Institute for National Strategic Security, National Defense University

Artificial Intelligence on the BattlefieldAuthor(s): Zachary Davis

Source: *PRISM*, Vol. 8, No. 2 (2019), pp. 114-131

Published by: Institute for National Strategic Security, National Defense University

Stable URL: https://www.jstor.org/stable/10.2307/26803234

**REFERENCES**
Linked references are available on JSTOR for this article:
https://www.jstor.org/stable/10.2307/26803234?seq=1&cid=pdf-reference#references_tab_contents
You may need to log in to JSTOR to access the linked references.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at https://about.jstor.org/terms

*Institute for National Strategic Security, National Defense University* is collaborating with JSTOR to digitize, preserve and extend access to *PRISM*

This content downloaded from
177.142.92.225 on Tue, 05 May 2020 16:02:51 UTC
All use subject to https://about.jstor.org/terms

For soldiers such as this one to dominate the battlefield, requires that acquisition professionals maintain the relevant skills and expertise, seek diverse career positions, and remain agile and adaptive to the changing acquisition 'battlefield,' emerging technologies, and fiscal constraints. (U.S. Army/ Shane Hamann)



All use subject to https://about.jstor.org/terms

# Artificial Intelligence on the Battlefield

## Implications for Deterrence and Surprise

By Zachary Davis

Artificial intelligence (AI) has burst upon the national security scene with a speed and an intensity surprising even the most veteran observers of the national policy discourse. Factors that have driven this spike of interest include the perception of AI as a revolutionary technology, on par with the discovery of fire, electricity, or nuclear weapons; the rapid absorption of nascent AI-based technologies into diverse sectors of the U.S. economy, often with transformative effects (as, for example, in the sciences and in social media); and the ambitions of potential U.S. adversaries.[1] Echoing the 19th-century naval strategist Alfred Thayer Mahan ("Whoever rules the waves rules the world"), Russian president Vladimir Putin has argued that the nation that rules in AI "will be the ruler of the world."[2] People's Republic of China President Xi Jinping is less outspoken on this matter, but he has committed China to become the dominant AI power by 2030.[3] There are mounting fears of a "Sputnik moment," which might reveal the United States to be woefully underprepared to manage the new AI challenges. If there is an AI arms race, what are the implications for U.S. security?[4] Could AI disrupt the strategic balance, as blue-water navies and nuclear weapons did in previous eras? Might it do so in a manner so severe that deterrence fails and leads to war? If war involving AI-guided weapons occurs, can we win?

This article will calibrate the potential risks and rewards of military applications of AI technologies and will explore:

- What military applications are likely in the near term?
- What are the potential consequences of these applications for strategic stability?
- How could AI alter the fundamental calculus of deterrence?
- How could AI-assisted military systems affect regional stability? Relatedly, what is the connection between regional stability and strategic deterrence?
- What are the risks of unintended consequences and strategic surprise from AI?

Dr. Zachary Davis is a Senior Fellow at the Center for Global Security Research at Lawrence Livermore National Laboratory and Research Professor at the Naval Postgraduate School. This article draws on a report published by the Technology for Global Security and the Center for Global Security Research on February 13, 2019 on "AI and the Military: Forever Altering Strategic Stability." The report was a collaboration of the Center for Global Security Research and the Lawrence Livermore National Laboratory.



This content downloaded from
177.142.92.225 on Tue, 05 May 2020 16:02:51 UTC
All use subject to https://about.jstor.org/terms

## AI, Big Data, and Machine Learning in Science and Business

Before answering the questions posed above, it is useful to recall the state of the art for AI in scientific and business applications. Much of the near-hysteria over AI stems from the fuzziness of our view of the technologies that combine to make AI. So far, at least, the national security community lacks a common language for discussing AI and a detailed appreciation of the different technologies and the timelines by which they might mature into militarily significant capabilities.

The term "artificial intelligence" is used to describe a range of loosely related phenomena that are generally associated with using computers to glean insight from "big data." Much as the generic term "cyber" is used in reference to everything from networks to hardware, software, automation, industrial controls, hacking, bullying, warfare, and all things social media, AI is used as a generic term that washes over meaningful distinctions between its different manifestations. This breeds confusion, especially regarding claims about its revolutionary effects.

For the vast majority of current applications, AI consists of algorithms that form the basis of pattern recognition software. When combined with high-performance computing power, data scientists are able to probe and find meaning in massive data collections. Neural networks supercharge the ability of the algorithms to identify and organize patterns in the data by "training" them to associate specific patterns with desired outcomes. Multiple layers of neural networks, known as deep learning neural networks, are what make current approaches to "machine learning," "supervised learning," and "reinforcement learning" possible.[5] However, the neural network approach portrays only a fraction of the advancements in AI methods. For example, AI also includes language processing, knowledge representation, and inferential reasoning, which are all increasingly possible due to advancements in software, hardware, data collection, and data storage. AI represents a quantum leap in the ability to find needles in data haystacks—as long as you know what you are looking for.

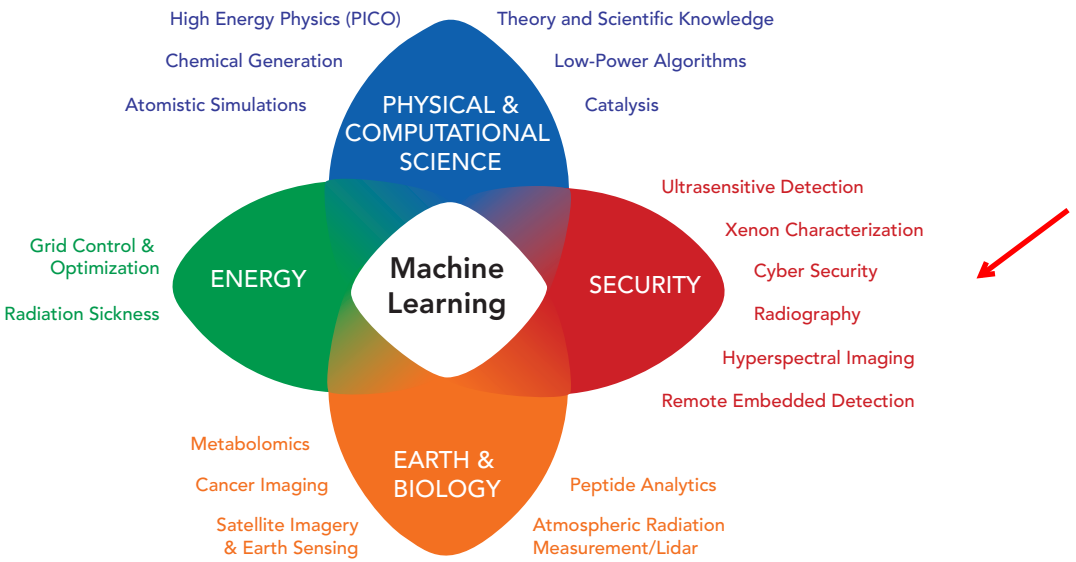
It is useful to distinguish between narrow and general applications of AI. Narrow AI encompasses discrete problemsolving tools designed to perform specific narrow tasks. General AI encompasses technologies designed to mimic and recreate functions of the human brain. The gap between the two is significant. Most experts appear to agree that the accomplishments of narrow AI, though quite significant, are a long way from the requirements of replicating human-like reasoning as envisioned by proponents of general AI. Although IBM's Watson, Google's DeepMind, and other such experiments have made breakthroughs in replicating human-like reasoning, they are far from being able to reliably replicate the performance of the human brain in its multiple dimensions. It is not surprising, however, that the human imagination has been captured by the prospect of what futurists have called "The Singularity"—a point in time when "we will multiply our effective intelligence a billion fold by merging with the intelligence we have created."[6] The quest for "superintelligence" notwithstanding, recent progress in brain enhancement for now mostly replenishes impaired functions[7] and has a long way to go before it is possible to equip citizens, soldiers, or robots with superhuman powers.[8]

