NOVEMBER 2018

ARTIFICIAL INTELLIGENCE AND NATIONAL SECURITY

THE IMPORTANCE OF THE AI ECOSYSTEM

PROJECT DIRECTOR

ANDREW P. HUNTER

LEAD AUTHOR

LINDSEY R. SHEPPARD

CONTRIBUTING AUTHORS

ROBERT KARLÉN

ANDREW P. HUNTER

LEONARDO BALIEIRO

A Report of the CSIS Defense-Industrial Initiatives Group

CSIS | CENTER FOR STRATEGIC & INTERNATIONAL STUDIES

NOVEMBER 2018

ARTIFICIAL INTELLIGENCE AND NATIONAL SECURITY

THE IMPORTANCE OF THE AI ECOSYSTEM

PROJECT DIRECTOR

ANDREW P. HUNTER

LEAD AUTHOR

LINDSEY R. SHEPPARD

CONTRIBUTING AUTHORS

ROBERT KARLÉN
ANDREW P. HUNTER
LEONARDO BALIEIRO

A Report of the CSIS Defense-Industrial Initiatives Group

CSIS | CENTER FOR STRATEGIC & INTERNATIONAL STUDIES

ABOUT CSIS

FOR OVER 50 YEARS, the Center for Strategic and International Studies (CSIS) has worked to develop solutions to the world's greatest policy challenges. Today, CSIS scholars are providing strategic insights and bipartisan policy solutions to help decisionmakers chart a course toward a better world.

CSIS is a nonprofit organization headquartered in Washington, DC. The Center's 220 full-time staff and large network of affiliated scholars conduct research and analysis and develop policy initiatives that look into the future and anticipate change.

Founded at the height of the Cold War by David M. Abshire and Admiral Arleigh Burke, CSIS was dedicated to finding ways to sustain American prominence and prosperity as a force for good in the world. Since 1962, CSIS has become one of the world's preeminent international institutions focused on defense and security; regional stability; and transnational challenges ranging from energy and climate to global health and economic integration.

Thomas J. Pritzker was named chairman of the CSIS Board of Trustees in November 2015. Former U.S. deputy secretary of defense John J. Hamre has served as the Center's president and chief executive officer since 2000.

CSIS does not take specific policy positions; accordingly, all views expressed herein should be understood to be solely those of the author(s).

ACKNOWLEDGMENTS

This report was made possible by the generous support of Thales USA, Inc.

Center for Strategic & International Studies
1616 Rhode Island Avenue, NW
Washington, DC 20036
202-887-0200 | www.csis.org

TABLE OF CONTENTS

ABSTRACT

ARTIFICIAL INTELLIGENCE (AI) has profound potential to affect the balance of power in both the global economy and in military competition. While AI has a long history, AI has begun to deliver results within the last decade, particularly with the recent rapid progress in machine learning and the increased availability of data and computing power. As impactful as the recent progress has been, AI remains highly problem-specific and context-dependent. It has proven extremely challenging to translate the progress in some fields to others, even those that are closely related. Enthusiasm in both the public and private sectors has obscured the importance of building the robust supporting capabilities for AI—an AI ecosystem—that are crucial to successful AI adoption. The AI ecosystem includes a skilled workforce and knowledgeable management; the digital capability for capturing, handling, and exploiting data; the technical foundation of trust, security, and reliability; and the investment environment and policy framework needed for AI to flourish. The government retains role in pursuing the harder areas of technology that do not deliver rapid returns on investment for the private sector; developing the tools required to establish AI reliability (including trust, explainability, validation, verification, and security) for critical government and national security applications; and developing and strengthening the AI ecosystem. This study presents the key steps to be taken to facilitate the successful integration of AI into national security applications based on an accurate understanding of where the AI field currently stands and what key factors are involved in successful AI adoption and management.

CH. 1

INTRODUCTION

ARTIFICIAL INTELLIGENCE (AI) has profound potential to affect the balance of power in both the global economy and in military competition. AI has a long history as a field with great promise, but for much of this history, AI's potential was unrealized. In the last decade, AI has begun to deliver results, particularly with the recent rapid progress in machine learning and the increased availability of affordable computing power. The enthusiasm this progress has generated is warranted in many respects. Yet it may also be obscuring some of the more important realities about AI and its implications for national security.

The recent progress of machine learning has created the impression that AI has fully arrived, focusing attention on how to control the technology while obscuring how early-stage and problem-specific the field remains. It has highlighted the tremendous capabilities of private sector companies to drive the field forward, but it has also obscured the criticality of government investment and involvement in its future success and in its national security applications. And it has obscured the importance of building robust supporting capabilities for AI—an AI ecosystem—that are crucial to successful AI adoption. The AI ecosystem includes a skilled workforce and knowledgeable management; the digital capability for capturing, handling, and exploiting data; the technical foundation of trust, security, and reliability; and the investment environment and policy framework needed for AI to flourish.

After decades of AI research being primarily the domain of government-funded researchers making slow progress, investment in AI has exploded since 2010. The vast majority of this recent increase in investment has come from private sector companies, particularly tech giants in the United States and China. Though the development of AI remains at a fairly early stage, this investment boom has paid off dramatically in areas such as speech recognition, image recognition, translation, and complex game play, and it is contributing to changes in the economy as AI increases productivity. As impactful as the recent progress has been, it remains highly problem-specific and context-dependent. It has proven extremely challenging to translate the progress in some fields to others, even those that are closely related.

Commercial firms can meet their business objectives by focusing on the techniques and problems where AI progress is most rapid, ignoring or biding their time in investing in those where much less progress has been made. While it is clear that commercial investment will continue to push the AI field forward, this investment is far from sufficient. The government retains a critical role in AI investment. This role includes pursuing the harder areas of technology that do not deliver rapid returns on investment for the private sector; developing the tools required to establish AI reliability (including trust, explainability, validation, verification, and security) for critical government and national security applications; and developing and strengthening the AI ecosystem.

While AI is something of a buzzword that promoters of all kinds like to advertise as something they are not only pursuing but implementing, it can be surprisingly

hard to get agreement among a group of experts on just how much "real" AI has actually been adopted in both the commercial and public sectors. In part this is due to definitional confusion about the AI field, and it is caused in part by the increased sensitivity to AI that results from predictions of its potential impact on human lives and livelihoods. For the public sector, another factor is the government's ongoing struggles with acquiring developmental software. All of these issues play a part, but a few key issues are most fundamental to the success of AI adoption.

For many potential AI users, there are two outstanding debts to be paid before successful AI adoption is likely. The first is workforce debt—a past failure to attract and retain the technical and management talent within the organization to successfully develop and implement AI in its systems. The second is technical infrastructure debt—the weakness of the organization's digital capability, i.e., its data and its computing and networking capabilities. Paying down these twin debts is critical to successful AI adoption. For the U.S. government, and particularly for the Department of Defense, these debts are major barriers to AI adoption.

It is also critical to understand that AI remains highly problem-specific and context-dependent. This means that AI performs narrow tasks and is embedded in larger systems where its impact can be hard to see. As an early stage technology, the actual improvement in capability delivered by AI can be marginal. This means that eager AI adopters are confronted with large upfront costs and often meager initial results.

Not only is AI hard to implement, it presents significant management challenges to any organization that seeks to harness it. Many AI users, particularly those whose missions involve substantial exposure of risk to human life or costly equipment, will require a high threshold of AI reliability before they truly commit themselves to depending on AI for the success of their missions. In the private sector, there are many under-explored legal issues associated with liability and intellectual property, and in the public sector, there are a profusion of critical missions where there is no clear path to establishing sufficient AI reliability. Successful AI management is likely only after a robust supporting AI ecosystem develops that can satisfy the majority of AI users in these key respects. While much of the AI ecosystem can and will develop in the private sector, this is a necessary but not sufficient factor for many government users, particularly for national security.

For all the challenges that AI presents, it is critical that the United States step up to the plate with the investment, management focus, and policy work required to succeed with AI adoption and AI leadership. The rest of the world is investing heavily in AI and has some advantages in pursuing the technology that can be met only with a major effort on the part of the U.S. government, working in coordination with private technology companies and in collaboration with partner and allied nations that share a desire to see AI result in more open and democratic societies.

CH. 2

CONCEPTUAL FRAMEWORK FOR ARTIFICIAL INTELLIGENCE APPLICATIONS

FOR THE PURPOSES of the discussions and policy recommendations of this study, artificial intelligence should be understood to mean a purpose-built, problem-specific algorithm or agent. AI is software; it is math and code in "algorithms that make decisions about data"[1] to implement the functionality of cognitive task execution in machines. While discussion of replicating human intelligence has occurred for centuries, including philosophical debates on the nature of intelligence itself, the advent of computing technology led to the modern understanding of AI. That is not to say that AI is "thinking" in a manner consistent with human intelligence or cognition; the way AI performs its tasks may bear no relation to what we would normally recognize as logical thought.

As a term, AI can be semantically problematic. Defining AI in terms of human intelligence sets us up to think of it incorrectly and gives a human-centric sense of what this technology is, suggesting it is focused on replication of human cognition. More generalizable capabilities that can be applied to widely varying problems, in the same manner as human intelligence, is beyond the capacity of today's systems and for most AI researchers is not even a near- or mid-term objective. A more technically relevant understanding of AI is necessary to move beyond the abstract discussion. As AI becomes more capable, it is gaining functionality in uncertain environments by introducing increased flexibility and adaptability to the technology, moving beyond rules-based code.

As it is primarily discussed in this report, AI is often referred to as "narrow AI." The focus on this understanding of AI is not meant to be an argument for or against any other conception of AI, but as a means for establishing the scope of the report's findings and recommendations. As this chapter will discuss, other definitions of AI exist in a variety of contexts. However, this report has chosen to focus on the issues that will be of primary significance to national policymakers and both defense and commercial implementers of AI in the next five to ten years.

Academically, AI is a field of study comprised of various loosely-connected disciplines spanning topics of knowledge abstraction, learning strategies, reasoning domain, and reasoning mechanisms. Texts categorize six disciplines under the AI umbrella:

- "Machine learning to adapt to new circumstances and to detect and extrapolate patterns;
- Natural language processing to enable successful communication in a given language;
- Knowledge representation to store information a machine knows and receives;
- Automated reasoning to use the stored information to answer questions and to draw new conclusions;
- Computer vision to perceive objects;
- Robotics to manipulate objects and move about."[2]

1 "Artificial Intelligence (AI)," skymind, https://skymind.ai/wiki/artificial-intelligence-ai.

2 Stuart Russell and Peter Norvig, "Artificial Intelligence: A Modern Approach," 3rd ed. (Harlow, UK: Pearson Education Limited, 2014).

Encompassing a wide range of tools, AI is narrowly applied math, code, statistics, and probability. Each AI tool and technique maps to different capabilities and functionalities, which may be effective or ineffective at different things. The same algorithm implementing machine learning and computer vision to classify objects may be optimized for one sensor input feed but relatively ineffective at performing the same task interpreting data from another source. Some AI tools, such as the zero-shot learning method of machine learning, are designed to operate in a data-austere environment lacking labeled training data.[3]

A 2018 discussion paper from McKinsey Global Institute, *Notes from the AI Frontier: Insights from Hundreds of Use Cases*, provides descriptions of the types of problems that AI, particularly those which leverage machine learning, is good at solving: anomaly detection, classification, clustering, continuous estimation, data generation, optimization, ranking, and recommendation.[4] For each AI application, the key is matching the AI technique to the task being performed. That said, the study team's conceptual framework of AI is defined in mission-agnostic terms. The study team does not discuss the application of specific algorithms to a specific problem set, mission, or problem-dependent set of criteria. Instead the analysis applies primarily at the level of the broader AI ecosystem, potential problem spaces, and technology research areas in mission-agnostic terms.

Further, it should be noted that along with the problem-specific nature of AI, many problems may also be solved with statistics and math, such as a linear algorithm, or optimization techniques, such as genetic algorithms, that would not be qualified as "artificial intelligence." There are gains to be had from deploying AI in certain contexts that would be lost by using traditional statistical techniques, particularly as a need for algorithmic flexibility and adaptability grows. AI is also distinct from autonomy, which is a description of task delegation. For this study, the team has focused specifically on those issues unique to AI, and not the broader challenge of companies competing on data analytics and data science.

Since AI has value to offer in informing critical decisions, it must be trusted and must be secure. Trust is a key issue that will drive or restrict advances in AI and the adoption and deployment of this technology. Reliability and trust in deploying AI technologies requires an understanding of model transparency, sometimes called "explainability." That is, the ability of a user to understand how and why an algorithm arrived at an outcome. Verification and validation (V&V), simulation, repeated use, and stress testing are mechanisms by which trust is built and model transpar-

3 Yongqin Xian, Bernt Schiele, "Zero-Shot Learning: A Comprehensive Evaluation of the Good, the Bad and the Ugly," *IEEE Computer Vision and Pattern Recognition* (CVPR) (2017), https://arxiv.org/pdf/1707.00600.pdf.

4 Michael Chui et al., *Notes from the AI Frontier: Insights from Hundreds of Use Cases*, discussion paper, McKinsey Global Institute, April 2018, https://www.mckinsey.com/~/media/mckinsey/featured%20insights/artificial%20intelligence/notes%20from%20the%20ai%20frontier%20applications%20and%20value%20of%20deep%20learning/mgi_notes-from-ai-frontier_discussion-paper.ashx.

ency is understood. Feedback can help to mitigate our uncertainty and build trust. However, the importance of trust can be circumvented by need or by layers outside of the AI that can provide a semblance of certainty or comfort.

Data is a valuable, necessary asset for the future of AI, and needs to be protected as such. Particularly in machine learning applications that require training on data sets, data very much impacts the value and quality of the algorithm's output. As the saying goes in computer science, "garbage in, garbage out." However, many national security applications are not data-rich fields. In addition to assuring and securing data in national security applications when we have it, the community must determine ways to leverage other elements to make up for the shortcomings of the data.

Successfully leveraging AI, however, requires more than data and algorithms. A skilled workforce, guided by ethical policies and standards, is necessary to understand the "ins and outs" of the data, the algorithms, and the problems to which AI is being applied. A digital foundation upon which the algorithms and data are built, such as database management and integration, is necessary. We need to consider AI in terms of ecosystems, complex networked supportive systems that include:

- Trusted and secure AI technology;
- The workforces to develop, use, maintain, and regulate it;
- The digital infrastructure and capability that enables AI technology;
- The policies and ethical standards that guide use.

WHY NOW? ADVANCES IN MACHINE LEARNING

The current excitement and drive to go "all in" on AI began with an advancement in the machine learning discipline used to identify patterns in data and calculate conclusions based on those patterns. While these terms are often used interchangeably, machine learning is a discipline within the AI field. Machine learning advances coupled with the availability of data and computing power mean that AI can answer more semantically useful questions at the speed of relevance. Machine learning was advanced and enabled by the increased availability and scalability of cloud computing, which allows AI to process massive data sets and execute algorithms that learn. Many of the current AI solutions implement machine learning along with data analysis, computer vision, or natural language processing to execute pattern recognition, particularly in areas where humans may struggle, such as rolling up various low confidence data streams, finding weak patterns consistently, and grappling with enormous data sets. These machine learning solutions exemplify one of the biggest advantages of AI–doing things humans do not have the time or capacity to do or do very poorly.

HINTON MAKES HEADWAY

Geoffrey Hinton has been at the forefront of AI neural network research for more than 30 years, trying to determine if a computer could process information more like a brain, using intuition instead of hard-coded rules. Brains are composed of networks of cells communicating with each other, and Hinton has been working to recreate this network using computers. In the 1950s, common knowledge said that neural networks were impossible to recreate, and that computers learned best through rules and logic. As proof of this, people pointed to a late-'50s project called the Perceptron, a machine designed to simply recognize images. This machine failed spectacularly, but Hinton continued, citing the ability of the human brain to learn without explicit programming. He believed the Perceptron simply lacked sufficient processing power.

Hinton received a PhD from the University of Edinburgh in 1972, with neural networks as his focus. Afterwards, he got a job at Carnegie Mellon University and later moved to the Canadian Institute for Advanced Research (CIFAR). It was there that he started the Learning in Machines and Brains program, which is where he would end up doing most of his work until the opening of the nonprofit Vector Institute in Canada.

In 2009, when computers finally had the processing ability to rapidly sift through massive amounts of data, neural networks started to outperform logic-based AI and thus caught the attention of U.S. tech giants. They started investing, and Google, when putting YouTube videos through its supercomputer, was eventually able to make the computer recognize a cat as a cat, something they never told it during training. This breakthrough is what put AI and Hinton and his colleagues in the spotlight, launching neural networks into mainstream research.[5] Hinton and his colleagues won the 2012 ImageNet Large-Scale Visual Recognition Challenges (ILSVRC). Their paper titled "ImageNet Classification with Deep Convolutional Networks" presented their breakthrough in machine learning. The convolutional neural network used to classify images achieved test error rates significantly lower than previously demonstrated.[6] The study is considered one of the most influential studies in the field of machine learning.

5 Katrina Onstad, "Mr. Robot," *Toronto Life*, January 29, 2018, https://torontolife.com/tech/ai-superstars-google-facebook-apple-studied-guy/.

6 Alex Krizhevsky, Ilya Sutskever and Geoffrey E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," NIPS'12 Proceedings of the 25th International Conference on Neural Information Processing Systems, vol 1. (December 2012): 1097-1105.

FRAMEWORKS FOR UNDERSTANDING ARTIFICIAL INTELLIGENCE

It is useful to use visual frameworks overlaying the forward march of time, the advancement of AI capability, and specific examples to understand the progression of the AI field. Most frameworks characterize the progress of AI to date in similar ways, beginning with the early rules-based reasoning that saw a period of development from 1989-2007 and progressing to the capability gains of recent years, the machine learning-enabled flexibility and adaptability in uncertain contexts. An example of early rules-based AI is a computer checkers game that plays against human opponents, such as Chinook.[7] The early instantiations of AI "reasoned" over

7 Chinook, https://webdocs.cs.ualberta.ca/~chinook/play/.

a very well-defined, narrow context using hardcoded rules and patterns and mathematical optimization techniques–the equivalent of solving a mathematical problem by "brute force." To compete against humans in checkers, the researchers programming Chinook tackled a search space of $5x10^{20}$ and leveraged hard-coded "tactical tables" capturing patterns in game play.[8] Limitations arise in rules-based reasoning as the design space becomes so large that it is not possible to hard-code the exhaustive solution set and then have a computer search and optimize over that set. The number of possible solution combinations to search, processing power, and processing time becomes infeasible. The evolution of AI agents successfully competing against humans in various board games is a good illustration of this point: in order to beat a human at Go, a game with a search space 10^{100} larger than chess, machine learning was necessary to enable the computer to "learn" from an initial training data set of 30 million moves.[9] Implementing statistical search space exploration through machine learning introduces flexibility and adaptability, allowing AI systems to answer increasingly semantically meaningful questions.

