The Unavoidable Technology: How Artificial Intelligence Can Strengthen Nuclear Stability

Jessica Cox & Heather Williams


To link to this article: https://doi.org/10.1080/0163660X.2021.1893019

Published online: 23 Mar 2021.

View related articles

View Crossmark data
Artificial intelligence (AI) is developing at a rapid pace—both the technology itself and its applications. It is becoming unavoidable in both the civilian and military domains and will soon impact numerous areas of civilian and military life. In July 2020, for example, countries such as Colombia and Russia applied facial recognition technology, a form of artificial intelligence, to combat the coronavirus by detecting whether someone was not wearing a mask or had a high temperature through the use of thermal cameras and sounding an alarm.¹ That same month, Open AI, a California company leading in AI technology, released the software GPT-3, a text generator that can mimic human creativity to write convincing essays, emails, or tweets.² Global AI-generated revenue is projected to climb from US$643.7 million in 2016 to US$36.8 billion in 2025, a factor of almost 60 times greater.³

Both the United States and its NATO allies have placed new emphasis on understanding the civilian and military applications of technological

---

The work of Jessica Cox is © 2021 North Atlantic Treaty Organisation
The work of Heather Williams is © 2021 Heather Williams
The Washington Quarterly • 44:1 pp. 69–85
https://doi.org/10.1080/0163660X.2021.1893019
advancements in the field of AI. This priority comes as both China and Russia have indicated increasing investments in AI as a force multiplier and enabler across society, and as countries like South Korea and Japan turn to AI to support vital sectors of their economies. But developments in AI are largely being driven by the private sector, as every major tech company competes to have the latest breakthroughs.

From a nuclear policy perspective, the majority of scholarship addressing the implications of AI, with a few important exceptions, argues that military applications of AI will have a destabilizing effect on strategic stability and global security. Areas regularly cited as problematic include the potential for deep fakes, human bias carried over into AI programming, cyber vulnerabilities, exacerbating great power rivalry, application to nuclear command and control, security dilemmas, and potential for loss of human control. This focus creates a rather defeatist paradigm, particularly for policymakers confronted with rapid advances in technology and limited thinking on how it might be applied in a stabilizing way. Whereas many of these studies treat AI as a monolith, we argue that the opportunities and risks associated with AI depend on its application, and considering their impact on nuclear weapons policies requires a greater degree of granularity.

Can some military applications of AI strengthen deterrence, for example? Might it facilitate advances in arms control, at a time when the international community desperately needs new ideas and new verification techniques?

Despite the potential risks inherent with adopting AI in nuclear systems, many of the potential benefits to both deterrence and arms control have been overlooked or ignored. Additionally, many in the nuclear policy community rely on the most extreme negative examples to make the case that AI is too dangerous to adopt into nuclear policies and practices. These views are short-sighted and ignore the reality that AI will increasingly play a role in all aspects of our civil and military policies—and that potential great power competitors such as China and Russia are investing heavily in AI-related military technology. The broader nuclear policy community should therefore start thinking now about how to harness the power of AI as a tool for greater stability, transparency, and security as well as steps that can be put in place to mitigate the most significant risk factors.

We argue that considering AI’s impact on nuclear weapons—both in the field of nuclear deterrence and for strategic arms control—requires a more nuanced approach, whereby it offers both opportunities as well as challenges for strategic stability. Because the risks of AI are increasingly familiar and discussed in
increasing scope and detail by experts and scholars, we choose here to focus our exploration on the potential benefits of AI on nuclear policy, which have received far less attention. We do not mean to diminish the potential risks associated with increasing reliance on AI, but rather to offer a more optimistic picture of how AI might also contribute to nuclear deterrence and arms control in unique and helpful ways in order to answer this question: how can decision-makers pursue potentially stabilizing applications of AI while managing the risks associated with AI and nuclear systems?

