Bipartisan Policy Center

CSET
CENTER *for* SECURITY *and*
EMERGING TECHNOLOGY

Artificial Intelligence and National Security

Artificial intelligence will have immense implications for national and international security, and AI's potential applications for defense and intelligence have been identified by the federal government as a major priority.

There are, however, significant bureaucratic and technical challenges to the adoption and scaling of AI across U.S. defense and intelligence organizations. Moreover, other nations—particularly China and Russia—are also investing in military AI applications. As the strategic competition intensifies, the pressure to deploy untested and poorly understood systems to gain competitive advantage could lead to accidents, failures, and unintended escalation.

The Bipartisan Policy Center and Georgetown University's Center for Security and Emerging Technology (CSET), in consultation with Reps. Robin Kelly (D-IL) and Will Hurd (R-TX), have worked with government officials, industry representatives, civil society advocates, and academics to better understand the major AI-related national and economic security issues the country faces. This paper hopes to shed more clarity on these challenges and provide actionable policy recommendations, to help guide a U.S. national strategy for AI. BPC's effort is primarily designed to complement the work done by the Obama and Trump administrations, including President Barack Obama's 2016 *The National Artificial Intelligence Research and Development Strategic Plan*,[i] President Donald Trump's Executive Order 13859, announcing the *American AI Initiative*,[ii] and the Office of Management and Budget's subsequent *Guidance for Regulation of Artificial Intelligence Applications*.[iii] The effort is also designed to further advance work done by Kelly and Hurd in their 2018 Committee on

June 2020

and Government Reform (Information Technology Subcommittee) white paper *Rise of the Machines: Artificial Intelligence and its Growing Impact on U.S. Policy.*[iv] Our goal through this effort is to provide the legislative branch with potential actions it can take to advance AI building on the work being done by the Trump administration.

I. Key Principles

1. Processes to develop and deploy defense and intelligence applications of AI systems must focus on human-machine teaming, trustworthiness, and implementing the DOD's Ethical Principles for AI.
2. The United States must work closely with allies and partners, while also seeking opportunities to cooperate selectively and pragmatically with competitors such as Russia and China.
3. The federal government should develop and refine metrics to evaluate foreign countries' AI sectors from both capabilities-based and conditions-based perspectives.
4. The federal government should invest in research, development, testing, and standardization in order to build and deploy more trustworthy cutting-edge AI systems.
5. Export and investment controls must be carefully targeted and strictly enforced in order to prevent the transfer of sensitive AI technologies to China.

II. Overview

The U.S. government has identified defense and intelligence as key areas where the United States must lead in the use of AI. In 2018, Congress stood up the National Security Commission on AI, which has released three reports to date on how the United States should be preparing for the national security implications of AI.[v]

The 2018 Department of Defense Artificial Intelligence Strategy posits that "the United States, together with its allies and partners, must adopt AI to maintain its strategic position, prevail on future battlefields, and safeguard [the free

and open international] order."[vi] Department of Defense AI efforts are being coordinated by the Joint Artificial Intelligence Center (JAIC), while the Army, Navy, Marines, and the Air Force have also stood up different task forces to advance their respective AI plans.

Similarly, the Intelligence Community's (IC) Augmenting Intelligence using Machines Strategy seeks to "secure and maintain strategic competitive information advantage for the IC through focused development and adoption" of artificial intelligence, process automation, and IC officer augmentation technologies.[vii]

II. Key Takeaways

PROCESSES TO DEVELOP AND DEPLOY DEFENSE AND INTELLIGENCE APPLICATIONS OF AI SYSTEMS MUST FOCUS ON HUMAN-MACHINE TEAMING, TRUSTWORTHINESS, AND IMPLEMENTING DOD'S ETHICAL PRINCIPLES FOR AI.

With the application of AI to defense and intelligence, the United States has an opportunity to secure a lasting competitive military advantage against its adversaries. Currently, DOD and IC AI research is underway in fields such as intelligence analysis, command and control, autonomous vehicles, logistics, weapons systems, and other areas.[viii]

Intelligence: AI may be particularly useful for intelligence because of the proliferation of sensors and the availability of large data sets AI algorithms can sift through to detect patterns, identify anomalies, and provide insights that improve situational awareness and support decision-making. Project Maven, for example, has demonstrated the value of using machine learning to support human analysts in analyzing large quantities of imagery from full-motion video collected by drones and identifying hostile activity during the counter-ISIL campaign.[ix] The speed and precision of AI-enabled intelligence analysis can provide U.S. forces an operational advantage against adversaries that do not possess similar capabilities.