Various frameworks exist for visualizing and exploring the evolution and increasing capability of AI. Examples of framework for understanding AI capability are the DARPA Waves,[10] IBM Broad AI,[11] or Dr. Marvin Minsky's Multi-Level Mind (focused primarily on capability and less on time-based development).[12] However, frameworks tend to differ in characterizing the path forward from the statistical, machine learning-based Narrow AI of today. In some, General Artificial Intelligence is painted as a distant theory, the logical extrapolation of where we are today but not something that will be attained any time soon. Some discussions convey that General AI, and the almost always dystopian future associated with it, are closer than we think. Regardless of the debatable inevitability of the end, progression forward in AI has meant moving up the value chain of tasking and problem solving. As a result, AI users expect higher performance. By moving up the value chain of tasking, AI implementations process bigger data sets faster with greater accuracy and precision. This requires computational creativity, flexibility, and adaptability as tasking becomes increasingly context-dependent higher up the value chain. The increasing need for understanding and responding to context is what drives the need for AI to interact with human operators and analysts. What this means for humans is a shift in allocation enabling tech-

8 Jonathan Schaeffer et al., "Chinook: The World Man-Machine Checkers Champion," *AI Magazine 17:1* (1996).

9 David Silver and Demis Hassabis, "AlphaGo: Mastering the Ancient Game of Go with Machine Learning," *Google AI Blog*, January 27, 2016, https://ai.googleblog.com/2016/01/alphago-mastering-ancient-game-of-go.html.

10 "DARPA Perspective on AI: Three Waves of AI," DARPA, https://www.darpa.mil/about-us/darpa-perspective-on-ai.

11 Arvind Krishna, "AI Learns the Art of Debate," IBM, June 18, 2018, https://www.ibm.com/blogs/research/2018/06/ai-debate/.

12 Marvin Minsky, *The Emotion Machine: Commonsense Thinking, Artificial Intelligence, and the Future of the Human Mind* (New York: Simon & Schuster, November 13, 2007).

nical capability growth. By removing the cognitive burden of simple or low-dimensionality tasks, AI enables humans to perform more highly dimensional or complex tasks.

In her 1994 Presidential Address to the Association for the Advancement of Artificial Intelligence (AAAI), Dr. Barbara Grosz highlighted the necessity of collaboration with users as well as AI systems to truly bring a capability to users.[13] Based on the technical maturity of AI and the tasking workload it is assuming, AI may require no interface with humans at all. It may require information transfer from AI to human or vice versa, as feedback impacts the processing and results. Most early AI operates apart from humans or adversarial to humans to maximize or optimize performance given a narrow context and objective (e.g., playing chess or jeopardy). These isolation assumptions result in AI agents "with fixed knowledge and a specified goal" that rely on human agents to formulate a problem statement that "would include background knowledge, a description of the state of some world, operators to use in that world, and a description of a desired state (a goal)."[14] The goal of increasing flexibility and adaptability in AI applications requires collaboration and learning from interactions. Higher capability in AI is progressing with a shedding of the traditional isolation assumptions.

HUMAN-MACHINE TEAMING FRAMEWORK

The shifting the workload between humans and machines changes and transforms the type of human involvement. An example of the transformation of the workload between humans and machines is in the evolution of radar and electronic warfare (EW) systems. Greater functionality of these systems was enabled through automation of the systems maneuvering the electro-magnetic spectrum (EMS), instead of human operators mechanically turning dials and knobs.[15] The value added is where AI and people are not competitors; where humans are freed up to execute more highly dimensional tasks.

"THE VALUE ADDED IS WHERE AI AND PEOPLE ARE NOT COMPETITORS; WHERE HUMANS ARE FREED UP TO EXECUTE MORE HIGHLY DIMENSIONAL TASKS."

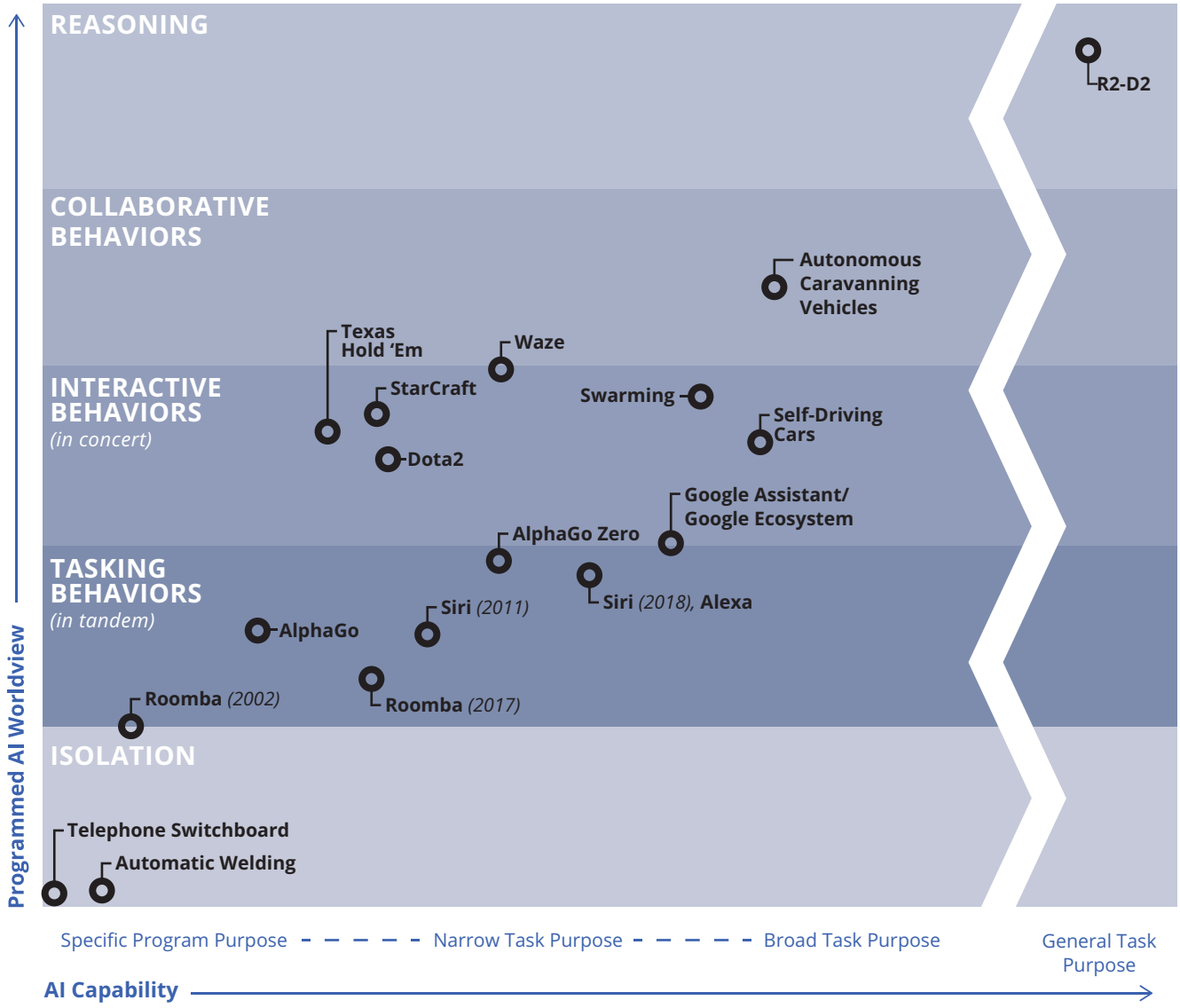
The study team proposes that, from the view of allocating work between humans and machines, AI is playing a significant role by moving up the "value chain" of tasks it can accomplish. Further, AI systems are becoming more collaborative as they assume a greater role in human-machine teaming. While many frameworks present a

13 Barbara J. Grosz, "Collaborative Systems: AAAI-94 Presidential Address," *AI Magazine 17(2)*, (1996): 67-85.

14 Daniel G. Bobrow, "Dimensions of Interaction," 1990, presented at the Association for the Advancement of Artificial Intelligence.

15 John Knowles, "The Pace of Change," *Journal of Electronic Defense* 37 (10), Oct. 2014, 6.

HUMAN-MACHINE TEAMING FRAMEWORK



visual representation of the progress of AI capability, the concept of teaming and the human element of the team introduces a new dimension. The combination of AI capability and the human element to the team results in a level of interaction between the entities. Previously, isolation assumptions in applying AI meant that human-machine teaming was relatively straightforward. The increasing capability of AI, however, allows more sophisticated implementations of machine technology. In fact, human-machine teams become more capable the more machines are integrated with the team, as detailed in David Mindell's work on ro-

bots and autonomy.[16] By introducing the human agent element, the resulting analytic framework provides a means to assess the resulting teaming behavior. The framework is agnostic to utility and does not focus on only commercial or only military uses. The goal here is not to create a desired state but to just track and chart current development. In terms of collaboration, the framework developed by the study team has focused on human-to-AI collaboration. As AI matures and multiple agents interact with each other, it will be just as important to focus on AI-to-AI collaboration.

The x-axis of the framework captures the AI capability, its ability to sense and interact with its environment, and the progression as technology advances. Many developments resulting from the integration of machine learning algorithms, processing data over greater contexts, drives AI capability further to the right on the x-axis. The y-axis captures the programmed "knowledge" of the environment, including humans or other AI agents. This worldview dictates the teaming behaviors as machines are integrated with humans in accomplishing tasks or missions. In her 1994 presidential address to AAAI, Dr. Grosz provides a thoughtful analysis on the nature of group behaviors between humans and machines. Beyond isolation, "interaction entails only acting on someone or something else, collaboration is inherently "with" others; working (labore) jointly with (co)."[17] Dr. Grosz also details an in-between degree of group behavior, contracting, which the study team has termed "cooperation" in order to avoid confusion with government contracting language. Moving up along the y-axis in the framework results from increasingly complex group behavior in the human-machine team.

As AI becomes more capable and human-machine teams become more cooperative, the increasing machine contact with humans requires building trust between the agents. Trust is required to develop beyond "isolation assumptions" with respect to human interaction, which inherently limits the scope of work AI may be allocated. Discussed further later in this report, trust in AI means understanding "How do we know that an AI agent is doing what we want?" and "How do we know it's not doing what we don't want?" For an effective human-machine team, at any level, the AI must be appropriately transparent, qualities of assurance and trust that are dependent on the application. A spam filter that utilizes natural language processing and machine learning, with feedback from the user on false positives and false negatives, requires less transparency than applications that may result in a risk to life.

Interacting or tasking behaviors may also not be tenable in high-stakes situation. We have a perception that the human is the "fail-safe." However, the model of self-driving cars returning control to humans suddenly and in situations requiring

16 David A. Mindell, *Our Robots, Ourselves: Robotics and the Myth of Autonomy* (New York: Viking, 2015).

17 Barbara J. Grosz, "Collaborative Systems: AAAI-94 Presidential Address."

significant decision making is proving to be an untenable model. A greater level of capability is required to either be fully isolated from the human or cooperate more effectively. Learning to provide insight and recommendations requires understanding value to users. Humans must decide what feedback matters and what feedback to give.

An incomplete understanding of the dynamic of human-to-AI trust is a significant limitation to increasing the capability of human-machine teams. AI enables the human-machine team to process information faster and over greater contexts. However, successfully applying AI technology requires trust in the algorithms, the data (including quality), and the outcomes. Further, unlocking the valuable functionality of an AI-enabled team requires a supportive AI ecosystem that extends beyond the technology itself.

CH. 3

INVESTMENT IN ARTIFICIAL INTELLIGENCE

FOR THE VAST MAJORITY of national security-related technologies, parallel tracks of investment flow from commercial and government investors. For instance, aviation technology is advanced by both private sector investment in high-efficiency engine technology and advanced control systems and Department of Defense investment in fighter technologies and electronic systems. For most dual-use technologies, the prominence of government investment has lessened in the past four decades, as the size of the defense market has shrunk and as the private sector has taken on by far the larger share of research and development in key twenty-first century technologies. This pattern has also been observed in AI development.

This landscape of parallel technology development requires the coordination, balancing, and integration of government and private sector investment.[18] To support such a balanced effort in AI, the U.S. government must identify where government investment plays a critical role that private sector investment cannot perform as efficiently, identify aspects of AI technology development most critical to national security missions, and understand how to enable the AI ecosystem needed to support national security capabilities. This section will analyze the various approaches commercial and government entities are taking to investment in the AI ecosystem. It should be noted that this chapter focuses primarily on investment activity within the United States. Global investment activity will be explored further in a later chapter.

Understanding the AI investment landscape is complicated by the fact that there is no standard method of data collection and no accepted taxonomy for categorizing these investments. Many applications and systems require some combination of the various disciplines, for example driverless cars, which require a combination of computer vision and machine learning to respond to a variety of road situations. The study team found that AI investment was reported using terminology associated with AI research disciplines (machine learning, robotics, knowledge representation, etc.) or the AI system (self-driving cars, virtual assistants, etc.), or a blend of the two approaches. For example, a 2017 discussion paper from McKinsey Global Institute, "Artificial Intelligence: The Next Digital Frontier?", categorizes AI investment by both academic discipline (machine learning, computer vision, natural language processing) and systems that would leverage developments across AI disciplines and other technologies (autonomous vehicles, smart robotics, virtual agents).[19] Academic researchers and commercial developers think about the AI disciplines differently, as academic researchers focus on furthering the fields through foundational research and commercial developers tend to focus on system-specific perspectives. Investing in a single academic discipline, such as machine learning, computer vision, or natural language process-

18 Ryan Crotty and Andrew Hunter, *Keeping the Technological Edge: Leveraging Outside Innovation to Sustain the Department of Defense's Technological Advantage* (Lanham, MD: Rowman & Littlefield, 2015), https://csis-prod.s3.amazonaws.com/s3fs-public/legacy_files/files/publication/150925_Hunter_KeepingTechnologicalEdge_Web.pdf.

19 Jacques Bughin et al., *Artificial Intelligence: The Next Digital Frontier?* McKinsey Global Institute, June 2017.

ing, will rarely deliver capabilities that integrate neatly with human operators and organizations to deliver functioning capability. The government acting as an AI investor must embrace both perspectives to effectively perform its roles as a user of and investor in AI.

COMMERCIAL INVESTMENT IN AI

Commercial investment has increased rapidly over the past decade as firms have witnessed the benefits of using machine learning to extract value from data. A 2017 McKinsey report estimated that the previous year's global commercial investment in AI from tech giants, including companies like Google and Baidu, was valued at $20 to $30 billion. Of this, an estimated 90 percent went to research and development (R&D) and deployment with the remaining 10 percent spent on acquisitions. While the tech giants are the largest AI investors, there are many smaller players as well. The same McKinsey report estimated that combined venture capital (VC) and private equity (PE) investment was $6 to $9 billion in 2016.[20] Writing for PitchBook, Dana Olsen states that global venture capital in AI and machine learning companies in 2017 was more than $10.8 billion, increasing from approximately $5.7 billion in 2016. These figures present a significant increase in value since 2010 when investment was estimated at less than $500 million.[21]

Clearly, AI potential has enticed global investment with a mix of internal and VC/PE investment spread in varying directions sector-wise, albeit growing in magnitude, as adoption is primarily limited in early stages, focusing on applying machine learning. The landscape is dominated by large tech companies that are poised to act as AI providers to private and public consumer bases. With respect to stock market performance, the U.S.-based FAANG companies (Facebook, Apple, Amazon, Netflix, and Alphabet's Google), along with the Chinese-based BAT companies (Baidu, Alibaba, and Tencent), are global leaders in AI valuation. Other companies have focused on internal investment as well: Salesforce on virtual agents and machine learning; BMW, Tesla, and Toyota on robotics and machine learning for driverless cars; ABB, Bosch, GE, and Siemens on machine learning and robotics; and IBM's Watson on cognitive computing services in the Internet of Things (IoT).[22]

Meanwhile, Moore's law has continued to decrease the cost of processing data in machine learning applications through the use of graphics processing units (GPUs). For example, NVIDIA's GTX 1080 GPU delivers nine teraflops for about $500. When adjusting for inflation, Mirhaydari notes that similar power output in 1961 would have cost about $9 trillion for a string of IBM 1620 computers. The democratization

20 Bughin et al., *Artificial Intelligence: The Next Digital Frontier?*

21 Dana Olsen, "2017 Year in Review: The Top VC Rounds and Investors in AI," PitchBook, December 20, 2017.

22 Bughin et al., *Artificial Intelligence: The Next Digital Frontier?*

of AI has been enabled through access to and collaboration with open source frameworks like Caffe, Google's TensorFlow, and Torch.[23]

Some sectors have led adoption of this technology, including the marketing, advertising, and financial industries due to the alignment of sector revenue generators with the functionality of machine learning algorithms, identifying patterns and trends and optimization in large data sets. While interest in the commercial potential of AI is clear, the immediate return on investment is often less clear in other arenas. Addressing the excitement among venture capital investors in AI, Anthony Mirhaydari at Pitchbook writes that "annual data generation is expected to hit 44 zettabytes (trillions of GB) by 2020, according to the IDC's Digital Universe Report…a CAGR [compounded annual growth rate] of 141% over just five years."[24] Despite this, only 33 percent of the data is useful for analytics, which makes data processing and data hygiene immensely important. Without "clean" data, AI systems will not be able to process the data at their full potential. In recognition of the importance of clean data, the U.S. Department of Defense began prioritizing data quality over quantity in FY15.

