuclear weapons would be largely peaceful, competition would be relegated to the margins, and nuclear superiority would be meaningless.[9] However, recent research suggests that this may not be the case. In fact, nuclear weapon states have sought to hold at risk the nuclear arsenals of their adversaries throughout the nuclear age.[10] Still, no country has attacked another's operational nuclear arsenal.[11] That is because the requirements for a disarming strike are considerably higher than for damage limitation or retaliation against nuclear attack. A state trying to disarm its opponent would have to be certain to destroy, or at least disable, all of its adversary's nuclear weapons. Even one operational nuclear warhead could inflict unacceptable damage in a retaliatory blow. The inability to be certain that a first strike could find and eliminate all enemy nuclear weapons has been the main roadblock for effective counterforce.[12] As the number of battlefield sensors grows and ISR capabilities improve, however, this roadblock is crumbling. Some expect it to be fully overcome with the help of AI.

---

[7] "Counterforce" is used to refer to a kinetic attack aimed at disarming an adversary's nuclear force. Per Glenn A. Kent and David E. Thaler, *First-Strike Stability: A Methodology for Evaluating Strategic Forces* (Santa Monica, CA: RAND Corporation, 1989), 2-3, "first-strike stability" refers to "a two-sided calculus of each side's cost of going first compared with its cost of striking second." Related to "crisis stability," which includes psychological and perceptional variables, first-strike stability solely "arises from the strategic force structure and the force postures within that force structure." Crisis stability and "arms-race stability," i.e. absence of incentives to build up a nuclear force, are generally understood to be the co-determinants of overall "strategic stability."

[8] Bernard Brodie, ed., *The Absolute Weapon: Atomic Power and World Order* (New York: Harcourt, 1946); Robert Jervis, *The Meaning of the Nuclear Revolution: Statecraft and the Prospect of Armageddon* (Ithaca: Cornell University Press, 1989).

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

[10] See Brendan R. Green and Austin Long, "The MAD Who Wasn't There: Soviet Reactions to the Late Cold War Nuclear Balance," *Security Studies* 26, no. 4 (2017), 606-41. For a more contemporary example, see Christopher Clary and Vipin Narang, "India's Counterforce Temptations: Strategic Dilemmas, Doctrine, and Capabilities," *International Security* 43, no. 3 (2019), 7-52.

[11] On the use of preventive military force to destroy another country's nuclear program prior to it achieving a weapons capability, see Rachel Elizabeth Whitlark, "Nuclear Beliefs: A Leader-Focused Theory of Counter-Proliferation," *Security Studies* 26, no. 4 (2017), 545-74.

[12] Jan Lodal, "The Counterforce Fallacy," *Foreign Affairs* 89, no. 2 (2010), 146.

On February 11, 2019, President Donald Trump issued an executive order on *Maintaining American Leadership in Artificial Intelligence*. Recognizing that "American leadership in AI is of paramount importance to maintaining the economic and national security of the United States and to shaping the global evolution of AI in a manner consistent with our Nation's values, policies, and priorities,"[13] it directed the Department of Defense (DoD) and other federal agencies to prioritize investments in AI research and development, high-performance computing, and an AI-versed workforce. DoD's own AI strategy subsequently detailed how the Pentagon views the particular risks and opportunities associated with advances in artificial intelligence.[14] DoD expects that AI will yield significant improvements for logistics, ISR, cyberspace and information operations, command and control, and semiautonomous and autonomous vehicles and weapon systems.[15] According to Deputy Secretary of Defense Robert Work, "what AI […] allows you to do is find the needle in the haystack."[16] This ability of AI to quickly analyze enormous amounts of data could help identify, locate, target, and ultimately eliminate an adversary's nuclear arsenal.

The remainder of this article proceeds in four parts: First, it outlines the theoretical underpinnings of nuclear deterrence and the military requirements of counterforce before assessing the potential of AI to enable an effective counterforce capability. Drawing on work by Keir Lieber and Daryl Press,[17] a fictional North Korea scenario anchors this discussion in the second section and suggests that AI could improve to the ability to find and eliminate time-critical targets, such as mobile nuclear-missile launchers. The third section argues that these improvements remain marginal, as serious technical limitations inherent to AI prevent it from providing results that suffice for the

[13] White House, *Executive Order on Maintaining American Leadership in Artificial Intelligence*, February 11, 2019, https://www.whitehouse.gov/presidential-actions/executive-order-maintaining-american-leadership-artificial-intelligence/.
[14] Department of Defense, *Summary of the 2018 Department of Defense Artificial Intelligence Strategy*, February 12, 2019, https://media.defense.gov/2019/Feb/12/2002088963/-1/-1/1/SUMMARY-OF-DOD-AI-STRATEGY.PDF; see also Department of Defense, *2018 National Defense Strategy Summary*, January 19, 2018, https://dod.defense.gov/Portals/1/Documents/pubs/2018-National-Defense-Strategy-Summary.pdf.
[15] See Sayler, *Artificial Intelligence and National Security*, 9-15.
[16] Phil Stewart, "Deep in the Pentagon, a Secret AI Program to Find Hidden Nuclear Missiles," *Reuters*, June 5, 2018, https://www.reuters.com/article/us-usa-pentagon-missiles-ai-insight/deep-in-the-pentagon-a-secret-ai-program-to-find-hidden-nuclear-missiles-idUSKCN1J114J. See also Department of Defense, *Memorandum on the Establishment of the Joint Artificial Intelligence Center*, June 28, 2018, https://admin.govexec.com/media/establishment_of_the_joint_artificial_intelligence_center_osd008412-18_r....pdf.
[17] Keir A. Lieber and Daryl G. Press, "The New Era of Counterforce: Technological Change and the Future of Nuclear Deterrence," *International Security* 41, no. 4 (2017), 9-49.

requirements of counterforce.[18] Policymakers cannot gain certainty that a counterforce strike would fully eliminate an adversary's retaliatory capability. Finally, the article suggests that even these limited improvements in counterforce capabilities might negatively affect international stability. At least in some situations, AI-infused ISR could provide a "good enough" counterforce capability. If leaders believe that an enemy nuclear attack is imminent, for example, they might conclude that a pre-emptive counterforce strike is warranted to limit damage to their country and military assets. Furthermore, adversaries would face greater pressure to hedge against continued interest in counterforce options and improving capabilities. Some of their countermeasures would likely be detrimental for first-strike stability.