**More Electricity than Lightbulb—What Is “AI”?**

Trying to define “AI” has become somewhat of a parlor trick. Is it software or hardware? Killer robots or drone swarms? Director of Perry World House, a center at the University of Pennsylvania, Michael Horowitz, offers a helpful framework wherein AI can either be applied broadly as “systems that select and engage targets on their own,” such as computer-guided precision weapons, or narrowly as “intelligent machines capable of cognitive judgements on par with humans.”

According to the US Air Force and NATO, AI “refers to the ability of machines to perform tasks that normally require human intelligence—for example, recognizing patterns, learning from experience, drawing conclusions, making predictions, or taking action—whether digitally or as the smart software behind autonomous physical systems.”

AI is often associated with lethal autonomous weapons (LAWS), which are weapons systems that can be activated to select and engage targets independently; however, we are interested in a broader range of AI military applications.

While AI systems may behave autonomously, they are not necessarily automatic. For decades, militaries have relied on automatic technologies, such as missile defense interceptors or perimeter security systems, that do not rely on AI. AI is also not synonymous with “unmanned” systems like unmanned aerial vehicles (UAVs) or underwater drones. What makes AI and machine learning unique from automated or unmanned systems is the ability to process massive amounts of data and make an independent analysis to achieve a desired objective. There are, of course, varying degrees of autonomy within AI, depending on a systems’ ability to move, take/provide direction, and make determinations. But whereas automatic systems are reactive, AI is more proactive. It can sift, sort, identify, and offer an initial conclusion on what something means. AI is particularly valuable in a society that creates ever-increasing amounts of data—for example, in 2020, the world was estimated to produce 44 trillion gigabytes of digital data, an amount humanly impossible to synthesize.
To be clear, AI, machine learning, and deep learning are distinct but often confused concepts that fall under data synthesis. AI is the umbrella term describing the capability of a machine to imitate intelligent human behavior, such as complex decision-making. Machine learning is a method of AI whereby computers have the ability to learn without being specifically programmed for a particular task, and deep learning is a subset of machine learning where a computer has the ability to learn more complex relationships between datasets.

Of course, AI goes well beyond military applications. Other applications include facial recognition, which has been used in China for years and was deployed for the first time in the UK in early 2020 against a watchlist of violent offenders. And law firms have begun exploring the use of AI for contract and document reviews to identify relevant information and reduce human error. Kevin Kelly, a leading technology expert and founder of Wired magazine, posits that technologies such as AI are “inevitable” and sees AI as akin to electricity. AI is not the light-bulb or a single application of electricity—it is the electricity itself, whereby it is a fundamental enabler of civilian and military systems and technologies. In short, AI is an unavoidable technology for governments, societies, and militaries.

Whether or not AI is as problematic as many people fear thus depends on politics more than the technology itself. In a recent article in The Washington Quarterly, researcher James Johnson cautioned that AI could create instability but also acknowledged that “AI is unlikely to be a strategic game-changer. Instead, it will likely reinforce the destabilizing effects of existing advanced conventional—especially counterforce—capabilities, thereby increasing the speed of warfare and compressing the decision-making timeframe.”

The impact of AI will also depend on its application in relation to other technologies. For example, if AI is adopted by a nuclear-armed nation with an aggressive nuclear neighbor, AI could have a stabilizing effect by improving early warning systems, reducing the pressure for early launches, and preventing miscalculations. Conversely, given the current lack of methods to test the safety and reliability of machine learning and other AI tools in nuclear weapons technology, integrating AI into nuclear systems may have a destabilizing effect, potentially undermining confidence in deterrence capabilities or creating additional vulnerabilities in nuclear systems and decision-making.
Nuclear Deterrence—A Role for AI in Strategic Stability

Traditionally, nuclear deterrence relies on a country having credible nuclear capabilities plus the political will or resolve to use them if necessary (and only if necessary), therefore persuading an adversary to not attack. Integrating AI into nuclear systems can enhance nuclear decision-making, improving perceptions of both capability and resolve.