Command and Control: AI can be used to collate and fuse information from various sensors and sources to develop a common operating picture for

decision-makers. AI-enabled command and control, as envisioned by programs such as the Defense Advanced Research Projects Agency's (DARPA) Mosaic Warfare and the Air Force's Multi-Domain Command and Control, can enhance the U.S. military's ability to coordinate forces and assets across domains. AI tools will become increasingly necessary for orchestrating effective human-machine teaming and coordinating between different intelligent agents and systems. Future AI systems will be able to provide decision-makers with courses of action based on real-time analysis. Such advances will help U.S. forces adapt to complex events and seize the initiative in high-stakes situations.

Logistics and Sustainment: Integration and scaling of enterprise AI applications to streamline back-office processes, personnel management (including troop rotations), equipment maintenance, and other logistics across DOD can help improve the functionality and longevity of military equipment; improve auditing and budgeting; increase efficiency and reduce costs. The Air Force, Army, and the JAIC in coordination with the Special Operations Command all have projects using AI-enabled approaches to predictive maintenance. The JAIC is also leading an effort to develop an AI-enabled flood and damage assessment to improve humanitarian response and reduce disaster impact. Recent assessments of DOD posture in AI suggest that enterprise AI applications present low-hanging fruit for DOD.[x] Yet, the challenges to scaling these efforts are not trivial.

> **Recommendation #1:** DOD should prioritize the integration and scaling of enterprise AI applications in logistics and sustainment as a matter of military readiness.
>
> **Recommendation #2:** The JAIC and each of the centralized AI service organizations should develop strategic plans outlining needs and relevant DOD-wide stakeholders to improve intraservice and interservice coordination and collaboration on the implementation of enterprise AI applications.

Autonomous Vehicles: The U.S. military is working to incorporate AI into a range of semi-autonomous and autonomous vehicles, including ground vehicles, naval vessels, fighter aircraft and drones.[xi] AI technologies in this space are used to perceive and map the environment, fuse sensor data, identify obstacles, plan navigation, and communicate with other vehicles. AI-enabled autonomous vehicles can help reduce risk to military personnel by undertaking dangerous and hazardous missions such as explosive ordnance disposal and route clearance. The U.S. Army, for instance, is interested in autonomous vehicle technology to reduce the number of service members needed to run resupply convoys in combat environments. While the technology for fully autonomous vehicles does not yet exist, different concepts for manned-unmanned teaming and human-AI collaboration are being developed across DOD. One example is Squad X, the DARPA experimental program that partners

infantry squads with AI and autonomous systems to make better decisions in complex and dynamic combat situations. In its most recent experiment, autonomous ground and aerial systems were used for sensing and surveillance to provide reconnaissance and improve situational awareness for infantry units moving through natural desert and mock city blocks.

The pace of progress in this field is uncertain. Commercial algorithms developed by the autonomous vehicles industry are not optimized for DOD uses. Additional challenges include the fragility and lack of robustness of algorithms, and differences between air, ground, and underwater combat environments.

> **Recommendation #3:** DOD should produce a report assessing existing data on realistic timelines and rationales for fielding AI-enabled autonomous capabilities in physical systems, from technology demonstration to testing and evaluation to deployment at scale.

Weapons Systems: Lethal autonomous weapons systems are an important and contested dimension of the public discourse about AI and national security. Many defense experts, for instance, argue that AI employed in automated, semiautonomous, or autonomous weapons systems can provide protection from incoming aircraft, missiles, rockets, artillery and mortar shells. The DOD AI strategy also posits that AI systems can help reduce the risk of civilian casualties and collateral damage by providing warfighters with greater situational awareness and enhanced decision support. Some technologists and ethicists, however, urge against using AI for military purposes, even calling for a preemptive ban on "killer robots."[xii] In February 2020, DOD adopted the five principles for ethical use of AI developed by the Defense Innovation Board which called for responsible, equitable, traceable, reliable, and governable AI for both combat and noncombat purposes.[xiii]

> **Recommendation #4:** DOD should continue working closely with industry and experts to develop and refine guidelines for implementing the ethical principles of AI throughout the entire life cycle of AI applications.

> **Recommendation #5:** Beyond existing outreach to industry and academia, DOD should assess the costs and benefits of a holistic communications strategy to engage and inform nontraditional audiences such as nongovernmental organizations, humanitarian groups, and civil society organizations to build public trust in DOD's commitment to ethical AI.