The discussions undertaken as part of this study suggest that strengthening the overall AI ecosystem of the U.S. government will be critical to ensuring that the United States does not cede an important advantage in AI. Presently, investment is in early stages with most investors still waiting for their bets on AI to pay off.[25] Given the relevance of speed in the digital era, there has been substantial debate about the existence and magnitude of the first-mover advantage in AI. While the private sector may see advantage in being first to develop AI techniques, how much of an advantage comes to users who implement AI solutions first? In conflict and on tight timelines, first-mover advantage can be key. However, there are other advantages in moving second, especially given the problem-specific nature of AI. If AI techniques are challenging to transition from one problem to another, it may not be possible to gain much enduring advantage by being the first to solve a problem.

U.S. GOVERNMENT INVESTMENT IN AI

Government investment in AI flows into a broad range of programs, budgets, and initiatives, some of which are explicitly focused on AI, while others are enablers of AI and other military capabilities, such as advanced networking technologies and computing capabilities. All of these investments form part of an ecosystem that makes implementing AI as an effective national security capability possible. Throughout this report, we reference this broader set of AI enablers as the AI ecosystem.

23 Mirhaydari, "Rise of AI Excites VC Investors, Challenges Society."

24 Anthony Mirhaydari, "Rise of AI Excites VC Investors, Challenges Society," PitchBook, October 12, 2017, https://pitchbook.com/news/articles/rise-of-ai-excites-vc-investors-challenges-society.

25 Bughin et al., *Artificial Intelligence, The Next Digital Frontier?*

Understanding the state of the U.S. government's investment in AI has been a priority for several years as both private and public entities seek to assess the U.S.'s position in the field. A 2018 Govini report, summarized in the following paragraphs, is a comprehensive resource on understanding and analyzing federal spending in artificial intelligence, big data, and cloud technology.[26] The report categorizes relevant Department of Defense spending into three main segments: learning and intelligence, advanced computing, and AI systems, totaling approximately $1.76 billion from FY2013 to FY2017.

Within each of these areas, FY2017 spending levels are listed below, including a comparison to the FY2013 baseline. (Not reflected in these numbers is the early September 2018 DARPA announcement of a planned $2 billion investment in AI across a variety of related technologies over the next five years.)[27]

- **Learning and Intelligence**
  - *Deep Learning* spending increased 9.4 percent to $158.3 million
  - *Machine Learning* decreased 3.5 percent to $154.4 million
  - *Natural Language Processing* decreased 4.7 percent to $38 million
  - *Data Mining* decreased 26.6 percent to $22.9 million
- **Advanced Computing**
  - *Super-computing* increased 16.1 percent to $356 million
  - *Neuromorphic Engineering* increased 21 percent to $126.9 million
  - *Quantum Computing* increased 9.3 percent to $68.5 million
- **AI Systems**
  - *Computer Vision* increased 11.2 percent to $395.4 million
  - *Virtual Reality* decreased 4 percent to $386.6 million
  - *Virtual Agents* decreased 6.1 percent to $56.7 million

Different Department of Defense entities focus investments in different areas. The United States Air Force, United States Army, and DARPA are the three largest spenders overall, outspending the next seven DoD components combined. A large chunk of this money is going to AI systems, though there are still significant amounts finding their way to Advanced Computing and to Learning and Intelligence. Within Learning and Intelligence, deep learning is the most competitive of any AI sub-segment. Much of this funding has gone to major defense contractors such as Leidos and Raytheon for AI systems, and Northrop Grumman for advanced systems and learning and intelligence. These are the top recipients of DoD AI-related investment funding, earning about $1.8 billion combined from FY13 to FY17. HP Enterprises, DLT Solutions, and Carahsoft Technology Corp. are by far the largest vendors of cloud services, earning $2.3 billion from FY13 to FY17. Cloud services provide an important enabling capability for machine learning due to the massive data sets and

26 Patrick Tucker, et al., The 2018 Federal Scorecard: The National Security Edition, govini and Defense One.

27 "AI Next Campaign," DARPA, https://www.darpa.mil/work-with-us/ai-next-campaign.

processing power involved. The DoD sees a clear advantage if they can be the first to get AI "right" and deploy it on the battlefield, as it would provide information superiority and allow them to make faster and more accurate decisions.

As noted in previous DIIG research, R&D has been dropping as a share of the DoD budget for eight years.[28] However, the share of the budget devoted to AI has increased over last four or five years in addition to the increase in the absolute dollar amount. A recent White House report, *Artificial Intelligence for the American People*, states that federal government investment in unclassified R&D for AI and related technologies has grown by more than 40 percent since 2015, as reflected in budgets requested for relevant government programs.[29] It can safely be assumed that increased sums likely exist for classified programs. The 2019 budget specifically identifies AI investment in basic research at the National Science Foundation and the National Institutes of Health and applied R&D at the Department of Transportation, Department of Defense, and the National Institutes of Health.[30] These agencies represent a wide variety of issue areas, which means the federal government recognizes the benefits AI can create across a range of topics. AI is not a tool for just one issue area but can be widely adopted to help solve problems.

Additional sourcing on investment in the AI ecosystem is available through the Networking and Information Technology Research and Development (NITRD) Program, a group of U.S. federal agencies supporting the development of Information Technology (IT) capabilities in the federal government.[31] It also includes both unclassified and classified R&D, with classified R&D generally being smaller than unclassified.[32] However, given the importance of infrastructure and digital capability, NITRD funding trends are problematic in that IT spending is decreasing while AI spending is increasing. The success of AI applications depends in part on having the right infrastructure to support access to data and computing and the productivity of the workforce. AI is grounded in basic computer science, so it is problematic and unsustainable for investment in foundational digital capability spending to decrease while AI spending is increasing.

From both the White House analysis and the Govini data, it is clear that U.S. government investment in AI and related technologies is already substantial. What is

28 Andrew Hunter et al., *Defense Acquisition Trends*, 2016: The End of the Contracting Drawdown (Lanham, MD: Rowman & Littlefield, March 2017), https://csis-prod.s3.amazonaws.com/s3fs-public/publication/170309_Ellman_AcquisitionTrends2016_Web.pdf?EOHx.4yzTSKOdaa9FMLs3KStHUSrIO5Q.

29 The White House, *Artificial Intelligence for the American People*, May 10, 2018, https://www.whitehouse.gov/briefings-statements/artificial-intelligence-american-people/.

30 The White House, *FY2019 Budget Request Administration R&D Priorities*, 2018, https://www.whitehouse.gov/wp-content/uploads/2018/02/ap_18_research-fy2019.pdf.

31 Note: NITRD is the "primary source of federally funded work on advanced information technologies (IT) in computing, networking, and software" in the United States; The Networking and Information Technology Research and Development Program, https://nitrd.gov.

32 National Science and Technology Council Committee on Technology, The Networking and Information Technology Research and Development Program Supplement to the President's Budget FY 2018, https://www.nitrd.gov/pubs/2018supplement/FY2018NITRDSupplement.pdf.

less immediately clear from these sources, however, is whether this investment is properly structured to meet critical government needs and integrate appropriately with private sector investments in AI.

ACCELERATING INVESTMENT IN THE AI ECOSYSTEM

Public sector investment in AI must include several critical elements. In addition to investing in the technologies necessary for foundational basic research and for adapting AI tech to government needs, systemic investments in the AI ecosystem are required to ensure that the government has the underlying digital capability to utilize AI. The reality is that this digital capability is underdeveloped or lacking in large parts of government, including in many parts of the Department of Defense. As will be discussed at greater length in the next chapter, shortfalls in the government's digital capability have resulted in a combination of technical and workforce debt that limits the potential for AI adoption. Investment is needed in things like network infrastructure, data collection, and data processing. Unfortunately, these investments are not glamorous, but strengthening the AI ecosystem in this way is critical to successful deployment of AI. There is a solid culture of experimentation, but this does not help if the underlying architecture cannot can translate the data from system to system. For example, the F-22 and F-35 fifth-generation fighter jets face interoperability challenges among themselves, let alone with fourth-generation fighters, as their underlying network architectures do not line up and the incoming data is not processed the same way.[33] Before the Department of Defense goes all in with AI on its current path, the underlying architectures need to be assimilated.

"INVESTMENT IS NEEDED IN THINGS LIKE NETWORK INFRASTRUCTURE, DATA COLLECTION, AND DATA PROCESSING."

When it comes to technology research, there are also multiple roles for government involvement. Government should act as an early stage investor, spreading its bets across multiple promising technological approaches rather than going all-in on any one technology. In addition, the government should focus investment in areas of AI research that industry players are less inclined to pursue. These are areas where the return-on-investment expectations involve potentially longer time horizons, which tend to be less attractive to short-term private sector investors focused on the adoption, proliferation, and implementation of mainstream applications. Some examples of this kind of investment include the verification and validation of large data sets to address vulnerabilities and bugs within the training and input data, AI systems that are fault tolerant and highly assured, and the integration of AI systems into human-machine collaborative teams.

33 Gregory Brook, "The F-22 and the F-35 Are Struggling to Talk to Each Other…And to the Rest of USAF," *Air Force Magazine*, March 2018, http://www.airforcemag.com/MagazineArchive/Pages/2018/March%202018/The-F-22-and-the-F-35-Are-Struggling-to-Talk-to-Each-Other---And-to-the-Rest-of-USAF.aspx.

Government application of AI, particularly for national security, will require a very high threshold of reliability for actual implementation. While commercial industry can select in AI development towards AI applications that are profitable and low risk, the government's mission set is not readily subject to modification. As a result, the government's need for AI reliability is likely to exceed that which will naturally be developed by in the commercial sector and should be a significant element in the government's research investment in AI.

For an AI capability to be reliable, it must deliver expected results consistently. As a result, there is a substantial need for investment in AI explainability and developing processes for the verification and validation of AI performance. Since AI leveraging machine learning is highly dependent on training data, data security is critical to AI reliability as the success of an AI program will only be as accurate as the data that it uses to learn. Key areas for government investment are areas that are unique to national security applications—explainability and transparency, robustness, verification and validation of data and models, and applications unique to military. The results may not be clear or tangible, but we may find more advancements through focus on the uniquely government areas. Given the space between many private sector applications and national security applications, like machine learning in the private sector being driven by applications to marketing, a mismatch in applicability may minimize the repercussions for missed investments.

While the government has a successful history of investing in basic research related to software technologies, including AI, the record of transitioning these capabilities into acquisition and operational use is much less successful. One key issue to ensuring effective AI investment is addressing the challenges in government software acquisition. This complex topic requires a study all its own, and several such efforts are underway, including a recent Defense Science Board study, an ongoing study at the Defense Innovation Board, and an upcoming CSIS study on acquisition of software-defined hardware-based systems.[34] A common theme of many of these efforts, and one raised by several experts the study team consulted for this study, is the need for greater flexibility in the government budget process for the development and acquisition of software. The rigidity of software acquisition presents barriers to the cooperative nature of software development, scaling, and repurposing. Software moves much faster than conventional acquisition, sometimes on the order of months, and the boundaries between concept development, production, and operation are less clear than with hardware. While the federal budget process seeks to create and enforce separation between procurement, research and develop-

34 Defense Science Board, "Design and Acquisition of Software for Defense Systems", Office of the Under Secretary of Defense for Research and Engineering, February 2018, Washington, D.C., https://www.acq.osd.mil/dsb/reports/2010s/DSB_SWA_Report_FINALdelivered2-21-2018.pdf; Jared Serbu, "Defense Innovation Board to Tackle DOD's Software Acquisition Problems, Using Software", Federal News Network, January 30, 2018, https://federalnewsnetwork.com/dod-reporters-notebook-jared-serbu/2018/01/defense-innovation-board-to-tackle-dods-software-acquisition-problems-using-software/

ment, and operations and maintenance activities, these activities inescapably blend together in software development. These problems are particularly acute for AI development, where so much of the foundational infrastructure for AI reliability has yet to be developed, but where AI implementation is already underway. Flexible contracting mechanisms, such as Other Transaction Authority contracts (OTAs) and flexible approaches to budgeting, are policy options to adapt to the growing software acquisition environment.

Education is a huge piece of ensuring capability in the long term. In looking at the dynamics in our cyber capabilities, the United States doesn't have enough people trained to do work with AI and investing in the people that use this technology is key. Since the Internet of Things is forcing the entire government to rethink how it does business, sufficient capability means having the technology and implementing it as needed. The U.S. government needs people who can develop this for the public sector, but the private sector can pay experts much more than the government can. Grants to universities can help bridge this gap, as the government does not own the advantage anymore; they have to be more strategic in their acquisitions.

Public sector national security work has thrived based on the exclusivity of the mission set. The government offers platforms and opportunities that the private sector cannot operate in. Flying and manufacturing high-performance jet aircraft is a good example of this. However, government work is not the prominent market force in AI like it is with other technologies. The DoD does not have exclusivity in application of AI, and it is certainly not a large market driver of it. The DoD must adopt commercial developments when applicable but cannot rely solely on the private sector to provide capability. Further, the DoD must compete for workforce talent and market share without relying on the allure of exclusivity.

CH. 4

ADOPTION OF ARTIFICIAL INTELLIGENCE

AI IS ONE OF POTENTIALLY many tools for competing in a data-driven age. The study team conducted research on how the public and private sectors are executing the adoption of AI to solve the problems of today and the near future. Narratives on the use of AI in the United States often follow one of two themes. One theme asserts that we are still waiting for AI, that it is coming but it is not here yet. This theme may be ominous in tone, invoking the language of Skynet and killer robots. Though untrue, this first theme is important to acknowledge because it sets a benchmark for AI capabilities that once reached, may be impossible to turn back on. The second theme states that AI is already here and is more prevalent than is immediately apparent. AI powers smartphones, optimizes the delivery of packages to customers across the country, secures points of entry at airports, and is leading cutting-edge developments in systems such as self-driving cars. There are many statements that capture both themes, sometimes referred to as the "AI effect," as noted by several researchers in the field: AI is whatever hasn't been done yet[35]; As long as you wait, AI is not new; AI is automation we are familiar with; As soon as we have seen a capability long enough, it is no longer AI; You've been waiting for artificial intelligence, but you're already using it.[36]

No narrative, however, captures the investment of time, energy, and money required to successfully apply AI in any sector, or the fact that the decision to deploy an AI technology depends on the problem to be solved rather than the application of AI merely for the sake of applying AI. The topic of AI adoption must be framed in a problem-specific manner by asking questions in relation to the functionality of the technology, like "What is AI doing?" and "Why was AI needed?" instead of "Was AI used or not?"

"AI" may refer to technologies as well as to the systems that employ these technologies. Particularly for the use of AI in the United States, the previously detailed advances in machine learning have excited organizations seeking to gain advantage through data and have driven the discussion on AI to prominence. Many applications of AI combine machine learning with other techniques such as computer vision (e.g., facial recognition, self-driving cars), natural language processing (e.g., personal virtual assistants, call center support), and data analytics (e.g., credit card fraud monitoring). Leveraging machine learning allows organizations to extract value in new ways out of source data, from text to images, video, and audio, to name a few.

Adopting AI requires more than just algorithms. Organizations must grow an AI ecosystem within the environment in which the algorithms live. This includes a capable workforce, foundational data practices and structures, computing and networking infrastructures, and a deployment strategy that articulates the applicability of AI to problem sets. As a result, AI has been minimally deployed at scale outside of the

35 Douglas Hofstadter, *Gödel, Escher, and Bach: An Eternal Golden Braid* (New York: Basic Books, 1979).

36 HubSpot, "What Is AI? You've Been Waiting for Artificial Intelligence, but You're Already Using It," YouTube video, 0:26, February 17, 2017, https://www.youtube.com/watch?v=UHb5XTh10Yw.

tech sector, where the AI ecosystem was relatively well developed.[37] When assessing the degree of AI applied throughout use cases, unlike autonomy, there is an absence of ready metrics and measures for evaluation of the "degree of AI" at the system level. Measures to assess the degree to which AI techniques have been incorporated into a technology solution would be a useful measure for both the public and private sector to assess which organizations are adopting AI, where AI is being deployed, and at what scale is AI being used. To infuse AI into an enterprise, organizations must not only adopt AI but also cultivate the AI ecosystem.

ADOPTERS IN COMMERCIAL SECTORS

Infusing algorithms into commercial sectors is not a phenomenon unique to the spread of AI. The finance sector has been deploying algorithms that work at hyper speeds since the 1980s. Algorithms operate as speeds faster than a human can keep up, with a significant majority of stock market activity being algorithms trading with other algorithms. Companies like UPS, Capital One, Marriott, GE, Proctor and Gamble, United Healthcare, and Walmart have implemented data analytics, sometimes including AI, at scale, requiring multi-year undertakings to transform their businesses. To increase efficiency and maximize the utilization of resources like fuel, trucks, and drivers, UPS began using predictive analytics in 2003 and in 2017 fully implemented their On-Road Integrated Optimization and Navigation route-optimization system.[38] The increasing use of AI is, in part, a continuation of developments in the Information Age: leveraging data and computing power to gain advantage in a domain.

Current AI adoption in the private sector is strongest in data-rich fields with well-defined and well-scoped problem sets, advocates or project champions to bring AI into an organization, and the longer timescales needed to support iteration and experimentation in human-machine teaming. Companies will either follow the progression of the market or identify and create a new market. AI technology is applied across a wide range of industries[39] across a spectrum of technological capability, from agriculture, pharmaceuticals, and travel, to name a few.

Within the private sector, two classes of actors are appearing on the market: suppliers of AI technology in the form of either tech giants or smaller data science firms, and consumers of AI technology who have a business case for using AI but rely upon the expertise of the tech industry to develop systems that meet their needs. Consumers of AI technology, both public and private, face the same challenges in cultivating an AI ecosystem and making the necessary investments in technology,

37 Irving Wladawsky-Berger, "Artificial Intelligence is Ready for Business' Are Businesses Ready for AI?", *Irving Wladawsky-Berger (blog)*, September 11, 2017, http://blog.irvingwb.com/blog/2017/09/artificial-intelligence-is-ready-for-business-are-businesses-ready-for-ai.html.

38 Thomas H. Davenport and Jeanne G. Harris, *Competing on Analytics: The New Science of Winning*, rev. ed. (Boston: Harvard Business Review Press, 2017).