Early Warning and Detection: Avoiding Unnecessary Use and Escalation

Most significantly, integrating AI into intelligence, surveillance, and reconnaissance (ISR) systems and using AI as part of the analytical tool kit for early warning and detection could improve target identification, prevent false positives or close calls, and increase understanding of adversary actions. AI can be used to increase a nation’s ability to understand the nature of its environment as well as speed of analysis and development of courses of action in a conflict, thus improving its overall capabilities in the nuclear space and reducing the risk of unintended escalation in the midst of a crisis with a nuclear dimension.

In a nuclear exchange, national leadership may have just minutes to assess a situation and determine whether to launch its own nuclear response. Nuclear weapons states such as the United States have sought to increase this decision-making time to reduce the potential for misunderstanding or miscalculation in a crisis, given the immense consequences of using nuclear weapons. Incorporating AI into early warning and decision-making analysis may improve both the speed and quality of information processing as well as eliminate potential biases from military leadership, which will be crucial in a crisis to allow time for de-escalation and tension reduction between sides.

AI can similarly enhance the ability to discriminate between real and false information, which is critical to preventing miscalculation or mistake in a crisis. A number of “close calls” have occurred throughout the years, in which nuclear weapons use was only avoided through human intervention based on uncertain information. In 1983, NATO undertook Command Post Exercise Able Archer to “practice command and staff procedures, with particular emphasis on the transition from conventional to non-conventional operations, including the use of nuclear weapons.” At the time, Soviet paranoia about a US nuclear first strike, with preparations under cover of a war game, was at an all-time high. When NATO began its Able Archer exercise, Soviet officials thought the exercise was real and put mobile intercontinental ballistic missiles
(ICBMs) on a three-minute alert. The same year, the Soviet early warning computer system signaled five incoming US Minuteman ICBMs. The watch officer on duty, Soviet Lt. Col. Stanislav Petrov, did not report the incident, concluding that it must be a false alarm. If he had reported that US nuclear missiles were inbound, the Soviets would have followed their nuclear doctrine and retaliated with no time to double-check or negotiate with the United States. Petrov ultimately made the right call—as the Soviet early warning system had mistaken sun shining off clouds for incoming missiles.

Even after the end of the Cold War, in 1995, a Norwegian civilian research rocket launch was detected by the radar crews from the Russian Missile Attack Warning System (MAWS) and mistaken for a US Trident II submarine-launched ballistic missile. Command and control procedures were enacted, including notification of President Boris Yeltsin and activation of the Russian “cheget” system—essentially its nuclear football. Russian authorities eventually determined that this was not a nuclear-armed ballistic missile launch after reviewing satellite and other intelligence information.

While these and other incidents were ultimately resolved successfully through the intervention of military experts and political leaders, they illustrate the dangers of imperfect systems that rely on human analysts with potentially inaccurate or inadequate information. As defense researcher Jaganath Sankaran argues, “future applications of AI to nuclear command and control should aspire to create an algorithm that could argue in the face of overwhelming fear of an impending attack that a nuclear launch isn’t happening.” Such an approach could both reduce the fog of war and reassure decision-makers that their course of action is correct in the face of uncertain information, either using nuclear weapons or refraining from it.

Although most nuclear-armed nations now use multiple systems to reduce the chances of a false positive warning signal, advancements in AI could significantly improve confidence in and functionality of these systems when integrated with human decision-making. AI tools could play a significant role in helping to identify patterns of life and reduce potential operator biases in conducting analysis. For instance, in order to be able to quickly discriminate between a country’s launch of a sounding/research rocket as opposed to an ICBM, which have very similar radar signatures, big data analytical tools could be used to collate and process massive amounts of electronic data—including signals, imagery, and open source collection—over time to identify patterns of behavior unique to each type of launch. Then, if there is an ambiguous or unexpected launch, these systems would be able to quickly determine whether or not the current circumstances more closely resemble one type of launch over the other—actually preventing future close calls or misinterpretations of data, as occurred in Able Archer in 1983. A human analyst would likely still need to make a final determination, but
through the use of AI and data analytics, they would have a more accurate, timely, and complete picture on which to base any decisions.