Counterforce in Theory and (Hypothetical) Practice

Much of scholarly thinking on nuclear weapons evolved from Bernard Brodie's observation in August 1945 that the chief purpose of military force would no longer be to win wars, but to avert them.[19] Once the Soviet Union had developed nuclear weapons, too, a "balance of terror" emerged.[20] And as the superpowers' nuclear arsenals expanded—to comprise of the triad of ground-, air-, and sea-launched weapons—Brodie argues, "no sensible opponent would try to eliminate our ICBMs in an initial attack unless he believed that he could at the same time with high confidence eliminate by far the major portions of our other retaliatory forces."[21] This insight led some scholars to conclude that international relations had been fundamentally transformed. In a seeming reversal of logic, offensive strategic nuclear weapons provided states with the ultimate defense.[22] By threatening unimaginable devastation, nuclear powers could deter their adversaries

[18] These come in addition to the challenges presented by AI's integration into the intelligence processes to gather, process, exploit, and disseminate products as well as general targeting-timeline constraints.
[19] See Brodie, *The Absolute Weapon*, 76.
[20] Even among the group of civilian nuclear strategists who gathered at the RAND Corporation in the 1950s and 1960s, which comprised of Brodie, Herman Kahn, Thomas Schelling, Albert and Roberta Wohlstetter, and others, no consensus emerged on how delicate this balance was. For a critique of the assumption that general thermonuclear war is extremely unlikely, see Albert Wohlstetter, *The Delicate Balance of Terror* (Santa Monica, CA: RAND Corporation, 1958), https://www.rand.org/pubs/papers/P1472.html. For an intellectual history of RAND's contributions to deterrence theory, see Austin Long, *Deterrence – From Cold War to Long War: Lessons from Six Decades of RAND Research* (Santa Monica, CA: RAND Corporation, 2008).
[21] Bernard Brodie, "The Development of Nuclear Strategy," *International Security* 2, no. 4 (1978), 71.
[22] Jervis, *The Meaning of the Nuclear Revolution*, 1; and Charles L. Glaser, *Analyzing Strategic Nuclear Policy* (Princeton, NJ: Princeton University Press, 1990), 94-9.

from aggression. During the Cold War, the United States' and Soviet Union's expansive and secure second-strike retaliatory capabilities produced the supposedly stabilizing and largely irrevocable condition known as "mutual assured destruction" (MAD),[23] contributing, according to historian John Lewis Gaddis, to the "long peace" of the later 20th century.[24] Yet, despite their considerable investments in survivable nuclear forces, the superpowers' nuclear doctrines did not conform to the prescriptions of nuclear revolution theory. "Rather than coming to grips with" MAD, they sought "to repeal the nuclear revolution"[25] and acquired nuclear warfighting capabilities beyond those necessary to threaten each other's population centers. When arms control treaties enshrined limitations and quantitative parity in the later stages of the Cold War, the superpowers competed for qualitative advantages.[26]

Deterrence remains only one of several roles that states envision for their nuclear forces. In order to hedge against future uncertainty and deterrence failure, U.S. doctrine, for example, assigns a damage-limitation role. This requires the ability to hold at risk adversary nuclear forces so that when deterrence fails, damage to the homeland can be minimized by destroying as many adversary nuclear weapons as possible.[27] Some have also suggested that such a posture might improve a

---

[23] See Spurgeon M. Keeny and Wolfgang K. H. Panofsky, "MAD versus NUTS: Can Doctrine or Weaponry Remedy the Mutual Hostage Relationship of the Superpowers?" *Foreign Affairs* 60, no. 2 (1981), 287-304.
[24] John L. Gaddis, "The Long Peace: Elements of Stability in the Postwar International System," *International Security* 10, no. 4 (1986), 99-142.
[25] Robert Jervis, *The Illogic of American Nuclear Strategy* (Ithaca, NY: Cornell University Press, 1984), 147. Joshua Rovner contends that rather than proving wrong the theory of the nuclear revolution, the apparent mismatch between historical record and theoretical prescriptions should motivate more nuanced theorizing: a more fine-grained "strategy-grand strategy distinction helps us understand the impact of nuclear weapons in the Cold War, and it provides a new way to evaluate the theory of the nuclear revolution." Rovner, "Was There a Nuclear Revolution? Strategy, Grand Strategy, and the Ultimate Weapon," *War on the Rocks*, March 6, 2018, https://waron-therocks.com/2018/03/was-there-a-nuclear-revolution-strategy-grand-strategy-and-the-ultimate-weapon/.
[26] Examples include Green and Long, "The MAD Who Wasn't There;" Austin Long and Brendan R. Green, "Stalking the Secure Second Strike: Intelligence, Counterforce, and Nuclear Strategy," *Journal of Strategic Studies* 38, no. 1-2 (2015), 38-73; and Niccolò Petrelli and Giordana Pulcini "Nuclear Superiority in the Age of Parity: US Planning, Intelligence Analysis, Weapons Innovation and the Search for a Qualitative Edge 1969–1976," *International History Review* 40, no. 5 (2018), 1191-209.
[27] According to the latest Nuclear Posture Review, "Every U.S. administration over the past six decades has called for flexible and limited U.S. nuclear response options, in part to support the goal of reestablishing deterrence following its possible failure. This is not because reestablishing deterrence is certain, but because it may be achievable in some cases and contribute to limiting damage, to the extent feasible, to the United States, allies, and partners. The goal of limiting damage if deterrence fails in a regional contingency calls for robust adaptive planning to defeat and defend against attacks […] In the case of missile threats from regional actors in particular, U.S. missile defense and offensive options provide the basis for significant damage limitation in the event deterrence fails." Department of Defense, *2018 Nuclear Posture Review*, 23. See also Department of Defense, *2019 Missile Defense Review*, 60, https://www.defense.gov/Portals/1/Interactive/2018/11-2019-Missile-Defense-Review/The%202019%20MDR_Executive%20Summary.pdf.

nuclear weapon state's position in crisis bargaining by providing the nuclear-superior state with an advantage in resolve.[28] Moreover, with growing sophistication, counterforce capabilities could negate the threat of nuclear retaliation by enabling a disarming first strike against a nuclear-armed rival.[29]

Whether a counterforce capability is supposed to facilitate crisis bargaining, damage limitation, or disarming first strike, a key requirement for holding at risk adversary nuclear forces is to know where they are at a particular point in time.[30] This is no easy task. During Operation Desert Storm in 1991, coalition forces conducted roughly 1,500 air strikes over the course of the 43-day campaign to destroy Iraq's mobile Scud launchers. Yet, the Gulf War Air Power Survey concluded:[31]

> The actual destruction of any Iraqi mobile launchers by fixed-wing coalition aircraft remains impossible to confirm. Coalition aircrews reported destroying about eighty mobile launchers. Special operations forces claimed another score or so. Most of these reports undoubtedly stemmed from attacks that did destroy objects in the Scud launcher area. But most, if not all, of the objects involved now appear to have been decoys, vehicles such as tanker trucks that had infrared and radar signatures impossible to distinguish from those of mobile launchers and their associated support vehicles, and other objects unfortunate enough to provide 'Scud-like' signatures.