Although general AI stimulates intriguing science fiction about cyborgs, space wars, and robot armies, narrow AI is already here—and has been for some time. In both business and science, AI has wide applications, primarily in data-rich research fields, including fundamental research (for example, in physics, chemistry, and biology) and applied sciences (medicine, aeronautics, and environmental studies). Data science is facilitating rapid advancements in every aspect of scientific discovery, even changing long-held methodological standards and



This content downloaded from
177.142.92.225 on Tue, 05 May 2020 16:02:51 UTC
All use subject to https://about.jstor.org/terms

**FIGURE 1. Disciplinary Areas of Deep Learning for Scientific Discovery at the Pacific Northwest National Laboratory.**




Source. Nathan Hodas, *Artificial Intelligence and Machine Learning to Accelerate Translational Research: Proceedings of a Workshop in Brief*, National Academies Press (July 2018), <http://nap.edu/25197>.

practices.[9] Figure 1 highlights some of the scientific areas where AI-fueled deep learning is having its greatest effect.

The crossover of AI into business applications has supercharged predictive analytics for market research, consumer behavior, logistics, quality control, and many other data-rich areas. The proliferation of cameras and sensors creates even more opportunities for data analysis. When combined with robotics, AI is ushering in a new industrial age, with far-reaching societal implications for labor and management.[10] For these types of applications, however, AI is more of a well-established, sustaining, and enabling technology than a revolutionary new disruptive technology in its own right. Data analytics is not new, but it is getting better.

For these scientific and business applications, AI is an enabling technology, a cross-cutting force multiplier when coupled with existing data-centric systems, such as the internet, health care, social media, industrial processes, transportation, and just about every aspect of the global economy, where recognizing patterns is the key to insight and profit. Growing interconnectivity, illustrated by the Internet of Things (IOT), is producing more data and providing more opportunity for AI algorithms to reveal hidden insights.

## What Military Applications are Likely in the Near Term? Tactical and Strategic Effects

Should we expect similarly important AI applications in the military field? Like so many technologies, AI is loaded with latent military potential.[11] Many see algorithmic warfare as the prime mover of a new revolution in military affairs.[12] AI was central to the so-called Third Offset Strategy pursued by the Department of Defense (DOD) in



This content downloaded from
177.142.92.225 on Tue, 05 May 2020 16:02:51 UTC
All use subject to https://about.jstor.org/terms

the second Obama Administration and thus was a principal focus of multiple government initiatives to accelerate the development of advanced technologies.[13] In June 2018, DOD established its Joint Artificial Intelligence Center and issued its Artificial Intelligence Strategy in February 2019.[14] The White House established its Select Committee on AI in May 2018 and released its Executive Order on Maintaining American Leadership in Artificial Intelligence in parallel with the DOD Strategy, also in February 2019.[15] DOD and Intelligence Community spending on AI has increased substantially.[16] For military applications with direct analogs in the civilian world, like logistics, planning, and transportation, AI-supported data analytics is already in use throughout the defense and intelligence communities.[17] These applications are separate and distinct from applications to warfighting, which tend to fall into one of two categories: ones having impact primarily at the tactical/operational level of war, and those that also have impact at the strategic level of war. Tactical or operational effects stem from the way wars are fought—including specific weapons and organizational concepts. We define "strategic" as "extraordinarily consequential actions capable of causing a shift in the balance of power."[18] The strategic level refers primarily to major conflict between great powers. It is possible, however, for actions at the operational level to spill over and have effects at the strategic level.

## AI Applications at the Tactical/Operational Level of War

The process of managing and making sense of the staggering amount of intelligence, surveillance, and reconnaissance (ISR) data involved in modern warfare is a natural fit for AI and is the objective of DOD's Project Maven, also known as the Algorithmic Warfare Cross Functional Team.[19] According to Lieutenant General Jack Shanahan, former Director of Defense Intelligence for Warfighter Support, Project Maven was conceived as "the spark that kindles the flame front for artificial intelligence across the rest of the department."[20] While Maven's initial mission was to help locate Islamic State fighters, its implications are vast. Multidomain warfare involves colossal amounts of heterogenous data streams that can only be exploited with the help of AI. Mirroring the proliferation of sensors in the civilian world, the multidomain, hybrid warfare battlefield has become a military version of the IoT, teeming with vital information for assessing tactical and strategic threats and opportunities. While the ability to manage this data colossus in real time portends tremendous advantages, failure to draw meaning from that information could spell disaster.

Being able to rapidly process the flood of information from varied platforms operating in multiple domains translates into two fundamental military advantages—speed and range. Moving faster than your adversary enhances offensive mobility and makes you harder to hit. Striking from farther away similarly benefits the element of surprise and minimizes exposure to enemy fire. These were central tenets of the previous revolution in military affairs that had its debut in the Gulf War. AI makes it possible to analyze dynamic battlefield conditions in real time and strike quickly and optimally while minimizing risks to one's own forces.

### *Omnipresent and Omniscient Autonomous Vehicles*

The new generation of autonomous vehicles is a high priority for military applications of AI, with much of the focus on navigation for a variety of unmanned land, sea, and air systems.[21] Space and undersea platforms will also benefit from AI-informed guidance systems. AI is at the heart of the so-called drone swarms that have been the subject of much attention in recent years.[22] AI-informed navigation software supported by ubiquitous sensors not only enables



This content downloaded from
177.142.92.225 on Tue, 05 May 2020 16:02:51 UTC
All use subject to https://about.jstor.org/terms

unmanned vehicles to find their way through hostile terrain, but also may eventually make it possible for complex formations of various types of drones operating in multiple domains, with complementary armaments to conduct sophisticated battle tactics, instantly adjusting to enemy maneuvers to exploit battlefield opportunities and report changing conditions. Autonomous vehicles and robotics are poised to revolutionize warfare.

As a recent Defense Science Board Study demonstrated, integrated battle management, command, control, communications, and intelligence (BMC3I) capabilities are well suited to finding and targeting deployed missile batteries, and thus could be the key to countering critical elements of the anti-access/area denial (A2AD) strategies of Russia and China.[23] These systems were designed to exploit vulnerabilities of U.S. land and sea assets in Europe and Asia. In addition to geolocating targets, AI-enabled BMC3I could help guide and coordinate kinetic effects involving multiple platforms, possibly providing a counter to current adversary A2AD. From this perspective, the cumulative effects of tactical-level AI could become a strategic-level game changer.

### *Big Data–Driven Modeling, Simulation, and Wargaming*

AI has steadily been increasing the power of simulations and gaming tools used to study nuclear and conventional weapons. From Samuel Glasstone's early calculations of nuclear effects to the extensive library of RAND studies on nuclear issues, quantitative methods have been integral to the development of nuclear weapons systems.

AI is enabling scientists to model nuclear effects to confirm the reliability of the nuclear stockpile without nuclear testing. Simulation and modeling is already a key part of the design process for nearly all major weapons systems, from jets and ships to spacecraft and precision-guided munitions.[24] Massive modeling and simulation will be necessary to design the all-encompassing multidomain system of systems envisioned for battle management and complex missions such as designing, planning, and managing systems for space situational awareness. On the production side, AI already informs quality control for novel production methods, such as additive manufacturing.[25]

AI is also enriching battlefield simulations and wargames involving multi-actor interactions. AI enables wargamers to add and modify game variables to explore how dynamic conditions (weapons, effects, allies, intervention, and so forth) could affect outcomes and decisionmaking. AI is used to analyze the results of such games.[26] These are examples of evolutionary learning that are unlikely to cause strategic surprise or undermine stability unless the results negatively influence decisionmaking.

