Technical Report

Artificial Intelligence and Strategic Trade Controls

Strategic Trade Research Institute

Center for International and Security Studies at Maryland

June 2020

SCHOOL OF
PUBLIC POLICY
CENTER FOR INTERNATIONAL &
SECURITY STUDIES AT MARYLAND

The authors of this report invite liberal use of the information provided, requiring only that the reproduced material clearly cite the source, using: Andrea Viski, Scott Jones, Lindsay Rand, Tucker Boyce, and Jonas Siegel, "Artificial Intelligence and Strategic Trade Controls," Strategic Trade Research Institute and Center for International and Security Studies at Maryland, June 2020.

Report Design and Layout by Andrea Viski

Printed in the United States of America

Acknowledgments

The authors would like to sincerely thank the participants of the dialogue on Emerging Technologies and Strategic Trade Controls held at the Stimson Center in Washington DC on March 14, 2019, as well as the participants of the dialogue on Artificial Intelligence and Strategic Trade Controls held at the Ronald Reagan International Trade Center on March 9, 2020. The authors would also like to thank Amy Nelson, Kevin Wolf, Carl Wocke, Aaron Mannes, Timothy Gildea, Aaron Arnold, Todd Perry, Nancy Gallagher, and Richard Cupitt for their support, comments, ideas, and feedback on this report.

About the Strategic Trade Research Institute

The Strategic Trade Research Institute was founded in 2017 and is an independent, international, board-governed non-profit organization dedicated to building networks of strategic trade research and practice through leadership, research, and innovation. STRI publishes the *Strategic Trade Review*, the leading peer reviewed journal dedicated to trade and security. STRI engages in quality research and capacity-building and is committed to promoting a diversity of voices and perspectives. To learn more about STRI, visit www.strategictraderesearch.org.

About the Center for International and Security Studies at Maryland

The Center for International and Security Studies at Maryland (CISSM) at the University of Maryland's School of Public Policy conducts research, education, and outreach about how powerful trends associated with globalization are affecting international security. It focuses on strategies to increase international cooperation, especially where powerful technologies—with both beneficial and dangerous uses—are becoming widely available to states and non-state actors. To learn more about CISSM, visit www.cissm.umd.edu.

Author Biographies

Andrea Viski

Andrea Viski founded the *Strategic Trade Review* in 2015 and has since served as its Editor-in-Chief. She founded the Strategic Trade Research Institute (STRI) in 2017, where she serves as Director. She cooperates with numerous organizations providing expertise and capacity for research, training, and project implementation. She is an adjunct professor at the Schar School of Policy and Government at George Mason University where she teaches a course on strategic trade controls and helped develop GMU's Master's Certificate of Strategic Trade. Viski is also a Research Associate at the Center for International and Security Studies (CISSM) at the University of Maryland and Nonresident Senior Fellow at the University of Georgia's Center for International Trade and Security (CITS). She has published extensively in the areas of export controls, nonproliferation of Weapons of Mass Destruction, trade sanctions, nuclear security, and international law.

Scott Jones

Scott Jones is Senior Advisor at the Strategic Trade Research Institute, Nonresident Fellow at the Stimson Center, Partner at TradeSecure, LLC., and an Affiliated Expert at CRDF Global. His areas of expertise include export controls and sanctions, international trade and investment policy, and emerging technologies. Previously, Dr. Jones served as Director at the University of Georgia's Center for International Trade and Security and as a foreign affairs analyst for the National Nuclear Security Administration and Los Alamos and Oak Ridge National Laboratories. His current activities and interests include United States and Chinese security and defense policy, the socioeconomic impact of exponential technology, foreign direct investment and national security, and the technology drivers of contemporary geopolitics.

Lindsay Rand

Lindsay Rand is a PhD student at the University of Maryland School of Public Policy and a Graduate Research Assistant at the Center for International and Security Studies at Maryland (CISSM). Rand's research is focused on the intersection of science and policy in the field of international security. Her doctoral research examines the role of science and technology in verification approaches for arms control agreements. Rand received an M.S. in nuclear health physics from Georgetown University, where her technical research included assessments of radiation detectors for the U.S. Navy and FEMA and the development of a lightweight radiation detection robot. Rand has a B.A. in physics and classical history from Carleton College.

Tucker Boyce

Tucker Boyce recently completed his Master of Public Policy program at the University of Maryland, where he has been a Graduate Assistant at the Center for International and Security Studies at Maryland. During graduate school, Tucker focused on international security topics and completed a capstone project on nonproliferation assistance in partnership with the Stimson Center. Prior to graduate school, Tucker supported an export control outreach program as post-bachelor's appointee at Los Alamos National Laboratory.

Jonas Siegel

Jonas Siegel is the Associate Director of the Center for International and Security Studies at Maryland (CISSM). Siegel leads research initiatives related to the Center's Nuclear Past, Present, and Future project, including efforts to evaluate the impact of nuclear energy trends on nuclear nonproliferation and nuclear security in East Asia; to identify ways to strengthen global nuclear governance systems; to analyze the nuclear proliferation implications of advanced nuclear energy developments; to understand the impact of nuclear materials transparency on security and proliferation; to develop minimum requirements for a comprehensive global nuclear material accounting system; and to investigate media coverage of Iran's nuclear program and the international response to it.

Executive Summary

Balancing the benefits and risks posed by artificial intelligence (AI), one of the most diffuse and rapidly evolving emerging technologies, is imperative when forming sound policy. This report analyzes the threats, trade linkages and mechanisms, and policy options in light of ongoing discussions regarding the prospects for applying export controls on artificial intelligence technologies and applications.

Using open source research, findings from organized dialogues, and expert interviews, the report authors identified policy options that go beyond export controls and encompass a coordinated, comprehensive, and technical approach to garnering the many benefits of artificial intelligence while mitigating its security risks. These approaches take into account both traditional nonproliferation strategies and ongoing debates concerning national security and economic competitiveness. Urgent, cross-sector action by governments and nongovernmental entities, including exporters, technology developers, academia, and civil society, is necessary to activate cooperative tools that mitigate the risks posed by AI. Lessons learned from strategic trade approaches to AI can be replicated, in certain situations, to other emerging technologies.

Risks

- The report authors organize AI-related risks according to risks to direct national security infrastructure, as well as indirect risks that stem from the development of the technology and nature of innovation;
- Risks to national security infrastructure include the potential of AI as an augmentation system for automation (decision-making and command and control), cyber capabilities, information and surveillance, and physical production;
- Indirect risks include asymmetric research and development progress, piecewise AI-relevant security measure implementation, poor or mismatched regulations over borders, and effects on norms,

governance, and credibility.

AI and Strategic Trade Controls

- The report analyzes six possible pathways to control AI-related transfers: software, data, computing power and associated hardware, services and deemed exports, end-use/end-user controls, and catch-all controls;

- U.S. policy with regards to developing, integrating, and applying AI is described in the report and determined to be increasingly implemented against a backdrop of concerns regarding security and economic risks;

- The report finds that early list-based export control policy efforts, particularly those justified by economic competitiveness arguments, are likely to be ineffective in most situations and could make it more difficult to mitigate whatever risks AI presents. In light of this, the report reevaluates the conceptual basis for imposing controls on emerging technologies where their dangers and military end-uses are not yet known.

Based on these analyses and findings, the authors assess and where appropriate recommend the following policy options:

- Outreach

- Interagency coordination and information-sharing

- Enforcement and licensing of catch-all

- Investment controls

- Development and implementation of research criteria

- Development of norms

- Private sector self-policing/role of competition

- Targeting intangible transfers of technology (ITT)

- Technology tracking

Table of Contents

Section I: Introduction

Section II: Technology Overview

Section III: AI/ML and Strategic Trade Controls

Section IV: Policy Options

This page intentionally left blank

I. Introduction

I.I Emerging Technologies and Strategic Trade Controls in Context

Technological superiority is often a determining factor of national security and economic competitiveness for major world powers. During periods of rapid technological change, a country's ability to stay at the forefront of research and development can determine its access to critical military power as well as control of markets. While the power established by technological advantage has always been a reality, exponential advances in specific technologies, whether building on existing capabilities or creating altogether new ones, has characterized the beginning of the 21st century and has redefined the scope of trade, security, and power.

In the United States, one dimension of maintaining an edge over adversaries in technological development has been to manage access to and development of certain emerging technologies through a combination of strategic trade management tools. This has included tightening investment controls on specific technological areas through the Foreign Investment Risk Review Modernization Act of 2018 (FIRRMA).[1] In addition, as part of its export control reform, the Department of Commerce published an Advanced Notice on Proposed Rulemaking (ANPRM) in 2018, seeking input from strategic trade stakeholders into potential new entries to be added to the United States control list on certain emerging technologies.[2]

The technologies at the forefront of United States strategic trade management efforts as delineated in the ANPRM comprise diverse and often overlapping areas:

1. Biotechnology
2. Artificial intelligence (AI) and machine learning
3. Position, navigation, and timing technology
4. Microprocessor technology

[1] "The Committee on Foreign Investment in the United States," United States Department of Treasury, <https://home.treasury.gov/policy-issues/international/the-committee-on-foreign-investment-in-the-united-states-cfius>.

[2] "Review of Controls for Certain Emerging Technologies," Bureau of Industry and Security, Department of Commerce, Washington, DC, October 11, 2018, <https://www.federalregister.gov/documents/2018/11/19/2018-25221/review-of-controls-for-certain-emerging-technologies>.

5. Advanced computing technology
6. Data analytics technology
7. Quantum information and sensing technology
8. Logistics technology
9. Additive manufacturing (e.g., 3D printing)
10. Robotics
11. Brain-computer interfaces
12. Hypersonics
13. Advanced materials
14. Advanced surveillance technologies

As the United States began to consider trade controls as a tool to manage emerging technologies, other countries also began to consider doing the same, whether through investment controls or unilateral controls on certain groups of technologies. The European Union (EU), in May 2019, adopted Regulation 2019/452 establishing a framework for the screening of foreign direct investments (FDI) and subsequent guidance on implementation of the regulation in March 2020.[3] In November 2019, the Japanese Diet passed an amendment to their Foreign Exchange and Foreign Trade Act (FEFTA) introducing new, more stringent controls on foreign investment.[4] While most countries already have some form of controls on FDI, many have chosen to tighten these laws over the last several years.[5] For years, the EU, individual EU Member States, and other countries have also been analyzing groups of technologies, such as additive manufacturing, to determine whether there is a basis for control in the multilateral export control regimes or on a state-level basis.

Using trade controls to manage the spread and use of new technologies is not an original development - as several experts have written, a number of attempts have been made in the past, in the United States and in other countries, to explore ways in which controls can be administered to new technologies that are still emerging to the extent that their potential

[3] "Regulation (EU) 2019/452 of the European Parliament and of the Council of 19 March 2019 Establishing a Framework for the Screening of Foreign Direct Investments into the Union," <https://eur-lex.europa.eu/eli/reg/2019/452/oj>.

[4] Sakon Kuramoto, Benjamin Miller, Hiroki Sugita, "Amendment to Japanese Foreign Exchange and Foreign Trade Act Regulations Expands Scope of "Restricted Businesses" to Include Some Information and Communications Technology Businesses," JD Supra, June 22, 2019, <https://www.jdsupra.com/legalnews/amendment-to-japanese-foreign-exchange-68547/>.

[5] For a full list of FDI legislation worldwide, see the Investment Policy Hub's website: <https://investmentpolicy.unctad.org/investment-laws>.

military end-uses and/or risks are not yet concretely established. Given that attempts to implement unilateral controls on new technologies in the United States are ostensibly rooted in the aim of establishing new entries in the control lists of multilateral export control regimes, so far attempts to do so have not been entirely successful with regards to new technologies whose conventional or WMD end-use is not clear or directly tied to a security threat.

For example, in 2013, the United Kingdom and France succeeded in passing a proposal in the Wassenaar Arrangement (WA), the export control regime dedicated to controlling conventional arms and dual-use goods and technologies, to control intrusion software and IP network communications surveillance systems.[6] As the United States tried to implement the new controls nationally, vehement industry opposition via comments on a public notice of the new rules, lobbying, and letters forced the United States to withdraw the controls it had proposed and implemented in its national legislation. The United States then renegotiated the controls within the WA in 2017, resulting in many more exemptions and narrower control of such technology. The final controls on intrusion software in the WA now reflect those negotiations.

The case of additive manufacturing (AM) is another useful example - while the WA introduced a control on a specific type of AM production equipment: "directional-solidification or single-crystal additive manufacturing equipment for the production of gas turbine engine blades, vanes and tip shrouds, as well as the associated software," the control was introduced more to "ensure coverage of equivalent technologies to prevent substitution for other already controlled production equipment," as noted by Kelley and Brockmann in 2018.[7] Other attempts to introduce controls on AM production equipment in the Missile Technology Control Regime (MTCR) in 2014 and in the Nuclear Suppliers Group (NSG) in 2016 did not succeed. While discussions continue in the various multilateral export control regimes about whether to introduce separate, specific controls on, for example, feedstock for AM printers or controls on technology transfer, regime

[6] Mark Bromley, Kees Jan Steenhoek, Simone Halink and Evelien Wijkstra, "ICT Surveillance Systems: Trade Policy, and the Application of Human Security Concerns," *Strategic Trade Review* (Vol. 2, No. 1), Spring 2016, <https://strategictraderesearch.org/wp-content/uploads/2017/11/STR_02.pdf>.

[7] Kolja Brockmann and Robert Kelley,"The Challenge of Emerging Technologies to Non-proliferation Efforts: Controlling Additive Manufacturing and Intangible Transfers of Technology," SIPRI, April 2018, <https://www.sipri.org/publications/2018/other-publications/challenge-emerging-technologies-non-proliferation-efforts-controlling-additive-manufacturing-and>.

members have not adopted new controls.[8]

These two examples highlight the difficulty with introducing new controls on emerging technologies in both the national and multilateral contexts. This challenge of introducing new controls is key to consider in more depth as the dilemma of controlling technologies whose military end-use is not yet crystallized, nor security threat and risk clearly established, contends with the very concept of "threat" and "security" – concepts that are at the crux of why certain materials, equipment, and technology are controlled at the multilateral level to begin with.