Across the aforementioned military functions and mission sets, AI tools and systems are being developed to augment human intelligence and allow for new forms of human-machine collaboration and teaming. Trust is critical for effective human-machine teaming. Yet, there is contested evidence about whether humans tend to not trust machines to perform effectively, especially in high-risk situations, or whether humans trust machines too much.

Recommendation #6: DOD should invest in research explicitly focused on trust in human-machine interactions, trust in AI systems, trust under stressful, dangerous, and high-stakes conditions in different domains, and how to train operators to place an appropriate level of trust in a given system.

DOD AI ethics principles establish that human beings are ultimately responsible for the development, use, and outcomes of AI systems. The issue of trust is therefore also pertinent to debates about safety and reliability standards as well as concerns about overreliance on AI systems leading to preventable mistakes and potentially catastrophic outcomes.[xiv]

Recommendation #7: DOD should consider incorporating an emphasis on safety, trustworthiness, and robustness across its human-machine collaboration and teaming research and development efforts.

Recommendation #8: The military law community should develop guidelines for how accountability and liability relating to the use of autonomous systems in warfare should be handled within the court-martial process.

China and Russia are also eager to field AI-enabled tools, weapons, and systems that will enhance their respective military capabilities and undermine U.S. technological and operational advantages. Balancing between the need to develop and adhere to AI safety standards and the urgency of remaining ahead of U.S. adversaries and competitors is not an easy feat.[xv] But leading in the responsible and ethical use and development of AI technologies is the best way to both realize AI's potential benefits for U.S. national security and limit the dangers posed by AI-enabled military systems for global stability.[xvi]

THE UNITED STATES MUST WORK CLOSELY WITH ALLIES AND PARTNERS, WHILE ALSO SEEKING OPPORTUNITIES TO COOPERATE SELECTIVELY AND PRAGMATICALLY WITH COMPETITORS SUCH AS RUSSIA AND CHINA.

As nations compete to gain relative advantage in AI, they will also need to cooperate to guard against potential dangers, mitigate risks, and realize the full benefits of AI for their citizens. Democratic nations face a two-fold challenge. On the one hand, they must deepen cooperation to improve the design and implementation of AI systems consistent with liberal democratic values. On

the other hand, they must find ways to cooperate selectively and pragmatically with competitors, such as with China and Russia.[xvii] Chief among the goals of international cooperation are to prevent unintentional use, lessen the risks of inadvertent escalation, and reduce the dangers of miscalculation and misperception involving AI-enabled systems and platforms.

Cooperation and competition are not mutually exclusive. It is important to specify the areas in which healthy competition fosters research breakthroughs for the common good and the areas in which cooperation is necessary to prevent accidents and manage escalatory risks. In doing so, democratic nations will need to balance the advantages of an open, stable international environment for scientific research with the imperative to protect sensitive technologies.[xviii]

To bolster cooperation on AI among democratic nations, the United States should leverage its alliances to promote democratic principles, foster research collaboration, and develop common standards.[xix] The Organization for Economic Cooperation and Development (OECD), which has developed a set of AI ethical principles, is one promising forum for international coordination on these issues.[xx]

Recommendation #9: The United States should take an active role in discussions of AI norms and standards in multilateral fora, including the UN, the OECD, and international standards-setting bodies.

Recommendation #10: The Department of State and the National Institute for Standards and Technology (NIST) should involve allies and partners in U.S. standards-setting initiatives related to AI, particularly NATO and EU allies, Japan, Australia, and South Korea, in order to ensure interoperability.

Recommendation #11: The National Science Foundation (NSF) should work with science funding organizations in allied countries to establish multilateral teams of AI researchers from the public and private sectors to promote talent development and foster partnerships on AI R&D.

Recommendation #12: DOD and the IC should work with allies to develop and exercise new plans and operational concepts for AI-enabled capabilities and systems, promote interoperability of military platforms and decision-making procedures, pool resources for cloud computing, and create centers of excellence for sharing nonsensitive data sets and developing common standards for test, evaluation, verification and validation.

Recommendation #13: To promote allied cooperation on national security related-AI, Congress, the White House and the Secretary of Defense should adopt the National Security Commission on AI's Q1 recommendations to expand and institutionalize AI-enabled warfighting and intelligence efforts. This includes creating a National Security Point of Contact and aligning AI adoption efforts starting with the Five Eye partners.

Recommendation #14: To promote multinational collaboration on AI R&D, the White House Office of Science and Technology Policy, in conjunction with Congress should:

- Organize and fund multinational innovation prize competitions. Such competitions could be modeled on DARPA's series of Challenges and the XPRIZE competitions, which have successfully tackled some of the toughest science and engineering problems, including in AI.
- Identify and fund opportunities for grants and loans to facilitate international personnel exchanges. Multilateral collaboration on AI would be particularly fruitful in areas such as AI safety and disease outbreak modeling.