### *Focused Intelligence Collection and Analysis*

With so many incoming streams of intelligence (human, signals, open-source, measurement and signatures, geospatial, electronic) being collected, all requiring analysis to be useful for policymakers, the Intelligence Community faces the challenge of information overload.[27] This is a data-centric problem for which AI and machine learning are well suited.[28] For example, a project at Lawrence Livermore National Laboratory uses neural networks to probe multimodal data sets (images, text, and video) in search of key indicators of proliferation activity. Machine learning also makes it possible to combine open-source trade and financial data with multiple forms of intelligence to glean insights about illicit technology transfers, proliferation networks, and the efforts of proliferators to evade detection.[29] These insights enable analysts to inform policymakers and support counterproliferation policy and actions.

Machine learning will be an important tool for all-source analysts who are increasingly required to take into account information from many sources,



This content downloaded from
177.142.92.225 on Tue, 05 May 2020 16:02:51 UTC
All use subject to https://about.jstor.org/terms

locations, and disciplines to understand today's global security environment. To the extent that better information leads to informed decisions, applying AI to these collection and analysis problems would benefit strategic stability.

## AI Applications with Implications for the Strategic Level of War

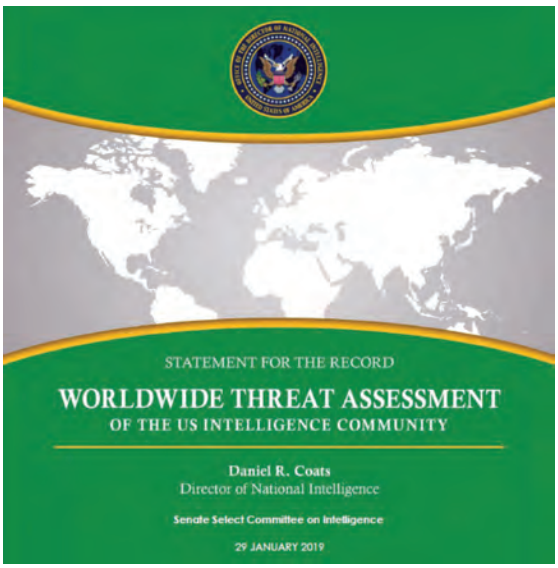
Some military applications of AI appear to have broader implications beyond the battlefield. AI that makes it possible to locate and target strategic assets could alter the logic of strategic deterrence.

### *A System of Systems Enabling Exquisite ISR*

For the military, object identification is a natural starting point for AI, as it requires culling images and information collected from satellites and drones to find things of military importance such as missiles, troops, and intelligence information. Accordingly, the National Geospatial-Intelligence Agency has led the charge in applying AI to military and intelligence needs.[30] But object identification is just the beginning. Intelligence, surveillance, and reconnaissance (ISR) is the key to multidomain situational awareness. This awareness is increasingly critical as the battlefield extends to all domains—sea, land, air, space, and cyber on a global scale.

### *Precision Targeting of Strategic Assets*

AI-empowered ISR that makes it possible to locate, track, and target a variety of enemy weapons systems raises the possibility of striking strategic assets, such as aircraft carriers, mobile missiles, or nuclear weapons. This capability, and perceptions of its existence, could disrupt long-held assumptions about deterrence stability, especially if it appeared possible to conduct a disarming counterforce strike against an adversary's retaliatory forces.[31] The combination of offensive weapons that can "find, fix, and finish" a significant portion of an adversary's strategic assets, with defensive systems that can shoot down remaining retaliatory capabilities, could challenge fundamental precepts of deterrence based on mutual vulnerability.[32]




The Director of National Intelligence's worldwide threat assessment in January asserted that "…The global race to develop artificial intelligence (AI)—systems that imitate aspects of human cognition—is likely to accelerate the development of highly capable, application-specific AI systems with national security implications." (DNI)

### *Effective Missile Defense*

Advancements in AI-enhanced targeting and navigation also improve prospects for a wide range of tactical and strategic defense systems, especially ballistic missile defenses, by empowering target acquisition, tracking, and discrimination.[33] The convergence of powerful new offensive and defensive capabilities has, however, rekindled fears of a surprise attack that could rattle strategic stability.

### *AI-Guided Cyber*

As an inherently digital domain, the cyber realm naturally lends itself to AI applications, as illustrated by the centrality of AI algorithms for social media titans such as Google and Facebook. The availability of enormous amounts of data in electronic formats is well suited to AI strengths. AI-guided probing,



This content downloaded from
177.142.92.225 on Tue, 05 May 2020 16:02:51 UTC
All use subject to https://about.jstor.org/terms

mapping, and hacking of computer networks can provide useful data for machine learning, including discovery of network vulnerabilities, identities, profiles, relationships, and other information that could be valuable for offensive and defense purposes.[34] Chinese applications of AI for surveillance purposes illustrate broad concerns about its implications for privacy and democracy.

On the offensive side, AI could help locate and target particular nodes or individual accounts for collection, disruption, or disinformation. Cyber attacks on national command infrastructure and networks, for example, could be catastrophic.[35] On the defensive side of the equation, AI can help detect such intrusions and search for debilitating anomalies in civilian and military operating systems.[36] AI will equally empower offensive and defensive measures, both of which could have positive and negative strategic effects.

## Potential Consequences of these Applications for Strategic Stability

AI has multiple potential applications in the military domain at both the operational and strategic levels of war. But at the strategic level, some of the implications may not be altogether positive, as already foreshadowed. Indeed, the disruptive effects of new technologies cannot be limited to the adversary. Some of those effects are potentially quite significant for strategic stability. How might this be so?

### *The Enemy Has AI Too*

No one country can gain all of the benefits of AI while denying them to potential adversaries. Competition to gain advantage will bring uncertainty about the future balance. Russia, China, and other nations' advancements in these same AI-enabled technologies have the potential to shift the strategic calculus as well, especially in regional contexts. For example, while Russian and Chinese A2AD systems designed to defeat U.S. regional forces may reduce U.S. allies' confidence in American security guarantees to protect them, the ability of the United States to defeat those A2AD systems with AI-accelerated ISR, BMC3I, defensive systems, and autonomous vehicles would demonstrate resolve and provide opportunities for joint U.S.-allied defense cooperation, thereby enhancing stability and deterrence. Reinforcing regional conventional deterrence is also an essential part of strategic stability.[37] However, even the *perception* of an imbalance that favors striking first can lead to misperception, miscalculation, and arms racing. Whatever advantages can be attained with AI are likely to evoke countermeasures that mitigate temporary unilateral advantages. Russian and Chinese interest in hypersonic vehicles and counterspace operations may fall into this category.

### *Data Is Fragile . . .*

AI systems are vulnerable to flawed data inputs, which can cause unintended consequences. In her book *Weapons of Math Destruction*, data scientist Cathy O'Neil demonstrates how AI logarithms distort reality and lead to incorrect, misleading, and unjust decisions.[38] Perhaps the biggest obstacle to increasing reliance on AI is the age-old problem of data reliability. AI can magnify the "garbage in, garbage out" problem.[39] Data comes from many places and is not always carefully collected or curated. Compounding the problems with the data itself leading to skewed results, AI often reflects human bias.[40] Computer vision—the AI-informed object and pattern recognition software behind Project Maven and many other applications—is easily fooled by misleading data.[41] Differentiating between similar objects is difficult and more challenging with denial and deception campaigns, such as the use of camouflage and decoys.[42] Even when data seems accurate, AI sometimes "hallucinates" things that do not exist.[43] Transferring these inherent problems of data reliability and interpretation onto the battlefield raises critical questions about



This content downloaded from
177.142.92.225 on Tue, 05 May 2020 16:02:51 UTC
All use subject to https://about.jstor.org/terms

the safety and reliability that come with the desirable qualities of speed and lethality. Accidentally hitting the wrong targets, for example, could have strategic consequences.