Iraq's adoption of shoot-and-scoot tactics, employing mobile transport-erector-launchers (TELs) to fire the missiles, contributed to the coalition's meager results. Additionally, the Scud crews relied on camouflage, concealment, and other deception techniques to keep their assets safe. This

---

[28] See Matthew Kroenig, "Nuclear Superiority and the Balance of Resolve: Explaining Nuclear Crisis Outcomes," *International Organization* 67, no. 1 (2013), 141-71.

[29] Lieber and Press, "New Era of Counterforce," 9. The question of whether the United States in the later Cold War seriously pursued a "splendid" first-strike capability, i.e., the ability to completely destroy an opponent's nuclear arsenal in a first strike, remains debated, although Fred Kaplan provides insights into deliberations among U.S. leaders of a first strike against Soviet nuclear forces in the context of the 1961 Berlin crisis. Kaplan, "JFK's First-Strike Plan," *The Atlantic*, October 2001, https://www.theatlantic.com/magazine/archive/2001/10/jfks-first-strike-plan/376432/.

[30] According to Lieber and Press, the new era of counterforce is being brought about by two compounding trends: increasing accuracy of nuclear delivery systems and improvements in remote sensing. These developments erode states' ability to enhance the survivability of their nuclear forces through hardening and concealment respectively. While this article focuses on the revolution in remote sensing, specifically AI-infused ISR, improved missile accuracy might be consequential: "As accuracy continues to improve, the effectiveness of conventional attacks on hard targets will continue to increase. Today, low-yield nuclear weapons can destroy targets that once required very large yield detonations. In the future, many of those targets will be vulnerable to conventional attacks." With fewer expected casualties, decisionmakers might be more willing to strike adversary nuclear arsenals. Lieber and Press, "New Era of Counterforce," 32.

[31] Thomas A. Keaney and Eliot A. Cohen, *Gulf War Air Power Survey, Summary Report* (Washington, D.C.: U.S. Government Printing Office, 1993), 83.

illustrates the staggeringly difficult challenge of finding and eliminating mobile, time-critical targets in even a largely uncontested environment. In the 2003 Iraq War, accordingly, U.S. forces went at great length to better perform in a repeat Scud hunt:[32]

> A tactical ballistic missile intelligence federation made up of fifteen different intelligence agencies and operational commands combined to do the intelligence preparation of the battlespace for Operation Iraqi Freedom. Potential launch areas or "Scud baskets" were identified. Geospatial data and analysis was generated to identify roads and paths Scud transporter-erector-launchers might traverse or potential hide sites. Intelligence on the potential Iraqi missile order of battle were combined with named areas of interest, coordinates were assigned, and "kill boxes" were identified and plotted. Then, had an engagement with Iraqi Scud transporter-erector-launchers taken place in Operation Iraqi Freedom (OIF), the triad of intelligence, surveillance, and reconnaissance, special operations forces, and attack platforms would have combined to attempt to destroy the Scud threat.

While these improvements were not put to the test as no Iraqi Scuds surfaced during the campaign, U.S. forces likely would have fared better than a decade earlier.[33] Additional capabilities that have since been integrated into the armed forces have further increased the ability to surveil the battlefield. In 2003, for example, unmanned aerial systems (UAS) were a relative novelty.[34] They since have matured and arguably become the keystone for U.S. operations worldwide.

However, with the enormous growth of sensors and data sources across all warfighting domains,[35] intelligence analysts today struggle with an overabundance of information. In the words of Deputy Secretary of Defense Robert Work, Project Maven, perhaps best-known example of the use of AI in a warfighting domain, was born out of this realization and intended to "reduce the human factors burden of [full-motion video] analysis, increase actionable intelligence, and enhance military decision-making" in support of the Defeat-ISIS campaign.[36] Frontline accounts also demonstrate the

---

[32] Barry R. Schneider, "Counterforce Targeting Capabilities and Challenges," USAF Counterproliferation Center Paper No. 22 (Maxwell AFB, AL: Air University, 2004), 18. See also Alan J. Vick et al., *Aerospace Operations Against Elusive Ground Targets* (Santa Monica, CA: RAND Corporation, 2001).
[33] Long and Green, "Stalking the Secure Second Strike," 58-60, point to six other factors making the 1991 Scud hunt an ill-fitting analogy and "distant data point from a technology perspective" for modern mobile ICBM scenarios.
[34] Schneider, "Counterforce Targeting," 19.
[35] Lieber and Press identify five trends that are "ushering in an age of unprecedented transparency" and provide analysts with ever growing amounts of data: increasingly diverse sensor platforms, a widening array of signals collected by such platforms, increasingly persistent observation, improvements in sensor resolution, and increasing data transmissions speeds. Lieber and Press, "New Era of Counterforce," 32-34.
[36] Department of Defense, *Memorandum on the Establishment of an Algorithmic Warfare Cross-Functional Team (Project Maven)*, April 26, 2017, https://www.govexec.com/media/gbc/docs/pdfs_edit/establishment_of_the_awcft_project_maven.pdf.

potential utility of AI as a force-multiplier. One former Marine artillery officer recounts the months spent developing a target list for strikes against ISIS weapon storage facilities in western Iraq in 2016: "Currently, both information collection and processing are manual, labor-intensive endeavors. AI can relieve human operators of much of that burden, performing the same tasks better and faster. […] and it will allow the military to increase its competitive advantage against both near-peer and non-state adversaries."[37] Following first successes with Project Maven, DoD announced that the effort to integrate AI for warfighting would be expanded to include additional sensor types and use cases.[38] With its requirement to find and identify mobile, time-critical targets the counterforce puzzle presents one such use case.

Counterforce in a Korea Contingency

To illustrate how advances in remote sensing could threaten the survivability of nuclear forces, Lieber and Press use a fictional scenario in which U.S. and partner forces attempt to find and track North Korean road-mobile missile launchers.[39] They argue that a combination of satellites and UAS could surveil almost the entirety of North Korea's road network, promising that its TELs would be detected by the sensors with high probability. They focus on three particular ISR

[37] Hans Vreeland, "Targeting the Islamic State, or Why the Military Should Invest in Artificial Intelligence," *War on the Rocks*, May 16, 2019, https://warontherocks.com/2019/05/targeting-the-islamic-state-or-why-the-military-should-invest-in-artificial-intelligence/.

[38] Mark Pomerleau, "What the Pentagon is Learning from its Massive Machine Learning Project," *C4ISRNET*, May 2, 2018, https://www.c4isrnet.com/intel-geoint/isr/2018/05/02/what-the-pentagon-is-learning-from-its-massive-machine-learning-project/.