Even as they improve cooperation, the United States and its allies and partners will need to selectively engage with Russia, China, and other competitors on AI safety and security.

Recommendation #15: The Department of State and U.S. allies should pursue carefully measured engagement with China and Russia to define shared concerns in AI safety and related concepts and terminology.

Recommendation #16: The Department of State and U.S. allies should promote track 1.5 and track 2 dialogues with government and nongovernment experts in China and Russia on AI safety, including exchanges to clarify relevant doctrines.

Recommendation #17: The Department of State and U.S. allies should explore with China and Russia the potential for confidence-building and crisis communications procedures to reduce the likelihood of unintentional use and mitigate the risks of escalation involving AI systems.

Given the uncertainties around AI and the rapid pace of change in this field, the United States and other major powers should not delay pragmatic action until a crisis emerges or accidents occur. The challenge is to shape the conditions of AI's safe and responsible development now, rather than wait for the day when these technologies are more advanced and more widely deployed. Policymakers will need to balance the competitive and cooperative pressures, while creating incentives for democracies to innovate and deepen collaboration in areas of mutual concern.

THE FEDERAL GOVERNMENT SHOULD DEVELOP AND REFINE METRICS TO EVALUATE FOREIGN COUNTRIES' AI SECTORS FROM BOTH CAPABILITIES-BASED AND CONDITIONS-BASED

PERSPECTIVES.

Major powers are competing to achieve strategic advantage in artificial intelligence. China spends several billion dollars a year to stimulate research in AI, expand its talent base, and assert technological leadership in the domain.[xxi] The European Union recently released its White Paper on AI.[xxii] More than 19 countries have formulated national AI strategies or plans, and the United States has put forward an Executive Order on maintaining American leadership in AI.[xxiii]

The flurry of AI strategies, plans, statements, and investments underscores the growing competitive pressures. Too often, however, the precise terms of competition are left under-specified. It is easy to assert a strategy of competition, but it is harder to define the means and ends of that competition. Toward what goal, with what tools, and over what timeframes are great powers competing in AI? How does one assess leadership in AI when the terrain is shifting so quickly? Which capabilities are most important for evaluating relative advantage?

Any assessment of global competitiveness in AI needs to recognize fundamental uncertainties. AI could exacerbate political tensions between democratic and authoritarian powers, but the policy and research communities need more refined metrics and methodologies for understanding how AI will relate to different regime types, or even whether regime type is the most salient characteristic. AI could upend the global economic and military balance of power, but estimates of its impact must be grounded in specific use cases and careful assessments of how countries and companies integrate AI into existing systems and platforms.

Evaluating global competitiveness in this field requires an appreciation of the core capabilities of AI and how those capabilities interact with the drivers of progress in science and technology. Broadly defined, there are two approaches to assessing global competitiveness in AI: a capabilities-based approach and a conditions-based approach.

Machine learning systems require data, sophisticated algorithms, and the computing power to run those algorithms so they can improve their performance. A capabilities-based approach focuses on such indicators as public and private sector funding for research and development, publications and patents, participation at top AI conferences such as the Neural Information Processing Systems annual meeting, and the semiconductor manufacturing equipment and fabrication plants needed for advanced chip design.

A conditions-based approach examines the innovation ecosystems and policy frameworks that enable the design, development, and deployment of AI capabilities. This approach considers the domestic pool of talent available to build modern AI systems, the support for faculty and curriculum design, and the educational systems necessary to produce high-end talent in computer science

and engineering. A conditions-based approach goes beyond raw quantities of data to consider the availability and accessibility of labeled data, the degree to which information is shared across public and private sector institutions, and the strength of data protection and privacy frameworks.

The private sector accounts for the majority of spending on AI. Whereas a capabilities-based approach looks at patents and publications, a conditions-based approach looks at national innovation ecosystems and the policy environments that support public-private partnerships in research and development. A capabilities-based approach focuses on semiconductor manufacturing capabilities; by contrast, a conditions-based approach focuses on the globalization of AI chip supply chains and the availability of test-range infrastructure and cloud computing for publicly funded research endeavors. A conditions-based approach also considers the degree of openness of a country and its integration into global research networks that are critical for staying at the forefront of advances in this field.

Both approaches are necessary to assess a country's relative competitiveness in AI. Data, algorithms, and computing power are the building blocks of AI systems. But those systems operate within policy and regulatory frameworks that can stymie or stimulate the adoption of AI at scale. Policymakers must formulate strategies with an eye toward strengthening the capabilities and conditions that will promote the security, economic prosperity, and core values of democratic societies.