### *. . . And Easily Manipulated*

Countering many AI applications can be simple and straightforward. Adversarial manipulation of data provides many opportunities for mischief and mistakes.[44] The fact that AI is easily deceived invites efforts to counter the sought-after military benefits.[45] By corrupting data in calculated ways, it may be possible to cause catastrophic equipment failures, miscommunication, confusion, logistical nightmares, and devastating mistakes in AI-reliant systems. The "black box" problem of not understanding how and why AI makes decisions also means that it would be hard to recognize if data had been compromised to produce inaccurate outcomes, such as hitting the wrong targets or misdirecting U.S. and allied forces. The vulnerability of data could be the Achilles' heel of AI.

### *Faster is Not Always Better*

Speedy decisionmaking and operational execution may not serve well the goals of effective crisis management. On October 19, 1962, only three days into the Cuban Missile Crisis, General Curtis LeMay counseled President John F. Kennedy, "I just don't see any other solution except direct military action right now."[46] Ten days later, the crisis was resolved diplomatically. If one of the advantages of AI is the speed it adds to decisionmaking, that same speed could be a disadvantage if it accelerates the escalation of conflict from crisis to war and even potential nuclear confrontation.[47] The battlefield advantages of AI-driven ISR and autonomous systems could shrink the time available for diplomacy to avoid or manage crises. As currently conceived, AI-driven battlefield systems would not include real-time reporting and analysis of national and international diplomatic efforts to avoid, control, contain, or end a conflict—violating Clausewitz's principle of war as "the continuation of politics by other means." In many cases, logic might dictate striking first, as General LeMay advised. Accelerated decisionmaking might have pushed the Cuban Missile Crisis toward different outcomes. In practice, slowing things down can be the key to victory, especially when the stakes involve nuclear weapons.

Many of the positive regional deterrence effects that could eventually result from an integrated ISR, defense, and battle management complex might not be attainable, at least not in the near term. The overarching architecture and strategy for complex new AI-guided ISR/battle management systems do not yet exist. In fact, a proliferation of AI systems may actually complicate one of the main problems confronting U.S. military forces—effective joint operations.

### *Systems of Systems of Systems*

AI-supported weapons, platforms, and operating systems operate according to custom-built software and hardware that is specifically designed for each separate system and purpose. There is currently no overarching mechanism to integrate scores of AI-powered systems operating on multiple platforms.[48] To achieve the desired effects of multi-domain ISR, it is necessary to integrate across scores of sensors, radars, weapons, and communications systems operating in multiple geophysical domains. If this were not challenging enough, those systems would be built and operated by different agencies, commands, and contractors, with different authorities, access, and procedures. Adding allies with their own AI systems to this landscape brings further complexity and risk. Without seamless integration, the hoped-for benefits of speed and lethality could be fleeting, and the credibility of such an unproven system of systems could be called into question. Massively complex and unproven capabilities could invite challenges that could be destabilizing.



This content downloaded from
177.142.92.225 on Tue, 05 May 2020 16:02:51 UTC
All use subject to https://about.jstor.org/terms

### *Strategic Warning Requires More than Data*

Big data and machine learning might not solve the challenge of strategic warning. Designing a multiplex of AI-informed platforms that have the ability to communicate in real time requires a new generation of data fusion, integrative software, and command architectures. Pulling all these pieces together to develop a holistic threat assessment that provides policymakers with strategic warning will not happen naturally. Instead, this task will require Herculean efforts to collect and analyze information "owned" by diverse stakeholders with different classification systems, analytic roles, and customer loyalties. Integrating and analyzing sensitive information from diverse sources is already a challenge, especially if it needs to be done quickly. Moreover, while machine learning, computer vision, and other techniques will help sort and prioritize the flood of intelligence information, analysts will still have to make judgments based on incomplete and sometimes unreliable information. Developing a fully integrated system capable of providing strategic warning will take many years.

### *AI Unpredictability*

The close operation and integration of multiple AI systems, as required on the battlefield, can be expected to have unanticipated results, some of which could have strategic consequences. The flip side of stovepiped systems not talking to each other is the issue of unexpected convergences. It is uncertain how separate AI-infused platforms might interact with one another, as various AI-guided systems operate in shared battlespace. Unknown outcomes resulting from friendly interactions are likely to be compounded by interactions with foreign AI systems. With so much uncertainty about the internal "black box" mechanisms that produce AI outcomes, AI-to-AI interactions are likely to produce unanticipated and unexplainable results— for example, choosing the wrong targets.[49] Lastly, we cannot anticipate how AI will converge with other technologies, such as quantum computing, electromagnetic pulses, Internet of Things, 5G, or blockchain/distributed ledgers. Potential convergences could produce strategic surprises that confuse and confound friends and foes alike, making the fog of war even more impenetrable and increasing the risks of escalation.

### *Who Is In the AI Loop?*

Whether or not there are humans in every part of the decisionmaking loop, that loop is getting crowded. The interface between humans and machines—where the proverbial "person in the loop" is supposed to exert human control— also raises critical questions about decisionmaking authority and organizational hierarchies.[50] Within the military, questions of rank, service branch, and responsibility for lethal actions can be contentious in the best of times, as illustrated by the debates over authority for U.S. drone strikes.[51] Deconflicting military and intelligence missions will not be made easier. With scores of AI-informed battlefield systems operating at breakneck speed, each connected to its own chain of command, coordination among the humans who are in the loop of fast-moving battlefield operations spanning multiple adversaries, domains, agencies, clearance levels, contractors, allies, and organizational cultures will be challenging, especially if the goal is to maintain offensive advantage via speedy decisionmaking. Budgets, reorganizations, accesses, personalities, and leadership changes may have as much influence over AI capabilities as the technology itself. There will be lots of men and women in the loop in lots of places, each influencing how AI contributes to separate and shared objectives. Achieving strategic effects will require extraordinary cooperation and communication.

### *Fake Nuclear News*

Public perception is a giant wildcard. AI algorithms are a central component of cyber influence



This content downloaded from
177.142.92.225 on Tue, 05 May 2020 16:02:51 UTC
All use subject to https://about.jstor.org/terms

operations aimed at shaping public perceptions. By now, it should be understood that the use and misuse of electronic media to manipulate public perceptions, including the use of fake news, cyber bots, and deep fakes, can affect strategic stability.[52] How the public views particular international conflicts can shape leadership decisionmaking and can build or undermine support for issues of war and peace, especially in democratic states. Decisions to escalate conflict could be influenced by public attitudes. AI-powered tools such as cyber bots and deep fake technology could enrage or pacify public opinion or mislead decisionmakers. Now that cyber conflict has become an ingrained feature of the international landscape, we should expect manipulation of public perceptions to affect crisis management, escalation, deterrence stability, and possibly nuclear decisionmaking.

### *Close Is Not Good Enough*

Decisions of war and peace cannot be left to predictive analytics. There are fundamental differences in the ways that data is used for scientific, economic, and logistic purposes and for predicting human behavior. Machine learning cannot reliably predict the outcomes of sports contests, elections, or international conflict, at least within acceptable margins of error for making big decisions involving questions of war and peace. Despite longstanding interest in predictive analytics that can tell decisionmakers what to expect before it happens, faith in the ability to predict incidents or outcomes of war and conflict based on big data machine learning is fraught with misplaced optimism.[53] Much like self-driving cars, where AI can correctly assess most—but not all—situations, a 90 percent success rate could mislead decisionmakers and put soldiers' and citizens' lives at stake. All of the potential dangers stemming from unreliable (outdated, biased, compromised) data, machine learning bias, and interpretation errors are magnified when human emotions, nonrational behavior, and inherent unpredictability cloud the data *and* the decisionmaking. The result is wider margins of error, which may be acceptable for research purposes but do not satisfy the practical and ethical demands of national security decisionmaking. Close is not good enough when it comes to war, especially where nuclear risks are involved.