[39] Lieber and Press, "New Era of Counterforce," 37-46. Per Department of Defense, *2018 Nuclear Posture Review*, 23, the United States is committed to "to locate, track, and target mobile systems of regional adversaries." For current U.S. doctrine for joint operations to counter air and missile threats, see Joint Chiefs of Staff, *Joint Publication 3-01: Countering Air and Missile Threats*, May 2, 2018, https://www.jcs.mil/Portals/36/Documents/Doctrine/pubs/jp3_01_pa.pdf?ver=2018-05-16-175020-290. For illustrative examples of how a crisis on the Korean Peninsula might escalate to and beyond the nuclear threshold, see Mark Bowden, "How to Deal with North Korea: There Are no Good Options. But Some Are Worse than Others," *The Atlantic*, July/August 2017, https://www.theatlantic.com/magazine/archive/2017/07/the-worst-problem-on-earth/528717/; Robin Wright, "What Would War with North Korea Look Like?," *The New Yorker*, September 6, 2017, https://www.newyorker.com/news/news-desk/what-would-war-with-north-korea-look-like; and Jeffrey Lewis, *The 2020 Commission Report on the North Korean Nuclear Attacks Against the United States: A Speculative Novel* (Boston, MA: Mariner Books, 2018).

platforms: satellites and standoff and penetrating UAS, equipped with synthetic aperture radar (SAR) and Ground Moving Target Indicator (GMTI) radar.[40]

North Korea is estimated to possess a stockpile of up to 60 nuclear warheads, with a capacity to produce an additional seven warheads per year.[41] It also possesses an array of TEL-based short-range, medium-range, intermediate-range, and intercontinental ballistic missile (ICBM) systems on which to mount these warheads.[42] With its antiquated air force and a submarine force tailored to coastal defense, infiltration, and espionage missions, North Korea adopted land-based ballistic missiles as the most cost-effective and survivable option among available delivery system for its nascent nuclear force.[43] A long-standing reliance on the Korean People's Army Ground Force as its main service branch and its experience with artillery and ballistic missiles also contributed to North Korea's decision to prioritize a land-based nuclear force. Its small territory and challenging geography, further, led it to opt for mobile delivery systems operating from underground facilities over fixed, hardened silos. While mobile missiles were an effective response to the increasing accuracy of precision-strike weapons in the later decades of the 20th century, mobile systems face several constraints, including geographical restrictions, outsized alert signatures, and high operational demands.[44] These are at risk of being exploited by the sensing revolution, making mobile systems, too, vulnerable to counterforce attacks.

While in peacetime North Korea's nuclear weapons are likely located in underground storage facilities, possibly even in a disassembled state to exert maximum central control by the supreme leader, the TELs would be dispersed in a crisis to enhance their survivability and facilitate North

[40] For information on the use of SAR and GMTI in military space radar systems, see Joseph Post and Michael Bennett, *Alternatives for Military Space Radar* (Washington, D.C.: Congressional Budget Office, 2007).
[41] Mary Beth D. Nikitin, *North Korea's Nuclear and Ballistic Missile Programs* (Washington, D.C.: Congressional Research Service, 2019), https://crsreports.congress.gov/product/pdf/IF/IF10472.
[42] Hans M. Kristensen and Robert S. Norris, "North Korean Nuclear Capabilities, 2018," *Bulletin of the Atomic Scientists* 74, no. 1 (2018), 41-51.
[43] See James Hackett and Mark Fitzpatrick, *The Conventional Military Balance on the Korean Peninsula* (London, UK: International Institute for Strategic Studies, 2018). On North Korea's efforts to diversify its nuclear launch platforms and operationalize a sea-based deterrent capability, see also Ankit Panda, "The Sinpo-C-Class: A New North Korean Ballistic Missile Submarine Is Under Construction," *The Diplomat*, October 18, 2017, https://thediplomat.com/2017/10/the-sinpo-c-class-a-new-north-korean-ballistic-missile-submarine-is-under-construction/. For why North Korea might be at a comparative disadvantage in developing a survivable sea-based deterrent, see Owen R. Coté, "Invisible Nuclear-Armed Submarines, or Transparent Oceans? Are Ballistic Missile Submarines Still the Best Deterrent for the United States?" *Bulletin of the Atomic Scientists* 75, no. 1 (2019), 30-5.
[44] See Paul Bracken, "The Cyber Threat to Nuclear Stability," Orbis 60, no. 2 (2016), 192-97.

Korea's "asymmetric escalation" strategy to deter invasion and/or compel concessions.[45] In recent years, nongovernmental analysts have used open-source and commercial satellite photographs to identify many of North Korea's missile bases and likely launch sites.[46] U.S. intelligence agencies almost certainly have even more detailed information about the deployment patterns of North Korean nuclear forces. In a crisis situation though, U.S. ISR assets would be critical for finding and tracking North Korea's TEL-based nuclear weapons in real time.[47]

Based on geospatial analysis, Lieber and Press argue that an array of 20 SAR satellites could provide coverage of at least 90 percent of North Korea's roads on as many as 50 passes per day. An additional four standoff SAR/GMTI UAS and four penetrating UAS, if positioned correctly, could provide persistent coverage of approximately 97 percent of the road network. Moreover, since the U.S. ISR arsenal includes many more assets, such as cyberspying/-intelligence, ground-based sensors, and satellites and UAS scanning other parts of the electromagnetic spectrum, they conclude that "concealment is under great duress."[48] However, their analysis excludes the process of generating actionable intelligence from data captured by the sensors. It instead assumes that North Korean TELs that are captured by satellites and UAS are also identified and classified as targets for counterforce strikes and therefore at risk. While imposing this restriction makes sense to simplify analysis and highlight the *potential* for advances in remote sensing to undermine concealment, it

[45] See Van Jackson, "What Is North Korea's Nuclear Strategy?" *The Diplomat*, May 28, 2015, https://thediplomat.com/2015/05/what-is-north-koreas-nuclear-strategy/; Vipin Narang, "Why Kim Jong Un Wouldn't Be Irrational to Use a Nuclear Bomb First," *Washington Post*, September 8, 2017, https://www.washingtonpost.com/outlook/why-kim-jong-un-wouldnt-be-irrational-to-use-a-nuclear-bomb-first/2017/09/08/a9d36ca4-934f-11e7-aace-04b862b2b3f3_story.html?noredirect=on&utm_term=.734b909d7c0d; Vipin Narang and Ankit Panda, "Command and Control in North Korea: What a Nuclear Launch Might Look Like, Like," *War on the Rocks*, September 15, 2017, https://warontherocks.com/2017/09/command-and-control-in-north-korea-what-a-nuclear-launch-might-look-like/; and Joseph Bermudez, Victor Cha, and Lisa Collins, "Undeclared North Korea: Missile Operating Bases Revealed," *Beyond Parallel*, Center for Strategic and International Studies, November 12, 2018, https://beyondparallel.csis.org/north-koreas-undeclared-missile-operating-bases/. On the evolution of North Korean thinking about nuclear strategy, see Joseph S. Bermudez, Jr., *North Korea's Development of a Nuclear Weapons Strategy* (Washington, D.C.: US-Korea Institute at SAIS, 2015).