Recommendation #18: The Intelligence Community should adopt and refine metrics for capabilities-based and conditions-based approaches to evaluating global competitiveness in AI.

Recommendation #19: The federal government should expand public-private partnerships, make available computing resources and testing infrastructure, and devise strategies for national R&D funding that leverage contributions from government, industry, academia, and philanthropy.

Recommendation #20: The federal government should establish a national science and technology analysis center to collect technical information, provide science and technology decision support, and disseminate relevant findings.

THE FEDERAL GOVERNMENT SHOULD INVEST IN RESEARCH, DEVELOPMENT, TESTING, AND STANDARDIZATION IN ORDER TO BUILD AND DEPLOY MORE TRUSTWORTHY CUTTING-EDGE AI SYSTEMS.

As progress in AI research shows promise in an increasing range of areas, defense and intelligence leaders naturally want to see it applied to the problems their organizations face. So far, however, a key limiting factor is that machine learning methods, which represent a large portion of modern AI research, are not reliable or secure enough for use in high-stakes defense and intelligence settings.

There are a number of reasons for this. Machine learning systems are currently brittle, malfunctioning in unexpected ways when given unfamiliar inputs, and opaque, processing information in ways that are very difficult even for experts to understand, let alone users.[xxiv] These properties make modern AI systems unsuitable for high-stakes defense and intelligence contexts in two ways: first, because they are vulnerable to intentional manipulation by adversaries, and second, because even in the absence of deliberate tampering they cannot be relied upon to handle novel situations appropriately. This is especially true of systems that use "deep learning," one of the most promising and most prevalent—but also most brittle and most opaque—types of machine learning.

High-stakes applications make up a substantial portion of potential use cases in defense and intelligence, including many listed above. A recent RAND study described three broad categories for defense applications of AI: "enterprise AI," "mission support AI," and "operational AI," in increasing order of how high-stakes failure would be and how uncontrolled the operating environment is.[xxv] Current machine learning systems are not at all suited for "operational" or even many "mission support" applications. The use of unreliable or insecure systems in critical settings could result in friendly fire, unintentional engagement or escalation, or other failures.[xxvi]

Recommendation #21: Defense and intelligence agencies should focus the development and deployment of machine learning systems on non-safety-critical settings, such as enterprise applications, until much higher standards of reliability, interpretability, and security can be achieved.

For previous generations of computer systems, DOD has developed a robust set of processes and tools for testing, evaluation, validation & verification (TEVV) to ensure the performance and reliability of any system before it is fielded. Unfortunately, these processes are poorly suited to machine learning systems. Unlike earlier AI paradigms, machine learning systems are not built around human-specified rules or procedures; instead, they process information using inscrutable numerical functions learned from data. Simplifying somewhat, the number of potential outputs a given system could produce is too large to be tested using established TEVV paradigms, which are based around exhaustive, up-front evaluation.

At present, the vacuum of appropriate TEVV methods is being filled by informal ad hoc testing without established best practices. This status quo will not be sufficient to reach a future of reliable, secure machine learning

systems. New paradigms and processes that take the unique challenges of machine learning systems into account are needed, such as the "CD/CI/CV" (continuous development, continuous integration, continuous verification and validation) approach advocated by DARPA's Information Innovation Office.[xxvii] Efforts to develop these paradigms and processes should be coordinated by an organization that can also draw on expertise in academia and the private sector, such as the National Institute of Standards and Technology.

> **Recommendation #22:** NIST should be resourced to lead and coordinate federal government efforts to develop best practices around TEVV for machine learning.

> **Recommendation #23:** Funding should be directed to the National Science Foundation and the Department of Energy to support basic research into the underlying science of TEVV for machine learning.

The United States must develop and implement the concepts, frameworks, and processes needed to assure AI for safety-critical settings. If successful, this will unlock whole new frontiers of how AI could be utilized. If unsuccessful, the likely result is accidents, failures, and sabotage.

EXPORT AND INVESTMENT CONTROLS MUST BE CAREFULLY TARGETED AND STRICTLY ENFORCED IN ORDER TO PREVENT THE TRANSFER OF SENSITIVE AI TECHNOLOGIES TO CHINA.