Considering the interplay between emerging technologies and strategic trade controls therefore may magnify deeper conceptual cracks in the nature, objectives, and use of controls in the modern security environment. Are these technologies being controlled, indeed, to keep certain conventional weapons and WMD out of the hands of "bad" actors? Or are they being controlled with the aim of developing an edge on the development and use of certain technologies? In the past, the answer was the former, not the latter. In the current environment, this has changed, and in fact while the expected and rather banal answer would be both, the answer as it appears to be forming from recent policy decisions over the last few years, at least in the United States, could be the latter.

One of the findings of this report is that attempts to control emerging technologies, and in particular artificial intelligence/machine learning (hereafter referred to as AI/ML), highlight deeper issues with using list-based export control solutions to ensure economic competitiveness. Indeed, this report finds that such measures are not only likely to be ineffective but that they also have the potential to make it harder to limit the malicious use of emerging technologies.

I.II Report Objectives and Structure

Certain technologies such as additive manufacturing (AM) have been analyzed in depth by researchers, policymakers, and industry from as far back as 2014 regarding their export control ramifications. Initial concerns over "3D printing the bomb" led to efforts to introduce control

[8] Ibid.

list classifications for certain related items in export control regimes and ultimately to the current effort by the United States Department of Commerce to determine what, if any, control list classifications could be implemented. However, to date, while the potential malicious uses of additive manufacturing technologies have been exhaustively analyzed, no country has thus far implemented comprehensive export controls on this technology. Other technology groups listed in the ANPRM, such as advanced surveillance technologies, have similarly been analyzed in varying contexts depending on not just strictly security, but human rights related end-use concerns.

The objective of this report is to focus on one technological area, artificial intelligence and machine learning, and investigate whether and how strategic trade control measures could be used to limit the potential malicious risks of these technologies. By looking closely at one particular technological area, the report also will highlight some of the qualitative differences in applying strategic trade controls on emerging technologies in general, taking into account the current security, economic, and political environment. The risks and threats section categorizes direct risks from AI/ML platforms and places these risks into this broader national security and economic context.

This involves outlining ways that AI/ML systems could augment and complement existing conventional and Weapons of Mass Destruction (WMD) platforms as well as putting AI risks in the context of broader national security and economic competitiveness concerns. The risks and threats section categorizes direct and indirect risks from AI/ML platforms and places these risks into the context of these national security and economic competitiveness foreign policy questions.

This report will also analyze how AI/ML may fit within the realm of strategic trade controls generally through analyzing the AI/ML transfer process, the nuances of listed versus non-listed items, and specific challenges regarding the potential use of export controls to manage AI/ML transfers. The report presents a view into how AI/ML and emerging technologies more broadly are challenging many of the core principles and practices taken as accepted practice in the strategic trade community. The authors will analyze how using traditional control tools may be reevaluated, if not the broader conceptual basis for controls themselves, based on the nature of certain evolving technologies. The authors also analyze recent U.S. efforts to apply controls on goods whose military end-use is not yet known.

The report then focuses in-depth on a multitude of policy options – both grounded in traditional export controls as well as more creative strategic trade management solutions – that can help policymakers determine sound policy. Finally, the report offers an outlook regarding what the future may hold for AI/ML in the context of strategic trade controls and the impacts of unexpected events on the trade dynamics surrounding AI/ML.

Given the U.S. lead in exploring strategic trade control measures as tools to mitigate the risks posed by emerging technologies, this report emphasizes United States efforts and policy and is meant to bolster current discussions and policymaking.

I.III Report Methodology

This report uses comparative analysis based on information available from other technologies, which are sometimes cross-sectional or magnified by AI/ML, case studies and examples, and empirical information to determine if, to what extent, and how, trade control tools can be applied to AI/ML. At the root of this endeavor is the goal of not just helping policymakers and other stakeholders determine policy options for AI/ML specifically, but to demonstrate how the AI/ML realm can be an example and provide lessons for the management of other emerging technologies in the present context and in the future.

The research, analysis, conclusions, and recommendations in this report stem from work done as part of a larger project implemented by the report authors on emerging technologies and strategic trade controls. An initial dialogue was held on March 14, 2019 that focused on this topic broadly, with attendees representing government departments, research, industry, and academia.[9] Based on feedback and output from the initial event, the project team chose AI/ML as the focus of a second dialogue on March 9-10, 2020. The event featured around 25 attendees with policy and technical expertise. The dialogue's objective was to foster the diffusion of information about how trade controls and other forms of governance affect the spread and use of AI/ML technologies and how they ought to be employed to mitigate the risks of the spread

[9] This event was organized and executed by the Strategic Trade Research Institute (STRI) and the University of Maryland's Center for International and Security Studies (CISSM) in cooperation with the Stimson Center. It took place at the Stimson Center in Washington, DC.

of capabilities with the most pernicious potential uses.

The authors, in addition to relying on conclusions from emerging technology and AI/ML-focused dialogues, also have conducted research regarding the technical aspects of AI/ML, scenarios and risks assessments of malicious uses of these technologies, and the level of technological development in certain areas in the United States and abroad. Finally, the authors have used existing research, policy developments, and expert knowledge to bolster the findings and conclusions in this report.

I.IV Literature Background

As part of the recent focus on AI/ML in policy circles, governments and a range of nongovernmental organizations—commercial businesses, research organizations, and nonprofits—have begun to analyze the broad impacts of these technological developments, including their economic and security benefits and risks and whether and how to govern their development and use. Indeed, the February 2019 executive order on "Maintaining American Leadership in Artificial Intelligence," directed senior administration officials to "inform the development of regulatory and non-regulatory approaches ... regarding technologies and industrial sectors that are either empowered or enabled by AI, and that advance American innovation while upholding civil liberties, privacy, and American values."[10]

Early studies of the international security implications of AI/ML focused on the quickly emerging yet poorly understood capabilities of autonomous systems; on the potential for AI/ML to collect and analyze large quantities of digital data; and on economic development opportunities associated with the technology.[11] Some of these studies emphasized the need for broad and thoughtful analysis of future AI/ML developments and the development of a range of approaches to ensuring that all of the benefits of the technology can be reaped and as

[10] White House, "Executive Order on Maintaining American Leadership in Artificial Intelligence," February 11, 2019, available at <https://www.whitehouse.gov/presidential-actions/executive-order-maintaining-american-leadership-artificial-intelligence/>.

[11] Greg Allen and Taniel Chan, "Artificial Intelligence and National Security," Belfer Center for Science and International Affairs, Harvard Kennedy School, July 2017; Andrew Ilachinski, "AI, Robots, and Swarms: Issues, Questions, and Recommended Studies," CNA, January 2017.

many of the risks of malicious use can be reasonably mitigated.

Despite the relatively broad focus of many early studies of AI/ML's ascendance, there was also an early focus on the need for the United States to gird for a prolonged and fierce competition in the development and application of AI/ML in economic and security contexts with foreign governments and firms, with Chinese AI/ML developments and applications being the most prominent.[12] This specific focus is motivated by policymakers and analysts who focus on China as a "near-peer" competitor and Russia as an unpredictable source of technological development—both of whom are seen to be using information technologies as means to counter dominant United States military capabilities.[13]

This set of concerns led to the 2018 United States National Defense Strategy stating the goal of investing in the "military application of autonomy, artificial intelligence, and machine learning, including rapid application of commercial breakthroughs, to gain competitive military advantages."[14] To maintain this military advantage, United States policy experts have recommended that the United States redouble its investment in both technological and workforce development and minimize the potential for foreign development by limiting "illicit technology transfers," establishing export controls on key AI-related technologies, and leading efforts to set global norms on the use of AI, among other proposals.[15]

In addition to focusing on the explicit integration of AI/ML into military-relevant technologies, security experts also began to consider the broader range of potential "malicious" uses of AI/ML.[16] This focus

[12] Elsa Kania, "Beyond CFIUS: The Strategic Challenge of China's Rise in Artificial Intelligence," Lawfare Blog, June 20, 2017, <https://www.lawfareblog.com/beyond-cfius-strategic-challenge-chinas-rise-artificial-intelligence>.

[13] United States-China Economic and Security Review Commission,"2017 Report to Congress," November 2018, available at <http://www.uscc.gov>.

[14] "2018 United States National Defense Strategy," United States Department of Defense, <https://dod.defense.gov/Portals/1/Documents/pubs/2018-National-Defense-Strategy-Summary.pdf>.

[15] Martijn Rasser et al.,"The American AI Century: A Blueprint for Action," Center for New American Security, December 2019, <https://www.cnas.org/publications/reports/the-american-ai-century-a-blueprint-for-action>.

[16] Miles Brundage, and Shahar. Avin, "The Malicious Use of Artificial Intelligence : Forecasting, Prevention, and Mitigation Report," Future of Humanity Institute, University of Oxford; Centre for the Study of Existential Risk, University of Cambridge; OpenAI, Oxford, February 2018, <https://arxiv.org/ftp/arxiv/papers/1802/1802.07228.pdf>.

has led to additional proposals to guard against the potential misuses by defining the types of threats posed by AI/ML technologies in ways that allow policymakers to direct their focus to the most harmful and destabilizing applications, and by outlining a new type of relationship between technologists, governments, and the public at large that could prevent the emergence of new dangerous applications.[17]

It is within this nascent yet contested policy space that this report addresses the potential application of strategic trade controls. In addition to relying on the various policy research and analyses completed to date, this report also builds on the presentations and papers developed specifically for the project's dialogues, the broader strategic trade literature, and private interactions that project team members have had with policymakers and experts in this space.

II. Technology Overview

II. I Overview of Different AI/ML Systems and Applications

Artificial intelligence (AI) is a term broadly applied to a host of computer science research and innovation branches that are currently promising to improve efficiency and functionality in nearly every industry. Together, AI/ML comprise just one of a series of advanced technology areas discussed in the United States Department of Commerce's emerging technology ANPRM; however, AI/ML systems can also act in concert with other advanced technologies.

Despite the widespread heralding and debate over AI implementation, a universal definition for the term has proven to be evasive; nearly every report, article, or op-ed offers its own definition.[18] In the most general sense, the term "artificial intelligence" is used to recognize the conference of human-level decision-making and cognition onto inanimate computation systems. Disagreements about the definition of AI beyond this very basic conception derive from, and in fact are emblematic of, the wide variation in the methodological bases and

[17] Ibid.

[18] Bernard Marr, "The Key Definitions of Artificial Intelligence (AI) That Explains Its Importance," *Forbes*, February 14, 2018, <https://www.forbes.com/sites/bernardmarr/2018/02/14/the-key-definitions-of-artificial-intelligence-ai-that-explain-its-importance/#200554de4f5d>.

applications of specific AI technologies.

While Section 238 of the FY2019 National Defense Authorization Act (NDAA) directed the Secretary of Defense to produce a definition of artificial intelligence by August 13, 2019, as of June 2020 no official United States government definition of AI exists.[19] The FY2019 NDAA does, however, provide a definition of AI for the purposes of Section 238:

1. Any artificial system that performs tasks under varying and unpredictable circumstances without significant human oversight, or that can learn from experience and improve performance when exposed to data sets.

2. An artificial system developed in computer software, physical hardware, or other context that solves tasks requiring human-like perception, cognition, planning, learning, communication, or physical action.

3. An artificial system designed to think or act like a human, including cognitive architectures and neural networks.

4. A set of techniques, including machine learning that is designed to approximate a cognitive task.

5. An artificial system designed to act rationally, including an intelligent software agent or embodied robot that achieves goals using perception, planning, reasoning, learning, communicating, decision-making, and acting.[20]

A number of key terms have been established to help delineate specific components or types of AI systems; for the most part, these categories

[19] Section 1051, meanwhile, establishes a National Security Commission on Artificial Intelligence as an independent establishment within the federal government for approximately two years, until October 1, 2020.

[20] Illustrative of the disparate definitions of AI, the Defense Innovation Board (DIB), classifies AI as "a variety of information processing techniques and technologies used to perform a goal-oriented task and the means to reason in pursuit of that task." See DIB "AI Principles: Recommendations on the Ethical Use of Artificial Intelligence by the Department of Defense – Supporting Document," November 2019, pp. 8- 10. In furtherance of a common definition for the United States government, the Trump administration on 11 February 2019 enacted an Executive Order on Maintaining American Leadership in Artificial Intelligence (AI)," which directs the National Institute of Standards and Technology (NIST) to create a plan for Federal engagement in the development of technical standards and related tools in support AI technologies.

detail the specific methodological basis used to achieve a given AI capability. Perhaps the most commonly connoted, and equally vague, AI application/branch is machine learning (ML), which refers to the capability of a system to learn based on experience *in addition to* explicitly coded instructions. This report looks at "AI/ML" in order to generalize the field of advanced programming and computing capabilities at large, although in developing practicable policies it will inevitably be necessary to consider smaller branches within the overarching categories. The panel for Stanford's One Hundred Year Study on AI identified 11 such emerging branches of AI research and technologies. These more specific categories include: large-scale machine learning, deep learning, reinforcement learning, robotics, computer vision, natural language processing, collaborative systems, crowdsourcing and human computation, algorithmic game theory and computational social choice, Internet of Things (IoT), and neuromorphic computing.[21] However, this is by no means an exhaustive list and in many cases there are overlaps among the different categories, where certain technologies are composed in part of other technologies - which has also contributed to the confusion over a universal AI definition.

Due to the fact that each branch of AI/ML relies on different programming architectures, different applications of AI/ML have been identified largely in relation to the infrastructure that would be most applicable. As the following list of examples indicates, many of these different infrastructure types can be combined in an iterative manner to harness the strengths and avoid the weaknesses of each individual system.[22] Examples of private industry and defense applications are also listed.