39 Michael Chui et al., *Notes from the AI Frontier: Insights from Hundreds of Use Cases*.

people, and digital infrastructure. A third kind of AI actor has not yet fully emerged, the AI integrator. An integrator doesn't just provide AI capabilities to consumers but incorporates AI and non-AI solutions into a coherent whole that is largely plug-and-play for the user. At this early stage of AI development, and given the highly problem-specific nature of today's AI, the absence of AI integrators who make AI adoption "easy" is not surprising. In this environment, AI consumers are largely required to act as their own integrators, a complex challenge that is made only more complex when the consumer has significant technical and workforce debt to pay down.

While many of the challenges of AI adoption in the private sector are present in the public sector as well, national security AI users confront additional challenges due to differing thresholds of reliability and incentives for the use of AI. While the private sector is motivated to maximize profits and decrease costs, the DoD mission requires prioritizing a complex mix of military and public interest objectives. The private sector will develop and deploy AI in ways that immediately support a business case by increasing profits or lowering costs. Government users will adopt AI mostly only after high levels of trust and reliability have been established, after issues of operational control have been worked out, and usually when existing methods of performing the same mission have been shown to be inadequate. However, certain parts of the private sector experience similar levels of risk and complexity as those present in many national security applications, offering either lessons learned or a pathfinder to deployment of AI systems. For example, deploying AI and predictive analytics to manage the maintenance and operation of a fleet of vehicles is very transferrable to the DoD and could be used to address the need for improving readiness levels through predictive maintenance and better fleet management.

ADOPTERS IN NATIONAL SECURITY

In translating discussion on AI to the language of national security, the algorithm that sits inside of a platform enables a capability or application; it is a tool or a new material. There may also be a maturation time during which the new technology is tested and iterated to finally improve the capability. As a result, it is difficult to immediately identify where AI is being applied. The narrative of "waiting for AI" is more prevalent in national security, particularly within the DoD and the intelligence community (IC). Research and discussion with subject matter experts reveals many more data-driven programs leveraging AI capabilities than are often discussed. Early adopters within the DoD and IC have been leading the way, carefully and usefully defining the requirements and starting with flexible programs set up to experiment, implement, refine, and iterate. These forward-leaners within DoD and IC can show results from programs initiated early in the AI hype cycle, often outside of traditional acquisition authorities.[40] Adopting and applying AI and machine learning to the

40 Daniel S. Hoadley and Nathan J. Lucas, *Artificial Intelligence and National Security*, Congressional Research Service, April 26, 2018.

national security mission has primarily been executed through hubs in the form of Centers of Excellence to the Services such as the Army Asymmetric Warfare Group, Navy Digital Warfare, Air Force Rapid Capabilities Office, and the Army Futures Command. DARPA has also served as a pathfinder for the DoD. As groups apply AI, it is being embedded within broader data and service platforms. Like the implementation of cloud computing in government, where the IC led the way for seven years while other agencies observed the demonstration before committing to cloud, AI traction is gained by demonstration and results.

Like commercial adoption of AI, application of AI in national security is occurring where components of the AI ecosystem presently exist or in fields such as cybersecurity,[41] where there is a significant need to operate at speeds and scales much greater than humans can manage. Programs are initiated and sustained in data-rich fields with well-defined problem sets, a workforce with data science skills, timescales conducive to the iterative nature of software development, and organizational advocates.

Programs mentioned here include those in production and use as well as those in the concept definition or demonstration phases. The Intelligence, Surveillance, and Reconnaissance (ISR) mission set provides a ready application space, particularly for machine learning techniques. At the heart of ISR is pattern recognition and identification of indicators or signals, often weak patterns distributed through a noisy and vast data set. Machine learning combined with image recognition or natural language processing acts as an amplifier for analysts combing through data sets. Palantir's All Source Information Fusion, the Distributed Common Ground System available to each of the armed services, National Geospatial-Intelligence Agency's (NGA) Boosting Innovative GEOINT, and the DoD's Project Maven implement data analytics, machine learning, and other AI techniques with varying levels of success in the ISR mission space.

Both the logistics and maintenance mission set present data-rich fields positioned to apply AI where pattern recognition and optimization are key. National security adopters are taking lessons learned from the private sector logistics and maintenance communities by applying AI in areas like the Stryker Logistics Support Activity (LOGSA) Proof of Concept,[42] the Autonomic Logistics Information System (ALIS) for the F-35,[43] and the U.S. Air Force Air Mobility Command's demonstration of a predictive maintenance system from Lockheed where AI optimizes fleet maintenance and sustainment.[44]

41 "Defense Department Sees Big Role for Artificial Intelligence in Cybersecurity," MeriTalk, February 16, 2017, www.meritalk.com/articles/defense-department-artificial-intelligence-cybersecurity-halvorsen-ibm-watson/.

42 Adam Stone, "Army Logistics Integrating New AI, Cloud Capabilities," C4ISRNET, September 7, 2017, https://www.c4isrnet.com/home/2017/09/07/army-logistics-integrating-new-ai-cloud-capabilities/.

43 Kris Osborn, "Air Force Chief Scientist confirm F-35 will include artificial intelligence," Defense Systems, January 20, 2017, https://defensesystems.com/articles/2017/01/20/f35.aspx.

44 Marcus Weisgerber, "The US Air Force Is Adding Algorithms to Predict When Planes Will Break," Defense One, May 15, 2018, https://www.defenseone.com/business/2018/05/us-air-force-adding-algorithms-predict-when-planes-will-break/148234/.

Robotics and unmanned systems programs aim to implement AI to grow the capability of rules-based autonomous systems or human-machine teams. The U.S. Air Force's Loyal Wingman program, the U.S. Marine Corp's Multi-Utility Tactical Transport program, and the Office of Naval Research's "Sea Hunter" Anti-Submarine Warfare (ASW) Continuous Trail Unmanned Vessel (ACTUV) (recently transitioned out of DARPA),[45] all require AI to support the capability goals and metrics of these programs.

Two areas with concepts in development are electromagnetic spectrum (EMS) activities, such as spectrum management and electronic warfare (EW),[46] and Command and Control (C2).[47] While the application of AI, particularly machine learning, to EMS activities is largely experimental,[48] the commercial telecommunications industry via the 5G cellular network and academic research at U.S. universities are demonstrating approaches to more effectively manage spectrum access through machine learning. As the EMS becomes increasingly congested and contested, innovation in and novel approaches to the underlying software and network architectures may be the key to success.

As AI is gaining traction through demonstration in early-adopting programs and commercial sector successes, the Defense Innovation Board has recommended catalyzing innovation in AI and machine learning, recognizing their impact across the spectrum of DoD operations.[49] Service-specific efforts are leading toward larger joint efforts supporting both AI innovation and the development of the AI ecosystem, through efforts such as the Joint Artificial Intelligence Center (JAIC), and the Joint Enterprise Defense Infrastructure (JEDI), and the Cloud Executive Steering Group (CESG) leading the way on the Pentagon's cloud computing upgrade.

It is no surprise then that the Pentagon established the Joint Artificial Intelligence Center (JAIC) on June 27, 2018, reporting to the DoD's chief information officer.[50]

45 "ACTUV 'Sea Hunter' Prototype Transitions to Office of Naval Research for Further Development," Defense Advanced Research Projects Agency, January 1, 2018, https://www.darpa.mil/news-events/2018-01-30a.

46 Mark Pomerleau, "What Is the Difference between Adaptive and Cognitive Electronic Warfare?" C4ISRNET, December 16, 2016, https://www.c4isrnet.com/c2-comms/2016/12/16/what-is-the-difference-between-adaptive-and-cognitive-electronic-warfare/.

47 Steve Hirsch, "Understanding Multi-Domain Command and Control," Air Force Magazine, January 9, 2018, http://www.airforcemag.com/Features/Pages/2018/January%202018/Understanding-Multi-Domain-Command-and-Control.aspx.

48 "Adaptive Radar Countermeasures (ARC)," DARPA, https://www.darpa.mil/program/adaptive-radar-countermeasures; "Behavioral Learning for Adaptive Electronic Warfare (BLADE)," DARPA, https://www.darpa.mil/program/behavioral-learning-for-adaptive-electronic-warfare; "Radio Frequency Machine Learning Systems (RFMLS)," DARPA, https://www.darpa.mil/program/radio-frequency-machine-learning-systems; "Spectrum Collaboration Challenge (SC2)," DARPA, https://www.darpa.mil/program/spectrum-collaboration-challenge.

49 Defense Innovation Board, Technology & Capabilities Recommendation 5: Catalyze Innovations in Artificial Intelligence & Machine Learning, Defense Innovation Board, https://innovation.defense.gov/Recommendations/.

50 Sydney Freedberg, "Joint Artificial Intelligence Center Created Under CIO," Breaking Defense, June 29, 2018, https://breakingdefense.com/2018/06/joint-artificial-intelligence-center-created-under-dod-cio/.

According to observers with knowledge of the center's mission, it is designed to facilitate the service's ability to identify standards for AI, gather the necessary talent, and deliver "AI-infused" solutions. The center's initial goal is to prove itself to policymakers by following the model of USAF's Project Maven, which helps ease workload in terms of classifying and identifying objects from surveillance missions, as well as explore issues on safe and ethical AI use. Like Project Maven, the DoD will rely on mostly third parties and contractors to develop AI capabilities that feed into larger DoD systems.[51]

Adoption is occurring in cycles of narrow use cases with the potential to build to broad use cases, spreading the AI bets throughout the ecosystem. Narrow projects offer the potential for speed, enable a pathfinder approach, and minimize risk while broader efforts have the potential to apply lessons learned from narrow projects across an enterprise or portfolio. Additionally, there are areas in national security ripe for adoption where there is data available, a relatively low risk, and time to iterate. Google has deployed machine learning to more efficiently cool its data centers resulting in a consistent reduction in the energy used.[52] Human resources, force management, and health care record management are areas that do not get much immediate attention. Fuel purchasing and supply chain management are areas where applying AI to logistics where DoD has substantial data that can be leveraged by AI.

Barriers and Enablers to AI Adoption in National Security

AI is difficult and will take time. Broader public discussions, particularly on topics of machine learning, place a premium on algorithm explainability, as if once the community addresses explainability, AI will be "easy." But explainability is not a panacea , and research explainability is not the same as operational explainability. Each application or problem-space dictates the need for model transparency. A daily email user has a lower threshold for spam filter model transparency (an application of machine learning and natural language processing) than an AI user in a safety-critical field like medicine or transportation. National security challenges are often wicked problems without a single answer that may be easily tested or corrected. There is much less control over various factors and incomplete knowledge and mistakes matter significantly.

"EXPLAINABILITY IS NOT A PANACEA."

This discussion is not intended to discourage the pursuit of algorithm explainability or adoption of AI technologies, but to identify the barriers to adoption across the AI ecosystem and identify enabling areas of work. Expectation management

51 Patrick Tucker, "The Pentagon is Building an AI Product Factory," Defense One, April 19, 2018, https://www.defenseone.com/technology/2018/04/pentagon-building-ai-product-factory/147594/.

52 Steve Ranger, "Google Just Put An AI In Charge of Keeping Its Data Centers Cool," ZDNet, https://www.zdnet.com/article/google-just-put-an-ai-in-charge-of-keeping-its-data-centers-cool/.

is critical and depends on a technically accurate understanding of the models and systems being developed and deployed. Significant data, training, and time goes into narrow AI systems. Like the concept of technical debt in software development, cultivating a healthy AI ecosystem will incur debts in time, money, and effort to bring together the "AI ingredients." A skilled workforce must be trained on new and sometimes unfamiliar technologies. Within this education and training, there is the serious concern that the evolution of required skills means that some people will lose jobs, which must be addressed. Building a foundational digital capability means investing in computing resources and business support systems, digitizing processes and information, and breaking down barriers around data, to name a few.

Data, data sharing, and access to processing power are not always readily available in areas where the DoD would like to be making operational, tactical, and strategic decisions. Unlike data-rich fields like personnel management, in many cases the data is not there by default and needs to be collected. This means that an investment must be made in the ecosystem on the subject of data. Individuals must understand the importance of gathering data, organizing data, and cleaning data. Organizations must consider data ownership and data rights in contracts with service providers and customers; deploying systems without owning the data will not be feasible for programs applying AI. For algorithmic approaches requiring iterative machine learning, contracts must mandate training data. While data is a key to AI, particularly for machine learning, some national security applications operate in data-austere environments. Human expertise and low-learning AI models may be able to be leveraged in these environments as the technology improves. In the national security context, this may require addressing policies that secure data, people, and computing infrastructure through isolation and silos. AI will generally not produce up-front savings or allow for personnel reductions in its initial implementation, but the up-front investment is a debt worth paying for the functionality gained from AI.

Public and Private Entities in the AI Ecosystem

Both public and private sector AI adopters must learn how to treat information as a strategic asset, deploy the right technology to leverage that information, integrate with the existing legacy systems, and connect data silos across functional units and across the enterprise.[53] Adopting AI technology requires the availability of data and computing power, enabled by a foundational digital capability upon which AI is deployed. A foundational digital capability includes things like database management, new requirements on data recording and reporting, competing needs for data access and data security, and hardware and software commensurate with use case needs.

53 *Fintech Vs. Traditional Trade: Surviving the Digital Transition*, Trade Finance, September 2016, https://www.baft.org/docs/default-source/default-document-library/trade-finance-sibos-2016.pdf?sfvrsn=0.

Further, the divergence of the AI supplier market from the traditional defense acquisition model presents challenges for an acquisition system already struggling to adapt to and support the acquisition of software. Given the significant interest and potential in AI in the commercial sector, the DoD is not the sole or dominant market as with exquisite defense systems. 2018 saw a tech giant willingly remove itself as a service provider to the DoD, a position that is sustainable due to the size and gravity of the commercial market. However, the democratization of AI has resulted in a supplier market with a more diverse set of actors capable of providing and integrating AI in a manner somewhat analogous to the defense systems integrator model. It remains that getting new entrants into the federal market is difficult. Hundreds of smaller data-driven firms, such as those in the In-Q-Tel portfolio, are capable of supplying AI technologies to address challenges in national security. AI integrators, the providers of data management platforms, are largely absent from the AI supplier market, though such integrators are not necessarily large firms.

Workforce and Organizations

A robust AI ecosystem requires robust cultural enablers and organizational structures. The adoption of AI in national security will be subject to the cultural characteristics and challenges of the armed services, DoD, and the IC. DoD acquisition culture expects new technologies to be delivered at a high level of technological maturity. Expected performance characteristics are highly defined. Before the technology is given to operational units, its effectiveness and suitability are expected to have been rigorously demonstrated. Machine learning and many other AI capabilities require iterative algorithm training. The challenge for the DoD is that much of the adoption will need to be done via supervised machine learning requiring significant subject matter expertise in the unit adopting AI. In the early stages of learning, unit performance could very well be degraded by the introduction of AI. For example, IBM's Watson debating tool took 6 years to develop a simple demonstration capability that itself was built on a capability developed over 10 to 15 years. Adopting this more evolutionary and loosely-defined approach to capability with operational users will take a sea-change in culture. More fundamentally, it is unclear if the testing and validation approaches developed for hardware-based capabilities are even meaningfully applicable to the verification and validation of AI. Hardware testing is based on test points that combine to form a model of predicted system performance for evaluation. With AI, particularly when it is unclear how the algorithm is performing its function and when the algorithm learns so that its performance is designed to vary over time even when confronted with the same conditions, it is not clear how any given test result necessarily relates to another. Existing test models can't predict the perfor-

mance of an AI system.[54] The difficulties associated with verification and validation of AI capabilities, particularly learning algorithms, will require the workforce adopting AI to be comfortable with processes and standards that are a complete departure from normal DoD practice.

A lack of understanding of and consensus on terminology and concepts in AI prevents progress on many fronts. Requests for Proposals (RFPs) and congressional language are not written by technology experts. As a result, requirements are not written for AI, though throughout 2018 individual programs are beginning to incorporate AI language into RFPs. The pull side of acquisition (e.g., "it would be really useful if my computer could respond to cyberattacks at speeds faster than humans") is as necessary as the push from private firms to further apply AI to national security problems.

An AI ecosystem lacking a knowledgeable workforce leads to difficulty in assessing machine learning processes and outcomes, be they successes or failures. While personnel are relatively willing to try new systems, a lack of clear ethical guidelines and norms results in an appropriate hesitancy out of an abundance of caution to connect emerging AI systems to systems that generate high-consequence outcomes. Any new technology or system requires a level of trust on the part of the human operator. This trust will only come from repeated use and human investment of time and resources. Building an understanding of what it will take to make AI systems viable requires consistent user engagement and an understanding of the operational goals AI supports. User engagement approached through a model of customer engagement is key at all levels and often missing. An already problematic separation between the systems being built and the user community is exacerbated when users are minimally aware of existing AI technologies and how these technologies might improve aspects of their work. Exacerbated by isolation from users, scaling software often adds complexity, proliferating attack surfaces and vectors, and creates coordination challenges.

"PREPARING THE ECOSYSTEM AND COORDINATING WITH INVESTMENT EFFORTS IS NECESSARY TO TAKE ADVANTAGE OF THE FUNCTIONALITY OF AI TECHNOLOGIES."

Risk-averse cultures cause organizations to back off from applying AI in the face of an underdeveloped and underinvested AI ecosystem. Rigid organizational policies can prevent access to basic software tools like MATLAB, let alone providing the environment of openness and experimentation necessary to recruit and retain a high-quality software workforce. A strong AI commercial market means that government entities must compete for top talent. Individuals who are not encouraged

54 David Tate, Rebecca Grier, Christopher Martin et al., A Framework for Evidence-Based Licensure of Adaptive Autonomous Systems (Alexandria, VA: Institute for Defense Analyses, 2016), http://www.dtic.mil/dtic/tr/fulltext/u2/1020297.pdf.

to pursue innovation will ultimately leave to find environments more supportive of the iteration and experimentation necessary for success in AI.

Where the AI ecosystem is robust enough to support the adoption of AI, adoption is affected by the nature of the mission. Commercial AI adoption success almost always happens in regimes where consequences are acceptable and fail-safes exist. Risk can be tolerated at an algorithmic level because there are controls outside of the system to manage mistakes and issues. Factors such as the severity of loss, timescale, reversibility, and the possibility of triggering events with unforeseen consequences all favor applications in the commercial sector. The DoD and IC operate in contested spaces that differ from the uncontested prototyping environment. Questions on whether or not the warfighter can use AI systems under the conditions of conflict remain unanswered. Confidence level and risk factors in most cases will be inputs but not necessarily the deciding factors.