Deliberate efforts by China to acquire and assimilate U.S. technologies are nothing new. Export controls and investment controls are two key policy tools for preventing this, and in recent years Congress has passed legislation updating both types of controls in order to deal with new challenges associated with emerging technologies: the Export Control Reform Act of 2018 (ECRA), and the Foreign Investment Risk Review Modernization Act of 2018 (FIRRMA). ECRA directed the Department of Commerce to develop new controls for emerging and foundational technologies. FIRRMA, in turn, modernized the Committee on Foreign Investment in the Unites States, one of the primary mechanisms that can prevent Chinese strategic investments in technology companies, by reviewing investment transactions involving companies that work on sensitive technologies such as AI.

ECRA and FIRRMA represent significant progress in preventing the transfer of sensitive technologies to China and other countries. One key area where additional action is needed, however, is in cutting-edge computer chips and the equipment used to manufacture them. State-of-the-art, specialized computer

chips are necessary to develop and deploy most advanced AI systems. For AI applications, a state-of-the-art, specialized AI chip can be over 10,000 times faster and more efficient than an older generation, unspecialized computer chip.[xxviii] With the computational costs for developing the most advanced AI systems running as high as tens of millions of dollars even with state-of-the-art, specialized AI chips, producing advanced systems is infeasible without them.

The small number of firms that can produce state-of-the-art chips are currently based in the United States, South Korea, and Taiwan. This creates an opportunity for these governments to synchronize in imposing end use and end user export controls on these chips to prevent misuse of AI by authoritarian actors.

> **Recommendation #24:** The Department of State and the Department of Commerce should coordinate closely with allies and partners, especially Taiwan and South Korea, to synchronize their export control regimes with existing U.S. export controls on advanced AI chips for Chinese and Russian military end uses and end users.

Partially to circumvent these types of export controls on computer chips, China is investing tens of billions of dollars to build up its own state-of-the-art chip production capacity. It has had some limited success with this. However, this success was only possible because China could import the chip manufacturing equipment it needed for its production facilities from abroad. China is currently unable to build the manufacturing tools necessary to produce these chips.[xxix] The large majority of the most specialized chip manufacturing equipment is produced in just three countries: the United States, the Netherlands, and Japan.

> **Recommendation #25:** The Department of State and the Department of Commerce should coordinate closely with Japan and the Netherlands to apply multilateral export controls for chip manufacturing equipment used to produce chips with feature sizes at or below the size thresholds listed in the Wassenaar Arrangement and U.S. Commerce Control List (CCL). This regime should include a presumptive denial of export licenses.

The threshold currently listed in the Wassenaar Arrangement and CCL for lithography equipment, a key type of chip manufacturing equipment, is 45 nanometers. If implemented consistently, this would restrict China to training and deploying AI systems with 65 nm chips, which would cost 25 times more than using state-of-the-art 5 nm chips.[xxx] Presumptive license denial is recommended because, despite the Wassenaar Arrangement and CCL's current 45 nm thresholds, most licenses have been approved under the current case-by-case licensing policy, allowing China to achieve production capacity as advanced as 14 nm. Notably, a presumptive denial policy would not necessarily reduce revenue for the companies involved. It would merely shift chip production, and therefore equipment sales, to countries outside China, which would then sell their chips to Chinese customers.

Conclusion

The federal government has rightly identified artificial intelligence as a key technology for U.S. national security and competitiveness in the decades to come. Congress has a vital role to play in coordinating a strategic approach to research, development, integration, and scaling of AI across the relevant agencies and departments. As DOD and the IC pursue ways to use AI to further their missions, considerations of trustworthiness, reliability, and ethics must be front and center. The Department of State should coordinate closely with allies and partners to develop standards, promote shared R&D, and ensure interoperability, in order to maintain U.S. leadership in AI. Concurrently, carefully measured engagement with Russia and China on issues such as AI safety, crisis communications, and escalation management will also be critical. In an increasingly complex global environment, the integration of AI into defense and intelligence will play a critical role in national security and international economic competition.

Endnotes

i Executive Office of the President. *The National Artificial Intelligence Research and Development Strategic Plan*, October 2016. https://www.nitrd.gov/pubs/national_ai_rd_strategic_plan.pdf

ii Executive Office of the President. *Executive Order 13859 Maintaining American Leadership in Artificial Intelligence*, Pub. L. No. 2019–02544, 84 FR 3967 E.O. 12859 3967 (2019). https://www.federalregister.gov/d/2019-02544.

iii Vought, Russell T., *Memorandum for the Heads of Executive Departments and Agencies: Guidance for Regulation of Artificial Intelligence Application*, January 7, 2020, https://www.whitehouse.gov/wp-content/uploads/2020/01/Draft-OMB-Memo-on-Regulation-of-AI-1-7-19.pdf.