### *Crowdsourcing Armageddon?*

Lastly, public–private partnerships shape the future of AI—but war remains the preserve of the state. As a quintessentially dual-use technology, AI is freely available to everyone. It is being developed and applied beyond the reach of governmental controls. Like many other dual-use technologies, governments rely on the private sector for the underlying research and development, software, hardware, and expertise required for AI to be used for military purposes. DOD and the Intelligence Community have deep ties to Silicon Valley and have doubled down on efforts to expedite the acquisitions process, especially for cyber and AI.[54] Competition among nations to secure AI talent could have strategic implications, especially with respect to counterintelligence, intellectual property, and respect for international norms of behavior.

### *America's Got Talent*

What this means in practice is that many countries will use the same experts, companies, and global supply chains to support their military AI aspirations, creating potential competitive conflicts of interest and security vulnerabilities related to sharing intellectual property. This dynamic is already evident in cyber markets, where Google and other companies have found it advantageous to accommodate Chinese government practices on censorship and surveillance while simultaneously expressing political opposition to supporting U.S. military AI projects such as Project Maven.[55] Global technology companies will have to weigh the costs and benefits of serving some



This content downloaded from
177.142.92.225 on Tue, 05 May 2020 16:02:51 UTC
All use subject to https://about.jstor.org/terms

national customers while keeping others at arm's length. The U.S. Government, however, has little choice but to remain heavily dependent on the private sector to develop and implement AI strategies.[56] Such dependence could have strategic implications if it interferes with our ability to compete for top talent and cutting-edge capabilities.

## How Could AI Alter the Fundamental Calculus of Deterrence?

In the classic Cold War movie *WarGames*, a young hacker breaks into a DOD supercomputer designed to use AI to plan and execute nuclear war plans. He engages the computer to play "Global Thermonuclear War" and accidentally triggers a simulated scenario of nuclear Armageddon, which is mistaken for the real thing. The computer ultimately learns that for nuclear deterrence, "the only way to win is not to play." If AI disrupts the central logic of nuclear deterrence as understood by the nuclear powers or fundamentally changes the underlying precepts that support it, the strategic consequences could be far-reaching, and the prospects that computers will learn "not to play" uncertain.

With these potential strategic impacts in mind, how could AI alter the fundamental calculus of deterrence? How might the convergence of the tactically and strategically relevant factors discussed above affect the strategic balance?

### *AI Is Changing Perceptions About the Threat of Surprise Attack*

At the top of the list of AI applications that could have true strategic significance for deterrence strategy is the threat of surprise attack. The combination of effective defenses with exquisite ISR that makes it possible to locate mobile targets and strike them with speed and precision raises long-held fears of an AI-guided "bolt from the blue" first strike. While the fundamental logic of deterrence is unchanged, perceptions that an adversary has sufficient intent and capability to conduct such a preemptive attack on vital assets can be expected to motivate a variety of countermeasures.

Evaluating the incentive to strike first evokes memories of Pearl Harbor, in which the United States underestimated Japan's risk calculus while fully recognizing Tokyo's military capacity to launch a cross-Pacific raid. AI contributions to military and intelligence capabilities do not override political considerations—with an important caveat added for the possibility of AI-fueled manipulation of public attitudes that could distort political judgment. Avoiding and deterring conflict remain a paramount responsibility for national leaders. Slightly improved odds of eliminating all but a few of an adversary's strategic weapons and shooting down any surviving retaliation with missile defenses still involves catastrophic risks and does not even begin to answer questions about the aftermath of such a conflict.

Nevertheless, possessing the theoretical capability to conduct a disarming first strike inevitably triggers a classic security dilemma, which is guaranteed to provoke countermeasures from those threatened by enhanced striking power. Further advances in defenses against counterforce strikes would be a predictable response, as well as hardening and camouflage to evade and confuse exquisite ISR. To the extent that AI influences perceptions of intent and capability and alters the calculus of risk and reward, it will inspire new thinking about possible offensive and defensive maneuvers in the evolution of nuclear strategy.[57]

### *Farewell to Mutual Vulnerability?*

Some may see AI as eroding mutual strategic vulnerability and thereby as increasing the risk of war. The combination of exquisite ISR with an effective defensive shield could make it tempting to conduct a disarming, decapitating, or blinding first strike at strategic targets, including nuclear command, control, and communications (NC3), early warning radars, or dual-capable missiles and aircraft.[58]



This content downloaded from
177.142.92.225 on Tue, 05 May 2020 16:02:51 UTC
All use subject to https://about.jstor.org/terms

Such a revision of deterrence logic could be highly destabilizing. Shared vulnerability and assured retaliation are central concepts of mutually assured destruction (MAD) deterrence theory. Switching the theoretical incentive from MAD to improve the odds of successfully conducting a disarming first strike could change the risk calculus that has formed the basis of strategic stability for decades.[59] Preventing such a revision of nuclear deterrence logic was the essence of Russia President Vladimir Putin's claim in March 2018 that his new weapons are "invincible against all existing and prospective missile defense and counter-air defense systems."[60] By evading *perceived* U.S. global strike and missile defense capabilities, Putin's claims about new AI-guided retaliatory forces were justified as efforts to preserve MAD.

### *AI Is Poised to Alter Regional Stability in Asia and Europe*

How could AI-assisted weapons systems affect regional stability, including U.S. allies? Widespread deployment of AI-supported ISR platforms is likely to affect regional stability in the five- to ten-year time frame. While the United States remains the leader in translating AI to currently deployed platforms, China and Russia are not far behind.[61] Many U.S. allies are rapidly advancing their own AI capabilities. Initially, the speed and lethality gained from AI-informed situational awareness and battle management systems are likely to provide the United States and its allies with options for countering Russian and Chinese A2AD.

The coming architecture of ISR, BMC3I, and defense systems appears well positioned to give net advantages to U.S. and allied regional security alliances. In addition to tactical military benefits, co-development of multidomain ISR provides opportunities for collaboration that directly addresses threats to allied security, especially with respect to extended deterrence relationships with key allies in Asia and Europe. Strengthening regional conventional deterrence and regional extended nuclear deterrence reduces incentives for risk taking and supports broader interests in strategic deterrence. AI applications that support these objectives will have beneficial effects for strategic stability.

### *AI Competition Could Also Benefit Strategic Stability and Bolster Deterrence*

Global competition in military AI is already heating up. An AI arms race is under way. Whatever advantages are possible in the near term, however, may be short-lived as U.S. allies, major adversaries, and a multitude of rising powers incorporate AI into their political and military strategies. In light of the rising tide that is advancing AI prospects around the world, temporary advantages are unlikely to yield lasting military predominance. For example, China and Russia will eventually possess their own versions of multidomain ISR coupled with precision strike and layered defenses. How will these capabilities influence Beijing's thinking about the U.S. role in the South China Sea, or Russian assessments of the North Atlantic Treaty Organization's defense of the Baltics?

These are not primarily technical issues. AI is enhancing the performance of many tactical and strategic systems but is not giving definitive unilateral advantage to any one. The nature of warfare is changing, and AI is fueling many of those changes, but the fundamental calculus of deterrence remains steady. Competition for military capabilities that retains a balance of power can be stabilizing.