[46] See, for example, the work conducted by Beyond Parallel at the Center for Strategic and International Studies, available at https://beyondparallel.csis.org/imagery/, and 38 North at the Henry L. Stimson Center, available at https://www.38north.org/topics/satellite-analysis/.

[47] On damage limitation/counterforce vis-à-vis North Korea, see John R. Harvey, "Negating North Korea's Nukes," *Defense News*, February 15, 2016, https://www.defensenews.com/opinion/commentary/2016/02/15/commentary-negating-north-koreas-nukes/; and Vince A. Manzo and John K. Warden, "The Least Bad Option: Damage Limitation and U.S. Deterrence Strategy toward North Korea," *Texas National Security Review* Policy Roundtable, February 07, 2018, https://tnsr.org/roundtable/policy-roundtable-good-choices-comes-north-korea/#essay6.

[48] Lieber and Press, "New Era of Counterforce," 48. See also Long and Green, "Stalking the Secure Second Strike," 60-4.

underappreciates the challenge of drawing conclusions from raw data. In an escalating crisis on the Korean Peninsula, U.S. forces would likely leverage more than just eight UAS and the occasional satellite overpass for intelligence preparation of the battlespace.[49] The input from the ever-expanding number of sensors, however, could quickly overwhelm analysts in their task to identify, classify, and cue potential time-critical targets.

Hold Your AI-Infused Counterforce Horses

AI promises to increase operational tempo and decrease uncertainty in combat environments where "speed is paramount, and the fog of war persists."[50] Similar to the way Project Maven helps analyze imagery data in support of the Defeat-ISIS campaign, AI could prove decisive in a North Korea contingency by improving the ability to promptly analyze sensor data for identifying and tracking North Korea's TEL-based nuclear force. Despite this broad and growing enthusiasm for military applications of AI—including AI-infused ISR for counterforce missions[51]—it is unlikely to prove revolutionary. Technical limitations prevent even AI-enabled counterforce from guaranteeing success. AI could also introduce new vulnerabilities into ISR systems for adversaries to exploit. Faulty assumptions derived from previous technological innovations, moreover, could lead decisionmakers to be overly optimistic about the improvements to counterforce yielded by AI, leading to misinformed decisions over the use of military force.[52]

Machine-learning algorithms for computer vision rely on input data for training and reliable object recognition. As discussed above, these can be provided by a multitude of different sensor platforms. The task of the AI would be to identify TELs among objects that are not mobile missile launchers. One challenge AI would encounter is the dataset imbalance—a challenge in

[49] For the potential nuclear-stability implications of other emerging technologies, particularly private-sector innovation in big-data analytics, such as pattern recognition, system visualization, and predictive analytics, see Bracken, "Cyber Threat," 197-99.
[50] Vreeland, "Targeting the Islamic State."
[51] See Stewart, "Deep in the Pentagon."
[52] See Rafael Loss and Joseph Johnson, "Will Artificial Intelligence Imperil Nuclear Deterrence?," *War on the Rocks*, September 19, 2019, https://warontherocks.com/2019/09/will-artificial-intelligence-imperil-nuclear-deterrence/. For a deeper technical analysis of the limitations of automated image recognition, see Joseph Johnson, *MAD in an AI Future?* (Livermore, CA: Lawrence Livermore National Laboratory, 2019), https://cgsr.llnl.gov/content/assets/docs/MAD-in-an-AI-Future.pdf.

classification problems called "class imbalance"—between "ground truth" pictures of TELs and pictures of other, similar military and non-military (non-TEL) vehicles. While relatively few would be available of the former, there is an abundance of the latter. An AI might then be incentivized "to increase its accuracy by rarely or never identifying a mobile launcher," producing false negatives.[53] Manually generating additional synthetic versions of TEL images to increase sample size, on the other hand, could lead to false positives as some non-TEL vehicles could be misclassified for their resemblance with synthetic TELs. More fundamental though, pictures are a poor representation for what really differentiates vehicles: their role and function.[54] While humans can induce function from the observable characteristics of a particular vehicle, AI's ability to do so remains limited. Relatedly, the "curse of dimensionality" prevents AI from reliably objects as the number of discernable features grows.[55] Attempting to compensate for the shortcomings of satellite imagery by increasing resolution or generating three-dimensional models of objects of interest would not only require exponentially more memory and running time, it would also make similar pictures seem increasingly dissimilar and *vice versa* because of AI's inability to discard the irrelevant information contained in images with higher resolution.[56]

Could future improvements to AI resolve these shortcomings? While not entirely implausible, machine-learning theory cautions against exaggerated optimism. For once, AI designers face a bias/variance tradeoff when deciding which data an algorithm should base its decisions on, leading to an irreducible error when working with imperfect—that is, most—measurements of reality.[57] Tradeoffs also exist between different AI algorithms. As of now, there is no one algorithm that can outperform all other algorithms, many with an infinite number of possible variations, at all possible problem sets.[58] Even if an algorithm performed perfectly in the past, for example, in TEL-hunt simulations, perfect performance in the future, confronted with previously unknown data or the real world, cannot be guaranteed.[59] No AI could therefore assure a fully effective counterforce strike in an operational context, some uncertainty would always remain. Moreover, the strategic

[53] Ibid., 4.
[54] Ibid., 5.
[55] Pedro Domingos, "A Few Useful Things to Know About Machine Learning," *Communications of the ACM* 55, no. 10 (2012): 82-3.
[56] Johnson, *MAD in an AI Future?*, 5-6.
[57] Domingos, "A Few Useful Things to Know," 81-2.
[58] See David H. Wolpert, "The Lack of A Priori Distinctions Between Learning Algorithms," *Neural Computation* 8 (1996), 1341–1390.
[59] Johnson, *MAD in an AI Future?*, 7.

nature of military affairs makes it difficult to tailor algorithms for better performance.[60] Adversaries are incentivized to try to beat the AI with creative concealment efforts. Alternatively, they could attempt to poison the AI's training data so that it produces false results in the field, which could lead to consequential misclassification of targets. Ultimately, the performance of any AI cannot be validated without application to the problem it is designed to tackle in the field.[61] However, inherent to validating AI-supported counterforce in the real world is the risk of nuclear escalation.