iv Hurd, Will, and Robin Kelly. "Rise of the Machines: Artificial Intelligence and Its Growing Impact on U.S." Subcommittee on Information Technology, Committee on Oversight and Government Reform: U.S. House of Representatives, Sep 2018. https://hurd.house.gov/sites/hurd.house.gov/files/AI%20White%20Paper%20Clean.pdf.

v National Security Commission on Artificial Intelligence. "Reports," 2020. https://www.nscai.gov/reports.

vi U.S. Department of Defense. "Summary of the 2018 Department of Defense Artificial Intelligence Strategy: Harnessing AI to Advance Our Security and Prosperity." 2018. https://media.defense.gov/2019/Feb/12/2002088963/-1/-1/1/SUMMARY-OF-DOD-AI-STRATEGY.PDF.

vii "The AIM Initiative: A Strategy for Augmenting Intelligence Using Machines." Washington, DC: Director of National Intelligence, 2019. https://www.dni.gov/files/ODNI/documents/AIM-Strategy.pdf.

viii Sayler, Kelley M., and Daniel S. Hoadley. "Artificial Intelligence and National Security." Washington, DC: Congressional Research Services, November 21, 2019. https://fas.org/sgp/crs/natsec/R45178.pdf.

ix Allen, Gregory C. "Project Maven Brings AI to the Fight against ISIS." *Bulletin of the Atomic Scientists*, December 21, 2017. https://thebulletin.org/2017/12/project-maven-brings-ai-to-the-fight-against-isis/.

x Tarraf, Danielle C., William Shelton, Edward Parker, Brien Alkire, Diana Gehlhaus, Justin Grana, Alexis Levedahl, Jasmin Leveille, Jared Mondschein, James Ryseff, Ali Wyne, Dan Elinoff, Edward Geist, Benjamin N. Harris, Eric Hui, Cedric Kenney, Sydne Newberry, Chandler Sachs, Peter Schirmer, Danielle Schlang, Victoria M. Smith, Abbie Tingstad, Padmaja Vedula, and Kristin Warren. "The Department of Defense Posture for Artificial Intelligence: Assessment and Recommendations." Santa Monica, CA: RAND Corporation, 2019. *https://www.rand.org/pubs/research_reports/RR4229.html*.

xi Feickert, Andrew, Jennifer K. Elsea, Lawrence Kapp, and Laurie A. Harris. "U.S. Ground Forces Robotics and Autonomous Systems (RAS) and Artificial Intelligence (AI): Considerations for Congress." Washington, DC: Congressional Research Services, November

20, 2018. https://fas.org/sgp/crs/weapons/R45392.pdf.

Martin, Bradley, Danielle C. Tarraf, Thomas C. Whitmore, Jacob DeWeese, Cedric Kenney, Jon Schmid, and Paul DeLuca. "Advancing Autonomous Systems: An Analysis of Current and Future Technology for Unmanned Maritime Vehicles." Santa Monica, CA: RAND Corporation, 2019. https://www.rand.org/pubs/research_reports/RR 2751.html

xii Stop Killer Robots. "Tech Workers: The World Needs You," n.d. https://www.stopkillerrobots.org/tech/.

xiii U.S. Department of Defense. "DOD Adopts Ethical Principles for Artificial Intelligence." February 24, 2020. https://www.defense.gov/Newsroom/Releases/Release/Article/2091996/dod-adopts-ethical-principles-for-artificial-intelligence/.

xiv Danzig, Richard. "Technology Roulette: Managing Loss of Control as Many Militaries Pursue Technological Superiority." Washington, DC: Center for a New American Security, May 30, 2018. https://www.cnas.org/publications/reports/technology-roulette.

xv Imbrie, Andrew, and Elsa B. Kania. "AI Safety, Security, and Stability Among Great Powers: Options, Challenges, and Lessons Learned for Pragmatic Engagement." Washington, DC: Georgetown University, Center for Security and Emerging Technology, December 2019. https://cset.georgetown.edu/wp-content/uploads/AI-Safety-Security-and-Stability-Among-the-Great-Powers.pdf.

xvi National Security Commission on Artificial Intelligence. "AI Commission Interim Report." 2019. https://www.epic.org/foia/epic-v-ai-commission/AI-Commission-Interim-Report-Nov-2019.pdf.

xvii Imbrie, Andrew, and Elsa B. Kania. 2019. "AI Safety, Security, and Stability Among Great Powers: Options, Challenges, and Lessons Learned for Pragmatic Engagement." Washington, DC: Georgetown University, Center for Security and Emerging Technology. https://cset.georgetown.edu/wp-content/uploads/AI-Safety-Security-and-Stability-Among-the-Great-Powers.pdf.