## Risks of Unintended Consequences and Strategic Surprise

Predicting the future of technology is a risky business. We know with certainty that AI is being incorporated into an array of military missions with the intent of improving our knowledge of the operational environment, adversary capabilities,



This content downloaded from
177.142.92.225 on Tue, 05 May 2020 16:02:51 UTC
All use subject to https://about.jstor.org/terms

and the speed and precision of offensive and defensive weapons. We can usefully speculate about how these developments are poised to change the face of modern warfare and how those changes might affect regional and strategic deterrence stability, based on our understanding of established political and military realities. More elusive, however, is a clear picture of how AI might converge with other technologies to produce unexpected outcomes, or "unknown unknowns." Nevertheless, here are a few possibilities that could have major strategic consequences and alter the underlying realities on which regional and strategic stability are founded:

- Distorted data could lead AI systems to take unintended actions, such as incorrectly identifying and striking the wrong targets. Data can be polluted intentionally via counter-AI methods or can occur naturally for many reasons. Such actions could hasten escalation and interfere with conflict management efforts.
- Compounding the problems of distorted data, AI makes mistakes with a frequency that could be untenable for decisions affecting strategic stability. Misinterpretations of data that lead to unintended actions could spark highly undesirable reactions, including escalation and retaliation.
- The convergence of AI and cyber presents several possibilities for unintended consequences and strategic surprise. AI-informed cyber attacks on NC3 could present the target of such an attack with a "use it or lose it" situation, prompting early resort to nuclear weapons.
- AI-supported cyber/information warfare, including use of fake news, deep fakes, and other methods could distort public and leadership perceptions of international events, inflaming passions and prompting escalation.
- Accelerated battle rhythm made possible by multidomain ISR could preclude diplomatic efforts to avoid or deescalate conflict. Even if AI works perfectly to increase the speed and lethality of warfare, moving at the speed of AI might not be optimal for all cases.
- Unpredictable AI interactions with foreign and friendly platforms could produce unwanted AI calculations that misrepresent human intentions. The "black box" underlying AI decisions is not well understood and could produce destabilizing results, such as striking the wrong targets.
- Unexpected convergences with other technologies, such as quantum computing and electromagnetic pulse, could confuse/distort offensive or defensive instructions and lead to undesirable results, such as striking the wrong targets.
- If it were eventually possible through a variety of AI-supported information gathering methods, emerging technologies, and analytic tools to track strategic assets such as submarines, the sanctity of assured retaliation could come into question. Such a strategic surprise could prompt a variety of destabilizing actions, including possible movement toward launch on warning postures.

## AI Is Part of a Bigger Challenge for Deterrence, Stability, and Strategy

Evolutionary changes in the logic of regional and strategic deterrence are not new, nor are they necessarily harmful to U.S. national security. Efforts to integrate AI-based technologies into U.S. defense and intelligence strategies illustrate the continued innovation and competitive advantages sought in support of U.S. national security policy. Applications of AI that support U.S. nuclear forces and infrastructure, such as command and control, logistics, and stockpile stewardship, serve to reinforce strategic deterrence by bolstering the survivability and credibility of our retaliatory forces.



This content downloaded from
177.142.92.225 on Tue, 05 May 2020 16:02:51 UTC
All use subject to https://about.jstor.org/terms

AI that bolsters tactical/battlefield applications can also support strategic deterrence, especially in a regional context. The connection between regional and strategic deterrence has always been important and appears to be even more tightly coupled as increased speed, precision, and lethality at the tactical level hold the potential to produce military outcomes that could escalate to the strategic level of conflict. Specifically, failure to deter Chinese or Russian aggression against U.S. regional allies that results in armed conflict may be hard to contain, especially if early victories on the battlefield leave one side facing a humiliating defeat. The United States and its allies still maintain conventional superiority, and AI is likely to extend those advantages in the near term to defeat Russian and Chinese A2AD. Rather than accept defeat, Russia or China might choose an "escalate to de-escalate" strategy that includes use of nuclear or other unconventional weapons to mitigate the technological advantages held by the United States and its allies, including AI-supported ISR, battle management, and defenses. For the military applications of AI to advance U.S. national security objectives, they must be integrated into a broader strategy that reinforces deterrence at the regional and strategic levels.

The rapid expansion of AI's military applications throughout the world merits a high level of focused attention to ensure maximum advantage for the United States and its allies, to minimize its negative impacts on strategic stability, and to prevent strategic surprise. PRISM

## *Notes*

[1] Anthony Cuthbertson, "What's Bigger than Fire and Electricity? Artificial Intelligence, Says Google Boss," *Newsweek*, January 22, 2018, <https://www.newsweek.com/artificial-intelligence-more-profound-electricity-or-fire-says-google-boss-786531>; "Elon Musk: Mark My Words—AI Is Far More Dangerous than Nukes," CNBC, March 13, 2018, <https://www.cnbc.com/2018/03/13/elon-musk-at-sxsw-a-i-is-more-dangerous-than-nuclear-weapons.html>; Peter Holley, "Stephen Hawking Just Got an Artificial Intelligence Upgrade, but Still Thinks AI Could Bring an End to Mankind," *Washington Post*, December 2, 2014.

[2] "Whoever Leads in AI Will Rule the World: Putin to Russian Children on Knowledge Day," RT News, September 1, 2017, <https://www.rt.com/news/401731-ai-rule-world-putin/.>

[3] Paul Mozur, "Beijing Wants A.I. to Be Made in China by 2030," *New York Times*, July 20, 2017, <https://www.nytimes.com/2017/07/20/business/china-artificial-intelligence.html>.

[4] Patrick Tucker et al., *The Race for AI: The Return of Great Power Competition Is Spurring the Quest to Develop Artificial Intelligence for Military Purposes* (Defense One e-book, March 2018), <https://www.defenseone.com/assets/race-ai/portal/>.

[5] Jürgen Schmidhuber, "Deep Learning in Neural Networks: An Overview," *Neural Networks* 61 (January 2015): 85–117, <https://doi.org/10.1016/j.neunet.2014.09.003>.

[6] Dom Galeon and Christianna Reedy, "Kurzweil Claims that the Singularity Will Happen by 2045," *Futurism*, October 5, 2017, <https://futurism.com/kurzweil-claims-that-the-singularity-will-happen-by-2045/>; "Artificial Intelligence and Life in 2030, One Hundred Year Study on Artificial Intelligence," report of the 2015 Study Panel, Stanford University, September 2016, <https://ai100.stanford.edu/sites/default/files/ai_100_report_0901fnlc_single.pdf>.

[7] Sara Reardon, "AI-controlled Brain Implants for Mood Disorders Tested in People," *Nature*, November 22, 2017; Antonio Regalado, "Reversing Paralysis: Scientists Are Making Remarkable Progress at Using Brain Implants to Restore the Freedom of Movement that Spinal Cord Injuries Take Away," *MIT Technology Review*, <https://www.technologyreview.com/s/603492/10-breakthrough-technologies-2017-reversing-paralysis/>.

[8] Sara Reardon, "The Pentagon's Gamble on brain Implants, Bionic Limbs, and Combat Exoskeletons," *Nature*, June 10, 2015; Annie Jacobsen, "Engineering Humans for War," *The Atlantic*, September 23, 2015, <https://www.theatlantic.com/international/archive/2015/09/military-technology-pentagon-robots/406786/>; Michael Joseph Gross, "The Pentagon's Push to Program Soldiers' Brains," *The Atlantic*, November 2018, <https://www.theatlantic.com/magazine/archive/2018/11/the-pentagon-wants-to-weaponize-the-brain-what-could-go-wrong/570841/>.

[9] David Weinberger, "Our Machines Now Have Knowledge We'll Never Understand," *Wired*, April 18,



This content downloaded from
177.142.92.225 on Tue, 05 May 2020 16:02:51 UTC
All use subject to https://about.jstor.org/terms

2017, <https://backchannel.com/our-machines-now-have-knowledge-well-never-understand-857a479dcc0e>.