- **Large-scale machine learning systems** are suited to the rapid analysis of large data sets. Machine learning is a method of analyzing data sets through a series of learned patterns and models. Large-scale machine learning systems are able to analyze data sets composed of complex variables with high dimensions. Given this ability to analyze large and complex data sets rapidly, machine learning is being applied in fields like finance, healthcare, retail,

[21] Peter Stone, Rodney Brooks, Erik Brynjolfsson, Ryan Calo, Oren Etzioni, Greg Hager, Julia Hirschberg et al, "Artificial Intelligence and Life in 2030," One Hundred Year Study on Artificial Intelligence: Report of the 2015-2016 Study Panel (2016): p. 52, <https://ai100.stanford.edu/sites/g/files/sbiybj9861/f/ai_100_report_0901fnlc_single.pdf>.

[22] Pei, Jing, Lei Deng, Sen Song, Mingguo Zhao, Youhui Zhang, Shuang Wu, Guanrui Wang et al. "Towards Artificial General Intelligence with Hybrid Tianjic Chip Architecture," Nature 572 (no. 7767), 2019: pp. 106-111, <https://www.nature.com/articles/s41586-019-1424-8>.

and transportation to identify patterns in data sets that allow for real-world problem solving on issues like fraud detection, traffic optimization, and targeted advertising. In the military realm, this capability may be applicable to noise-signal detection or geospatial surveillance data, for example.[23]

- **Deep learning** is a specialized version of machine learning that is suited to perception and recognition tasks that require prediction and pattern tracing, such as computer vision, activity identification, and natural language processing. Deep learning systems use labeled data sets and systems with large computing power capabilities to employ a classification model (typically composed of neural networks) capable of operating with immense accuracy and adaptability. Given these capabilities, deep learning has been deployed across industries to perform complex, automated tasks, such as driverless automobile operation, hearing and speech translation, and cancer cell detection. In the defense and security field, deep learning could be applied to complex image recognition tasks, like action/response optimization based on sensor analysis.[24]

- **Reinforcement learning** is another specialized version of machine learning that applies frameworks to undertake decision-making in novel environments. Reinforcement learning may be accomplished through a number of mathematical approaches (including through deep learning methodologies) and signifies that a computer system is taking in data in test trials, responding, monitoring results, and learning optimized responses based on environmental patterns. Given these capabilities, reinforcement learning is suited towards strategy and optimization problem-solving.[25] Private industry has become increasingly interested in reinforcement learning for a variety of tasks, ranging from the optimization of cloud and network services to video game simulation and enhancement. In the defense industry, reinforcement learning is being applied in areas such as swarm/UAV cluster task scheduling and war-

[23] "Machine Learning: What it Is and Why it Matters," SAS, <https://www.sas.com/en_us/insights/analytics/machine-learning.html>.

[24] "What is Deep Learning?,"Mathworks, <https://www.mathworks.com/discovery/deep-learning.html>.

[25] "MachineLearningAlgorithmsUse,SAS,<https://blogs.sas.com/content/subconsciousmusings/2017/04/12/machine-learning-algorithm-use/>.

gaming.[26,27]

II.II World AI Outlook

Regardless of the ambiguity around any given AI/ML-labeled technology, or perhaps because of it, AI/ML has become a touchstone in emerging technology research in both the private and public sector. As briefly surveyed above, researchers are exploring the possibility for AI/ML to dramatically change operations across a range of industries, including in manufacturing, healthcare, transportation, and finance, as well as in areas that could benefit government or military operations. A number of factors have been attributed to this rapid resurgence of AI/ML research and implementation, including: the mass availability of data, speed and storage infrastructure developments (i.e. the Cloud), and general computer science research improvements.[28]

AI/ML has found tremendous interest in the private sector. A recent MIT Technology Review insight report found that 72% of organizations in a survey had deployed AI by 2018 and 87% had by 2019.[29] The top drivers of this transition were identified to be quality control, customer care, and cybersecurity applications.[30] Below is a detailed overview of the different applications for which industries are seeking to employ AI/ML. In addition to the specific benefits gained from discrete AI/ML applications, AI/ML also emblemizes the future, promising the ability to navigate ever-growing troves of data and to harness the power of the digital evolution.[31] However, the exact pace at which AI will be implemented has become a point of growing uncertainty.[32] One

[26] Jun yang, Xinghui You, Gaoxiang Wu, Mohammad Mehedi Hassan, Ahmad Almogren, and Joze Guna, "Application of Reinforcement Learning in UAV Cluster Task Scheduling," *Future Generation Computer Systems* (Vol. 95), 2019, <https://www.sciencedirect.com/science/article/abs/pii/S0167739X18325299>.

[27] Glenn Moy and Slava Shekh, "The Application of AlphaZero to Wargaming," *Australasian Joint Conference on Artificial Intelligence*, 2019, <https://link.springer.com/chapter/10.1007/978-3-030-35288-2_1>.

[28] Babak Hodjat, "The AI Resurgence: Why Now?" *Wired*, 2015, <https://www.wired.com/insights/2015/03/ai-resurgence-now/>.

[29] "The Global AI Agenda," MIT Technology Review, March 26, 2020, <https://mittrinsights.s3.amazonaws.com/AIagenda2020/GlobalAIagenda.pdf>.

[30] Ibid.

[31] Ibid.

[32] Erin Winick, "Every Study We Could Find on What Automation Will Do to Jobs, in One Chart," *MIT Technology Review*, 2018, <https://www.technologyreview.com/2018/01/25/146020/every-study-we-could-find-on-what-automation-will-do-to-jobs-in-one-chart/>.

hindrance is the lack of knowledge about AI; a recent study found that 83% of senior business leaders were "unfamiliar" with the term.[33]

Examples of AI/ML Application by Sector

Healthcare[34]

□ Analytics: All levels of healthcare and medical practitioners are finding uses for AI/ML analytic power for tasks ranging from administrative oversight of medical records and worker optimization to drug research and diagnosis technique improvements.

□ Technology Intersection: In combination with other emerging technologies, AI/ML is proving to be an effective way to improve robotics in healthcare. AI/ML inclusion has the potential to increase accuracy and maneuverability of robotics for surgery and general patient care, and thus to reduce associated risks and costs. Furthermore, robot-assisted surgery, via AI/ML augmentation, can increase the number of surgeries that can be performed using "minimally invasive" operations, which in turn reduces the required duration of a patient's hospital stay.

Engineering and Construction[35]

□ Analytics: AI/ML analytic power has been projected to dramatically improve the efficiency of all stages of engineering and construction projects (including design, preconstruction, construction, operations, and asset management) through supply chain and production task management.

□ Technology Intersection: In combination with other emerging

[33] Darrell West and John Allen, "How Artificial Intelligence is Transforming the World," The Brookings Institute, (2018), <https://www.brookings.edu/research/how-artificial-intelligence-is-transforming-the-world/>.

[34] Bernard Marr, "How is AI Used in Healthcare – Five Powerful Real-World Examples that Show the Latest Advances," *Forbes*, July 27, 2018, "https://www.forbes.com/sites/bernardmarr/2018/07/27/how-is-ai-used-in-healthcare-5-powerful-real-world-examples-that-show-the-latest-advances/#7440478e5dfb>.

[35] Jose Blanco, Steffen Fuchs, Matthew Parsons, and Maria Ribeirinho, "Artificial Intelligence: Construction Technology's Next Frontier," McKinsey and Company Article, April 4, 2018, <https://www.mckinsey.com/industries/capital-projects-and-infrastructure/our-insights/artificial-intelligence-construction-technologys-next-frontier>.

technologies, AI/ML has the potential to increase safety and efficiency of engineering and construction projects. In addition to the obvious benefits of construction robots automated using AI/ML, AI/ML embedded within computer vision systems can allow for better surveillance of construction environments, as well as better quality control during fabrication and defect detection during operation.

Business and Marketing[36]

□ Analytics: The business and marketing sectors are finding a wide variety of tasks for AI/ML analytic application, including consumer behavior forecasting, high precision personalized advertising and targeting marketing, supply chain management, and staffing optimization.

□ Technology Intersection: When used in conjunction with other types of emerging technologies, such as computer vision and natural language processing, AI/ML can also greatly improve customer interaction capabilities through "chatbots."

Transportation and City Planning[37]

□ Analytics: AI/ML has the potential to drastically minimize the difficulties of long and short-term city planning, especially for transportation and traffic management. AI/ML methods can be applied to optimize smart pricing for HOVs on highways and bridges, to "dynamically adjust" speed limits, and to improve public transportation flow scheduling.

□ Technology Intersection: In conjunction with other technologies such as drones, microwave sensors, radars, and even cars themselves, AI/ML implementation could lead to the automation of routine tasks such as law enforcement and personal driving.

[36] Michael Chui, Nicolaus Henke, and Mehdi Miremadi, "Most of AI's Business Uses Will Be in Two Areas," *Harvard Business Review*, July 20, 2018, <https://hbr.org/2018/07/most-of-ais-business-uses-will-be-in-two-areas>. See also "Business Applications for Artificial Intelligence: An Update for 2020," Harvard Professional Development Blog, March 18, 2019, <https://blog.dce.harvard.edu/professional-development/business-applications-artificial-intelligence-what-know-2019>.

[37] Peter Stone, Rodney Brooks, Erik Brynjolfsson, Ryan Calo, Oren Etzioni, Greg Hager, Julia Hirschberg, Shivram, Kalyanakrishnan, Ece Kamar, Sarit Kraus, Kevin Leyton-Brown, David Parkes, William Press, Anna Lee Saxenian, Julie Shah, Milind Tambe, and Astro Teller, "Artificial Intelligence and Life in 2030." One Hundred Year Study on Artificial Intelligence: Report on the 2015-2016 Study Panel, Stanford University, 2016, <https://ai100.stanford.edu/sites/g/files/sbiybj9861/f/ai100report10032016fnl_singles.pdf>.

A growing contingent of countries have also been devoting resources to progress domestic AI/ML research in order to reap the economic benefits from the activities listed above, as well as the strategic/security benefits from military/government implementation. Given the dispersed and varied nature of AI/ML investment and resource allocation, as well as the high dependence on specific applications, there is some discrepancy over which countries are "leading" the race to develop and implement AI/ML. With respect to the weaponization of AI/ML, China, Russia, and the United States have unsurprisingly emerged as the most committed states.[38] However, with respect to broader implementation across industries, for the sake of economic and technological gains, a number of countries have exerted energy and resources in order to assert themselves as AI/ML leaders.

In a report published in 2020, Cognilytica found France, Israel, the United Kingdom, and the United States to be leading in technical dominance, with China, Germany, Japan, and South Korea close behind. In terms of sheer funding size, Cognilytica found the United States and China to have the lion's share.[39] In looking at private industry, a 2018 report published by CB Insights found the United States, Europe, Israel, Canada, and Japan, to be the geographies with the greatest private investment. Notably, China is absent from this list because much of its investment is through the government, with private sector investment data being unreliable.[40]

Using yet another set of metrics, Oxford Insights found Singapore, the United Kingdom, Germany, the United States, Finland, Sweden, Canada, France, Denmark, and Japan to be the top ten leaders in "AI readiness." Here, an AI readiness score is determined using 11 metrics grouped under four overarching categories: governance; infrastructure and data; skills and education; and government and public services. In

[38] Tom Semonite, "For Superpowers, Artificial Intelligence Fuels New Global Arms Race," *Wired*, September 2017, <https://www.wired.com/story/for-superpowers-artificial-intelligence-fuels-new-global-arms-race/>.

[39] Kathleen Walch, "Why the Race for AI Dominance is More Global Than You Think," *Forbes*, February 9, 2020, <https://www.forbes.com/sites/cognitiveworld/2020/02/09/why-the-race-for-ai-dominance-is-more-global-than-you-think/#72c8cac6121f>.

[40] "Top AI Trends to Watch in 2018," CB Insights - Reports, 2018, <https://www.cbinsights.com/reports/CB-Insights_State-of-Artificial-Intelligence-2018.pdf>.

this index, China is ranked 20th with respect to AI readiness.[41]

Finally, in a 2019 report, Deloitte focused on specifically early AI adopters, including China, Germany, the United Kingdom, the United States, Canada, and Australia. The approach of analyzing early adopters in this report was taken in order to examine the variance of AI maturity rate.[42]

II. III Mapping World AI Development

With the goal of improving current efforts to monitor and manage global dual-use emerging technology development, the authors of this report are establishing a methodology to track rising technologies across sector by geographical area. Using open source data, the authors have undertaken a mapping of quantum information technologies, positioning, navigation, and timing capabilities (PNT), and computer vision technologies.[43] The goal of this type of methodology is to develop a dynamic tracking capability that helps predict the controllability of technologies at various stages. For example, by providing an indication of the relative dispersion of a certain component, the exact types of export control approaches can be tailored to the stage of technology development.

This kind of mapping research can be useful for a variety of purposes. Not only can it help policymakers focus resources and effort, but it can also drive capacity-building and cooperation priorities. Mapping can further give stakeholders the ability to disassociate hype from reality by demonstrating what real “chokepoint” technologies are as opposed to technologies that already have widespread foreign availability. Especially when it comes to considering policy options such as adding new controls to existing control lists, visa vetting, or others, access to clear information regarding risks and threats is key. Therefore, some of the work done as part of the mapping component of the broader

[41] “Government Artificial Intelligence Readiness Index 2019,” International Development Research Center and Oxford Insights, 2019, <https://ai4d.ai/wp-content/uploads/2019/05/ai-gov-readiness-report_v08.pdf>.

[42] Jeff Loucks, Susanne Hupfer, David Jarvis, and Timothy Murphy, “Future in the Balance? How Countries are Pursuing AI Advantage,” Deloitte Insights, 2019, <https://www2.deloitte.com/us/en/insights/focus/cognitive-technologies/ai-investment-by-country.html>.

[43] Report forthcoming in the fourth quarter of 2020, published by the Strategic Trade Research Institute (STRI) and the University of Maryland’s Center for International and Security Studies (CISSM).

emerging technology project executed by the report authors informs parts of this report.