Similarly, these same types of tools might eventually be able to discern and discriminate between nuclear and conventional weapons systems in an incoming attack—preventing the risk of unintended escalation in a conflict between nuclear armed adversaries. More countries are developing dual-capable missile systems—that is, missiles that can carry either conventional or nuclear warheads on the same missile airframe—which a country on the receiving end would not be able to distinguish between until it detonates on the ground. AI may be able to quickly assess such technical information as telemetry and flight trajectory to determine whether the payload is nuclear or conventional. Alternately, “pattern of life” analysis—that is, regular monitoring of activities and behaviors through intelligence so that anomalies can be quickly and easily identified—may reveal signs in advance of a launch that would signal the type of warhead on the missile system, enabling military officials to calibrate responses in a crisis. Either way, big data analytical systems could cull through large quantities of ISR, open source, and other sensor data to mitigate problems with single data sources or sensor malfunction—and do so within the timelines necessary in a wartime environment, where there will be imprecise information, short timelines, and chaos in the information space.

AI-enabled systems might also be critical to fuse various forms of data that are individually ambiguous, but, when synthesized quickly enough, collectively indicate that a missile has in fact been launched. Jason Matheny, a member of the National Security Commission on AI, notes that this is particularly important for “different scenarios where you may not have a definitive warning from a single sensor. And particularly as missile designs are meant to evade certain classes of missile warning systems, we’re going to have to use some form of sensor fusion that’s AI enabled in order to alert the human operator that a missile has likely been launched.”

Incorporating AI into early warning systems may be particularly stabilizing for countries that do not have the advanced satellites, sensors, and forward-deployed radar systems that the United States and Russia have developed over the years to ensure any missile launch is detected and assessed for threat potential within minutes. For instance, incorporating AI into early warning and detection systems may provide China with more transparency and confidence into America’s and its allies’ activities in the Asia-Pacific, reducing tensions and potential escalation dynamics.

**Military Planning and War Gaming**

The development of better war games, conflict scenarios, and training exercises is another area where AI can potentially be beneficial for nuclear deterrence by
helping decision-makers better understand potential nuclear-armed adversaries, manage escalation in a crisis, and bolster credibility and resolve regarding nuclear use if necessary. Western civilian and military decision-makers often exhibit biases inherent in exercises and analysis based on their own similar experiences, background, and thinking. War games are regularly used by the military to foster training and education as well as to refine operational plans, test various strategies, and identify areas for future capability development. However, military war games and training exercises have significant limitations—most notably, organization and confirmation biases from scenario developers and military strategists. In the most problematic cases, DoD will fund a game “with a prior idea about what they want the outcome of the game to be … to justify new authorities or new doctrine.”

The US military has recognized the limitations of traditional war game methodologies since the late 1970s and early 1980s and sought to use automation and AI to examine a wide range of scenarios, change underlying variables in a simulation, and explore more real-world situations. The use of AI and machine learning algorithms will help military planners better design and manage war-games, trainings, and exercises in a way that is more realistic and challenging, allowing participants to better prepare against adversaries that may be unpredictable, capable of thinking dynamically, and acting in ways that may be counterintuitive to Western military doctrine or principles. In the realm of nuclear deterrence, where it would be critical to understand adversary actions and send clear signals in order to control escalation and ultimately prevent all-out nuclear war, having war games and other tools that are unpredictable and developed from a variety of operational concepts, historical backgrounds, and diverse perspectives will be critical for creating successful decision-makers. Because there is so little history of nuclear use scenarios in real life, testing out strategies and concepts in a war game against intelligent and diverse competitors long before there is an actual crisis is even more important than in conventional warfare.