[10] Darrell West, *The Future of Work: Robots, AI, and Automation* (Washington, DC: Brookings Institution Press, 2018); Molly Kinder, "Learning to Work with Robots: AI Will Change Everything. Workers Must Adapt—or Else," *Foreign Policy*, July 11, 2018, <https://foreignpolicy.com/2018/07/11/learning-to-work-with-robots-automation-ai-labor/>.

[11] Zachary Davis and Michael Nacht, eds., *Strategic Latency: Red, White and Blue, Managing the National and International Security Consequences of Disruptive Technologies* (Berkeley, CA: Lawrence Livermore National Laboratory, 2018).

[12] Frank G. Hoffman, "Will War's Nature Change in the Seventh Military Revolution? Exploring War's Character and Nature," *Parameters* 47, no. 4 (Winter 2017–18).

[13] "Deputy Secretary: Third Offset Strategy Bolsters America's Military Deterrence," Department of Defense, October 31, 2016, <https://dod.defense.gov/News/Article/Article/991434/deputy-secretary-third-offset-strategy-bolsters-americas-military-deterrence/>.

[14] "Summary of the 2018 Department of Defense Artificial Intelligence Strategy," February 2018, <https://media.defense.gov/2019/Feb/12/2002088963/-1/-1/1/SUMMARY-OF-DOD-AI-STRATEGY.PDF>; Memorandum from the Deputy Secretary of Defense, "Establishment of the Joint Artificial Intelligence Center," June 27, 2018.

[15] The White House, "Summary of the 2018 White House Summit on Artificial Intelligence for American Industry," May 10, 2018; The 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/>.

[16] DARPA, "DARPA Announces $2 Billion Campaign to Develop Next Wave of AI Technologies," September 7, 2018, <https://www.darpa.mil/news-events/2018-09-07>.

[17] Daniel Hoadley and Nathan Lucas, *Artificial Intelligence and National Security* (Washington, DC: Congressional Research Service, April 26, 2018); Marcus Weisgerber, "The Pentagon's New Artificial Intelligence Is Already Hunting Terrorists," *Defense One*, December 21, 2017, <https://www.defenseone.com/technology/2017/12/pentagons-new-artificial-intelligence-already-hunting-terrorists/144742/>; Matt Leonard, "Army Leverages Machine Learning to Predict Component Failure," *Defense Systems*, July 2, 2018, <https://defensesystems.com/articles/2018/07/03/army-vehicle-predictive-maintenance.aspx>.

[18] "Strategic Latency and Warning: Private Sector Perspectives on Current Intelligence Challenges in Science and Technology," Report of the Expert Advisory Panel Workshop, Lawrence Livermore National Laboratory, January 8, 2016. *Strategic warning* describes the goal of alerting decisionmakers of impending threats of a strategic nature. *Strategic surprise* describes the failure to provide adequate warning of such threats.

[19] Kelsey Atherton, "Targeting the Future of the DOD's Controversial Project Maven Initiative," *C4ISRNET*, July 27, 2018, <https://www.c4isrnet.com/it-networks/2018/07/27/targeting-the-future-of-the-dods-controversial-project-maven-initiative/>.

[20] Jack Corrigan, "Project Maven Uses Machine Learning to Go Through Drone Video Feeds, but That's Just the Beginning, Air Force Lt. Gen Shanahan Said," *Nextgov*, November 2, 2017, <https://www.nextgov.com/cio-briefing/2017/11/three-star-general-wants-artificial-intelligence-every-new-weapon-system/142225/>.

[21] National Academies of Science, *Autonomy in Land and Sea and in the Air and Space, Proceedings of a Forum*, 2018, <http://nap.edu/25168>.

[22] National Academy of Sciences, *Counter-Unmanned Aircraft System (CUAS) Capability for Battalion and Below Operations, Abbreviated Version of a Restricted Report*, 2018, <www.nap.edu/read/24747/chapter/1>.

[23] Defense Science Board, *Study on Countering Anti-access Systems with Longer Range and Standoff Capabilities: Assault Breaker II*, 2017 Summer Study on Long Range Effects, Office of the Under Secretary of Defense for Acquisition, Technology, and Logistics, June 2018.

[24] Lisa Owens Davis, "Moving at the Speed of S&T: Calibrating the Role of National Laboratories to Support National Security," in Davis and Nacht, *Strategic Latency: Red, White and Blue*.

[25] "Machine Learning to Prevent Defects in Metal 3D Printed Parts in Real Time," Lawrence Livermore National Laboratory, Newsline, September 13, 2018, <https://webcenter.llnl.gov/myllnl/faces/oracle/webcenter/portalapp/pages/top-story-wrapper.jspx?articleId=52535&_afrLoop=77869951013468&_afrWindowMode=0&_afrWindowId=blank#%40%3F_afrWindowId%3Dblank%26_afrLoop%3D77869951013468%26articleId%3D52535%26_afrWindowMode%3D0%26_adf.ctrl-state%3Dt66qlfya5_65>.

[26] Andrew Reddie, Bethany Goldblum, Kiran Lakkaraju, Jason Reinhardt, Michael Nacht, Laura Epifanovskaya, "Next Generation Wargames" *Science*, December 21, 2018; Magy Seif El-Nasr, Anders Drachen, Alessandro Canossa, editors, *Game Analytics: Maximizing the Value of Player Data*, (London: Springer, 2013)



This content downloaded from
177.142.92.225 on Tue, 05 May 2020 16:02:51 UTC
All use subject to https://about.jstor.org/terms

[27] Marc Pomerleau, "Can the Intel and Defense Community Conquer Data Overload?" *C4ISRNET*, September 5, 2018, <https://www.c4isrnet.com/intel-geoint/2018/09/05/can-the-intel-and-defense-community-conquer-data-overload/?utm_source=Sailthru&utm_medium=email&utm_campaign=daily%20brief%209/5/18&utm_term=Editorial%20-%20Daily%20Brief>.

[28] Marc Pomerleau, "Here's How Intelligence Agencies Will Take Advantage of Machine Learning and AI," *C4ISRNET*, May 1, 2018, <https://www.c4isrnet.com/intel-geoint/2018/05/01/heres-how-intelligence-will-take-advantage-of-machine-learning-and-ai/>.

[29] "Deep Learning to Advance Nuclear Nonproliferation," *LLNL Newsline*, August 21, 2018.

[30] Ben Conklin, "How Artificial Intelligence Is Transforming GEOINT," GCN, April 18, 2018, https://gcn.com/articles/2018/04/18/ai-transform-geoint.aspx; Sandra Erwin, "NGA official: Artificial Intelligence Is Changing Everything, We Need a Different Mentality," *Spacenews*, May 13, 2018, <https://spacenews.com/nga-official-artificial-intelligence-is-changing-everything-we-need-a-different-mentality/>.

[31] Edward Geist and Andrew Lohn, *How Might Artificial Intelligence Affect the Risk of Nuclear War?* (Santa Monica, CA: RAND, 2018); Paul Bracken, "The Intersection of Cyber and Nuclear War," *The Strategy Bridge*, January 17, 2017, <https://thestrategybridge.org/the-bridge/2017/1/17/the-intersection-of-cyber-and-nuclear-war>.

[32] Jeremy Hsu, "AI Can Help Hunt Down Missile Sites in China," *Wired*, November 21, 2017, <https://www.wired.com/story/ai-can-help-hunt-down-missile-sites-in-china/>.

[33] Jen Judson, "Hyten: To Address Russian and Chinese Missile Threats, It's All About the Sensors," *Defense News*, August 7, 2018, <https://www.defensenews.com/digital-show-dailies/smd/2018/08/07/hyten-to-address-russian-and-chinese-missile-threats-its-all-about-the-sensors/>.