Despite these challenges, faulty assumptions about AI continue to drive decision making. These result primarily from misconceptions about the technical maturity of machine "reasoning" and misplaced expectations about the scalability of AI solutions. Confronted with impressive results, observers regularly ascribe human-level intelligence to algorithms.[62] This ignores the fundamentally different processes by which AI and humans acquire knowledge. While intuition, compositionality, and causal models (and learning-to-learn) allow humans to arrive at deep levels of understanding from relatively few data points,[63] even state-of-the-art machine learning algorithms "are not learning the true underlying concepts."[64] In narrowly defined problems AI has sometimes exceeded human performance. Lacking deep understanding, however, AI results mostly remain non-transferable to much different problem sets. This closely relates to the illusion of scalability. Unlike defense innovation in the physical world, which usually follows linear growth rates, tackling increasingly complex problems in the computational world often requires resource investments to grow exponentially.[65] To scale up AI solutions, accordingly, efficiency must be attained

[60] Yann LeCun, "Generalization and Network Design Strategies," *Connectionism in perspective* 19 (1989), 143, observes "that good generalization performance on real-world problems cannot be achieved unless some *a priori* knowledge about the task is built into the system." The strategic and dynamic nature of the counterforce problem, however, limits the ability to gain *a priori* knowledge and therefore hinders AI tailoring for better performance.
[61] Johnson, *MAD in an AI Future?*, 8-9. On the difficulty of validating the real-world performance of military software technology, see also David Lorge Parnas and Danny Cohen, *SDI: Two Views of Professional Responsibility* (San Diego, CA: Institute on Global Conflict and Cooperation, 1987), 4-6.
[62] For example, Stuart Armstrong, Kaj Sotala, and Seán S. Ó hÉigeartaigh, "The Errors, Insights and Lessons of Famous AI Predictions–and What They Mean for the Future," *Journal of Experimental & Theoretical Artificial Intelligence* 26, no. 3 (2014), 317-342, identify a general overconfidence among experts concerning future AI developments.
[63] See Brenden M. Lake et al., "Building Machines that Learn and Think Like People," *Behavioral and Brain Sciences* 40 (2017), 15-31.
[64] Ian J. Goodfellow, Jonathon Shlens, and Christian Szegedy, "Explaining and Harnessing Adversarial Examples," *arXiv preprint arXiv:1412.6572* (2014), 2.
[65] Danko Nikolic, "The Challenge of Scaling AI Technology," *DXC.technoloy Blog*, April 27, 2017, https://blogs.dxc.technology/2017/04/27/the-challenge-of-scaling-ai-technology/.

at the expense of accuracy. An efficient solution could therefore not be guaranteed to be the "best" solution, and the best one not to be particularly efficient.[66]

In sum, technical limitations will constrain AI, for the foreseeable future, to operate at levels of confidence insufficient for problems as complex and dynamic as nuclear counterforce. Still, faulty assumptions derived from earlier military-technological innovation in the physical world continue to drive decision making for the computational world. For competitive dynamics in international relations, however, perceptions of the U.S. pursuit of AI for military operations might be more important than their actual capabilities. From the perspective of adversaries, after all, the United States has an expressed interest in counterforce options and is investing substantial sums to improve its ability to hold at risk adversary arsenals, including through AI-infused ISR. The next section assesses some of the options adversaries like North Korea might pursue to hedge against improving U.S. counterforce capabilities.

Hedging Against the Ghost of Counterforce Future

AI's independent impact on the ability to find and destroy nuclear weapons will remain limited.[67] Much of defense planning, however, is conservative and based on worst-case scenarios because of the high-risk nature of military affairs, especially in the nuclear realm.[68] The United States maintains the ICBM-leg of its nuclear triad partly out of concern that its submarine-launched weapons might one day become vulnerable.[69] China and Russia express concerns about U.S. ballistic missile defense not for the current state of these systems, but because improvements could eventually facilitate a surprise first strike and "mop up" any remaining warheads that were not destroyed in their silos. Russia in particular explains its investments in hypersonic and other advanced weapon systems in response to future U.S. missile-defense capabilities.[70] This is all despite a general

[66] Johnson, *MAD in an AI Future?*, 10.
[67] For an assessment of the relative significance and contextual dependencies of emerging technologies in defense transformation more broadly, see Colin S. Gray, "Technology as a Dynamic of Defence Transformation," *Defence Studies* 6, no. 1 (2006), 26-51.
[68] See Michael Fitzsimmons, "The Problem of Uncertainty in Strategic Planning," *Survival* 48, no. 4 (2006), 136-7.
[69] Brad Roberts, *The Case for U.S. Nuclear Weapons in the 21st Century* (Stanford, CA: Stanford University Press, 2015), 269. See also Department of Defense, *2018 Nuclear Posture Review*, 43.
[70] George Lewis and Frank von Hippel, "Limitations on Ballistic Missile Defense—Past and Possibly Future," *Bulletin of the Atomic Scientists* 74, no. 4 (2018), 203-4; and Fiona S. Cunningham and M. Taylor Fravel, "Assuring

understanding that missile defense as it stands today is unable to perform this role. Ultimately, adversaries should be expected to hedge against the dangers of a future in which AI will play a role in nuclear strategy, no matter how unlikely this application is today.

Facing improving counterforce capabilities and an expressed U.S. interest in counterforce options, what measures could North Korea take to enhance the survivability of its nuclear weapons? While some possible nuclear-posture adjustments would be unsuitable for a land-based force like North Korea's, others are more feasible.[71] Of these, however, the most effective measures are also those most likely to undermine first-strike stability, and opportunity costs abound.

The two Cold-War superpowers relied to a great extent on hardening to make their land-based nuclear forces more survivable. They placed their ICBMs in hardened silos and their strategic nuclear bombers in hardened hangers built to withstand surface and air bursts from near misses of enemy nuclear weapons. However, while hard to destroy, ICBM silos are easy to find.[72] All else being equal, against a smaller country, like North Korea, relatively fewer nuclear weapons would suffice to destroy the same number of targets because of closer proximity between the silos. Additionally, improvements in weapons accuracy have made hardened weapons increasingly vulnerable as well, even to non-nuclear precision-strike attacks.[73] Accordingly, facing a much more expansive and sophisticated U.S. arsenal, moving toward a silo-based ICBM force would be a poor choice for North Korean leaders to make.