II. IV Risks and Threats of Globalized AI on the National Security Infrastructure

While AI/ML systems are being aggressively pursued due to all the benefits discussed above, they also pose a number of security risks and threats. This section discusses the risks posed by AI/ML technologies and their applications. The increased use of AI/ML systems introduces a number of possible risks and threats, regardless of whether systems are used by state or non-state actors. The risks discussed in this section relate to "narrow" AI, which has limited application outside its original intent. This is opposed to the more advanced, futuristic "general" AI.[44] The importance of these threats depends on a number of factors, including how advanced a technology is, how the technology can be used for malicious purposes, the intention of the possessing actor, the defenses possessed by the target, and the infrastructure, expertise, and capabilities possessed by the actor.

What makes many of these AI systems dangerous is their interaction with existing technologies, or "convergence."[45] AI often functions as an "augmentation system," enhancing existing resources rather than acting as a resource on its own.[46] This section outlines major categories of AI risks, including examples of particular technologies AI might improve. These four categories draw upon examples and categories in the existing literature on how AI could interact with other systems. These categories are arranged thematically and each could include activities by state and non-state actors in WMD-related and non-WMD domains. It is particularly difficult to separate the risks associated with

[44] Paul Scharre and Michael Horowitz, "Artificial Intelligence: What Every Policymaker Needs to Know," Center for a New American Security, June 19, 2018, <https://www.cnas.org/publications/reports/artificial-intelligence-what-every-policymaker-needs-to-know>.

[45] For more on convergence, see Natasha Benjema, "The Future of Defense Innovation: Removing the Silos between the Warfighters and Innovators," National Defense University, Research Paper No. 2, May 2018, <https://wmdcenter.ndu.edu/Portals/97/Documents/Publications/Articles/EC%20research%20paper%20no%202%20-%20Bajema%20-%20FINAL.pdf?ver=2018-05-08-135700-113>.

[46] Dialogue on Artificial Intelligence and Strategic Trade Controls, organized by the Strategic Trade Research Institute and the Center for International and Security Studies at Maryland on March 9, 2020, in Washington DC.

AI-use related to both types of weapons systems in certain categories, particularly if those systems can be entangled with each other.[47]

Categories:

1. Automation: decision-making and command and control
2. Enhanced cyber capabilities
3. Information and surveillance
4. Physical production

While these categories are defined thematically by use, another way to divide them could be by type of outcome; this could include "misuse," "accident," and "structural risks." Distinguishing risks in this way is useful because the intent of an actor is not necessarily malicious.[48]

1. *Automation: Decision-Making and Command and Control*

   AI systems can aid humans in making national security and defense decisions.[49] While there are possible benefits of automating systems, there are also a number of risks decision-makers may encounter. One former United States government official divides the military application risks of AI into "characteristics" and "applications," where the latter includes how militaries would actually apply the technology in practice.[50]

   Automation could raise risks for nuclear systems, including if early warning systems are aided by automation and if states

[47] For an in-depth discussion of conventional/nuclear entanglement in general see James Acton, "Escalation Through Entanglement: How the Vulnerability of Command-and-Control Systems Raises the Risks of an Inadvertent Nuclear War," Carnegie Endowment for International Peace, August 8, 2018, <https://carnegieendowment.org/2018/08/08/escalation-through-entanglement-how-vulnerability-of-command-and-control-systems-raises-risks-of-inadvertent-nuclear-war-pub-77028>.

[48] Dialogue on Artificial Intelligence and Strategic Trade Controls, organized by the Strategic Trade Research Institute and the Center for International and Security Studies at Maryland on March 9, 2020, in Washington DC.

[49] This analysis does not assume AI has complete autonomy over decision-making, as humans will likely have a range of control over decisions as they interact with AI capabilities.

[50] Larry Lewis, "Killer Robots Reconsidered: Could AI Weapons Actually Cut Collateral Damage?" Bulletin of the Atomic Scientists, January 10, 2020, <https://thebulletin.org/2020/01/killer-robots-reconsidered-could-ai-weapons-actually-cut-collateral-damage/>.

adapt weapons delivery systems with automated components.[51] Automation for decision-making can also be further integrated to nuclear command and control. Some scholars have even argued that an AI-enabled command and control system would benefit the United States. They argue systems like the one currently researched by Defense Advanced Research Projects Agency (DARPA) may aid the United States when the nuclear response window is so narrow.[52] Command execution is predicated upon information type and processing speed. Currently, information available to military decisionmakers arrives in diverse formats from multiple platforms, often with redundancies and/or unresolved discrepancies.

In most cases, disparate data must be manually processed. AI could expedite the command and control process. For example, the United States Air Force is developing a system for Multi-Domain Command and Control (MDC2) which intends to centralize planning and execution of air-, space, cyberspace-, sea, and land-based operations. In the immediate future, AI may be used to fuse data from sensors in all of these domains to create a single source of information, also known as a "common operating picture," for decision makers.[53] However, the risks of automating command and control systems in such a way includes yielding too much power to AI capabilities, where humans could end up "surrendering their judgement."[54] Theoretically, this automation could be applied to a variety of defense use cases and multiple levels of military hierarchy.

2. *Enhanced Cyber Capabilities*

Machine learning and other AI capabilities can aid in both cyber defense and offense. Depending on who is using these capabilities

[51] Michael C. Horowitz, Paul Scharre, and Alexander Velez-Green, "A Stable Nuclear Future? The Impact of Autonomous Systems and Artificial Intelligence." December 2019, pp. 4-5.

[52] Adam Lowther and Curtis McGiffin, "America Needs a "Dead Hand," War on the Rocks, August 16, 2019, <https://warontherocks.com/2019/08/america-needs-a-dead-hand/>. See also Bryce Farabaugh, "Bad Idea: Integrating Artificial Intelligence with Nuclear Command, Control, and Communications," CSIS Defense 360, December 3, 2019, <https://defense360.csis.org/bad-idea-integrating-artificial-intelligence-with-nuclear-command-control-and-communications/>.

[53] Major General (ret) Tim Zadalis, "United States Air Force Multi-Domain Command and Control: Maintaining Our Asymmetric Advantage," *Journal of the Joint Air Power Competence Center* (Edition 26), Spring/Summer 2018, pp. 10-15.

[54] Michael C. Horowitz, Paul Scharre, and Alexander Velez-Green, "A Stable Nuclear Future? The Impact of Autonomous Systems and Artificial Intelligence," December 2019, p. 4.

and how they are employed, AI-enhanced cyber capabilities could be either beneficial or threatening. AI could also make more advanced cyber capabilities increasingly accessible to people or groups with less resources.[55] Analysts have noted AI-enabled cyber capabilities have the ability to both increase the strength of cyber operations and decrease the ability for a target to attribute an attack. This also has nuclear domain implications where AI capabilities could enable a malicious actor to attack command and control systems.[56]

3. *Information and Surveillance*

AI capabilities could improve information operations and accelerate the effectiveness of propaganda. A bright spot is that AI could help detect these very forms of information operations, but the systems could also make propaganda more effective through the use of capabilities like natural-language processing and ML-guided targeting. This includes the production and dissemination of fraudulent content like deep-fakes and bots which spoof real human interactions.[57] In addition to information operations, surveillance techniques may be enhanced and democratized by AI due to the capabilities of the systems and the relatively small resource requirements.[58]

The information and surveillance category of AI applications could have implications on WMD issues. They could enhance intelligence gathering, targeting, or information operations. For instance, laboratories that study AI have noted the capabilities can easily enhance disinformation by having computers write realistic, human-sounding narratives.[59] Russia has undertaken disinformation efforts in Syria for example, including concerning

[55] Roman Yampolskiy, "AI is the Future of Cybersecurity, for Better or Worse," Harvard Business Review, 2017, <https://hbr.org/2017/05/ai-is-the-future-of-cybersecurity-for-better-and-for-worse>.

[56] James Johnson and Eleanor Krabill, "AI, Cyberspace, and Nuclear Weapons," War on the Rocks, January 31, 2020, <https://warontherocks.com/2020/01/ai-cyberspace-and-nuclear-weapons/>.

[57] Michael C. Horowitz, Paul Scharre, Alexander Velez-Green, "A Stable Nuclear Future? The Impact of Autonomous Systems and Artificial Intelligence," <https://arxiv.org/abs/1912.05291>, p. 5.

[58] Ibid, p. 3.

[59] Cade Metz and Scott Blumenthal, "How A.I. Could Be Weaponized to Spread Disinformation," New York Times, June 7, 2019, <https://www.nytimes.com/interactive/2019/06/07/technology/ai-text-disinformation.html>.

chemical weapons.[60] Analysts have also suggested that Russian efforts to create and disseminate "fake news content" could theoretically be aided and expanded by AI capabilities.[61] This category is a prime example of AI increasing the pervasiveness and speed of an existing harm.

4. *Physical Production*

A fourth and smaller category of risk is the role AI could play in the design of physical objects including weapons components. Generative design, a machine learning process where systems can independently iterate on physical design, could aid a user without advanced technical knowledge in creating weapons components through 3-D printing.[62] There is already a body of literature on national security risks stemming from additive manufacturing.[63] Additive manufacturing aided by AI capabilities could accelerate these risks.

II.V Risks and Threats of Globalized AI in Non-Security Contexts

Beyond the immediate risks that AI capabilities could be integrated directly into national security infrastructures, a more subtle possibility is that specific AI developments could introduce non-military risks as a function of general trendlines in the development of the technology and the nature of technology innovation today. The asymmetric globalized development of AI-related technologies; poor or mismatched regulations across borders; and piecemeal implementation of AI-related security measures could each prove destabilizing. The effects of this destabilization could ripple across the realms of economics, infrastructure security, health and safety, and state governance.

[60] Center for Strategic and International Studies, "Russian Disinformation in Syria," Podcast, <https://www.csis.org/podcasts/babel/russian-disinformation-syria>.

[61] William Drozdiak, "Europe's Challenges in an Age of Social Media, Advanced Technologies, and Artificial Intelligence," Hoover Institution, February 4, 2019, <https://www.hoover.org/research/europes-challenges-age-social-media-advanced-technologies-and-artificial-intelligence>.

[62] Matthew Gault, "3-D Printers Could Help Spread Weapons of Mass Destruction," Scientific American, September 10, 2019, <https://www.scientificamerican.com/article/3-d-printers-could-help-spread-weapons-of-mass-destruction/>.

[63] For an overview of additive manufacturing risks see Deloitte Insights. "3D Opportunity for Adversaries," August 22, 2017, <https://www2.deloitte.com/us/en/insights/focus/3d-opportunity/national-security-implications-of-additive-manufacturing.html>.

Asymmetric Research and Development Progress

Asymmetric research and development progress, especially among global superpowers, has triggered concern that even civilian-purposed AI could result in monopolization of the industry or components - and thus produce economic shifts. The United States has focused primarily on the potential for China to outpace other countries in global development. These concerns gained momentum as Chinese leadership announced plans for Chinese superiority in AI technology, including the proclamation of its "New Generation Artificial Development Plan" (AIDP) and its "Made in China 2025" plan, both of which outline China's plans to boost investment in AI development in hopes of gaining technical supremacy.[64] Specific fears over China overtaking the United States and other world leaders with respect to research and development were also voiced in a 2019 analysis that looked at AI talent concentration changes over time.[65] If China, or any one country for that matter, ends up monopolizing technology development or if economic profits from AI-related developments are concentrated in a single country, additional fractures in the global economy could open up. These types of economic developments could lead to increased insecurity, particularly if a leading country chooses to restrict the dissemination of the technology.

Piecewise AI-Relevant Security Measure Implementation

As actors increase their AI expertise related to cybersecurity, both states and non-state actors could leverage AI capabilities for both offensive and defensive purposes. AI-enhanced cyber capabilities could aid in hypothetical offensive cyber actions on a variety of systems and critical infrastructure. On the other hand, AI could also enhance cyber defense capabilities, which could help a group decrease the vulnerability to attacks. The exact implications of these AI-enabled cyber technologies then depend on how quickly the technology spreads and what actors use the technology for.

[64] Gregory Allen, "Understanding China's AI Strategy," Center for New American Security, July 6, 2019, <https://www.cnas.org/publications/reports/understanding-chinas-ai-strategy>.

[65] Sarah O'Meara, "China's Ambitious Quest to Lead the World in AI by 2030, *Nature*, August 22, 2019, <https://media.nature.com/original/magazine-assets/d41586-019-02360-7/d41586-019-02360-7.pdf>.

Poor or Mismatched Regulations over Borders

Another source of concern is that regulations and standards governing AI development and use established by different countries or regions will be inconsistent or potentially conflicting. Many countries are pursuing early attempts to establish standards and regulations governing AI use with respect to algorithm transparency, ethical application of AI, and permissible data acquisition and application. For example, the U.S. has issued a list of basic principles related to AI use, the European Union has developed its "General Data Protection Regulation" which includes AI use, and China has announced the Beijing AI Principles, established by the Beijing Academy of Artificial Intelligence.[66,67] Among these early efforts, discrepancies have already arisen, especially with respect to data privacy, domestic policing application, and lethal autonomous weapons application.[68]

This incongruence has the potential to create apprehension about how different countries' AI regulations will affect development and use globally, especially as technologies are moved across borders. It also could lead to regulatory arbitrage as companies seek out looser regulations in order to accelerate unfettered research and development on new technologies. With respect to security, a lack of uniform ethical standards guiding the development and use of AI/ML technologies translates to an increased risk of nefarious use of AI/ML that could ultimately lead to broader escalation.

Effects on Norms, Governance, and Credibility

How state actors and international organizations choose to approach governance of AI development and use will lay the foundation for global norms for generations to come. The advantages in shaping norms that come from being an early mover in this space have been a driver

[66] Cameron Kerry, Joshua Meltzer, and Alex Engler, "The U.S. and EU Should Base AI Regulations on Shared Democratic Values," Brookings Institution, March 2, 2020, <https://www.brookings.edu/blog/techtank/2020/03/02/the-u-s-and-eu-should-base-ai-regulations-on-shared-democratic-values/>.