**Improving Perceptions of Capability and Resolve to Act**

These examples highlight only some of the ways AI and its subsets of machine learning, big data analytics, and predictive analysis might be able to increase confidence in nuclear systems, support better political resolve to act, and build a more secure nuclear deterrent.

As we think about the potential benefits of AI for nuclear deterrence policy and approaches, there are a number of other areas that might benefit from more reliance on AI. For instance, AI might be able to increase the security of nuclear systems—including elements of command and control networks—by
identifying unforeseen vulnerabilities or areas ripe for exploitation. The Defense Advanced Research Projects Agency (DARPA) is already looking into how AI can be used to identify vulnerabilities in conventional military systems, and it would make sense that the same methodologies could be used to detect and patch potential vulnerabilities in nuclear systems as well. In short, the benefits that AI may bring to nuclear deterrence can increase stability, credibility and security, and nuclear policymakers should seek to do more in this arena.

**AI Applications in Arms Control**

Arms control has historically been a tool for managing certain types of weapons or their applications, and it is one of several tools available to limit competition and increase stability between nuclear-armed adversaries. Discussions about AI and arms control typically veer toward attempts to manage the technology itself, such as the Campaign to Stop Killer Robots. Other efforts include the UN Convention on Certain Conventional Weapons (CCW) Group of Government Experts (GGE) on LAWS. The GGE’s chair, Amandeep Gill, observed in 2019 that “AI will have an impact on the conduct of warfare, bring new capabilities into being, and alter power equations. It will challenge traditional arms control thinking and raise questions about compliance with international law. New tools and competences will be required, and inventive ways to build trust and confidence among users will be needed.” The most common recommendation for AI arms control is to develop international norms, or rules of the road. These norms might include a commitment to always keep a “human in the loop” in nuclear command and control.

But there is another way to think of AI and arms control: rather than be the subject of an arms control agreement, AI could also provide solutions to many contemporary arms control challenges. Given Russia’s poor track record of arms control compliance, China’s lack of transparency, and concerns about nuclear and missile programs in North Korea and Iran, effective verification will be a major priority for future arms control and non-proliferation efforts. While arms control compliance is ultimately a political issue, technical solutions can provide opportunities to improve implementation, increase confidence, and enhance stability. To explore the potential benefits of AI to arms control, we look at two specific applications: object identification and pattern recognition.
Turning first to object identification, a major challenge for intelligence analysis is both sifting through large amounts of data and accurately identifying potential threats as outlined in several of the nuclear deterrence cases above. In terms of arms control, AI could be particularly beneficial, for example by improving object identification in satellite imagery of missile production facilities or test ranges. RAND has been exploring the use of AI to track mobile missiles, for example. This tracking would provide an additional verification tool for future strategic arms control agreements. For example, during the Senate ratification debate of the New START in 2010, many experts and policymakers alike expressed concern that the Treaty would not be able to verify production, deployment, or movement of new Russian mobile missiles.

Similarly, the work of open source intelligence experts, such as the Open Nuclear Network of One Earth Future, could be buttressed by applying AI technology that could recognize and analyze missile activity in North Korea. Arms control with North Korea still seems unlikely as the Kim regime has not expressed an interest in relinquishing its nuclear weapons, but in the meantime, improved information about North Korea’s nuclear activities and plans can improve situational awareness and understanding of their nuclear program and processes, potentially providing the groundwork for denuclearization and cooperative arms control at a later time. Human involvement would still be essential to provide important contextual information, but these developments might facilitate arms control verification by identifying treaty-prohibited items or activities in record time and provide convincing intelligence to the international community.