North Korea could also attempt to make its nuclear force harder to find by improving mobility and concealment. However, TELs will remain largely confined to roads. After all, ICBMs are bulky. Russia's road-mobile SS-27 "Sickle B," for example, measures around 22 meters in length with a

---

Assured Retaliation: China's Nuclear Posture and U.S.-China Strategic Stability," *International Security* 40, no. 2 (2015), 16-9. On President Trump's ambition to "detect and destroy any missile launched against the United States — anywhere, anytime, anyplace," see White House, *Remarks by President Trump and Vice President Pence Announcing the Missile Defense Review*, January 17, 2019, https://www.whitehouse.gov/briefings-statements/remarks-president-trump-vice-president-pence-announcing-missile-defense-review/. For a critique thereof, see Laura Grego, "No, Missile Defense Will Not Work 97% of the Time," *All Things Nuclear*, Union of Concerned Scientists, October 13, 2017, https://allthingsnuclear.org/lgrego/missile-defense-will-not-work-97-percent.

[71] Per Vipin Narang, *Nuclear Strategy in the Modern Era: Regional Powers and International Conflict* (Princeton, NJ: Princeton University Press, 2014), 4, "nuclear posture" here refers to the capabilities, doctrine, and command-and-control procedures that constitute a state's nuclear weapons capability.

[72] See Bracken, "Cyber Threat," 191.

[73] Ibid., 192. See also Lieber and Press, "New Era of Counterforce," 18-32.

diameter of approximately 2 meters, weighing 47,000 kg.[74] North Korea's Hwasong-15, which was flight-tested on November 29, 2017, is estimated to be of similar proportions.[75] The 9-axle TEL this missile is transported on is unable to navigate much of the country's difficult terrain; "The belief that mobile missiles can be transported off road and on, and can be made operational in a short time, is an illusion."[76] This funnels TELs into identifiable areas of operation—i.e., the North Korean road network. Their logistics tails further reduce mobile launchers' ability to move fast and undetected. Until ICBMs become significantly lighter and smaller, substantial mobility increases will remain out of reach.

Concealment may be the most promising path to make North Korea's land-based nuclear forces more survivable, and one that could leverage AI-specific countermeasures. For example, mobile missile teams can rely on presurveyed launch and hide sites equipped with camouflage tarps to protect their launchers from multispectral aerial reconnaissance.[77] To fool an AI-infused ISR capability it might even suffice to simply have alerted TELs look different from the imagery used to train the AI system on a particular mobile launcher by obscuring its observable characteristics or building "TEL-shells" to distract the AI with decoy vehicles resembling mobile-missile launchers. Because of the above-discussed bias/variance tradeoff, the AI might then not be able to reliably distinguish a TEL from a commercial truck. Adversaries might also attempt to implant faulty data in AI training datasets to introduce bias in an algorithm's developmental stage.[78] Imperceptible to humans, such inputs could dramatically alter the performance of AI, making it fail in unexpected

[74] Missile Defense Project, "SS-27 'Sickle B' (RT-2PM2 Topol-M)," *Missile Threat*, Center for Strategic and International Studies, June 15, 2018, https://missilethreat.csis.org/missile/ss-27/.
[75] Missile Defense Project, "Hwasong-15 (KN-22)," *Missile Threat*, Center for Strategic and International Studies, June 15, 2018, https://missilethreat.csis.org/missile/hwasong-15-kn-22/; and Michael Elleman, "The New Hwasong-15 ICBM: A Significant Improvement That May be Ready as Early as 2018," *38 North*, Stimson Center, November 30, 2017, https://www.38north.org/2017/11/melleman113017/.
[76] Bracken, "Cyber Threat," 194.
[77] See Li Bin, "Tracking Chinese Strategic Mobile Missiles," *Science and Global Security* 15, no. 1 (2007), 10.
[78] See, for example, Christian Szegedy et al., "Intriguing Properties of Neural Networks," *arXiv preprint arXiv:1312.6199* (2013); and Ian Goodfellow, Patrick McDaniel, and Nicolas Papernot, "Making Machine Learning Robust Against Adversarial Inputs," *Communications of the ACM* 61, no. 7 (2018), 56-66. On cybersecurity risks and the hacking of nuclear weapons systems more broadly, see, for example, Andrew Futter, *Cyber Threats and Nuclear Weapons: New Questions for Command and Control, Security and Strategy* (London, UK: Royal United Services Institute, 2016), https://rusi.org/sites/default/files/cyber_threats_and_nuclear_combined.1.pdf; and Beyza Unal and Patricia Lewis, *Cybersecurity of Nuclear Weapons Systems: Threats, Vulnerabilities and Consequences* (London, UK: Chatham House, 2018), https://www.chathamhouse.org/sites/default/files/publications/research/2018-01-11-cybersecurity-nuclear-weapons-unal-lewis-final.pdf.

and uninterpretable ways. As of now, there are few effective countermeasures to adversarial inputs.[79]

Concealment measures, too, come with costs, however. Extensive preparations and security at designated launch and hide sites might draw scrutiny. Moreover, command and control of nuclear forces requires communication, for mobile systems more so than for fixed ones. While communication can be reduced to a minimum, if a launch order is to be transmitted, some open channel must be maintained. Its logistics tail further expands a TELs footprint. "Tells" of launch preparations or heightened alert status would likely be picked up by a variety of sensors.[80] Even if concealment can reliably fool AI-supported imagery analysis, other sensor platforms would still collect signals. Furthermore, attempts to implant adversarial data, for example through offensive cyber operations, risk detection and research efforts are underway to address AI's vulnerability to adversarial inputs.[81] Given the power and resource imbalance between those who can marshal counterforce capabilities and those who hedge against them—like the United States and North Korea respectively—and the fact that leadership in sensing requires familiarity with potential countermeasures, however, the balance is likely to shift further in favor of great powers with counterforce ambitions.[82]

In addition to these technical adjustments to improve survivability, North Korea might also revise its nuclear weapons employment doctrine. Changes in employment doctrine are not so much about ensuring a retaliatory capability should North Korea suffer from a first strike as they are about enhancing the credibility that it would use its nuclear weapons in a conflict with the United States.[83] Because it cannot assure mutual destruction, North Korea would likely have to escalate to the nuclear level early in such a conflict. Its asymmetric escalation doctrine would draw on short-, medium-, and intermediate-range nuclear assets to stave off a conventional invasion, while holding in reserve ICBMs to deter nuclear retaliation by threatening the U.S. homeland.[84] With improving U.S. counterforce capabilities, however, North Korea faces growing use-it-or-lose-it