[67] Will Knight, "Why does Beijing suddenly care about AI ethics?" MIT Technology Review, May 31, 2019, <https://www.technologyreview.com/2019/05/31/135129/why-does-china-suddenly-care-about-ai-ethics-and-privacy/>.

[68] "Regulations of Artificial Intelligence in Selected Jurisdictions," The Law Library of Congress, January 2019, <https://www.loc.gov/law/help/artificial-intelligence/regulation-artificial-intelligence.pdf>.

for some states to acquire AI/ML capabilities.[69] This race to technical supremacy and leadership in norm development has grown especially tight among states that fall along the divide between democratic and authoritarian governments.[70]

Even on the national governmental level, AI development, technological, and political uncertainties have led to opacity regarding the objectives of norms and how they should be implemented. For example, in the United States, the Defense Innovation Board has proposed a series of principles to guide "ethical" AI development and use; however, some of these principles have already been deemed contradictory.[71] China has emphasized the importance of cooperation in AI development to reduce threats, yet it continues its technology development unabated.[72] There is also uncertainty over the extent to which norms on AI development and use will impact other areas of governance, for example, digital security, human rights, or nuclear security.

In light of the multitude of security challenges posed by the development of AI/ML, policymakers have been keen to find solutions to mitigate security risks and threats while advancing technological progress. Due to the rapid diffusion of AI/ML and the diverse composition of state and non-state actors involved in its development, controlling the flow of technology has become an increasingly attractive policy option, particularly in the United States. Controlling flows of a technology that is so diffused, largely intangible in nature, and whose military end-uses are still emerging, however, is a daunting task. While list-based export controls appear to be an easy solution to inhibit the spread of this technology to actors who may seek to use them to threaten security (or economic) interests, this potential solution must be weighed and thoughtfully evaluated against other options that broadly fit within the strategic trade management realm. The next section analyzes the ways in which AI/ML fits within the export control field and, more broadly,

[69] Jessica Newman, "The New AI Competition is Over Norms," *The Hill*, April 22, 2019, <https://thehill.com/opinion/technology/439973-the-new-ai-competition-is-over-norms>.

[70] William Cohen, Leon Panetta, Chuck Hagel, and Ash Carter, "America Must Shape the World's AI Norms - or Dictators Will." *Defense One*, February 27, 2020, <https://www.defenseone.com/ideas/2020/02/america-must-shape-worlds-ai-norms-or-dictators-will/163392/>.

[71] Morgan Dwyer, "AI Principles and the Challenge of Implementation." Center for Strategic and International Studies, November 19, 2019, <https://www.csis.org/analysis/ai-principles-and-challenge-implementation>.

[72] Gregory C. Allen, "Understanding China's AI Strategy: Clues to Chinese Strategic Thinking on Artificial Intelligence and National Security." Center for a New American Security, February 2019, pp. 4-5.

strategic trade, and evaluates current efforts to apply list-based controls on specific AI/ML technologies.

III. AI/ML and Strategic Trade Controls

III. I A Basis for Control?: AI/ML Transfers

There are a number of AI/ML technology areas and components that could theoretically be targeted by export controls. A review of different types of technical and policy-oriented AI/ML research and feedback from different stakeholders at the in-person dialogue held in March 2020 on AI and strategic trade controls helped the report authors identify general categories of technologies and transfer processes.[73] The following analysis facilitates thinking through how such controls would be identified and organized. This analysis is key in order for policymakers to have a strong basis from which to evaluate whether export controls are an appropriate and/or effective tool for managing AI/ML risks.

There are four broad categories of AI processes that could be a subject of export controls or other restrictions: software, data, computing power and hardware, and services. Each of these categories may have implications for specific export control approaches, particularly deemed exports and intangible technology areas due to the lack of physical goods involved in general AI/ML development.[74]

Software

Software is one avenue considered for controls on AI/ML exports. Most AI/ML systems are composed of different types of software. Usually there is a base-level operating software and an end-use specific package added onto this software, if not built in. Because software is intangible, it is highly transferable and suppliers can circumvent most established trade control verification measures. Furthermore, base-level software programs are already widely available. For this reason, researchers in

[73] As mentioned in Section I, this dialogue took place on March 9-10, 2020, in Washington DC. The dialogue included around 25 experts that discussed the topic under Chatham House Rules.

[74] "Deemed Exports," Bureau of Industry and Security, Department of Commerce, <https://www.bis.doc.gov/index.php/policy-guidance/deemed-exports>.

the field believe that any export control regulations placed on general purpose AI software (the base-level programming) will be ineffective and may damage United States competitiveness.[75] Application-specific AI software is considered to be similarly difficult, though because it is linked to a specific outcome, may be easier to track and therefore enforce.[76]

Data

Data is used to train ML and AI algorithms to achieve specific objectives. Accurately labeled data is needed in an extremely large quantity in order to train and verify the functioning of these algorithms. In some cases, it is also used as a means to accomplish a task as well by providing necessary information in and of itself. For these reasons, data is another branch that could conceivably be regulated. Similar to software, data is intangible (if digitally transmitted) or incredibly small (if transmitted on a physical thumb drive), and thus would be challenging to regulate thoroughly. Some forms of technical data are already subject to export control regulations in the status quo though, and there was a recent United States Department of Commerce ruling on the proper storage and encryption of that data.[77] It would then be forbidden to transfer these types of export-controlled data for use in an AI/ML system. However, large amounts of data including high-volume datasets are already available online through open source data sets.

Computing Power and Associated Hardware

Although AI itself consists entirely of software, a certain computing power is needed in order to ensure the AI program is able to run. Furthermore, the amount of computing power that is needed to drive AI and ML programs tends to scale with the power/finesse of the program. Significant research has been done in recent years to map out what type of computing power will be needed for AI and ML programs of different levels and with varying capabilities. These computing power needs have been mapped onto physical hardware components that would be

[75] Carrick Flynn, “Recommendations on Export Controls for Artificial Intelligence,” CSET Issue Brief, February 2020, <https://cset.georgetown.edu/wp-content/uploads/Recommendations-on-Export-Controls-for-Artificial-Intelligence.pdf>.

[76] Ibid.

[77] Braumiller Law Group, LLC, “Storing Export Controlled Data in the Cloud – What’s the Latest?” August 16, 2016,<https://www.braumillerlaw.com/storing-export-controlled-data-in-the-cloud-whats-the-latest/>.

required in order to be achieved. Thus, another option for export controls would be to target individual hardware components that would allow for AI/ML programs of extremely high strength. Certain computers and associated equipment, often falling under the term "high-performance computing," have been on the Commerce Control List (CCL) for decades. Current controls on computers include specifications about the types of computing equipment and temperatures at which they operate. If certain computing hardware is closely associated with the use of advanced AI/ML, these could be a possible way to control the dispersion of AI/ML systems.

Services and Deemed Exports

In addition to the transfer of AI/ML-associated data, hardware, and software, transfers will also likely take place intangibly as deemed exports. These transfers could occur when American citizens and foreign persons collaborate on research or trade services that harness AI/ML capabilities. As with potentially controlled information in other technology areas, transfers of AI/ML information could be facilitated through research collaborations or tacit knowledge sharing. When commercially-motivated firms export physical goods associated with AI/ML, they may share tacit knowledge concerning the operation and maintenance of these systems. International research teams may share findings across borders, with or without associated transfers of physical goods.

In addition to collaborations and tacit knowledge transfers, AI/ML is emerging in the service sector. There is a burgeoning sub-industry of companies advertising AI and ML "as a service." This provision of AI/ML as something groups can purchase without in-house AI/ML expertise is projected to be a multi-billion dollar industry in the next five years and includes service initiatives launched by major American technology companies.[78] The deemed export implications of this emerging service industry are important to consider as the market grows, especially if these services are provided seamlessly across borders.

End-Use/End-User Considerations

Given that AI has been considered to be a type of augmentation system,

[78] Daniel Newman, "Why AI As A Service Will Take Off In 2020," Forbes, January 7, 2020, <https://www.forbes.com/sites/danielnewman/2020/01/07/why-ai-as-a-service-will-take-off-in-2020/#4c19ef413366>.

another opportunity for export controls is to target components that allow for end-use application. Most national and all multilateral export control regimes envision the application of controls beyond control lists in order to provide maximal regulatory flexibility. End-use controls, for example, focus on specific and prescribed applications of select commodities, thereby requiring licensing for otherwise benign transactions.

Controls can also be exercised over high-risk destinations and/or users. The latter are typically identified by various United Nations and national lists or by published "red flags" or similar risk indicators. In the United States, several agencies – namely the Department of State, Department of Commerce, and Department of Treasury - consolidate denied individuals and organization under various United States sanctions lists. Export to these end-users is forbidden or may require an authorization. In the case of AI/ML, these lists could facilitate effective regulation through end-user controls without the externality of limiting potentially peaceful, lawful trade.

End-use and end-user controls in the AI/ML context would identify activities that the export could be applied to and apply targeted control over this use, either through control of the end-use activity or control of technologies that allow for this end-use application. This appears to be the approach that the United States Government is leveraging most in developing new export controls for AI/ML. For example, on January 6, 2020, the United States Department of Commerce's Bureau of Industry and Security (BIS) announced a regulation requiring a license for software specifically meant to aid geospatial imagery analysis.[79]

The drawback of this approach is that it is predicated on the utility of the specific technology external to the end-use specified. For example, if the specific software used for geospatial imagery end-use could also be applied for a variety of other uses, then the technology may be transferred extensively regardless of the control, under the pretense of the other applications. Given that AI/ML has attracted interest from a large number of industries, specific software types are likely to have numerous applications. In the case of the end-use control announced

[79] "Addition of Software Specially Designed To Automate the Analysis of Geospatial Imagery to the Export Control Classification Number 0Y521 Series," United States Government - A Rule Proposed by the Industry and Security Bureau, January 6, 2020, <https://www.federalregister.gov/documents/2020/01/06/2019-27649/addition-of-software-specially-designed-to-automate-the-analysis-of-geospatial-imagery-to-the-export>.

by BIS on January 6, 2020, this type of software is simultaneously applicable in medical imaging and gaming technologies.[80]

Use of Catch-All Controls

If the end-use/end-user is deemed to be a primary way to control the dispersion of AI systems, catch-all controls that would require a license for export of even unlisted AI components is one possible approach. Existing catch-all measures aim to control unlisted items that would reach a proliferation-related end use/end-user.[81] The use of catch-all controls is now standard good practice worldwide for export control implementation. However, the use of these controls can create challenges for exporters who prefer clear and transparent guidance on what does and does not need authorization. The regulatory ambiguity associated with these controls then risks deterring businesses and entrepreneurs from entering certain markets. Catch-all controls also pose challenges for enforcement as they can be difficult to prove violations of catch-all clauses during prosecutions. While these factors make catch-all implementation challenging in all export control situations, it is even more daunting when applied to emerging technologies where end-use is even more nebulous, such as in the AI/ML context.

III.II United States AI/ML Policy and Outlook

The United States has become particularly concerned about the potential risks of AI/ML-related technology over the last decade. On one hand, policymakers understand the potential for AI/ML to generate immense economic benefits that could accelerate productivity and growth, although these benefits would potentially be distributed unevenly in societies. On the other hand, policymakers see the potential for AI/ML to be widely applied in the defense sector.[82] Consistent with AI's revolutionary potential across economic sectors, major national defense establishments are already exploring and to a limited extent applying AI systems as force multipliers. In particular, military AI research is

[80] Dave Aitel, "We Need a Drastic Rethink for Export Controls on AI," *Council on Foreign Relations*, January 21, 2020, <https://www.cfr.org/blog/we-need-drastic-rethink-export-controls-ai>.

[81] United States State Department, "Catch-All Controls," <https://2009-2017.state.gov/strategictrade/practices/c43179.htm>.

[82] "United States National Defense Strategy," Department of Defense, 2018, <https://dod.defense.gov/Portals/1/Documents/pubs/2018-National-Defense-Strategy-Summary.pdf>.

underway in the fields of intelligence collection and analysis, logistics, cyber-enabled operations, command and control, and in a variety of semi-autonomous and autonomous platforms.

The type of radical shift in the symmetry of military power that is possible with AI and related technologies is driving the United States Department of Defense to see the maintenance of a "national security innovation base" — of which the capability to develop and apply military AI/ML technologies is a good example — as central to gaining competitive military advantage. Senior United States policymakers have placed particular emphasis on "protecting [United States] critical AI technologies from acquisition by strategic competitors and adversarial nations," once they are acquired.[83]

These dual United States priorities of integrating AI/ML technologies into military applications and building and protecting a vibrant research and development base (particularly the private-sector component) to support advances in underlying technologies give United States policymakers, particularly defense officials, wide latitude to fund and develop policies and approaches.

Beyond specific military applications of AI systems, AI is both defining and accelerating a reformulation of national security, blending economic innovation with traditional definitions of national security. For example, the Trump Administration in 2018 articulated a "new organizing principle" for strategic policy: economic security is national security. In particular, innovative technologies are accorded priority in this new calculation due to their perceived revolutionary impact on economic development, driving economies into the so-called Fourth Industrial Revolution. AI, as an enabling platform, is central to this accelerated economic growth model, with associated industries and services dependent on AI applications. As such, mastery and assimilation of AI will simultaneously confer both strategic and economic advantages. Therefore, the articulation of national policy to support the development and export control of AI/ML will be fundamental to realizing United States national security and economic objectives.

[83] Executive Order on Maintaining American Leadership in Artificial Intelligence, February 2019, <https://www.whitehouse.gov/presidential-actions/executive-order-maintaining-american-leadership-artificial-intelligence/>.

Recent legal and regulatory changes in the United States are indications of the wide and urgent agreement among United States policymakers regarding emerging technologies. AI/ML export controls are in many respects a bellwether heralding the emerging contours of next generation controls, or, more precisely, the problematizing of this policy instrument as a practical tool.