xviii Hannas, Wm. C., and Huey-meei Chang. 2019. "China's Access to Foreign AI Technology: An Assessment." Washington, DC: Georgetown University, Center for Security and Emerging Technology. https://cset.georgetown.edu/wp-content/uploads/CSET_China_Access_To_Foreign_AI_Technology.pdf.

xix Imbrie, Andrew, Ryan Fedasiuk, Catherine Aiken, Tarun Chhabra, and Husanjot Chahal. 2020. "Agile Alliances: How The United States and Its Allies Can Deliver A Democratic Way of AI." Washington, DC: Georgetown University, Center for Security and Emerging Technology. https://cset.georgetown.edu/wp-content/uploads/CSET-Agile-Alliances.pdf.

xx OECD. "OECD AI Policy Observatory." n.d. https://oecd.ai/.

xxi Acharya, Ashwin, and Zachary Arnold. 2019. "Chinese Public AI R&D Spending: Provisional Findings." Washington, DC: Georgetown University, Center for Security and Emerging Technology. https://cset.georgetown.edu/wp-content/uploads/Chinese-Public-AI-RD-Spending-Provisional-Findings-2.pdf.

xxii "White Paper on Artificial Intelligence - A European Approach to Excellence and Trust." 2020. COM(2020) 65 Final. Brussels:

European Commission. https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf.

xxiii Campbell, Thomas A. 2019. "Artificial Intelligence: An Overview of State Initiatives." Evergreen, CO: FutureGrasp. http://www.unicri.it/in_focus/files/Report_AI-An_Overview_of_State_Initiatives_FutureGrasp_7-23-19.pdf.
Executive Office of the President. Executive Order 13859 Maintaining American Leadership in Artificial Intelligence, Pub. L. No. 2019–02544, 84 FR 3967 E.O. 12859 3967 (2019). https://www.federalregister.gov/d/2019-02544.

xxiv Yuan, Xiaoyong, Pan He, Qile Zhu, and Xiaolin Li. 2018. "Adversarial Examples: Attacks and Defenses for Deep Learning." Cornell University, July. https://arxiv.org/abs/1712.07107.
"The Building Blocks of Interpretability." n.d. Distill. https://distill.pub/2018/building-blocks/.

xxv Tarraf, Danielle C., William Shelton, Edward Parker, Brien Alkire, Diana Gehlhaus, Justin Grana, Alexis Levedahl, Jasmin Leveille, Jared Mondschein, James Ryseff, Ali Wyne, Dan Elinoff, Edward Geist, Benjamin N. Harris, Eric Hui, Cedric Kenney, Sydne Newberry, Chandler Sachs, Peter Schirmer, Danielle Schlang, Victoria M. Smith, Abbie Tingstad, Padmaja Vedula, and Kristin Warren, "The Department of Defense Posture for Artificial Intelligence: Assessment and Recommendations." Santa Monica, CA: RAND Corporation, 2019. https://www.rand.org/pubs/research_reports/RR4229.html.

xxvi Scharre, Paul. 2016. "Autonomous Weapons and Operational Risk." Washington, DC: Center for a New American Security. https://www.cnas.org/publications/reports/autonomous-weapons-and-operational-risk.

xxvii Defense Advanced Research Projects Agency. "Information Innovation Office." n.d. https://www.darpa.mil/about-us/offices/i2o/more. See also recommendation #4 of Defense Science Board, "Report of the Defense Science Board Summer Study on Autonomy." June 2016. https://apps.dtic.mil/dtic/tr/fulltext/u2/1017790.pdf.

xxviii Khan, Saif M. 2020. "AI Chips: What They Are and Why They Matter." Washington, DC: Georgetown University, Center for Security and Emerging Technology. https://cset.georgetown.edu/research/ai-chips-what-they-are-and-why-they-matter/.

xxix Flynn, Carrick. 2020. "Recommendations on Export Controls for Artificial Intelligence." Washington, DC: Georgetown University, Center for Security and Emerging Technology. https://cset.georgetown.edu/wp-content/uploads/Recommendations-on-Export-Controls-for-Artificial-Intelligence.pdf.

xxx Khan, Saif M. 2020. "AI Chips: What They Are and Why They Matter." Washington, DC: Georgetown University, Center for Security and Emerging Technology. https://cset.georgetown.edu/research/ai-chips-what-they-are-and-why-they-matter/.


Bipartisan Policy Center


CSET
CENTER for SECURITY and
EMERGING TECHNOLOGY