[79] Goodfellow, McDaniel, and Papernot, "Making Machine Learning Robust," 65.
[80] Bracken, "Cyber Threat," 196.
[81] Goodfellow, McDaniel, and Papernot, "Making Machine Learning Robust," 66.
[82] See Lieber and Press, "New Era of Counterforce," 46-7.
[83] See Peter D. Feaver, "Correspondence: Proliferation Pessimism and Emerging Nuclear Powers," *International Security* 22, no. 2 (1997), 185–207.
[84] Narang, "Kim Jong Un Wouldn't Be Irrational." On the functional logic of asymmetric escalation, see also Narang, *Nuclear Strategy*, 19-21.

pressure.[85] In a crisis, North Korea would be incentivized to employ its weapons before a counterforce strike could degrade its nuclear capability.[86] North Korean concerns about the efficacy of U.S. missile defenses would further spur preemptive and massive launch; the larger the volley of incoming North Korean ICBMs, the greater the chance that at least one warhead would evade defenses.[87] While North Korea could not hope to make a dent in the U.S. nuclear retaliatory capability, its leaders might perceive a slim chance that the destruction of a major city might shatter U.S. resolve and compel it to stop fighting. The perception of being confronted with an AI-improved counterforce capability would shorten North Korea's timeline for effective nuclear employment. It would have to rely more extensively on pre-delegating launch authority, co-locating operational warheads and delivery systems, and maintaining a state of high alert to maintain a credible deterrent. However, in addition to growing risks of deliberate nuclear use, this would increase the probability of inadvertent and accidental employment. An otherwise containable crisis on the Korean peninsula could thus quickly spiral toward nuclear disaster.[88]

Conclusion

[85] To assert, as Vince Manzo and John Warden do, that nuclear employment "is in no way a 'dominant strategic move' for North Korea" and therefore irrational, is an overly generous assumption about crisis decision making and the ability to control escalation in an ongoing conflict. See Manzo and Warden, "Want to Avoid Nuclear War? Reject Mutual Vulnerability with North Korea," *War on the Rocks*, August 29, 2017, https://warontherocks.com/2017/08/want-to-avoid-nuclear-war-reject-mutual-vulnerability-with-north-korea/. For a more complete picture, see, for example, Robert Jervis, "The Political Effects of Nuclear Weapons: A Comment," *International Security* 13, no. 2 (1988), 80-90; Barry R. Posen, *Inadvertent Escalation: Conventional War and Nuclear Risks* (Ithaca, NY: Cornell University Press, 1991); and Mark S. Bell and Julia Macdonald, "How to Think About Nuclear Crises," *Texas National Security Review* 2, no. 2 (2019), 40-65, https://tnsr.org/2019/03/how-to-think-about-nuclear-crises/.

[86] For first-strike stability in the context of India-Pakistan and U.S.-China relations, see, respectively, Clary and Narang, "India's Counterforce Temptations;" and Caitlin Talmadge, "Would China Go Nuclear? Assessing the Risk of Chinese Nuclear Escalation in a Conventional War with the United States," *International Security* 41, no. 4 (2017), 50-92.

[87] Ankit Panda and Vipin Narang, "Deadly Overconfidence: Trump Thinks Missile Defenses Work Against North Korea, and That Should Scare You," *War on the Rocks*, October 16, 2017, https://warontherocks.com/2017/10/deadly-overconfidence-trump-thinks-missile-defenses-work-against-north-korea-and-that-should-scare-you/.

[88] Moreover, improving counterforce capabilities might spur arms races in peacetime with externalities reaching beyond the U.S.-North Korean nuclear competition. If the United States continued to seek the capabilities to hold at risk North Korean nuclear weapons and North Korea expanded its strategic nuclear arsenal to maintain survivability, its expanding arsenal would come to resemble that of China and Beijing might begin to perceive a threat to its nuclear deterrent. See Bin, "Tracking Chinese Strategic Mobile Missiles," 26.

This article provides an assessment of whether AI enables states to conduct effective counterforce strikes against adversary nuclear arsenals. An effort was made to minimize abstraction by drawing on recent U.S. experiences with the application of AI to the Defeat-ISIS campaign, technical assessments of the vulnerabilities and limitations of current machine-learning technologies, and open-source material on North Korea's nuclear program, arsenal, and force posture. While this necessarily makes for an incomplete picture, it nevertheless suggests some preliminary conclusions about future, AI-infused counterforce capabilities and their effect on international stability.

This analysis has shown that AI can play a critical role in improving ISR for military operations. Since the failed Scud hunt of the 1991 Gulf War, U.S. armed forces have invested heavily in improving their ability to find, fix, and finish mobile missile launchers. More recently, AI has been leveraged to better make use of the expansive full-motion video imagery provided by UAS operating in the skies above Syria and Iraq. While in the past counterforce suffered from an inability to identify and track enemy TELs, these advances in AI-supported remote sensing could finally enable an effective counterforce capability. Should a crisis erupt on the Korean Peninsula, the United States' unparalleled military power and ISR capabilities, equipped with AI, would be more able than ever to find and destroy North Korea's road-mobile nuclear weapons. Or so some argue.

AI's true potential to revolutionize counterforce remains hampered by inherent flaws. These lie in the shortcomings of the data available to AI for both training and operationalization. Because of inherent limitations illustrated by machine-learning theory and adversary incentives to fool algorithms, future improvements should not be expected to perfect AI either. Faulty assumptions about the inner workings of artificial intelligence, however, lead policymakers to continue to overestimate the impact and potential of AI in military affairs and overlook its real limitations.

Thus, while there is demand for AI-infused ISR capabilities to improve target identification and elimination, including of adversary nuclear forces, supply will not satisfy the necessarily high requirements for perfection in counterforce. Yet, expressed U.S. interest in damage limitation and counterforce options as well as AI's contribution to a marginally improving counterforce capability provide powerful incentives for adversaries, like North Korea, to hedge by increasing the survivability of their nuclear forces and adjusting their employment doctrines. Such measures,

particularly the pre-delegation of launch authority and co-location of operational warheads and delivery systems, however, would heighten the risk of inadvertent and accidental nuclear use during a crisis.

U.S. leaders should want to mitigate such risks. Yet, AI-infused ISR and greater effectiveness of non-nuclear weapons are tremendously useful for future military operations in the context of great-power competition. Neither can these efforts be siloed out of the counterforce complex, particularly from the point of view of adversaries. Even if U.S. counterforce capabilities are expressly not aimed at China or Russia but rather the product of damage-limitation requirements vis-à-vis North Korea, efforts to keep up with a gradually expanding North Korean nuclear arsenal might eventually encroach upon China's relatively small number of weapons as well. After all, states in competitive relationships care more about what others can *do* to them than what others *say* about their intentions. Thus, chances are that we have indeed entered a new era of counterforce, resulting in greater instability among nuclear-armed states. However, contrary to the expectations of some, this is fueled less by the technological change brought about by AI than by the perception of threats and other's intentions in a world of ever-evolving military capabilities.