III.III The U.S. and Export Controls on AI/ML

Since the end of the Second World War, governments have sought to manage the transfer of strategic items and technologies simultaneously through national and multilateral export controls. Technologies deemed relevant to WMD or conventional items were placed on control lists, thereby establishing a basis upon which to practice trade controls. Confronted with the current range of "emerging technologies," governments are likewise attempting to develop list-based controls.

In 2018, the United States Congress enacted the Export Control Reform Act of 2018 (ECRA). Section 1758 of ECRA instructs that: "The President shall establish and, in coordination with the Secretary, the Secretary of Defense, the Secretary of Energy, the Secretary of State, and the heads of other Federal agencies as appropriate, lead, a regular, ongoing interagency process to identify emerging and foundational technologies that A are essential to the national security of the United States." The Act, however, did not identify emerging or foundational technologies.

In late 2018, the United States Department of Commerce (DOC) published an Advanced Notice on Proposed Rulemaking (ANPRM) seeking public comment on criteria for identifying emerging technologies. The ANPRM included fourteen broad representative categories of technology, including AI and machine learning, from which the government seeks to determine whether if and which emerging technologies are important to United States national security for which effective export controls should be implemented. Over 85% of the responses to the ANPRM from a wide variety of stakeholders concerned AI and suggested that crafting associated export controls would be very challenging.[84] One ANPRM commentator succinctly

[84] See Scott Jones, "Regulating the Future: Concerns Over Defining 'Emerging Technologies,'" World Export Control Review, Issue 79, March 2019.

captured this dilemma:

> "The ANPRM notes that, "Certain technologies, however, may not yet be listed on the CCL or controlled multilaterally because they are emerging technologies. As such, they have not yet been evaluated for their national security impacts." These two sentences are at the heart of the problem of defining emerging technology within an export control framework. The uncertainties and ambiguities around emerging technology make them difficult if not impossible to govern from an export control perspective, and yet this is exactly what the process to be established through this ANPRM is tasked to do."[85]

Traditionally, technology controls were circumscribed on the basis of threat and/or weapons system. Experts, usually in one of the multilateral export control regimes, in each technology area work backwards from the identified threat to describe the technical characteristics of commercial items necessary for the development, production, or use of such items. Regulators would then work to add the items to their respective national control lists. The current approach to working from technologies forward to weapons or to "national security concerns" is unprecedented.[86]

The United States and other governments are keenly aware of the inherent challenges in developing viable controls for emerging technologies. The recent United States effort, for example, aspires to define "specific emerging technologies that are important to the national security of the United States for which effective controls can be implemented that avoid negatively impacting United States leadership in the science, technology, engineering, and manufacturing sectors." ECRA further requires that controls over emerging and foundational technologies take into account:

> "1) The development of emerging and foundational technologies in foreign countries; 2) the effect export controls

[85] "Comment for the Department of Commerce ANPRM on "Review of Controls on Certain Emerging Technologies," Samuel Evans, Research Fellow in the Program on Science, Technology, and Society at Harvard University's Kennedy School of Government. Source, <t.ly/DE95l>.

[86] Kevin Wolf, Testimony before the Senate Committee on Banking, Housing, and Urban Affairs "Confronting Threats from China: Assessing Controls on Technology and Investment" June 4, 2019, p. 3.

> imposed may have on the development of such technologies in the United States; and 3) the effectiveness of export controls imposed on limiting the proliferation of emerging and foundational technologies to foreign countries."[87]

To date, the United States has not identified new additions to the control list based on any emerging technologies associated with the ANPRM. While the United States has recently identified a narrow AI export control, the control is not part of the emerging technologies process. In January 2020, the Bureau of Industry and Security (BIS) published an interim final rule on "Software Specially Designed to Automate the Analysis of Geospatial Imagery."[88] BIS established this control under a relatively obscure provision of the Export Administration Regulations (EAR), the "0Y521" classification series.[89] To ensure that emerging technologies of concern were captured and appropriately controlled, the 0Y521 process was established in 2012 "[A]s a mechanism for situations in which an item that warrants control is not controlled yet (e.g., as with an emerging technology) this rule proposes the addition of a new, miscellaneous ECCN to the CCL, similar to USML Category XXI (Miscellaneous Articles)."[90] An Export Control Classification Number (ECCN) refers to the five character alpha-numeric designation used by the Commerce Control List (CCL) to categorize dual use items for export control purposes. Unlike routine list entries, 0Y521 designations would not be determined technically. Although described as a classification, the decision to identify an item as included in an 0Y521 ECCN "would be a foreign policy determination, not a technical classification," and the government could publish the designation as a final rule immediately.[91]

[87] Section 1758, '"Export Control Reform Act of 2018."

[88] Review of Controls for Certain Emerging Technologies: A Proposed Rule by the Industry and Security Bureau, Federal Register, November 19, 2018. See also, Ashton Carter, "Shaping Disruptive Technological Change for Public Good," Belfer Center, Harvard University, August 2018.

[89] See Proposed Revisions to the Export Administration Regulations (EAR): Control of Items the President Determines No Longer Warrant Control Under the United States Munitions List (USML) A BIS Proposed Rule by the Industry and Security Bureau on 15 July 2011 Revisions to the Export Administration Regulations (EAR): Export Control Classification Number 0Y521 Series, Items Not Elsewhere Listed on the Commerce Control List (CCL) BIS Final Rule 13 April 2012 "significant regulatory action."

[90] See Proposed Revisions to the Export Administration Regulations (EAR): Control of Items the President Determines No Longer Warrant Control Under the United States Munitions List (USML) A Proposed Rule by the Industry and Security Bureau on 07/15/2011, <t.ly/KB9yr>.

[91] See Kevin Wolf and Scott Jones, "0Y521 and Section 1758: Emerging Technologies by any Other Name?" World Export Control Review, Issue 89, May 2016.

The current "AI" control covers geospatial imagery software specially designed for training a Deep Convolutional Neural Network (CNN) to automate the analysis of geospatial imagery and point clouds.[92] The actual application of control is dependent on the interpretation of "geospatial imagery and point clouds," and, as such, potentially problematizes exports of software designed to train AI systems in image recognition. In one public comment to the interim final rule, Uber provided the following critique:

> "Uber believes that it is crucial to accurately define the scope and meaning of the terms used in this Interim Final Rule, in order to reduce current ambiguity that could lead to unintended control on other software and have negative impacts on both the economy and technology advancement of the United States (United States)."[93]

The lack of specificity is consistent with the current drift of United States controls that increasingly conflate economic and national security, an approach that subverts the previous control methodology predicated on identifying a weapons system before articulating composite items and technologies.[94] For example, recent AI controls comments by the Silicon Valley Leadership Group noted that "geospatial imagery subject to export review should be explicitly linked to specific applications that are required for the development or use of conventional weapons or weapons of mass destruction." Until, and if, a new export control methodology is specified, we can expect continued conceptual and, therefore, application challenges.[95]

Since its inception as a policy tool, dual-use trade controls have been premised on the concept of absolute and relative sources of supply,

[92] A point cloud is a collection of data points defined by a given coordinate system. A point cloud is also known as a digital surface model.

[93] See AH-89 Public Comment 21 <https://www.regulations.gov/document?D=BIS-2019-0031-0022>.

[94] One scholar recently observed, "AI seems much more akin to the internal combustion engine or electricity than a weapon. It is an enabler, a general-purpose technology with a multitude of applications. That makes AI different from, and broader than, a missile, a submarine, or a tank." See, Michael Horowitz, "Artificial Intelligence, International Competition, and the Balance of Power," *Texas National Security Review* (Volume 1, Issue 3), May 2018.

[95] See AH-89 Public Comment 11 <https://www.regulations.gov/document?D=BIS-2019-0031-0012>.

the latter being managed through like-minded regimes.[96] Increasing foreign availability obviates the practical effectiveness of otherwise unilateral controls. In tandem with the rapid growth and diffusion of global value chains, production off-shoring, and the internationalization of graduate degree programs, basic and advanced technological capabilities are widely disseminated across economies.[97] With a few notable exceptions (e.g., semiconductor manufacturing equipment and jet engine hot sections), dual-use commodities are widely produced. Unilateral controls, therefore, can undermine control effectiveness, harm domestic producers, and, in some cases, accelerate indigenization of targeted technology.[98]

Artificial intelligence-enabling technologies and hardware are widely available.[99] For example, many of the data-training platforms (e.g., Pytorch and TensorFlow) are open source and could run on as or on a cloud-based service. While the United States could control the export of semiconductor manufacturing equipment (SME), or control the products resulting from United States-origin SME, export control over semiconductors would be limited over time and would require multilateral coordination. For example, Chinese AI developers could, while suboptimal, use commercial rather than dedicated AI chipsets

[96] Arguably, export controls have also been policy tools directed at political adversaries, rather than as general trade instruments. The practice of controls becomes considerably more complicated when the target is economically intertwined with the controller. For example, James Lewis observes of the current technology control dilemma between the United States and China: "United States export controls were not designed for a hostile power with which the United States has an exceptionally close economic relationship." See, James Lewis, "Managing Semiconductor Exports to China, Center for Strategic and International Studies, May 5, 2020, https://www.csis.org/analysis/managing-semiconductor-exports-china>.

[97] Increasing sources of supply are insuperable challenge for United States technology control efforts. In particular, "[T]he end of the Cold War and globalization of national economies has led to a competitive market. Despite some liberalization over the years, the United States still maintains the strictest unilateral export control regime on dual-use technologies. Such bans have had unintended consequences: namely, to drive global customers to foreign competitors at the expense of United States suppliers since the United States is not the only source of dual-use products or technologies." See, Belay Seyoum, "Export Controls and International Business: A Study with Special Emphasis on Dual-Use Export Controls and Their Impact on Firms in the United States" *Journal of Economic Issues* (Vol. 5 No. 1), 2017, pp. 45–72.

[98] On the indigenization of technology as a response to export controls, see Michael J. Noble, "Export Controls and United States Space Power," *Astropolitics*, (Vol.6 No. 3), 2008, pp. 251-312 and Kemp, R. S, "The Nonproliferation Emperor Has No Clothes." *International Security*, (Vol. 38 No. 4), 2014, pp. 39–78. In particular, Kemp notes: "the technologies needed to make nuclear weapons have remained static, whereas the indigenous capabilities of states have steadily grown over the last half-century. What was once exotic is now pedestrian, and nuclear weapons are no exception."

[99] See, for example, "Artificial Intelligence Index Report 2019," AI Steering Committee, Stanford University, < https://hai.stanford.edu/sites/default/files/ai_index_2019_report.pdf>.

if denied access.[100] The delay afforded by ring-fenced controls would require the articulation of and coordination with a national AI policy in order to achieve a strategic advantage.

The United States and other governments are trying to rapidly assimilate, while reconciling with the national security implications of, emerging technologies. In contrast with the other emerging technologies, AI/ML has received the majority of policy focus. However, the current United States efforts to control the export of AI/ML related technologies have been extremely limited in scope and of questionable efficacy in application. Simultaneously, many governments have revised their national security review processes for foreign direct investment (FDI) into rapidly re-defined "strategic sectors," which include emerging technologies. These trends suggest that export controls are necessary but not sufficient policy instruments and that a range of policy tools will be necessary to meet otherwise contradictory national objectives.

IV. Policy Options

With the objective of minimizing the risks of AI/ML being used for malicious end-uses while maintaining competitiveness on the world stage, the authors propose a series of options for both regulators and the AI/ML sector. While several recommendations are specific to AI/ML, many can be applied to any of the emerging technological areas considered by national governments as potentially amenable to list-based controls in the future, and even applied to current list-based export-controlled sectors as a complement to export authorizations.

IV.I Outreach

Exporter awareness and buy-in is key to effective trade policy. For years, competent authorities managing trade controls, especially in countries with advanced controls, have launched outreach efforts to

[100] There is some debate as to using "AI chips" versus commercial chips for AI systems. While dedicated, or AI chips, do realize higher efficiencies and speeds, they are not necessary for a range of AI applications. As noted in a recent report on AI chips, experts disagree on the need for leading nodes for AI chips." See Saif M. Khan and Alexander Mann, "AI Chips: What They Are and Why They Matter, An AI Chips Reference," Center for Security and Emerging Technology, Georgetown University, 2019, < https://cset.georgetown.edu/wp-content/uploads/AI-Chips%E2%80%94What-They-Are-and-Why-They-Matter.pdf>.

educate and underscore the importance of compliance to their exporters, with a varying degree of success. In recent years, outreach focus has broadened from strictly industry-based efforts to also include academic and research institutions due to increased awareness of the intangible ways that sensitive, controlled transfers can take place – through publications, online communications, travel, conversations, or other intangible means.[101]

The methods used to increase awareness within industry and academia, to be most effective, cannot focus solely on the law and possible punishments. Effective compliance outreach must focus on the ultimate objective – be it security, or competitiveness, or both – to instill a *culture* of compliance. For dual-use goods, for example, reinforcing to exporters that their seemingly innocuous exports could be used by bad actors in a nuclear, chemical, biological, or missile program introduces an ethical/moral obligation that goes beyond threatening punishment for breaking an administrative obligation.

In the context of emerging technologies whose military end-use is not always concretely known and therefore under a clear-cut export license obligation, the role of outreach in focusing on potential malicious use is still key and one of the most fundemental recommendations suggested by the report authors. For AI/ML, exporters represent not just large multinational companies who may already be aware of catch-all controls and other effective compliance practices, but also small and medium-sized companies, research institutes and academia, and smaller, informal groupings in maker, DIY, and other communities.

Communication and relationship-building between regulators and AI/ML technology holders and developers is perhaps the most critical activity at this stage of technology development. Without concrete regulations managing transfers of AI/ML technologies, the onus to act responsibly falls directly on exporters. Efforts by regulators to increase self-policing through reinforcement of ethical standards in the AI/ML community will serve to increase compliance and minimize the possibility of bad actors acquiring sensitive technologies for malicious end-uses.