Early success stories demonstrate the importance of the AI ecosystem. In a DoD context, the application of AI in areas like logistics and predictive maintenance may serve as a bridge to build trust and demonstrate where and how AI can work before coming in to contribute to irreversible decisions. Preparing the ecosystem and coordinating with investment efforts is necessary to take advantage of the functionality of AI technologies.

CH. 5

MANAGING OPERATIONAL ARTIFICIAL INTELLIGENCE

SUCCESSFULLY MANAGING THE USE OF AI TECHNOLOGIES requires incorporating AI capabilities into existing organizational control structures while also addressing the new challenges in ethical use and trust. The recognition that managing AI is more complex than simply buying new software is slowly coming. In informal terms, managing operational AI means asking questions of "How do you make AI do what you want?" and "What do we want AI to do for us?" The challenge of answering these questions is directly tied to the context of the AI application. While an amazing advance in technology, AlphaGo Zero presented no concerns for AI management. The machine was able to make mistakes with zero costs and complete knowledge of the impact of its decisions on other actors (which were limited to its opponent).

Much debate in AI management focuses on the edge cases unique to the DoD mission set. These edge cases require high accuracy (i.e., minimal acceptance of false positives and false negatives) and low latency (i.e., tight timelines with minimal time for thoughtful review or consideration, particularly by a human). There can be a large gap between how some organizations address operational risk by focusing on low-probability, high-risk issues and how AI adopters think about these issues by focusing on high-probability, low-risk issues. Even for cases that are low-accuracy and high-latency, managing operational technology must still be addressed for implementation. AI stakeholders need to better coordinate approaches to and discussions on AI to address the full spectrum of operational concerns.

LEVELS OF MANAGEMENT

Managing AI presents challenges to organizations at every level, including at the strategic, operational, and tactical levels. Feedback on AI management runs in both directions. Lessons learned from AI practice inform policy moving from the tactical level up to the strategic level through the operational level, while policy mandates inform changes in AI practice moving from the strategic level to the tactical level through the operational level.[55]

At the strategic level of the AI ecosystem, society and governments shape the overall ecosystem. These high-level organizations and entities institute the laws, policies, procedures, guidelines, standards, and best practices that govern the use of AI technologies. Norms of use and ethical standards are established that shape the ecosystem, guiding formal institutions and behavior at the operational and tactical levels.

The operational level includes organizations or entities that apply AI technology to their domains or problem spaces. Moving from the strategic level to the operational level, policy provides vision and direction to organizations in applying the tech-

55 Adapted from The National Implementation Research Network's Active Implementation Hub Practice-Policy Feedback Loops, https://implementation.fpg.unc.edu/module-5/topic-3-practice-policy-feedback-loops.

nology to meet strategic priorities. Feedback from the operational level up to the strategic level informs application areas and the establishment and refinement of standards and best practices.

The tactical level of the ecosystem is people-oriented. This level contains the end users, the managers, and the researchers. Actors at this level implement the technology within an organizational functional area or capability space. Moving from the operational level down to the tactical, the ecosystem enables application through incentives and executes risk reduction. Tactical level activities then provide feedback to the operational level as it informs on the functions and limitations of AI technology, as well as identifying new research areas to advance the capability.

Managing operational AI also means addressing risk introduced into the ecosystem across each level. The ability to adapt to new contexts is an attribute that users strive for in AI technologies, but adaptability is also what most worry about, placing a high premium on effectively managing operational AI. Negligence in large distributed networks and complex systems can have significant and cascading effects**.** As AI lives on the network, network security practices must be considered and potentially refined. A multi-level ecosystem requires a multi-level risk acceptance policy.

MANAGING AI AT THE STRATEGIC LEVEL

Strategic level management provides structure and guidance on investment, application, and continued maintenance of AI technologies. This activity is primarily government or societal in nature. It includes formal policy and guidance, ethical guidelines and norms, legal and regulatory frameworks, standards for verification and validation (V&V) and operational test and evaluation (T&E), training and doctrine, and tactics, techniques, and procedures (TPPs).

In many of the sectors or applications explored in previous sections, the operational use of AI may require a regulatory approach to establish standards and guidelines. National security, health care, and financial sectors all work with sensitive and consequential information. Strategic-level guidance is necessary in areas of AI application where abuse of privacy and sensitive information presents concerns. IoT security[56] and facial recognition technology[57] are examples of such areas; others include medical devices, border security, toys, and automobiles.

Previous CSIS research highlights the emerging discussion on the necessity of reforming federal data policies, workforce development, risk management, and test-

56 Rishi Bhargava, "Your IoT is Probably Not A-OK," Forbes, July 16, 2018, https://www.forbes.com/sites/forbestechcouncil/2018/07/16/your-iot-is-probably-not-a-ok/#3c889de3763d.

57 Victoria Cavaliere, "Microsoft Wants Regulation of Facial Recognition Technology to Limit 'Abuse,'" CNNMoney, July 14, 2018, https://money.cnn.com/2018/07/14/technology/microsoft-facial-recognition-letter-government/index.html.

ing and evaluation to meet the needs of software-driven products.[58] Strategic-level management of AI does not mean starting from scratch or with a blank slate: many existing software, information security, and network risk management policies and guidance may serve as a foundation or starting point for effectively integrating AI into national security solutions. The technology itself is integrated with existing hardware and software, a simple example being deploying a machine learning-based analytic platform on existing computers, networks, and data sets. An overview of the most pertinent documents is provided in the following section.

U.S. Policy Approaches to Artificial Intelligence

The United States currently manages AI technologies through existing software and hardware policies as AI-specific policies are being developed. One area of early progress is in managing the implementation of software and certain weapons systems that operate in the high-accuracy, low-latency regime. Further, a bipartisan AI caucus was formed in the House of Representatives in May 2017. Successfully managing AI at the strategic level means identifying requirements for developing, securing, and deploying AI capabilities that may be captured in best practices, procedures, and guidelines. Standards for accountability and reliability suggest clear guidance of acceptance and rejection. However, an expertise gap persists in technical authorities at the strategic level, with non-specialists often issuing technical decisions. As stated in a recent Congressional Research Service (CRS) report, "No independent entity in the commercial sector or inside government is charged with validating AI system performance and enforcing safety standards."[59]

Appendix A provides an overview of some of the relevant policies.

Liability, Accountability, and Model Transparency

The concept of model transparency, or explainability, is closely tied to the concepts of accountability and liability in the conversations on AI, particularly in applications that leverage machine learning techniques. Accountability and transparency are, like AI, problem-dependent. It is not immediately clear how or where to assign liability when AI systems produce unfavorable outcomes, particularly for systems where the algorithm or data has been developed by an outside entity. Legal questions around big data, machine learning, and artificial intelligence are a burgeoning field, growing beyond software questions such as licensing rights.[60] Questions like

58 William Carter, *A National Machine Intelligence Strategy for the United States* (Washington, DC: Center for Strategic and International Studies/Booz Allen Hamilton, March 1, 2018), https://csis-prod.s3.amazonaws.com/s3fs-public/publication/180227_Carter_MachineIntelligence_Web.PDF?CLlXGgQQQoc78akgCk.2StKO7NsrC2J1.

59 Hoadley and Lucas, *Artificial Intelligence and National Security*.

60 Xavier Seuba, Christophe Geiger, and Julien Penin (eds), *Intellectual Property and Digital Trade in the Age of Artificial Intelligence and Big Data*, Global Perspectives for the Intellectual Property System, CEIPI-ICTSD, Issue 5, 2018, https://www.ictsd.org/sites/default/files/research/ceipi-ictsd_issue_5_final_0.pdf.

"How do we prevent failure?" and "Who owns the failure?" are at the heart of liability, accountability, and transparency.

"IT IS NOT IMMEDIATELY CLEAR HOW OR WHERE TO ASSIGN LIABILITY WHEN AI SYSTEMS PRODUCE UNFAVORABLE OUTCOMES."

Managing AI liability from a hardware mindset presents challenges. It is not certain that approaches for liability used in hardware are applicable for software-driven products. Manufacturing defects place liability on the supplier; however, misuse on the part of the buyer may place liability on the user. Terms of service approaches to liability require a high level of competence on the part of the user or adopting organization when the provider is not liable if the user incorrectly uses a product. Software systems are updated regularly with very clear terms of service. This is a very different kind of service and set of expectations from hardware.

Accountability has legal implications as well as organizational change management implications. When things go wrong, accountability is the mechanism by which current failures are addressed and future failures are prevented. Accountability in systems blending hardware, software, and AI, such as self-driving cars, is challenging traditional operational liability. For traditional hardware systems, there is a culture that passes operational liability on to the user. Particularly in cases where the consequence of failure may be high, there will be significant challenges in the areas of agency and accountability with AI systems. Regulators at the strategic level must be convinced of system reliability and believe that the instances of catastrophe will be rare enough to make the benefits worth the risk. The current approach to system reliability requires exhaustive testing that most companies cannot afford. Presently, accountability and reliability statistical concepts require system testing to ensure an acceptable number of failures per quantity of uses. This presents challenges in machine learning techniques when testing for a quantity of failures biases the representation of failures in data. The statistical concept of reliability does not hold nicely with AI systems, posing challenges for traditional understandings of accountability and liability. It is extraordinarily difficult for testing and evaluation to address events that almost never happen. With AI, regulators, technology suppliers, and end users will have to find mechanisms to ensure that the system works without exhaustive testing.

Intellectual Property

A gap exists in approaches to address the intellectual property (IP) of data, the models, and the outputs of AI applications. AI requires both a data set as well as the algorithmic model to produce an output. In the national security context, for government agencies purchasing AI solutions from the private sector, it is not yet clear how much of the IP the agency owns or should own. Concerns about IP apply to the data and models in AI. Both are concepts that are relatively new and not well-understood. As a 2018 Congressional Research Service report on AI recounts, "Members of DoD

leadership also cited the tech sector's insistence on preserving intellectual property rights as a stumbling block [to cooperation]. This is particularly challenging when it comes to AI, because many companies are not selling code to the department along with the AI application, which makes it difficult to gain a deeper understanding of how the system will perform."[61]

If an organization invests time, money, and effort into an AI model, protecting the IP, the intellectual product of that investment, is key. While some companies have figured out a way to watermark models, comprehension of how to protect IP with AI systems is very limited. IP is critical for tech companies. There are many cases in the data services sector where the model is a company's IP. If the acquisition of an AI model requires source code access by the purchaser, many companies are unwilling or unable to meet that requirement. On a commercial basis, companies almost never sell IP. On the defense side, to a customer such as DoD, some companies will sell IP if it is developed at government cost.

Under current practice, data related to AI may not be considered intellectual property. However, data is the most important part of AI and a source of value to organizations implementing AI solutions. Similarly, how one uses the data is more important than the data different organizations possess.

Model transparency as a mechanism for trust is challenged by the concept of IP. One method for users to build trust in a model is by exploring and interrogating various aspects of the model to ensure proper functioning. This type of activity requires access to the model's source code. To that point, the test and evaluation community is reaching the conclusion that it will be impossible not to open the black box and look at the underlying code.

Protecting the IP of both data and models also requires understanding the capabilities of reverse engineering. By analyzing the outputs of a model, it is possible to reverse engineer the model itself–a possibility that calls into the question the practice of using "black box models" to protect model IP. It is also possible to reverse engineer aspects of the data set by interrogating the model or outputs. For DoD applications where the data is highly sensitive, reverse engineering to learn about the underlying data poses challenges for security standards and practices.

MANAGING AI AT THE OPERATIONAL AND TACTICAL LEVELS

Managing AI at the operational and tactical levels addresses how organizations and their people effectively use AI technologies. Structure and command delegation impact the risk organizations are willing to accept and therefore the AI they are willing to adopt and apply. While strategic-level guidance can set the conditions, needs, and

61 Hoadley and Lucas, *Artificial Intelligence and National Security*.

FOCUS AREAS FOR MANAGING AI

priorities, the operational and tactical levels must deliver the ingredients to deploy and manage the technical solutions. This means AI must demonstrate success in a realistic data environment, answering realistic operational questions. Further, it is necessary to build a cloud computing architecture, reduce barriers to data, provide realistic and accurate training data, and build cross-functional teams that merge subject matter expertise with data science expertise.

Organizations may develop and deploy AI capabilities to meet their own internal needs and provide capabilities to external entities. Each personnel role has unique needs for understanding the functions and limitations of AI technology. While organizations may be highly structured, or relatively flat, we have identified four roles present in organizations: individual contributors, first line managers, middle- or second-line managers, and senior leaders. Each role has needs that must be met to effectively manage AI technology.

At the individual contributor level are the power users, the experts on tools and data sets. These experts differ from daily users, as many platform solutions have daily

users that live at the management levels. At the individual contributor level, personnel and teams understand the functionality and limitations of the underlying data, models, and outputs. For analytic platforms, they synthesize insights that are then passed up the chain to the first line managers.

First line managers oversee the individual contributors and must understand the methodology, data sources, and limitations. Personnel at this level speak to the "what" and "why" as well as the high-level "how" when briefing findings and insights.

Second line managers, or middle management, focus on the big picture context of "what" and "why" and are further removed from the algorithms and data at the individual contributor level. Personnel at this level must look up and down through an organization, supporting senior leadership and first line management.

Senior management are outcome-oriented for internal products, aligning business capabilities with strategic goals.

AI may be integrated into an organization through an enterprise capability, across various business units within an organization, or as a business capability, supporting one unit or functional group. How and where within an organization AI is deployed has implications for who needs to understand what aspects of the technology. At each level of an organization, successfully deploying AI means understanding how much information is necessary to understand an outcome, operating at an acceptable level of risk, and having people trust the assessment. Often, these are new issues and new systems that have not been dealt with before. Each level has differing knowledge needs to understand how to use the product, brief results, and explain and demonstrate value. Senior leaders identify doors to open relying on mid-level leaders to execute that vision within the organization. Often senior and middle leaders have matured in a qualitative methods-based work environment and may be new to the software-intensive systems that leverage AI technologies. Personnel in these positions increasingly have to manage quantitative products, projects, and workforces. The problem-specific nature of AI also requires subject matter experts and capability providers to work in cross-domain settings. As an example, a mid-level manager who wants to leverage machine learning and natural language processing to parse news articles for key words and phrases must work with machine learning scientists, data scientists, and political scientists. This means that leaders must understand the metrics of professional success across these domains, measures that often differ significantly from discipline to discipline.

Developing the AI ecosystem means investing time and resources into growing the right talent. Change management is required to manage operational AI as the proficiencies and literacies are lacking throughout organizations in many application areas. A mix of cultural barriers and limited understanding of the technology itself present challenges in effectively managing AI. Non-digital natives and digital natives alike must all adapt to new technologies. Within the DoD, there is a significant cultural difference and barrier in how military personnel approach new technolo-

gies. Where younger personnel are willing to work through bugs and try to improve new technologies, older personnel may be tempted to cast aside new systems at the first indication of a problem. Further, many in the workforce harbor concerns of being replaced by new technologies, particularly technologies that serve an automating function. Addressing these human factors and cultural elements are necessary to AI efficacy.

It is necessary to recognize that AI is inhuman and as such it is unwise to hold AI to a human standard of accomplishing a given task. That is not to say AI should be held to a higher or lower standard comparatively. It is important to set technically relevant and measurable success criteria, such as the success rate of an image recognition algorithm, that are operationally effective and make things operationally better for the team. With new technology adoption, it is difficult to recover from initially fielding technology at the wrong threshold. Deploying a new technology often requires the assumption that reliability and trust will increase with time and familiarity, while taking the appropriate steps to prove safety, functionality, and so on. Ideally, the appropriateness of the trust also increases over time. However, there is a perception that the human element is a fail-safe, though humans are often the weaker element. Public opinion tends to view AI differently, believing that systems need to be perfect.[62] This presents a barrier to operationalizing AI by setting machine expectations too high.

Successfully managing operational AI means growing the talent and the expertise within an organization. Contrary to some narratives, AI does not necessarily result in a workforce reduction or the reduction of the cognitive load on employees. This means that while some tasks may now be executed computationally, or "automated", the human involvement required to operate AI may be new to organizations. As David Mindell writes, "Human-factors researchers and cognitive scientists find that rarely does automation simply 'mechanize' a human task; rather, it tends to make the task more complex, often increasing the workload (or shifting it around)…Automation changes the type of human involvement required and transforms but does not eliminate it."[63] For example, manufacturing jobs have often said to be most at risk for human job loss to machines, an instance of implementing early rules-based AI. The machine replaces the mundane task of assembling the given product, and the original human is out of a job. However, the implementation of the machine creates a need for a host of new jobs, including software development and machine maintenance. These new jobs have new technical skills that were not required beforehand, so leadership must identify the tools, training, talent, and environment each level of the organization requires to be successful.

62 Nidhi Kalra and David G. Groves, "The Enemy of Good: Estimating the Cost of Waiting for Nearly Perfect Automated Vehicles" (Santa Monica: RAND, 2017), https://www.rand.org/pubs/research_reports/RR2150.html.

63 David A. Mindell, *Our Robots, Ourselves: Robotics and the Myth of Autonomy* (New York: Viking, 2015).

PERSONNEL ROLES AND CORRESPONDING NEEDS IN ORGANIZATIONS

INDIVIDUAL CONTRIBUTORS

Power users, experts on tools and/or datasets

Understand functionality and limitations

For analytic platforms, synthesize insights

FIRST LINE MANAGERS

Understand methodology, data sources, and limitations

Speak to the "what" and "why", high level "how"

SECOND LINE MANGERS, MIDDLE MANAGEMENT

Big picture context of "what" and "why"

Must look up and down through an organization, supporting Senior Leadership and 1st line management

SENIOR MANAGEMENT

Outcome oriented for internal products, aligning business capabilities with strategic goals

Securing and Assuring Data and Algorithms

Managing operational AI is not just about the algorithms and models. Equally important–arguably more important–is the of data upon which the AI relies. Assuring the quality and accuracy of the data, the verification and validation of data sets, is a new concept. What make AI training possible is the availability of data. Data availability allows AI to get better algorithms. Good training data makes for good outcomes. Data must be considered a protected, strategic asset. However, data must also be accessible. One significant challenge of managing AI will be in balancing data security and accessibility.