Pattern-recognition applications look for anomalies in massive amounts of data. Intelligence expert James L. Regens flags the particular value of AI in pattern recognition whereby it filters out anomalies in an actor’s behavior and is particularly valuable in “[eliminating] a large amount of noise to detect whether signals exist before even attempting to find the right ones.” The CIA already applies AI for various pattern-recognition tasks, such as geo-locating images without the associated metadata, creating 3-D models from satellite imagery, and inferring a building’s function based on “pattern-of-life” analysis. In the context of arms control, the use of AI for pattern recognition could be particularly useful to identify cheating behavior under an arms control agreement or assess the nature of new military activities that may undermine strategic stability.

These and other applications will contribute to national technical means (NTM) of verification. Non-interference in NTM for arms control verification protocols is a well-established tool of most major arms control treaties and is also used unilaterally by states to observe others’ activities and behaviors. With advances in AI, states may invest more heavily in NTM, which will have positive implications for future arms control agreements. First, it could give an advantage
to countries with more sophisticated AI and NTM capabilities. The United States, for example, would have an advantage over Russia in this area, which might make Russia more interested in traditional verification activities in future arms control agreements.41 Second, states will have to go to greater lengths to hide any cheating activities. And finally, it might reduce the need for on-site inspections, which could make participation in arms control agreements with “closed” societies like China, North Korea, and Russia more attractive than relying primarily on the current “boots on the ground” procedures. That said, even with enhanced AI monitoring, there remain important benefits of “boots on the ground” inspectors, which have proven to be particularly valuable both to observe compliance and build trust between parties to an arms control agreement. But monitoring particularly sensitive facilities with AI-enhanced NTM, like nuclear warhead storage sites, might be especially useful when nations have traditionally been reluctant to allow on-site inspections.42

Mitigating AI/Nuclear Risks

While there are clear benefits for nuclear deterrence and arms control from developments in AI, the continued integration of AI into nuclear-related systems does have risks that policymakers should understand and take steps to mitigate as technologies evolve. We believe that as policymakers seek to responsibly integrate AI into nuclear systems, they should also act now to prevent some of the most dangerous risks associated with AI to promote confidence and enhance awareness of the beneficial uses of AI.

While AI may provide additional benefit for many aspects of the nuclear deterrence mission, there should always be a “human in the loop” for targeting and employment of nuclear weapons. We cannot outsource nuclear decision-making.43 The tremendous consequences of any nuclear employment requires that accountable humans make the moral and ethical calculus to take the drastic step of nuclear use as “the burden of possible fearsome consequences looms over both sides, not one alone.”44 The principles of discrimination and proportionality under the law of armed conflict would be difficult for AI systems to properly evaluate, given the subjective and situation-specific considerations to determine if nuclear use is warranted. Furthermore, and particularly for Western democracies, it is essential that there is accountability for the use of nuclear weapons, given the gravity of such a decision. The Soviet “Dead
Hand” system, which was designed to automatically launch Soviet ICBMs if an incoming nuclear attack was detected, is one such automated and autonomous system. Without a “human in the loop,” that could have unintentionally led to all-out nuclear war.

Given that AI is not widely integrated into nuclear targeting and employment currently, nations with nuclear weapons should act now to codify these principles into international law. Similar to the 1967 Outer Space Treaty—which prohibits, inter alia, the stationing of nuclear weapons in space—the UN General Assembly could mandate the Conference on Disarmament (CD), which includes all nations with nuclear weapons, to negotiate a new treaty that outlines such a prohibition on the use of AI in nuclear employment. This would also serve to break the decades-long stalemate at the CD over the program of work. As a first step and until a formal treaty is in place, the P5 should affirmatively implement a moratorium on incorporating AI into nuclear employment decisions. But even with meaningful human control, as AI systems become more complex, they may also become more susceptible to accidents or failures, so meaningful human control alone is insufficient when considering the risks associated with AI.