Both sides can benefit, as described below:

[101] "A Resource on Strategic Trade Management and Export Controls: Controls Tangible/Intangible," United States Department of State, <https://2009-2017.state.gov/strategictrade/practices/c43180.htm>.

Benefits to Regulators

- Inform government authorities of the latest AI/ML technological developments and state-of-play;
- Provide information regarding the composition of those in the field. As outreach efforts are grounded in a mapping of the national technological base in order to determine to whom to conduct outreach, the outcome will be an accurate and constantly updated view of AI/ML sector, which can also show industry trends and development. This information is useful not just for security, but technological competitiveness vis-à-vis other countries;
- Develop trust and open lines of communication between regulators and the AI/ML sector. Similarly, as for controlled and listed items, a strong relationship between regulators and exporters means that exporters have more incentive to be compliant as well as ask questions and report suspicious behavior without fear of getting punished or receiving increased attention;
- Develop more nuanced and realistic understanding of the effect of potential controls over specific technology areas, and how they may affect industry/research R&D and output;
- Better prepare regulators to handle license applications in the case that controls over specific AI/ML technologies are eventually introduced;
- Outreach can help strengthen norm-setting regarding uses and strengthen ethical standards.

Benefits to Technology Holders – Current and Potential Exporters

- Technology holders can be better informed and aware of the potential security-related applications of their technologies;
- Technology holders can contribute to the design and implementation of any regulations;
- Technology holders can be better prepared to comply with any eventual export control authorizations necessary or can spot situations where catch-all controls may apply;

- Attending events can create networking opportunities for technology holders and learn about how other organizations are handing emerging technology challenges.

Key in any outreach effort is the focus not just on large multinational companies but on the multitude of smaller players in the field. In the United States, Technical Advisory Committees (TACs) advise the Department of Commerce on the technical parameters for export controls applicable to dual-use commodities and technology and on the administration of those controls. In 2018, the Department of Commerce established a TAC on Emerging Technologies (ETTAC) composed of academia, industry, National Laboratories, and United States government departments and agencies. This TAC should include diverse voices from the United States AI/ML community – not just leading United States multinational companies, but also small and medium-sized enterprises, university research departments, and even, potentially, members of the maker community. Similar forums can be replicated in other contexts and countries in order to create a bridge between the private and public sector when formulating policy and forecasting future policy needs challenges.

IV.II Interagency Coordination

The ways in which emerging technology sectors evolve are of interest to a multitude of government agencies for reasons not just relating to potential security risks, but also of economic competitiveness, absorption of specific applications by the national defense establishment, and more. AI/ML-related policy and research activities are spread across many government agencies with few points of integration at the technical level, although the government has launched efforts in four areas under a broad Executive Order and amalgamated these initiatives on a public website: "AI for Innovation," "AI for Industry," "AI for the American Worker," and "AI with American Values."[102]

With many agencies within government departments working on AI/ML issues, there is a risk of duplicating efforts, producing counterproductive results, and not having a common approach using all available information and resources to produce effective policies.

[102] "AI for American Innovation," United States White House, <https://www.whitehouse.gov/ai/ai-american-innovation/>.

This report, therefore, recommends stronger coordination at the agency working level to share information and coordinate efforts in the AI/ML realm, particularly at the policymaking, research, and outreach level.

For example, the United States Department of Defense has several agencies that play different roles in developing, promoting, and harnessing AI/ML. The Joint Artificial Intelligence Center (JAIC) was formed in 2018 as an Artificial Intelligence Center of Excellence providing expertise to help the Department "harness the game-changing power of AI."[103] The Center conducts outreach to academia and industry with the goal of identifying partnerships and is part of a larger effort, Project Maven, whose goal it is to use AI/ML algorithms to "turn the enormous volume of data available to DOD into actionable intelligence and insights."[104] In 2019, the DOD adopted an Artificial Intelligence Strategy, and in 2020, adopted AI Ethical Principles.[105] The DOD also has other agencies, such as the Defense Technology Security Administration (DITSA), DARPA, and the Defense Innovation Unit (DIU), working on AI/ML issues. While some, such as JAIC, are focused on how to integrate AI/ML innovation, others focus on maintaining the edge on AI/ML and tracking potential threats to security from AI/ML being integrated for military purposes outside of the United States.

Many other agencies similarly follow AI/ML for various reasons: within the Department of Energy, State, Commerce, Homeland Security, etc. The coordination between different agencies with varying degrees of expertise of AI/ML could be strengthened. The expanse of AI/ML policy-making warrants sharing of information between different initiatives where the information could be used for a variety of policy objectives.

Increased coordination could be facilitated through regular meetings and dialogues (both informal and formal), and these meetings should include interagency stakeholders working on industry outreach and technical experts specializing in different emerging technology areas.

[103] "Joint Intelligence Center," <https://www.ai.mil/index.html>.

[104] "Project Maven Industry Day Pursues Artificial Intelligence for DOD Challenges," United States Department of Defense, October 27, 2017, <https://www.defense.gov/Explore/News/Article/Article/1356172/project-maven-industry-day-pursues-artificial-intelligence-for-dod-challenges/>.

[105] "DOD Unveils Its Artificial Intelligence Strategy," United States Department of Defense, <https://www.defense.gov/Explore/News/Article/Article/1755942/dod-unveils-its-artificial-intelligence-strategy/>; "US DOD Announces Adoption of AI Ethical Principles," Army Technology, February 25, 2020, <https://www.army-technology.com/news/us-ai-ethical-principles/>.

To avoid the duplicative efforts we warned of, these forums could also be a chance for agencies to share both results of ongoing engagements and contacts in a diverse set of industry communities.

For example, when JAIC conducts industry and academic days, their objective may be to form partnerships for the DOD to be able to incorporate the latest technology in their operations. These events could also be useful for a wide variety of purposes: identifying the structure of AI/ML initiatives - as in whether they are start-ups, number of employees, foreign investment, etc. - the kinds of technologies they are focusing on, their competitors, their geographic spread across the country, and more. This data is useful, for example, to the Department of Commerce's Technical Advisory Committee on Emerging Technologies, as well as BIS' overall task of identifying possible controls on emerging technologies such as AI/ML or target exporter outreach.

There are inevitably other agencies similarly interested in that kind of information, or may have their own useful informaton to share. Therefore, coordinated information-sharing and establishing effective interagency processes and systems should be a priority in the overall effort to balance the benefits and risks from AI/ML.

IV.III Enforcement and Licensing of Catch-All

While front-end efforts such as outreach can play a role in minimizing the risks posed by emerging technologies, policymakers should also prioritize the role of enforcement, especially with an emphasis on enforcement of catch-all violations. Given that most AI/ML-related goods and technologies will not end up on a control list in the short-term, it will be increasingly important to use information gathered through interagency communications – and to a large extent intelligence - to identify security threats relating to intangible or tangible transfers and act on them.

In addition, licensing departments should emphasize that certain end-uses can fall under catch-all controls and therefore require a license – and that exporters that do not request authorization be punished. While the catch-all mechanism inevitably causes uncertainty for exporters, entrustment of judgement to them to make the right decision in the case of a potential catch-all case is more desirable than a new control that can impact trade to a larger degree.[106]

IV.IV Investment Controls

Like export controls, foreign direct investment (FDI) controls seek to prevent the acquisition of strategic tangible and intangible assets by foreign entities. Until recently, national security reviews of FDI were focused on "traditional" resources (e.g., defense sector and key infrastructure). With the rapid emergence of China as a technology disruptor and the increased awareness that technologies such as AI/ML can be harnessed for offensive purposes, many governments are rapidly adjusting the scope of and constituent definitions for restricted sectors.

In the United States, the Foreign Investment Risk Review Modernization Act (FIRRMA) passed as part of the 2018 NDAA. FIRRMA reforms the Committee on Foreign Investment in the United States (CFIUS) process currently used to evaluate and address national security-related concerns related to foreign investment into the United States.[107] FIRRMA's most substantial change was to the scope of "covered transaction," which defines much of CFIUS's jurisdiction, to include "critical technologies." As defined in ECRA, critical technologies include "emerging and foundational technologies." Therefore, the United States has essentially merged its export control and FDI review regimes into a seamless front against the acquisition of emerging technologies.

The expanding scope trend is not unique to the United States, however. For example, the Organization for Economic Cooperation and Development (OECD) notes:

> "Over the past ten years, a number of countries have introduced for the first time or significantly amended policies specifically tailored to address national security concerns stemming from foreign investment. This policy making activity is in part driven by a re-evaluation of what national

[106] This sentiment reflects views from several company compliance professionals from the first dialogue on emerging technologies organized by STRI, CISSM, and the Stimson Center in March 2019.

[107] For additional information on FIRRMA and CFIUS reform, see Stephanie Zable, "The Foreign Investment Risk Review Modernization Act of 2018," Lawfare, August 2, 2018, <https://www.lawfareblog.com/foreign-investment-risk-review-modernization-act-2018>.

> security encompasses and in which ways it can be threatened. This re-evaluation has broadened the scope of sensitive sectors."[108]

Much of the AI/ML ecosystem is powered by venture capital and other transnational equity investment activities. To the degree that governments adjust their FDI review concepts and procedures, they will likewise have to manage economic expectations and requirements. Nevertheless, investment controls are a powerful and, to date, seldom deployed policy means to address acquisition risks.

IV.V Research Criteria

Previous efforts to assess and mitigate the potential for AI/ML-related technologies have emphasized the need for those engaged in AI research to adopt new approaches to evaluating their work (and the resulting capabilities) for potential misuse or harmful application and possibly to restrict its transfer. Recognizing the need for researchers to play an active role in ensuring the responsible use of the technology they help to develop is an important step in preventing the development and dispersion of specific technologies and capabilities, as no such standards yet exist.

Developing and implementing criteria and methods by which to assess fundamental and applied AI research for societal impact and to restrict its transfer is a far more fraught proposition. The incentive structures for the involved parties--for individual researchers, commercial entities, supporting governments, oversight bodies, etc.--make the prospect of influencing the course of research and commercial application, and the distribution of scientific findings, daunting. For instance, a researcher or commercial entity could downplay the potential harm that could be caused by a machine learning algorithm with the hope of benefiting from the profound good that the same technology could lead to. Similarly, it would be easy for an oversight body to be overly cautious in its review of potential technology applications and impose conditions on research that limit its positive or neutral contributions.

[108] Wehrlé, F. and J. Pohl (2016), "Investment Policies Related to National Security: A Survey of Country Practices", OECD Working Papers on International Investment, No. 2016/02, OECD Publishing, Paris, <https://www.oecd-ilibrary.org/finance-and-investment/investment-policies-related-to-national-security_5jlwrrf038nx-en>.

The generally accepted scientific ethos that research findings should be shared widely to advance public understanding and the application of scientific knowledge for good is also a considerable hurdle to oversight in research enterprises.

The report authors recommend widely accepted and applied standards that set a baseline for the minimum level of consideration and evaluation to which AI research and applications ought to be subjected. Rather than preventing all potential for harmful application or misuse, which would be neither possible nor desirable, a system of research oversight should aim to encourage a minimal level of careful consideration that most, if not all, potential technologies and application would be required to meet. Such a system would resemble the institutional review requirements (that aim to protect human subjects and vulnerable populations) of most research entities more so than existing nonproliferation measures, such as export controls or safeguards.

Such a process of technology review would create greater transparency about work that is under development and signal the thoughtfulness with which researchers are pursuing their work, without compromising confidentiality or sharing trade secrets. Indeed, an optimal research review process could serve as a “pre-patent” filing, a “down payment” on securing the protections on technology development that most researchers desire. An appropriately calibrated review system would also allow for additional levels of review, should that be requested or deemed necessary, and allow for the development of additional governance mechanisms--including strategic trade controls.

For AI/ML research of concern, which could have military applications or potentially cause societal harm, limited constraints on the transfer of related data and findings--similar to those adopted by the biotech community and certain journals--could minimize the potential harm caused by publication or dispersal of some AI/ML research.

If a research review and restraint system is to function as a necessary gateway in the development of AI-related technologies and applications, then it is important for it to have certain design characteristics:

- It should be simple and efficient for AI researchers to submit research into the review process;

- It should not be dependent on AI researchers identifying potentially harmful uses of the research but instead require description of research intent and methods;

- The data collected should include information about high-level intentions--what problems the research is intended to solve and the general approach being taken in technology development--as well as information about specific technological approaches that distinguish the research from other work in the field. Governments should also explore ways of making this information available in the form of datasets online;

- The data collected should include information about enabling technologies or limiting factors to research and development;

- The imposition of restrictions on the transfer of related data or research should be limited to cases where there is a clear potential for malicious use or accident. This and other criteria for restriction should be developed in concert with and endorsed by the appropriate academic, research, and industry associations.

IV.VI Development of Norms

The development of a research review system would be an important step in setting norms of behavior for researchers and entities involved in the development of AI technologies and their integration into specific applications. But additional norm development efforts will be needed to limit the ability of AI-related technologies to cause harm.

The Defense Innovation Board released a report on the "Ethical Use of Artificial Intelligence by the Department of Defense" in October 2019 aimed at laying the groundwork for the necessary military norm development.[109] While all of the report's recommended ethical principles are valuable to explore in regards to defense technology development and use, one in particular warrants additional emphasis: the goal that the Defense Department aim to make its use of AI systems "governable." Defense AI systems "should be designed and engineered

[109] U.S. Defense Innovation Board, "AI Principles: Recommendations on the Ethical Use of Artificial Intelligence by the Department of Defense," 2019, <https://media.defense.gov/2019/Oct/31/2002204458/-1/-1/0/DIB_AI_PRINCIPLES_PRIMARY_DOCUMENT.PDF>.

to fulfill their intended function while possessing the ability to detect and avoid unintended harm or disruption, and for human or automated disengagement or deactivation of deployed systems that demonstrate unintended escalatory or other behavior," the report states.