How the data is used remains the most valuable part of AI research. Data may be reverse engineered from the model used to process it; both must be secured appropriately given this technical reality. For example, in intelligence scenarios, the compromise of exquisite data through the models could mean the compromise of sources and methods. When AI is introduced, new attack vectors are introduced, such as deep learning spoofing and data spoofing. If the training data is known or manipulation of data is too predictable, adversaries can easily anticipate and predict actions and outcomes. Adversaries can spoof sensors and thus the data collected by those sensors without needing to mess with the underlying model code. Put simply, adversaries do not need to know what is in the box to exploit the box.

Not all data management concerns are introduced into the system by an adversary. Incomplete understanding of the technology and models by those employing them is also a source of error. Like non-AI statistical models, overfitting is a problem. Machine learning techniques run the risk of being overfit to the training data set or one operational data set, limiting broader functionality. Managing operational machine learning means bringing AI ingredients together–the data, the right model and model tuning, field expertise, and computing power.

There is limited recognition that software and AI may fail during operation without harboring an identifiable defect. As David Tate and David Sparrow at the Institute for Defense Analyses state, "Defense acquisition is predicated on the assumption that system unpredictability has been ironed out by the time it passes through the procurement cycle."[64] Further, the cost of developing software is increasing at a time when hardware is becoming cheaper.[65] However, the complex and unpredictable environments in national security and defense certainly have the potential to induce unanticipated and fringe failure modes in AI.[66] It is important to remember that AI systems can often produce adverse outcomes. There are a growing number of cases of AI and machine learning systems getting odd outputs from algorithms during normal operation. AI systems will cheat a lot and go about solving problems in counterintuitive or false ways.

Aspects of software testing and continuous improvement will need to change for AI-enabled technologies. DoD policy is to treat reliability as a statistical concept in which organizations statistically verify how often a system fails. For software, reliability is currently measured by the number of bugs. There is no recognition of a failure due to the chaotic behavior of the software system that might take you into a place where you don't want to be. Further, as detailed previously, introducing failure modes into data for training purposes may bias the data towards failures modes. Verification and validation (V&V) is very new in the AI space. V&V reliability and its applicability to AI algorithms and data is a concern that reaches across the strategic level through the operational and tactical levels. It is a technically difficult challenge that will be a major limiter in the ability to field AI models in our systems.

64 David M. Tate and David A. Sparrow, "Acquisition Challenges of Autonomous Systems," in *Acquisition Research Symposium* (Monterey, CA: Naval Postgraduate School, 2018).

65 David M. Tate, "Acquisition Cycle Time: Defining the Problem," Institute for Defense Analyses, October 2016.

66 Hoadley and Lucas, *Artificial Intelligence and National Security*.

CH. 6

INTERNATIONAL ACTIVITY IN ARTIFICIAL INTELLIGENCE

WHILE THE COMMON NARRATIVE surrounding AI focuses primarily on activity in the United States and China, the international market in AI is diverse in the number of actors, investment amounts and types, and areas of focus within the AI discipline. There is a clear divide between a focus on developments in robotics and developments in machine learning, as well as a divide between military and social applications.

In 2018, the global AI and robotics defense industry was valued at $39.22 billion. With a projected compounded annual growth rate (CAGR) of 5.04 percent, the market is expected to be valued at $61 billion by 2027.[67] Market Forecast attributes this valuation and growth to investment in new systems from countries such as the United States, Russia, and Israel as well as procurement of systems by countries such as Saudi Arabia, India, Japan, and South Korea. According to Market Forecast, the share of expenditure and market shares in order of highest proportion are as follows:

- Robotics (primarily due to ongoing procurement);
- Computer vision;
- Natural language processing;
- Speech recognition;
- Social media analysis, multi-agent systems, and knowledge representation and reasoning.

It should be noted that Market Forecast did not identify machine learning specifically as an area of growth or investment in the materials available to the team.[68]

Though multiple companies in multiple nations are developing AI at the same time, not everyone is developing it on the same terms. There is a cultural aspect to how these countries approach problems and solutions that includes morality, democratic principles, basic individual rights, the role of the state, and agreement on the laws of warfare.

RUSSIAN FEDERATION

In a 2017 speech, Russian President Vladimir Putin said artificial intelligence is the future of mankind and that whoever becomes the leader in this sphere will become the ruler of the world.[69] Russia's most recent AI plan could lead to 30 percent of their military force being remote controlled and populated with autonomous robotic plat-

67 "Global Artificial Intelligence & Robotics for Defense, Market & Technology Forecast to 2027," Market Forecast, January 18, 2018, https://www.marketforecast.com/reports/global-artificial-intelligence-robotics-for-defense-market-technology-forecast-to-2027-1058.

68 Ibid.

69 Radina Gigova, "Who Vladimir Putin Thinks Will Rule the World," CNN, September 2, 2017, https://www.cnn.com/2017/09/01/world/putin-artificial-intelligence-will-rule-world/index.html.

forms by the year 2030, but it has also confirmed that humans will still be involved in the decision making when it comes to lethal force.[70]

Russian military experts have documented interests in areas including cruise missiles, unmanned underwater vehicles, unmanned ground vehicles, electronic warfare, and cybersecurity (especially blockchain). They also plan to create a so-called "library of goals" to help weapon systems with target recognition and navigation guidance.[71]

Russia has also been actively creating avenues for AI adoption and development. Era, an innovation technopolis, opened in September and is predicted to become a 50-acre city by 2020.[72] They also created the Foundation for Advanced Studies, a government organization similar to DARPA. Writing for the Congressional Research Service, David Hoadley and Nathan Lucas express skepticism that Russia will achieve all of their AI goals, if any. This comes from the fact that there is a known lack of sophistication in their technology industry and that there was a 7 percent decrease in defense research funding announced in 2017, with additional projected decreases of 3.2 percent and 4.8 percent in 2018 and 2019, respectively.[73] Potential shortfalls in investment funding are likely to constrain Russia's ability to challenge for AI leadership globally, but Russian investment may still be sufficient to take leadership in niche applications, particularly those focused on national security. Russia may also make progress on many of the organizational and non-technical issues surrounding the AI ecosystem that could allow for significant progress in AI implementation despite a lack of investment funding.

PEOPLE'S REPUBLIC OF CHINA

China thinks information is power and that power should be controlled by the state. The institutional alignment of Chinese government, military, financial institutions, and corporations enables an ease of access to technology by the military and other security forces. The top-down coordination tends to put private sector innovation in close coordination with government goals and objectives. In China, generally, commercial companies, university research laboratories, the military, and the central government routinely work together closely. As a result, the Chinese government has a direct means of guiding AI development priorities and

70 Greg Allen and Taniel Chan, *Artificial Intelligence and National Security*, Harvard Kennedy School, Belfer Center, July 2017, https://www.belfercenter.org/sites/default/files/files/publication/AI%20NatSec%20-%20final.pdf.

71 Samuel Bendett, "In AI, Russia is Hustling to Catch Up," Defense One, April 4, 2018, https://www.defenseone.com/ideas/2018/04/russia-races-forward-ai-development/147178/?oref=defense-one_today_nl.

72 Patrick Tucker, "China, Russia, and the US Are All Building Centers for Military AI," Defense One, July 11, 2018. https://www.defenseone.com/technology/2018/07/china-russia-and-us-are-all-building-centers-military-ai/149643/?oref=d_brief_.

73 Hoadley and Lucas, *Artificial Intelligence and National Security*.

principles.[74] Many times, China has shown a tendency to adopt a technology first and figure out its privacy implications later. This is evident in the adoption of their health-care system, self-driving cars, traffic management, and facial recognition for payment authentication.[75] When it comes to AI application in their defense industry, the Chinese have focused on cybersecurity, social governance (facial recognition and surveillance), cruise missiles,[76] and unmanned systems.[77] However, the distinction between defense and civilian applications is not always clear cut. In Chinese policy there is little civil-military divide with respect to targets; civilian infrastructure is a valid target.

China is aiming to become a world leader in AI by 2030 and cultivate a domestic AI industry worth about $150 billion,[78] though as Elsa Kania describes in her congressional testimony, the Chinese government has not committed to investing $150 billion directly, as has sometimes been assumed.[79] Research reveals a difficulty in clearly delineating between government dollars and corporate dollars for investment given the Chinese economic structure, but a translation of the text indicates a goal of a core market value of $150 billion.[80]

The demand for AI talent in China is far outpacing the availability of skilled researchers.[81] The shortage of skilled programmers and computer scientists specializing in AI has spurred several educational ventures to increase the number of programmers in China. The PLA National University of Defense Technology has added an institute for intelligent sciences,[82] Tsinghua University is uniting civilian and military R&D in an advanced laboratory specializing in military intelligence,[83] and Kai-Fu

74 Hoadley, *Artificial Intelligence and National Security*.

75 Andy Chun, "China's AI Dream is Well on Its Way to Becoming a Reality," *South China Morning Post*, April 22, 2018.

76 Ben Blanchard, "China Eyes Artificial Intelligence for New Cruise Missiles," *Reuters*, August 19, 2016. https://www.reuters.com/article/us-china-defence-missiles/china-eyes-artificial-intelligence-for-new-cruise-missiles-idUSKCN10U0EM.

77 Stephan De Spiegeleire, Matthijs Maas, and Tim Sweijs, *Artificial Intelligence and the Future of Defense: Strategic Implications for Small and Medium-Sized Force Providers*, The Hague Centre for Strategic Studies, 2017.

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

79 Elsa Kania, China's Threat to American Government and Private Sector Research and Innovation Leadership, Testimony before the House Permanent Select Committee on Intelligence, July 19, 2018, 2.

80 Graham Webster, Rogier Creemers, Paul Triolo, and Elsa Kania, "China's Plan to 'Lead' in AI: Purpose, Prospects, ad Problems," *New America*, August 1, 2017, https://www.newamerica.org/cybersecurity-initiative/blog/chinas-plan-lead-ai-purpose-prospects-and-problems/.

81 "Top AI Trends to Watch In 2018," CBInsights, PDF.; Chen Yu, "New Plan to Harness AI Talent," *China Daily*, April 4, 2018, http://www.chinadaily.com.cn/a/201804/04/WS5ac40c20a3105cdcf65161a9.html.

82 Tucker, "China, Russia, and the US Are All Building Centers for Military AI."

83 Tom Simonite, "Ex-Google Executive Opens a School for AI, with China's Help," *Wired*, April 5, 2018, https://www.wired.com/story/ex-google-executive-opens-a-school-for-ai-with-chinas-help/.

Lee, former head of Google's China operations and founder of Sinovation Ventures, is leading a new AI/machine learning program out of Peking University.[84] In addition to AI and computer science programs at Chinese universities, many U.S. universities have a growing presence in areas like Shanghai bringing additional education opportunities.[85] And it has been working; many foreign universities in from around the world have been moving to China because of the Chinese government's serious commitment to supporting them.[86]

China also released its AI Development plan on July 20, 2017; this is the source of the claim about a Chinese plan to invest $150 billion in AI.[87] The Chinese strategic objectives and timelines are as follows (conversions to dollars added):

2020

- Overall technology and application of AI will be in step with globally advanced levels
- Cultivate the world's leading AI backbone enterprises
- *"The scale of AI's core industry will exceed 150 billion RMB (~$21.7 billion), and exceeding 1 trillion RMB (~$150 billion) as driven by the scale of related industries."*[88]

2025

- Major breakthroughs in basic theories for AI, such that some technologies and applications achieve a world-leading level and AI becomes the main driving force for China's industrial upgrading and economic transformation
- *"The scale of AI's core industry will be more than 400 billion RMB (~$58 billion), and the scale of related industries will exceed 5 trillion RMB (~$726 billion)."*[89]

2030

- China will be the world's primary AI innovation center, achieving visible results in intelligent economy and intelligent society applications and laying an important foundation for becoming a leading innovation-style nation and an economic power
- *"The AI core industry scale will exceed 1 trillion RMB (~$150 billion), with the scale of related industries exceeding 10 trillion RMB (~$1.5 trillion)."*[90]

84 Simonite, "Ex-Google Executive Opens a School for AI, with China's Help."

85 David Baroza, "Berkeley Reveals Plan for Academic Center in China," *New York Times*, Nov. 16, 2011, https://www.nytimes.com/2011/11/17/world/asia/cal-berkeley-reveals-plan-for-engineering-center-in-china.html; Lia Zhu, "US Schools Setting Up Campuses in China," China Daily, Jan. 25, 2018, http://usa.chinadaily.com.cn/a/201801/25/WS5a699e74a3106e7dcc136a42.html.

86 Barboza, "Berkeley Reveals Plan for Academic Center in China."

87 Graham Webster, Rogier Creemers, Paul Triolo, and Elsa Kania, "China's Plan to 'Lead' in AI: Purpose, Prospects, ad Problems," *New America*, August 1, 2017, https://www.newamerica.org/cybersecurity-initiative/blog/chinas-plan-lead-ai-purpose-prospects-and-problems/.

88 Ibid.

89 Ibid.

90 Ibid.

China-based companies are significant players in AI development. Baidu, Alibaba, and Tencent, also known collectively as BAT, are the big tech players in China actively investing inside the country and out. A CBInsights briefing on "The State of AI in China" quantifies that 44 percent of equity deals backed by BAT went to companies based in the United States from 2014 to 2018. Another 46 percent remained inside China, with 8 percent going to Israel and 3 percent to Canada. Further, the government announced that the first wave of open AI platforms will rely on Baidu for autonomous vehicles, Alibaba for cloud services in support of smart cities, and Tencent for health care.[91] China is also investing abroad as a government. A DIU study in January 2018 charted the increase of Chinese investment in U.S. AI companies since 2010, showing not only the valuation of the deals but the relative increase in activity from $1.5 million in 2010 to $353.6 million in 2016.[92]

Further, Chinese tech companies are deploying different strategies when competing for market share in third-market countries. Chinese conglomerates are buying stakes in local firms and weaving them together into complex tapestries of services.[93] U.S. firms, on the other hand, have transplanted their services broadly to other markets; Amazon has pledged more than $5 billion to replicate its offerings in India, for example.

FRANCE

France has pledged more than $1.85 billion over five years to advance the country's position in AI research. President Macron laid out his plan during an early 2018 speech announcing a new national strategy to catch up to the world leaders in AI, namely China and the United States, and make France a leader in its own right.[94] The funding will help in executing the new strategy, particularly in the health care and autonomous vehicles sectors, with the aid of policy proposals aimed at mitigating the challenges ahead, like recruitment and retention of AI talent. France recognizes that the goals set out in its new AI strategy depend on the country's ability to attract foreign and French researchers abroad to France. Macron hopes to entice researchers by allowing publicly-funded scientists to work 50 percent of their time in private

91 Deepashri Varadharajan, *State of AI in China*, CBInsights, 2018.

92 Michael Brown and Pavneet Singh, "China's Technology Transfer Strategy: How Chinese Investments in Emerging Technology Enable a Strategic Competitor to Access the Crown Jewels of U.S. Innovation," Defense Innovation Unit Experimental, 2018.

93 "FAANGs v BATs, America's Tech Giants Vie with China's in Third Countries," *Economist*, July 5, 2018, https://www.economist.com/leaders/2018/07/05/americas-tech-giants-vie-with-chinas-in-third-countries.

94 Nicholas Thomas, "Emmanuel Macron Talks to Wired About France's AI Strategy," Wired, March 31, 2018, https://www.wired.com/story/emmanuel-macron-talks-to-wired-about-frances-ai-strategy/; Tania Rabesandratana, "Emmanuel Macron Wants France to Become a Leader in AI and Avoid 'Dystopia,'" Science Magazine, March 30, 2018, http://www.sciencemag.org/news/2018/03/emmanuel-macron-wants-france-become-leader-ai-and-avoid-dystopia.

companies instead of the previous 20 percent.[95] French companies have also been receiving some attention, such as big data company Saagie, which has used AI to offer companies in the financial and insurance industries a big data analytics platform.[96]

GERMANY

The German government has proposed close cooperation with other countries as part of its AI strategy. Germany hopes to link data centers with France, for example, and establish bilateral research programs.[97] Cooperation is also happening in the private sector through cooperation or investment at the corporate level. After the United States, Germany is the second most popular investment destination for Chinese firms investing in advanced technology sectors, such as automobiles and robotics.[98] German automotive supplier Continental AG has teamed up with China's Baidu to jointly develop technology for self-driving cars. Chinese electrical appliance manufacturer Midea bought German robot maker KUKA in 2016.[99]

UNITED KINGDOM

The United Kingdom has agreed to cooperate with France in several research areas including AI. Together, they hope to work on advancing AI capabilities to improve digital services and train future talent to work in the field. UK companies have also been making use of AI and its applications for the private sector. Companies such as Streetbees and Peak, for example, have enjoyed success by leveraging AI solutions to address private needs for consumer intelligence and business analytics.[100]

ISRAEL

In Israel, the government has found opportunities for cooperation on military drones with countries like Japan and the United States. Given Japan's focus on the use of unmanned systems for security, Japan and Israel have announced joint research on

95 Rabesandratana, "Emmanuel Macron Wants France to Become a Leader in AI and Avoid 'Dystopia.'"

96 Deepashri Varadharajan, *State of AI in 2018*, CBInsights, Presentation, 2018 https://www.cbinsights.com/research/early-stage-enterprise-artificial-intelligence-startups/?utm_source=CB+Insights+Newsletter&utm_campaign=912c7ee39d-TuesNL_06_26_2018&utm_medium=e-mail&utm_term=0_9dc0513989-912c7ee39d-90995085.

97 Tim Stockschläger, "The New Federal Government is Planning an AI Center with France," Hannover Messe, Feb. 18, 2018, http://www.hannovermesse.de/en/news/the-new-federal-government-is-planning-an-ai-center-with-france-70848.xhtml.

98 Abishur Prakash, "International AI Developments Include New Partnerships, Police Robots," Robotics Business Review, June 2, 2017, https://www.roboticsbusinessreview.com/security/international-ai-developments-include-new-partnerships-police-robots/.

99 Tom Green, "Weekend Huddle Seals Deal as China's Midea Buys KUKA," Robotics Business Review, June 28, 2016, www.roboticsbusinessreview.com/manufacturing/weekend-huddle-seals-deal-chinas-midea-buys-kuka/.