Nuclear nations should therefore strongly consider banning or limiting fully autonomous, nuclear-armed weapons systems that cannot be recalled or redirected in a crisis. Several Russian systems in development today, such as its Poseidon nuclear-armed autonomous underwater drone, have similar automaticity challenges with redirection or recall once launched. The Poseidon system is designed to navigate autonomously and travel continuously after being launched from a submarine. Given the secrecy surrounding the program, it is unclear whether or not Russian command-and-control authorities will be able to systematically control and direct it once launched at sea. Additionally, because the Poseidon is entirely managed by software, there is a risk of hacking or malicious or incorrect code leaving the system vulnerable or unpredictable.

To mitigate this risk, nations should pledge to include the ability to alter or cancel a mission for any autonomous systems developed and/or include a “mission abort” feature that could be implemented after launch. This option would give decision-makers more flexibility, more time, and greater ability to control escalation. It would also prevent the accidental or unintended launch of nuclear systems either due to system failure or human error.

There are a number of risks associated with AI generally that will have even greater consequences for nuclear policy, if not managed. Issues such as deep fakes, image spoofing, cyber hacking, and the limited understanding of AI are risks—widely discussed in other literature—inherent to any future use of AI. In the nuclear domain, these challenges can significantly undermine
confidence in deterrence capabilities and decision-making as well as transparency over nuclear armaments necessary for arms control verification. Image spoofing and “deep fakes” are already being deployed by states, such as Russia, as well as non-state actors and bot networks. AI tools might struggle to differentiate real from fake images, for example, which could frustrate pattern recognition applications. Cyber hacking presents a similar challenge. It will be critically important that nuclear policymakers in Washington, Moscow, Beijing, and elsewhere understand these challenges and can take appropriate steps to mitigate the risks or at least be cognizant of the potential problems and implications.

This means that the nuclear community needs to put a premium on recruiting new talent with technical expertise and knowledge of AI, cyber systems, and the potential implications—both positive and negative—for nuclear policy. In particular, states with nuclear weapons should invest in educational initiatives that bridge the gaps between different technical domains and encourage scholarship that is multidimensional and forward-leaning. This investment should include partnering more closely with private companies that are driving innovation in AI and providing more incentives to recruit into government from the private sector. Furthermore, the P5 should consider adding the benefits and risks of emerging technologies such as AI as a standing topic on its annual agenda. The P5 should also seek to engage India and Pakistan in these discussions as well as NATO allies, in order to develop best practices and common understanding of threats and risks.

Other potential future risks for including AI in nuclear policy have been the subject of much scholarship over the past three to five years. However, we believe many of these risks are unlikely to manifest in the near term, and AI technology may never mature to the level that some of these authors claim. That is why we propose focusing on achievable and significant steps that can be taken now to improve nuclear policy and mitigate some of the most pressing challenges. These steps may be difficult to achieve in the current security environment, plagued by great power competition and opacity in many states’ strategic ambitions and capabilities, but they are highlighted here to further demonstrate the double-sided nature of AI: it is not necessarily the technology itself that is risky, but rather it depends on its application. Therefore, as nuclear nations look to integrate AI and related technologies into nuclear deterrence operations and strategic arms control, they will have an incentive to address and mitigate potential risks so as to take full advantage of the benefits of AI.

AI itself is not necessarily risky, but it depends on the application.
Notes


10. For a helpful discussion on the nature of AI competition, see Ben Buchanan, “The U.S. Has AI Competition All Wrong,” Foreign Affairs, August 7, 2020, https://www.foreignaffairs.com/articles/united-states/2020-08-07/us-has-ai-competition-all-wrong.


Jessica Cox and Heather Williams


48. See, Boulanin et al., Artificial Intelligence, Strategic Stability and Nuclear Risk, 38.


50. See, for example, Boulanin et al., Artificial Intelligence, Strategic Stability and Nuclear Risk; Johnson, “Artificial Intelligence in Nuclear Warfare”; Maas, “How Viable Is International Arms Control.”