This goal suggests that in addition to allowing for human intervention in an automated process, an AI system can have capabilities that allow it to "detect and avoid unintended harm or disruption." In other words, AI systems should be designed and developed so that they have the inherent capability to recognize unintended effects and do something about it. Because no machine is yet capable of making the complex and nuanced judgement about whether an unintended function is directly harmful or disruptive—or could initiate a sequence of events that could prove harmful or disruptive--fulfilling the goal of governability for all AI applications (not just defense-related applications) would require that an AI system be able to communicate with human operators about *all* unintended outcomes.

This might seem like a monumental task considering that AI/ML systems are valued precisely because they can operate autonomously under varied and unpredictable circumstances. But if understood as a prompt for developers and users of all AI systems to have a clear understanding of what the system is intended to do and to ensure that the system is capable of detecting and avoiding anything that is outside defined operations, then this goal could be particularly useful in helping human operators and regulators identify and mitigate the hard-to-define risks of specific AI/ML technologies and capabilities. In the end, the suggestion that AI-systems be "governable" could be an essential element of developing the broader means to avoid the spread of AI/ML technologies and applications that could cause deliberate or unintended harm.

The process of defining and ensuring the adoption of ethical principles, like governability, will lead to a deeper exploration and understanding of the harm that could be caused by AI systems. It will also help to avoid the development of systems whose capabilities could veer widely from a prescribed purpose and lead to AI development approaches that limit the potential for unintended effects. Adopting this type of norm-setting approach to technology development could have profound, cascading effects on the field—allowing for the unfettered use of beneficial technologies and avoiding potential harm.

IV.VII Private Sector Self-Policing/Role of Competition

In many industries, companies and research organizations have developed compliance best practices and self-monitoring protocols for a variety of legal or ethical requirements. For example, the International Forum of Sovereign Wealth Funds (IFSWF) established the Santiago Principles which promote transparency, good governance, accountability, and prudent investment practices whilst encouraging a more open dialogue and deeper understanding of SWF activities. Similarly, in the absence of explicit laws or norms, AI developers could be encouraged to develop *de facto* best practices on AI controls.

Increased competition over AI/ML business gains may also organically incentivize the private sector to self-police/self-regulate the export market through the establishment of in-house best practices, intra-industry trading norms, or through fostering the development of a more severe IP/Patent regulatory environment. With respect to individual business decisions on the export of technologies, companies may be compelled to give more scrutiny to recipients of certain technology exports.

One potential pathway to private sector self-regulation would be the decision of firms to selectively export based on a perceived market advantage of keeping certain technologies, processes, and information in-house. Another potential pathway to self-regulation would be the establishment of norms within certain industries to ensure that technologies are narrowly applicable to certain applications within that industry only. This could be driven if industries perceive that they will be able to maintain an advantage from universal AI/ML developers by producing highly specified technologies that are superior for certain applications. However, this also would depend on the superiority of this market strategy over that of universal AI/ML developers. Finally, private industry could receive business incentives to bolster IP/patent legal institutions, or to only export to countries that maintain strict IP standards/policing.

There is a lot of interest and uncertainty over this avenue for AI/ML export regulation development. Due to the concern that government-imposed regulations may unnecessarily disadvantage private industry innovation and profit compared to the global market, there is a growing contingent that believes self-regulation by private industries could be the best solution. However, due to the external development of these practices, should they come to fruition, there is little certainty over how compatible their developments will be with government security export goals.

IV.X Targeting Intangible Transfers of Technology (ITT)

Establishing enhanced controls on the access that certain individuals or groups of individuals can exercise with regards to specific sensitive sectors is a commonly used nonproliferation strategy. While controls on so-called deemed exports do not exist worldwide, most countries do have controls on intangible technology transfers (ITT), and many do screen foreign nationals who come to their countries to conduct studies or research in sensitive domains. In the United States, rigorous export controls extend to foreign nationals and an authorization is necessary in order to share controlled information with them. While the policy is generally clear in the case of listed items, for non-listed technologies that could have security-related end-uses, the issue of deemed exports, or from a broader perspective visa vetting, is less clear and certainly more controversial.

With regard to deemed exports in the United States, a license would be required to share information regarding "emerging technologies," presumably, if a) the technology was within the scope of existing control list parameters b) there would be a risk of the technology falling under catch-all controls c) the technology would be shared with a restricted end-user. Because in the case of AI/ML, the latter two would be the most likely scenarios (in the absence of existing broad controls on AI/ML), it would be wise to increase awareness of catch-all controls and sanctions, and how they could apply to AI/ML, in research and scientific communities. Therefore, the authors recommend outreach on catch-all controls, deemed exports, ITT, and restricted end-users – and especially emphasize these topics for the AI/ML research and maker communities.

Visa vetting poses a more difficult challenge. As stated in a significant number of stakeholder responses to the Department of Commerce's ANPRM, companies, research organizations, and academia see foreign nationals as essential to the United States' leadership and realized potential in AI/ML sectors. Yet, valid concerns exist among the policy community regarding the theft, exploitation, or malicious use of valuable technology by foreign nationals with ill intent. The question hinges, therefore, on analyzing the cost/benefit - indeed, the risk assessment in both cases - of a stringent visa vetting policy on foreign nationals coming to work or study in the U.S in AI/ML sectors.

IV.XII Technology Tracking

Policymakers and government leaders will also have to determine the amount of resources they will invest in tracking AI/ML technology use, to the extent that it is allowed to propagate freely. Depending on the extent to which policymakers allow AI/ML to be disseminated, they may wish to establish methods to track usage of the technologies, including who is using the technologies and to what activities they are applying the technologies (and may require exporters to do this work). With respect to actors, policymakers may be interested in the division between governmental, military, and civilian use of a certain technology. With respect to application, policymakers may be interested in determining the frequency of use (e.g. continuous vs instantaneous), as well as the specific applications of the technology.

A number of approaches could achieve a base level of technology tracking but will require early investment in data collection/processing capabilities and infrastructure development. A central decision will be how data is collected on AI/ML use. Will governments use secretly/covertly acquired data, open/transparent data, or leverage data collected by private industry? With respect to the type of information acquired, will this intelligence seek to identify computing power, type/architecture of technology, data processing speed/volume? In determining what the data will be used for, the government will need to identify what types of infrastructure will need to be built to acquire data that cannot be done with industry buy-in. This will likely have unilateral and multilateral dimensions, where global transparency norms and agreements could be instituted to fill in gaps in unilateral data acquisition capabilities. Ironically, AI/ML itself will likely have to be applied for unilateral data acquisition. One thing is certain, considerable international cooperation will be needed to implement any technology tracking system.

V. Outlook

The existing and potential security risks of AI/ML technologies and future research, development, and deployment should be urgently addressed. While the growth and diversification of AI/ML technology and integration over the coming decades will likely lead to even more economic and social efficiencies and benefit health and well-being, these developments are also likely to increase the potential for harmful use.

In the meantime, the tools available to limit or mitigate risks of AI/ML technology development and use will also change over time, with some of the policy options listed in this report becoming infeasible and others becoming more attractive. The more investment in technology and integration of capabilities, and the more scientists and engineers involved in the enterprise, the harder it could be to track AI/ML developments, identify specific uses and applications, monitor investments and movements of people, organize effective outreach, or develop effective norms of research or use. In particular, the foreign availability of specific technologies and capabilities and the size of the global marketplace for specific legitimate goods could limit even further the potential for some forms of trade control. As limited as the imposition of list-based controls for certain emerging technologies that rely on readily available information technologies or intangible technologies is at present, list-based controls could become even more irrelevant.

As advances in AI/ML technologies and uses continue, there could be new technologically-savvy options for risk mitigation. Some of these approaches could even benefit from the very AI/ML advances in capability that regulators seek to monitor. For example, AI/ML capabilities could make it more cost-effective to monitor research trends and stay up-to-date with technological and capability development. It is also possible to imagine these capabilities being used to identify suspect suppliers or users of goods, even if the goods themselves are not controlled; or to increase the review throughput of trade-related documents.

What policy options need to be pursued in the short-term, if they are to be pursued at all? Which could have a multiplying effect on risk mitigation efforts? And is it possible to affect the developmental drivers of those AI/ML capabilities and applications that will have the greatest impact on the governance options available to officials and non-governmental parties?

Use them or lose them.

Officials are tempted to impose list-based controls on specific emerging technologies or to restrict specific users and uses of exported technologies now, because of the likelihood that it will be more difficult to do so over time as the foreign availability of specific technologies and capabilities and the profusion of the workforce capable of contributing to AI/ML development expands even further. There is some truth to this argument, as controlling exports on technologies and capabilities that are widely available for export from other countries weaken the effectiveness of those very controls - and potentialy drive research, development, and manufacturing capacities overseas as businesses seek out environments where they can freely participate in the global marketplace. This should not be read as an endorsement of list-based controls to stem the proliferation of potentially harmful technologies. Instead, it should be seen as an explanation for why the rush to mitigate the risks of emerging technology proliferation is currently so skewed toward an emphasis on list-based controls.

If norms are to further emerge to set specific expectations about how AI/ML development and use should proceed globally, then they need to be set early and often. It is difficult, if not impossible, to formulate a specific norm of behavior if a different pattern of behavior has already been established. Imagine convincing private and public research organizations to stop specific types of research, production, and exports that they have already been doing for years and that are likely to continue in other parts of the world. It was possible to do this in the case of biological and chemical weapons that had been developed and used for decades before a norm of non-development and use was introduced and adopted. But these norms were established as part of multilateral processes that resulted in specific legally binding prohibitions—albeit with limited verification and enforcement measures. A prohibition on AI/ML development and use is neither possible nor desirable, and norms of development and use that would affect specific underlying components, technologies, and intangible technologies would be difficult to negotiate in a timely fashion in a multilateral context.

The development of norms and ethical standards promises to be an important component of any attempts to mitigate the risks from AI/ML development and use, primarily because the causes of potential harm are so undefined and emergent that an underlying set of principles is needed to enable adaptation of risk mitigation over time. As such, it is important to build a process parallel to technology development that ensures norm development keeps pace and that opportunities for the negotiation of multilateral, legal restrictions on specific capabilities can be identified and pursued before they are past due.

As part of the norm development process, it has been suggested that the entities involved in AI/ML be conscripted into the process of mitigating risk. While it is essential for researchers, industry, and internal oversight bodies to be involved in the adoption and strengthening of specific norms of behavior, including those related to research oversight and product development, it is unlikely that such "self-policing" is a sustainable path to risk mitigation. Not only are economic incentives misaligned in the short run, but unless such efforts are part of a systematic, multinational effort (say, as part of an international professional association), self-policing will be too uneven and idiosyncratic to provide much confidence that even simple norms of behavior are being adhered to.

Multiplying effect.

When anticipating the maturation and proliferation of AI/ML capabilities, it is valuable to consider what risk mitigation policy options could have a positive and multiplying effect on other risk mitigation efforts. By developing those options and capabilities in the short run, it might ease the adoption of additional risk mitigation efforts or obviate the need for others. For example, governments and private-sector actors could act now to increase their abilities to track AI/ML investments and research and to understand the implications of specific AI/ML capabilities and applications—as, or even before, they are fully developed. Doing so would extend the time and scope within which the above referenced policy options could be relevant.

If governments and nongovernmental organizations are able to be somewhat or completely transparent about their investments and goals, it might also reduce overreactions and destabilizing competition, particularly between government efforts to develop and integrate specific AI/ML military capabilities. For example, if the United States government better understood how the Chinese military is investing in and integrating specific AI/ML capabilities into its command and control systems, military planners would have greater certainty about when and where the United States could adjust accordingly, rather than basing planning decisions on worst-case scenario thinking.

Having more clarity into the composition and intention of foreign AI/ML investment and development could also help tremendously in the processes of research oversight and norm development. If a national government can develop as accurate an understanding as possible of its own vulnerabilities to specific AI/ML related risks, and the risks are shared by other governments, then these governments might be more willing to engage in international norm development or even negotiate specific legal constraints on capabilities. If greater research and investment reporting and transparency increased confidence in participants' willingness to abide by those constraints, then tracking and disclosing these types of data could have even broader effect on risk mitigation efforts.

Effects of circumstance.

It is fair to assume that demand for specific AI/ML capabilities and applications will be the major force in determining the scope of technological development and therefore technological risk. Yet governments, governmental organizations, private-sector actors, and scientific and expert communities will be reticent to temper demand for AI/ML innovation—in general or in terms of specific applications—even if they anticipate potential harmful uses. The uncertainties in what constitutes potentially harmful technologies are sufficient enough that entities will hesitate to forgo or prevent many, if any, forms of research and development in the fear that they will miss out on potential benefits or be outdone in development by others.

Circumstance, however, could affect the shape and intensity of demand for technological development. For instance, the current global context—a disease pandemic—could lead policymakers to become aware of or reframe the benefits and risks of unconstrained international competition to develop and use AI/ML technologies with little thought about the downstream effects on security. Global connections between commerce, social well-being, and national security are front and center as governments, nongovernmental, and international organizations attempt to limit disease spread, mitigate the effects of infection and panic on populations and economies, and develop scientific and social systems to limit its future spread. While the pandemic has heightened political tensions between major economies—particularly between the United States, Europe, and China—in the short run, the long-term impact could be greater coordination and cooperation to stave off future disease outbreaks and mitigate their effects.

Similarly, if and when governments realize that they have limited abilities to prevent intentional and accidental harm from some civilian and military AI/ML technologies and applications, then they are likely to seek ways to mitigate, rather than eliminate, technological risks. The various discussions and efforts to develop AI/ML governance mechanisms referred to or put forward in this report suggest that governments and civil society have reached this point. More likely, however, additional harmful accidents or uses of these capabilities will be necessary to orient policymakers in this way. Acceptance of vulnerability does not sit well for most governments or officials. When this realization sinks in, though, demand for relatively unconstrained innovation and integration of AI/ML technologies could diminish. When this happens, policymakers and expert communities will be called upon to implement the types of trade controls and governance measures presented here. The time to fully think through likely pathways of technology development, proliferation, and risk mitigation is now.