100 Varadharajan, *State of AI in 2018*.

unmanned surveillance systems.[101] U.S.-Israeli cooperation on military drones has been given new importance as the United States included a section on its National Defense Authorization Act addressing cooperation between the two countries to counter unmanned aerial systems.[102] Israel has also independently developed drones for military use, such as the Harop loitering munition created by the Israeli Aerospace Industries (IAI)[103] and the Guardium, a fully automated UGV deployed by IDF for patrolling the Gaza border.[104] These kinds of unmanned systems could take on AI capabilities in future iterations for the purposes of manned-unmanned teaming or other applications.

SAUDI ARABIA

Saudi Arabia has grand ambitions for the potential of AI, outlined in the kingdom's Vision 2030 plan, including building a city from scratch. NEOM, as the futuristic city is to be called, will rely heavily on AI. The Saudi plan calls for an investment of $500 billion to make the project a reality.[105] Saudi Arabia has also shown interest in the use of AI for military purposes, particularly in UGVs, autonomy, and robotics.[106] In 2016, the country agreed to buy an undisclosed number of Chinese Pterodactyl planes, also known as the Wing Loong, the unmanned aerial vehicles (UAV) manufactured by state-owned Chengdu Aircraft Industry Group.[107] Saudi Arabia has even gone so far as to award Hanson Robotics' "Sophia" citizenship, a first for any nation as countries seek to define norms and standards around AI and robotics.[108]

101 "IAI-Elbit, Mitsubishi-Fuji in Israel-Japan Joint Unmanned Research Pact," DefenseWorld.net, July 1, 2016, http://www.defenseworld.net/news/16485/IAI_Elbit__Mitsubishi_Fuji_In_Israel_Japan_Joint_Unmanned_Research_Pact#.W75L-2hKhpg.

102 Seth Frantzman, "New Defense Budget Bill Foresees US-Israel Counter-drone Cooperation," Defense News, Aug. 13, 2018, https://www.defensenews.com/unmanned/2018/08/13/new-defense-budget-bill-foresees-us-israel-counter-drone-cooperation/.

103 Thomas Gibbons-Neff, "Israeli-made Kamikaze Drone Spotted in Nagorna-Karabajk Conflict," *Washington Post*, Apr. 5, 2016, https://www.washingtonpost.com/news/checkpoint/wp/2016/04/05/israeli-made-kamikaze-drone-spotted-in-nagorno-karabakh-conflict/?noredirect=on&utm_term=.7dd74630d203.

104 Andrew Tarantola, "This Unmanned Patroller Guards Israeli Borders for Days on End," Gizmodo, September 14, 2012, https://gizmodo.com/5943055/the-g-nius-guardium-guards-golan-like-a-golem.

105 Meghan Han, "AI as the New Oil: Saudi Arabia's $500 Billion Smart City," Medium.com, May 17, 2018, https://medium.com/syncedreview/ai-as-the-new-oil-saudi-arabias-500-billion-smart-city-f7b63f7c9423.

106 Ed Clowes, "Interest in Robotic Warfare 'High' in Saudi Arabia," ZAWYA, February 23, 2017, https://www.zawya.com/mena/en/story/Interest_in_robotic_warfare_high_in_Saudi_Arabia-ZAWYA20170224080630/.

107 "Saudi Arabia Buys High-tech China Drones," Arab News, Sept. 1, 2016, http://www.arabnews.com/node/978446/saudi-arabia.

108 Meghan Han, "AI as the New Oil: Saudi Arabia's $500 Billion Smart City," Medium, May 2017, https://medium.com/syncedreview/ai-as-the-new-oil-saudi-arabias-500-billion-smart-city-f7b-63f7c9423.

ESTONIA

Estonia has been at the forefront of digital initiatives in government to make it more efficient and responsive to the needs of society in setting norms, responding to threats, and securing data.[109] From transportation to information security, the Estonian government continues to improve and expand on the capabilities of its electronic services and is now looking for ways to incorporate AI into its efforts to improve e-government initiatives.[110] Estonia's private sector is also making significant contributions to the military use of AI. Milrem Robotics, based in Tallinn, Estonia, has built the world's first fully modular hybrid unmanned ground vehicle (UGV), the Tracked Hybrid Modular Infantry System (THeMIS).

JAPAN

Japan is seeking to boost its defense capabilities through closer bilateral cooperation in AI and robotics. In 2018, the Indian government announced that it would seek closer bilateral cooperation with the Japanese government in AI and robotics to jointly develop unmanned ground vehicles (UGVs).[111] Japan also expressed interest in joint drone research with Israel in 2016.[112] Japanese investment in AI is dominated by the private sector. The 2018 Japanese budget allotted $720 million for AI while Japan's private sector is expected to contribute about $5.4 billion in comparison.[113]

UNITED ARAB EMIRATES (UAE)

The UAE government released its first national AI strategy in 2017. With the AI market in the UAE expected to reach $50 billion by 2025,[114] the Emirates have a goal to make the gulf nation a leader in AI investment in the region and the world. The strategy aims to improve several sectors through the incorporation of AI solutions, including transportation, health care, space, energy, education, and technology. The UAE is turning itself into a hub for investment in unmanned systems and has also

109 Adam Janofsky, "Estonian President, Eyeing Bigger U.N. Role, Urges Government Action on Cybersecurity," *Wall Street Journal*, April 5, 2018, www.wsj.com/articles/estonian-president-eying-bigger-u-n-role-wants-to-raise-cybersecurity-on-global-agenda-1522874300.

110 Federico Plantera, "'Artificial Intelligence is the Next Step for E-governance in Estonia,' State Adviser Reveals," Estonia.com, September 2017, https://e-estonia.com/artificial-intelligence-is-the-next-step-for-e-governance-state-adviser-reveals/.

111 "India, Japan to Introduce AI, Robotics in Defence Sector," *Times of India*, Jan. 22, 2018, https://timesofindia.indiatimes.com/india/india-japan-to-introduce-ai-robotics-in-defence-sector/articleshow/62597018.cms.

112 Abishur Prakash, "Japanese Military Drones, Robotics Develop in Response to U.S.-China Pivot," Robotics Business Review, February 17, 2017, https://www.roboticsbusinessreview.com/unmanned/japanese-military-drones-robotics-develop-response-u-s-china-pivot/.

113 "Japan's Budget for AI to be Less than a Fifth of that Planned by U.S. and China," *Japan Times*, February 25, 2018, https://www.japantimes.co.jp/news/2018/02/25/business/tech/japanese-government-spending-ai-less-20-u-s-china/#.W5FuzEZKiCh.

114 Han, "AI as the New Oil: Saudi Arabia's $500 Billion Smart City."

expressed enthusiasm for unmanned ground vehicles, including the Estonian THeMIS system mentioned above.[115]

SOUTH KOREA

South Korea has long been a world leader in technology and is looking to improve its position in AI. In 2018, the government announced an investment plan calling for $2 billion over five years towards the application of AI solutions in defense, life sciences, and public safety. The plan also included calls for the education of 5,000 AI experts over the next five years.[116] While advanced, South Korea has experienced some controversy as well. The Korea Advanced Institute of Science and Technology (KAIST) faced international backlash for its AI and national defense program over "killer robot" concerns.[117] However, development continues in the private sector. The South Korean-based XBRAIN has received recognition for its cloud-based machine learning assistant that helps developers and scientists build and deploy machine learning models.[118]

INDIA

India allocated $477 million in 2018 for its Digital India program, the government's initiative to "promote AI, machine learning, 3D printing, and other technologies."[119] The program does not limit itself to governance and services, but also extends to the military sector. In 2010, the Indian military expressed the goal of using autonomous systems to conduct 50 percent of military operations and, in 2013, the Indian Defense Research and Development Organization (India's equivalent of DARPA) was said to be developing robots with a high level of intelligence.[120]

AUSTRALIA

The Australian government is looking to build its AI and machine learning capabilities to improve business innovations across several sectors in the country, includ-

115 Muhammad Aamir, "Milrem Intends to Bring Production Unit in UAE," Emirates News Agency, Feb. 22, 2017, http://wam.ae/en/details/1395302599180.

116 Abishur Prakash, "Robotics & Geopolitics: U.S., South Korea Seek AI Leadership with New Efforts," Robotics Business Review, May 18, 2018, www.roboticsbusinessreview.com/regional/u-s-south-korea-ai-leadership-efforts/.

117 Benjamin Haas, "'Killer Robots': AI Experts Call for Boycott over Lab at South Korea University," The Guardian, April 5, 2018, https://www.theguardian.com/technology/2018/apr/05/killer-robots-south-korea-university-boycott-artifical-intelligence-hanwha.

118 Varadharajan, State of AI in 2018.

119 Ananya Bhattacharya, "India Hopes to Become an AI Powerhouse by Copying China's Model," Quartz India, February 13, 2018, https://qz.com/india/1198182/modi-government-pushes-for-artificial-intelligence-like-china-but-is-india-ready-for-it/.

120 Abishur Prakash, "Defense Automation Leads to New Capabilities, Worries," Robotics Business Review, May 29, 2017, https://www.roboticsbusinessreview.com/unmanned/defense-automation-leads-new-capabilities-worries/.

ing agriculture, health care, energy, mining, and cybersecurity. To that end, it has earmarked in its 2018-2019 budget $29.9 million over four years.[121] In the private sector, companies like Sydney-based Hyper Anna are helping the financial services industry with on-demand AI solutions targeting supply chain management, forecasting, and other applications.[122]

PAKISTAN

In Pakistan, the government has announced a plan to invest $3.3 million over three years on AI research and capabilities. The project will be carried out by the Pakistani Higher Education Commission which has selected six universities to host the program. The Pakistani government hopes to address its relative deficit compared to other countries' use of AI to improve capabilities in industrial sectors, warfare, and surveillance.[123]

INTERNATIONAL ORGANIZATIONS, PARTNERSHIPS, NORMS

Several international organizations have also started to develop strategies to deal with artificial intelligence and the potential implications of its use. Privacy standards and adherence to the laws of war are at the heart of the international discussion on ethics and norms in AI. The United Nations has its own AI and Robotics program operated by the UN Interregional Crime and Justice Research Institute (UNICRI), established in 2015. And, in 2016, the UN announced the creation of the Center on Artificial Intelligence and Robotics. The new office's objectives are to monitor developments in AI and robotics to increase understanding of the risks and benefits of these technologies and allow member nations to better coordinate and share information on the topic.[124] Regional intergovernmental organizations are also taking action. For example, the European Union (EU) rolled out a new plan in April 2018 calling for member countries to focus on European advantages, such as research and industry, to improve AI capabilities while taking into consideration its socioeconomic impact.[125] Canada is also drafting ideas for responsible AI usage. The final version of the Montreal Declaration for Responsible AI is expected in late 2018. The declaration is similar in nature to corporate responsibility documents. It

121 Rohan Pearce, "Budget 2018: Government Seeks to Boost Australian AI Capabilities," Computerworld, May 8, 2018, https://www.computerworld.com.au/article/640926/budget-2018-government-seeks-boost-australian-ai-capabilities/.

122 Varadharajan, State of AI in 2018.

123 Aqsa Khunshan, "Pakistan Govt Announces 1.1 Billion Rupees Fund for Artificial Intelligence Projects," Tech Juice, Apr. 21, 2018, https://www.techjuice.pk/pakistan-govt-announces-1-1-billion-rupees-for-artificial-intelligence-projects/.

124 UNICRI Centre for Artificial Intelligence and Robotics, UNICRI, http://www.unicri.it/in_focus/on/UNICRI_Centre_Artificial_Robotics.

125 "Everything You Need to Know about the EU Artificial Intelligence Strategy," Futuribile, May 3, 2018, https://futuribile.org/2018/05/03/everything-need-know-eu-ai-strategy/.

is meant to bring together stakeholders to discuss the way forward regarding the ethical and socially-responsible development of AI.[126]

Defense and security coalitions such as NATO and the Five Eyes information sharing agreement provide a data sharing and environment of commonality that may support AI application. The United States has a unique position of being able to leverage its technological leadership through partnerships and alliances in developing AI technologies as well as establishing norms and policies to guide application by first prioritizing partners with common data and common platforms. To maximize the benefit of AI to U.S. national security, the United States should take the lead in the conversation on ethics and norms as debates on lethal autonomous weapons and AI in general become more prominent in the international arena.

IMPLICATIONS OF DIFFERENT APPROACHES

The diversity of nations investing in AI and the many areas of focus and specialization they are pursuing shows that while interest in AI is relatively uniform, the understanding of what advantages AI generates varies widely. Democratic nations approach AI differently from more authoritarian regimes. The United States and other democratic nations need to be aware of the complete global picture of AI investment and the differing intended uses for AI in different countries. Differences in values, like the value of life and privacy, may lead nations to different advancements and would not restrict their development and implementation of AI in the same way. In some respects, paying less attention to security, privacy, and reliability concerns may lead to rapid initial advancement of AI in certain security applications providing certain advantages to development of AI capabilities for countries that devalue those considerations or are willing to sacrifice them in the short term. At the same time, these nations may engender long-term costs from such decisions that may make it harder to compete globally in the commercial marketplace.

Ethics and values are often reflected in law which can guide development of the technology, but when it comes to AI it can be the other way around. Within the United States, the development of AI technology, and software more broadly, often outpaces the legislative and regulatory processes for establishing standards and guidelines. In other countries, the technology is allowed to develop relatively freely. For example, China has successfully deployed facial recognition technology at scale, but at a cost of privacy in the biometric data it takes from its citizens. These capabilities can be exported and may be attractive to similarly-minded regimes. Consider China's CloudWalk Technology, a Guangzhou-based start-up, which has

126 Montreal Declaration for Responsible AI, University of Montreal, https://www.montrealdeclaration-responsibleai.com.

deployed mass facial recognition programs in Zimbabwe.[127] Second order effects of implementing control and surveillance, and how governments use AI to consolidate and centralize social control will decide if it offers advantages or disadvantages in the long run.

These compromises of privacy by government actors have not gone without controversy or debate in the U.S. as AI technologies are developed. As AI becomes more widespread and accessible, access to a powerful technical capability must be balanced with considerations of privacy and bias in both AI as well as predictive analytic solutions.[128] The pressure will remain on nations of this mindset to join together to establish international norms for AI privacy and security.

CREATING ADVANTAGE IN ARTIFICIAL INTELLIGENCE

How do the societies of other nations around the world create a comparative advantage in AI and machine learning? Given the democratization of AI and the availability of open-source code and algorithms, for machine learning applications in particular, as well as the problem-specific nature of most AI applications, the primary hurdle for initial AI implementation comes from building a robust AI ecosystem. Comparative advantage will be attained not just in having data but in the quality and use of the data as well as the other elements of the AI ecosystem including digital capability and a capable AI workforce. As a result, there is a bit of an obsolescence trap for continuing to neglect the AI ecosystem.

Various nations will begin with different strengths and weaknesses in their AI ecosystems. Consider the example of China: its comparative advantage is in having access to significant volumes of data given the nation's population size and ability to centralize data. AI may also disrupt operations and shift tactical level advantage; examples include biometric tracking, extensive monitoring, or the ubiquitous facial recognition technology in the United Kingdom led to the conviction of Russian operatives in the Skripal poisoning case. Comparative advantage also comes from processing power. The speed at which calculations can be done by even a basic computer today is infinitely faster than humans will ever be capable of unaided. Additionally, the speed at which processing power is maximized can also create large advantages. These comparative advantages can be undermined, however, by weaknesses elsewhere in the AI ecosystem. Enduring advantage then is likely to go to the nations that work comprehensively to foster and strengthen the entire AI ecosystem.

127 Amy Hawkins, "Beijing's Big Brother Tech Needs African Faces: Zimbabwe is Signing Up for China's Surveillance State, But Its Citizens Will Pay the Price," Foreign Policy, July 24, 2018, https://foreignpolicy.com/2018/07/24/beijings-big-brother-tech-needs-african-faces/.

128 Dakin Andone, "Police Used Facial Recognition to Identify the Capital Gazette Shooter. Here's How it Works." CNN, June 29, 2018, https://www.cnn.com/2018/06/29/us/facial-recognition-technology-law-enforcement/index.html; Dan Hurley, "Can an Algorithm Tell When Kids Are in Danger?" *New York Times Magazine*, January 2, 2018, https://www.nytimes.com/2018/01/02/magazine/can-an-algorithm-tell-when-kids-are-in-danger.html

When discussing AI advantage in the international context, Russian President Vladimir Putin's comments about the leader in AI dominating the world immediately come to mind. AI is a collection of technologies, one that requires more precise language. Putin's comments are of a general nature, a different scope of AI compared to the specificity of computer vision and self-driving cars. It is not necessarily true that the first entity to adopt self-driving cars will rule the world, but the suite of technologies that self-driving cars require means that the creators have reached a critical point in development where the technologies are advanced enough that they would be able to be adapted to other uses. Although technology may be transferrable from one use to the next, data may not be, leaving developers always looking for more varieties of data. The advantages of a robust AI ecosystem, however, are readily transferrable among different applications.

In many ways, the archetype of the successful AI ecosystem exists today in Silicon Valley. Here companies have ready access to a technically skilled workforce, investment capital, large volumes of data, robust networking, inexpensive computing power, and experienced technology executives. While replicating these dynamics in other countries, or in the U.S. government, is not simple, it is also not impossible. Silicon Valley is a good example of how a robust AI ecosystem can create an enduring advantage through a combination of money, talent, and terms of the trade all in one geographic location, already being mimicked in Shenzen, China which is known as the Silicon Valley of China.

"THE PRIMARY HURDLE FOR INITIAL AI IMPLEMENTATION COMES FROM BUILDING A ROBUST AI ECOSYSTEM."

Another aspect of creating an advantage in AI comes from the fact that groups of actors around the world can use the technology–not just a few nations or companies. The technology will also be handled by individual consumers. For example, commercial drones have proliferated throughout the world through a variety of vendors. While easily accessible, these commercial platforms incorporate features like live camera views, obstacle avoidance, and the ability to follow pre-determined flight paths, that pose a legitimate national security risk in the right circumstances.[129] These private sector experimenters can develop new ways of utilizing AI for purposes both beneficial and detrimental to national security.