[34] Jack Corrigan, "DARPA Wants to Find Botnets Before They Attack," *Defense One*, September 12, 2018, <https://www.defenseone.com/technology/2018/09/darpa-wants-find-botnets-they-attack/151199/?oref=defenseone_today_nl>.

[35] "Nuclear Weapons in the New CyberAge: A Report of the Cyber-Nuclear Weapons Study Group," Nuclear Threat Initiative, September 2018, <https://www.nti.org/media/documents/Cyber_report_finalsmall.pdf>.

[36] Michael Sulmeyer and Kathryn Dura, "Beyond Killer Robots: How Artificial Intelligence Can Improve Resilience in Cyber Space," *War on the Rocks*, September 6, 2018, <https://warontherocks.com/2018/09/beyond-killer-robots-how-artificial-intelligence-can-improve-resilience-in-cyber-space/>.

[37] Dave Johnson, *Russia's Conventional Precision Strike Capabilities, Regional Crises, and Nuclear Thresholds*, Livermore Papers on Global Security no. 3, February 2018; John Warden, *Limited Nuclear War: The 21st Century Challenge for the United States*, Livermore Papers on Global Security no. 4, July 2018.

[38] Cathy O'Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy* (New York: Broadway Books, 2017).

[39] Hillary Sanders and Joshua Saxe, "Garbage In, Garbage Out: How Purportedly Great Machine Language Models Can Be Screwed Up by Bad Data," Proceedings of Blackhat 2017, July 26–27, 2017, Las Vegas, NV.

[40] Jesse Emspak, "How a Machine Learns Prejudice," *Scientific American*, December 29, 2016, <https://www.scientificamerican.com/article/how-a-machine-learns-prejudice/>; ProPublica, "Machine Bias," May 23, 2016, <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>; Will Knight, "Forget Killer Robots, Bias Is the Real AI Danger," *Technology Review*, October 3, 2017, <https://www.technologyreview.com/s/608986/forget-killer-robotsbias-is-the-real-ai-danger/>.

[41] Louise Matsakis, "Researchers Fooled a Google AI into Thinking a Rifle Was a Helicopter," *Wired*, December 20, 2017, <https://www.wired.com/story/researcher-fooled-a-google-ai-into-thinking-a-rifle-was-a-helicopter/>.

[42] Daniel Cebul, "Differentiating a Port from a Shipyard Is a New Kind of Problem for AI," *C4ISRNET*, September 18, 2018, <https://www.c4isrnet.com/intel-geoint/2018/09/18/differentiating-a-port-from-a-shipyard-is-a-new-kind-of-problem-for-ai/?utm_source=Sailthru&utm_medium=email&utm_campaign=Daily%209/19&utm_term=Editorial%20-%20Daily%20Brief>.

[43] Anna Rohrbach et al., "Object Hallucination in Image Captioning," Cornell University Library, <https://arxiv.org/abs/1809.02156>.

[44] Sandia National Laboratory, *Counter Adversarial Data Analytics*, SAND2015-3711, May 8, 2015.

[45] Defense Science Board, "Memorandum for Chairman, Terms of Reference, Defense Science Board Task Force on Counter Autonomy," June 18, 2018, <https://www.acq.osd.mil/dsb/TORs/2018_TOR_CounterAutonomy_18Jun2018.pdf>.

[46] Tim Weiner, "Word for Word, The Cuban Missile Crisis: When Kennedy Faced Armageddon, and His Own Scornful General," *New York Times*, October 5, 1997, <https://www.nytimes.com/1997/10/05/weekinreview/



This content downloaded from
177.142.92.225 on Tue, 05 May 2020 16:02:51 UTC
All use subject to https://about.jstor.org/terms

word-for-word-cuban-missile-crisis-when-kennedy-faced-armageddon-his-own.html>.

[47] Paul Scharre, "A Million Mistakes a Second," *Foreign Policy*, September 12, 2018, <https://foreignpolicy.com/2018/09/12/a-million-mistakes-a-second-future-of-war/>.

[48] Lawrence Livermore National Laboratory, "Building a Network of Collaborative Autonomous Machines," Science and Technology Review (June 2018); Mark Pomerleau, "To Win Future Conflicts, Combatant Commands Must Be Integrated," *C4ISRNET*, August 15, 2018, <https://www.c4isrnet.com/show-reporter/dodiis/2018/08/14/to-win-future-conflicts-combatant-commands-must-be-integrated/?utm_source=Sailthru&utm_medium=email&utm_campaign=Daily%208/15&utm_term=Editorial%20-%20Daily%20Brief>.

[49] Will Knight, "The Dark Secret at the Heart of AI," *Technology Review*, April 11, 2017, <https://www.technologyreview.com/s/604087/the-dark-secret-at-the-heart-of-ai/>.

[50] Michael Piellusch and Tom Galvin, "Is the Chain of Command Still Meaningful?" *War Room*, U.S. Army War College, September 6, 2018, <https://warroom.armywarcollege.edu/articles/chain-of-command/>.

[51] Stimson Center, *An Action Plan on U.S. Drone Policy: Recommendations for the Trump Administration*, 2018, <https://www.stimson.org/sites/default/files/file-attachments/Stimson%20Action%20Plan%20on%20US%20Drone%20Policy.pdf>.

[52] Herb Lin, "Developing Responses to Cyber-Enabled Information Warfare and Influence Operations," *Lawfare*, September 6, 2018, <https://www.lawfareblog.com/developing-responses-cyber-enabled-information-warfare-and-influence-operations>.

[53] Kori Schake, "Why We Get It Wrong: Reflecting on the Future of War," book review of Lawrence Freedman, *The Future of War: A History, War on the Rocks*, August 10, 2018, <https://warontherocks.com/2018/08/why-we-get-it-wrong-reflections-on-predicting-the-future-of-war/>; Richard Danzig, *Driving in the Dark: Ten Propositions About Prediction and National Security*, Center for a New American Security, October 2011.

[54] Frank Gac, Timothy Grayson, and Joseph Keogh, "What Works? Public-Private Partnerships for Development of National Security Technology," in Davis and Nacht, *Strategic Latency*.

[55] Suzanne Nossel, "Google Is Handing the Future of the Internet to China," *Foreign Policy*, September 10, 2018, <https://foreignpolicy.com/2018/09/10/google-is-handing-the-future-of-the-internet-to-china/.>

[56] Laura Seligman, "Why the Military Must Learn to Love Silicon Valley," *Foreign Policy*, September 12, 2018, <https://foreignpolicy.com/2018/09/12/why-the-military-must-learn-to-love-silicon-valley-pentagon-google-amazon/.>

[57] Lawrence Freedman, *The Evolution of Nuclear Strategy* (New York: St. Martin's Press, 1981).

[58] James Acton, "Escalation through Entanglement: How the Vulnerability of Command and Control Systems Raises the Risks of an Inadvertent Nuclear War," *International Security* 43, no. 1 (Summer 2018).

[59] Kier Lieber and Daryl Press, "The New Era of Counterforce: Technological Change and the Future of Nuclear Deterrence," *International Security* 41, no. 4 (Spring 2017).

[60] August Cole and Amir Husain, "Putin Says Russia's New Weapons Can't Be Beat. With AI and Robotics, They Can," *Defense One*, March 13, 2018, <https://www.defenseone.com/ideas/2018/03/putin-says-russias-new-weapons-cant-be-beat-ai-and-robotics-they-can/146631/>.

[61] Elsa Kania and John Costello, *Quantum Hegemony: China's Ambitions and the Challenge to U.S. Innovation Leadership* (Washington, DC: Center for a New American Security, September 12, 2018), <https://www.cnas.org/publications/reports/quantum-hegemony>.



This content downloaded from
177.142.92.225 on Tue, 05 May 2020 16:02:51 UTC
All use subject to https://about.jstor.org/terms