THE VALUE OF MOVING FIRST

The problem-specific nature of AI conditions the ability of AI to provide a specific first-mover advantage. It limits the scope and scale for what can be gained from introducing AI to individual military capabilities. Countries that develop their AI ecosystems will have the ability to apply AI to the data their systems produce whether

129 Divya Joshi, "Here are the World's Largest Drone Companies and Manufacturers to Watch and Invest In", Business Insider, July 18, 2017, https://www.businessinsider.com/top-drone-manufacturers-companies-invest-stocks-2017-07.

or not they receive assistance from outside sources. Advantage depends on a few factors: the sector in which AI is being deployed, the size and breadth of the deployment, and the timeline of action. In areas such as cybersecurity or defense, where humans may not be capable of responding at the necessary speeds, AI could present a big advantage to a country that successfully applies it in those applications first. Electronic warfare is similar, in that the complexity of frequency-hopping emissions means that both attack and defense must constantly move across the spectrum. In cybersecurity, AI technologies can be used with bot nets to overwhelm defenses. As cyber defense capabilities use more AI technology, defenders are more able to operate at the speed and scale of the attackers.

The intense interest and investment in AI from both U.S.- and Chinese-based tech giants suggests that these firms may be seeking a first-mover advantage. The size and breadth of commercial sector companies allows them to acquire promising and innovative smaller firms and to leverage their access to the huge volumes of customer data they routinely collect. The largest of these firms may be able to establish an entire AI ecosystem within its own organization. For these firms, ensuring they have the scope and scale in their AI ecosystem to be able to innovate in AI without dependence on their rivals provides a strong incentive to invest and grow. However, a second-mover advantage exists in those areas with longer timelines; this strategy lets the pioneer find all the pitfalls while setting up the competition to execute effectively. A good example of this is Facebook being able to capitalize on the shortcomings of Myspace.

In general, first-mover advantages for AI appear likely to be limited and short-lived. For short timelines, a first-mover or first strike advantage may be more significant. However, the behavior of corporate tech giants suggests a strong incentive to have an AI ecosystem robust enough to allow for AI independence from outside actors that may not be reliable partners.

IMPLICATIONS FOR THE UNITED STATES

The case for increased U.S. investment in AI in response to international competition is compelling. Potential competitor nations are enacting international AI investment strategies, including both military and economic competitors. While the United States has an advantage in the existing robust AI ecosystem in Silicon Valley, it would be foolhardy to believe that this advantage will be sufficient or long enduring without substantial investment and research. Not everything about the international picture is negative, however. The United States and its technology firms are attractive partners for others looking to participate in the AI game and U.S. partners and allies are making substantial investments that can potentially be leveraged to the benefit of the United States. Furthermore, since AI is likely to provide the foundation for progress in both military and economic spheres, it is a compelling area of focus for every nation with the technical ability to pursue it.

While the United States has a robust AI ecosystem in Silicon Valley, this capability is no substitute for building the AI ecosystem within the U.S. government, particularly within the national security and intelligence agencies. Only with a robust U.S. government AI ecosystem will DoD be in a position to capitalize on new AI capabilities as they emerge and quickly transfer them to military applications. Speed of adoption is key.

International development and investment in U.S.-based AI companies and start-ups also poses risks. These companies are being aggressively pursued by international competitors. This presents a two-sided threat. On the one hand, if an adversary succeeds in acquiring technology from a U.S. firm or start-up through investment, that is a detriment to national security. On the other hand, if the United States refuses to work with companies that have Chinese investment or funding because of the national security risk that this investment might pose, these companies could lose access to significant investment potential and the U.S. government also loses a potential partner for new ideas and capabilities. Both the United States and international competitors, therefore, could consider strategic investments in other nations' AI ecosystems to gain advantage. However, the open nature of the U.S. economy, in contrast to many potential competitor nations, makes the risk of strategic AI investing more acute for the United States. The recent strengthening of the investment review processes in the United States and many of its partners and allies is an important corrective.

CONCLUDING REMARKS

THE PROCESS OF UNDERSTANDING AI's implications for national security, just like the AI field itself, is still in an early stage. A major goal of this report is to establish what key steps need to be taken to facilitate the successful integration of AI into national security applications based on an accurate understanding of where the AI field currently stands and what key factors are involved in successful AI adoption and management. It makes sense then to conclude with a discussion of the some of the future analytic work required to further this objective.

A central finding of this effort is the importance of building a robust AI ecosystem in government, addressing the twin issues of workforce debt and technical infrastructure debt. There is substantial work to be done in understanding and mitigating each of these debts. The John S. McCain National Defense Authorization Act for Fiscal Year 2019 included important reforms to the Department of Defense personnel management system, providing personnel tools for increasing workforce specialization that can be applied to recruiting and retaining the technical talent required for AI implementation. Further study can help guide the Department of Defense's use of these authorities to provide the incentives most valued by the targeted workforce, for which government will continue to compete with the private sector. Further study can also help identify the critical areas of technical infrastructure debt that must be tackled to advance the government AI ecosystem and how this debt can most efficiently be paid, such as access to cloud computing, workforce development and education, and other capabilities.

The development of AI requires significant improvement in the way the Department of Defense approaches the acquisition of software. Here, further study is already underway. The Defense Science Board recently completed a study on this topic and a further study is underway at the Defense Innovation Board. CSIS has a related study on the acquisition of software-defined hardware-based systems. One area where there is likely to be room for further study is in identifying a successful, enduring business model for iterative software development that properly aligns incentives between the government and its private sector software development partners. Just as fundamental to successful acquisition of AI is the work of understanding how to

test AI so that the verification and validation of AI performance can be demonstrated. The Institute for Defense Analyses has done important research on this question, but the work required to succeed on this front has far to go on both the theoretical and practical levels.

Significant legal issues associated with AI remain very much in doubt. There is currently little understanding of how to assign liability for failures in AI performance and adverse outcomes from AI usage. There is also little clarity on the ownership of intellectual property (IP) created by AI, both data and algorithms. The answer to questions of IP ownership will be critical to the incentive structure created for AI developers. And finally, the question of ethical development and employment of AI remains a vital issue for further study. This not only includes the critical issue of how to ensure that AI is not used in violation of the laws of war and respects the related protections of human rights, but also the protection of individual privacy and human autonomy.

Now is the right time to study these and other questions surrounding AI. This is important so that policymakers are appropriately prepared to react to the development of AI, but also to shape this development in furtherance of strategic objectives.

SUMMARY OF KEY FINDINGS AND RECOMMENDATIONS

A summary of the recommendations presented throughout this report, as well as their corresponding aspect of the AI ecosystem, is presented below.

Trust

- The importance and necessity of AI transparency is application-specific.
- Trust must be met across algorithms, data, and outcomes.
- Users must understand the mechanisms by which systems can be spoofed.

Security

- Robust and resilient digital capability requires balancing development, operations, and security.
- A culture of network risk management and cybersecurity ownership throughout and across organizations is critical.

People

- Applying AI requires a skilled and educated workforce with domain expertise, technical training, and the appropriate tools.
- Organizations must cultivate a culture of data excellence.
- Success for users in machine learning requires iteration, experimentation, and learning through early sub-optimal performance.

Digital Capability

- An organization must build the foundational digital capability to successfully apply AI technologies (e.g., database management, information integration). This is necessary to pay down the tech debt.
- Gaining competitive advantage through information and analytics is an enterprise-wide endeavor from headquarters to the deployed warfighter.

Policy

- Ethical policies and standards must guide the application and implementation of AI technologies.
- The U.S. government must strengthen its own AI ecosystem through the following steps:
    - *Reform hiring authorities and security clearance processing to support bringing in key government and access for contractor personnel.*
    - *Improve the government's ability to acquire and iterate developmental software by changing budgeting practices for software development.*
    - *Engage industry broadly and spread bets, utilizing small- to medium-sized data science firms in addition to the tech and defense industry giants, because the problem-specific nature of AI and the early stage of the field mean it is impossible to know where the breakthroughs will come from.*
    - *Invest in early stage research and development, specifically those areas requiring federal support that may be less commercially viable.*
    - *Develop tools for AI trust, security, explainability, validation, and verification that can address the high threshold for AI reliability that many government applications will require.*
- Leveraging AI capability means structuring organizations to support the right mix of technical knowledge and domain expertise.
- The U.S. government must recognize the implications of international activity in AI and move to:
    - *Protect the robust private sector AI ecosystem in the United States and partner nations from attacks and detrimental investment; and*
    - *Leverage partner nation resources by working first with those partners with common objectives, equipment, and data-sharing agreements while building that commonality with additional partners.*

APPENDIX A

Appendix A provides an overview of existing software and hardware policies that impact AI systems.

Department of Defense Directive 3000.09, Autonomy in Weapons Systems

Effective November 21, 2012, the Department of Defense Directive (DoDD) on Autonomy in Weapons Systems addresses various policy concerns surrounding the development and use of autonomous and semi-autonomous weapons. This directive explicitly requires autonomous and semi-autonomous weapons "to minimize the probability and consequences of failures" by codifying regulatory limitations and assigning human control and responsibility. As defined by the DoD, an autonomous system is one that once activated, can select and engage targets without further intervention by a human operator. However, this does not exclude the ability for a human operator to override any decisions the system makes. A semi-autonomous system is one that only engages targets after they are selected by a human operator, which includes "fire-and-forget" systems.

The directive has two parts. The first establishes guidelines for the development and use of autonomous and semi-autonomous functions in weapon systems, and the second assigns responsibilities to the Under Secretary of Defense for Policy (USD(P)), Under Secretary of Defense for Acquisition, Technology and Logistics (USD(AT&L)), and the Chairman of the Joint Chiefs of Staff (CJCS), intended to minimize the probability and consequences of failures in these weapon systems that could lead to unintended or incorrect engagements. Autonomous systems are still in the early stages of development and the DoD cannot risk having a weapon mistaking a civilian contingent for a military one, or even losing control of the system. A significant part of this directive lies in how weapon systems are to be used and developed. The directive states:

> *Persons who authorize the use of, direct use of, or operate autonomous and semi-autonomous weapon systems must do so with appropriate care and in accordance with the law of war, applicable treaties, weapon system safety rules, and applicable rules of engagement.*[130]

130 U.S. Department of Defense, Directive, *Autonomy in Weapons Systems*, Number 3000.09, November 21, 2012, http://www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodd/300009p.pdf.

Under this section, any use of these weapon systems must fall in line with established national and international norms and regulations. Article 36 of the Geneva Conventions, which the United States has ratified, aims to prevent weapons that would violate international humanitarian law (IHL), also known as the law of war. The remainder of the directive assigns roles and responsibilities to members of the DoD to ensure autonomous and semi-autonomous weapon systems are being developed in a way that complies with IHL. All levels of development are addressed in the directive so that development can proceed cautiously and deliberately.

Department of Defense Instruction 5000.02, Operation of the Defense Acquisition System

The Department of Defense Instruction (DoDI) 5000.02 was initially released on January 7, 2015 and revised in August 2017 to include additional acquisitions milestone models that may be used to smooth purchase of AI systems, including the Defense Unique Software Intensive model, Incrementally Deployed Software Intensive model, and a Hybrid (Software Dominant) model.

The Department of Defense Instruction (DoDI) 5000.02 provides models that serve as examples of defense program structures tailored to the type of product being acquires or the need for accelerated acquisition. Three of these models can help smooth the process of purchasing AI systems: Defense Unique Software Intensive Model, Incrementally Deployed Software Intensive Model, and the Hybrid (Software Dominant) Model.

The Defense Unique Software model is one that is dominated by the need to develop a complex, usually defense unique, software program that will not be fully deployed until several software builds are completed. These builds are central to this model as a series of testable subsets of the overall capability that , together with a clearly defined decision criteria, ensures adequate progress is being made at each step. Examples of this product include command and control systems and significant upgrades to combat systems found on major weapon systems. Many AI systems will likely fall into this category since their inherent design makes them a major upgrade. The ability to analyze vast amounts of data and make decisions infinitely faster than a humans will increase the pace of combat to a point where having human analysts at every level might even slow decision-making in an extremely time-sensitive environments.

The Incrementally Deployed Software Intensive model differs from the previous model by in the rapid delivery of capability through multiple acquisition increments, each of which provides part of the overall required capability, compared to requiring full capability before deployment. Several builds and deployments may still be necessary to satisfy the requirement for an increment of capability. This model allows for improvements in a system as new capabilities are discovered. Learning AI systems are always evolving and becoming more complex, and as new capabilities are discovered and required by defense systems, it is important to have a standardized method to incorporate that into existing systems.

The Hybrid (Software Dominant) model is a combination of the previous two models: how a new system can be released incrementally as well as include intermediate software builds. New discoveries with AI systems are likely to be marginal improvements as well as huge breakthroughs. Regardless, it is important to understand and have a method of incorporating progress into current deployments.

Department of Defense Instruction 5000.75, Business System Requirements and Acquisition

Department of Defense Instruction (DoDI) 5000.75, on Business Systems Requirements and Acquisition, presents acquisition instruction specifically for information technology systems, which may include AI given the software nature of AI technologies. The changes from 5000.02 allow a wider variety of options to follow when assessing acquisition requirements and allow for more variation. Effective February 2, 2017, DoDI 5000.75 establishes policy for the use of the business capability acquisition cycle requirements and acquisition and supersedes DoDI 5000.02 for all business acquisition programs that are not designated as Major Defense Acquisition Programs (MDAPs) according to DoDI 5000.02. The reason for the change is that in practice, tailoring the models from 5000.02 took too much time and effort, making it difficult to justify the benefits it produced. Changes from 5000.02 include alignment of acquisition, function, infrastructure, and IT investment governance to streamline decision making and creating an information-centric approach to evaluating programs (rather than relying on acquisition and requirements documentation).

Foreign Investment Risk Review Modernization Act of 2017 (H.R.4311)

H.R. 4311 / S. 2098, the Foreign Investment Risk Review Modernization Act of 2018 (FIRRMA), was introduced November 8, 2017, addressing concerns of the Committee of Foreign Investments in the United States (CFIUS) about foreign investment in tech firms.

House Bill 4311, introduced by Representative Robert Pittenger, and Senate Bill 2098, introduced by Senator John Cornyn, are identical bills with the goal of making changes in the way the CFIUS operates to minimize national security concerns. On the U.S. Treasury website, CFIUS is described as an interagency committee authorized to review foreign investment transactions in domestic companies to determine the effect of said transactions on United States national security.[131]

Under the original CFIUS review process, reviews were only triggered by covered transactions, which essentially meant a full takeover of the company by a foreign entity. If, during the review process, CFIUS found a threat to national security, it could impose certain conditions of acquisition before allowing the deal to proceed. The review process now covers joint-venture projects, minority position investments, and real estate transactions near military bases or other sensitive facilities.

131 U.S. Department of the Treasury, *The Committee on Foreign Investment in the United States* (CFIUS), https://home.treasury.gov/policy-issues/international/the-committee-on-foreign-investment-in-the-united-states-cfius

Regardless of intentions of these foreign investments, FIRRMA widens the scope of CFIUS review to protect U.S. national security interests.

This update to CFIUS is especially relevant to AI. If foreign investors are using the previously mentioned methods to gain access to technology developed on U.S. grounds, they can bring it back to their home soil to further develop it. When it comes to AI, there is increased recognition that it is the next big thing, so when the United States loses access to research and knowledge to foreign countries, it affects the ability of the U.S. government to develop it for domestic gain and use it to protect U.S. national security. Senator Richard Burr, a cosponsor of S. 2098, says, "the CFIUS process is key to proactively identifying and mitigating foreign efforts to acquire critical U.S. technology and know-how through investment."[132]

Fundamentally Understanding the Usability and Realistic Evolution of Artificial Intelligence Act of 2017 (S. 2217)

Senate Bill 2217 was introduced on December 12, 2017 by Senator Maria Cantwell aiming to establish the Federal Advisory Committee on the Development and Implementation of Artificial Intelligence. Before fleshing out the duties of the newly-established advisory committee, the bill acknowledges the importance of AI to the well-being of the country, noting that understanding and preparing for the ongoing development of AI is critical to the economic prosperity and social stability of the Unites States. This bill also defines for the advisory committee AI as "any artificial system that performs tasks under varying and unpredictable circumstances, without significant human oversight, or that can learn from their experience and improve their performance. ... They may solve tasks requiring human-like perception, cognition, planning, learning, communication, or physical action." [133]

No later than 540 days after the bill is enacted, the advisory committee must submit a report to Congress with findings and recommendations for the following conditions, among others:

- *The competitiveness of the United States, including matters relating to the promotion of public and private sector investment and innovation into the development of artificial intelligence.*
- *Ethics training and development for technologists working on artificial intelligence.*
- *Accountability and legal rights, including matters relating to the responsibility for any violations of laws by an artificial intelligence system and the compatibility of international regulations.*

132 "Cornyn, Feinstein, Burr Introduce Bill to Strengthen the CFIUS Review Process, Safeguard National Security", Nov. 11, 2017, https://www.cornyn.senate.gov/content/news/cornyn-feinstein-burr-introduce-bill-strengthen-cfius-review-process-safeguard-national

133 U.S. Congress, House, *FUTURE of Artificial Intelligence Act of 2017*, H.R. 4625, 115th Cong., 1st sess. Introduced in House December 12, 2017, https://www.congress.gov/bill/115th-congress/house-bill/4625.

- *How to create a climate for public and private sector investment and innovation in artificial intelligence.*
- *Whether and how networked, automated, artificial intelligence applications and robotic devices will displace or create jobs and how any job-related gains relating to artificial intelligence can be maximized.*
- *How the privacy rights of individuals are or will be affected by technological innovation relating to artificial intelligence.*
- *How existing laws, including those concerning data access and privacy, should be modernized to enable the potential of artificial intelligence.*
- *How the Federal Government utilizes artificial intelligence to handle large or complex data sets.*

The bill also explicitly lays what groups should be represented on the advisory committee, including members from the academic and research community and civil liberties groups. The committee is also given the abilities to hold hearings and conferences, issue reports, and enter cooperative agreements with third-party experts to further the understanding the future of AI. This bill highlights the U.S. government's foray into AI and the role it can and will play in society in the near future. The more lawmakers and the public know about AI, the easier it will be to implement and apply it in a safe and effective manner. The bill was referred to the Committee on Commerce, Science, and Technology on December 12, 2017 and has yet to be passed out to the Senate floor.[134]

National Institute of Standards and Technology (NIST) Risk Management Framework (RMF)

The National Institute of Standard and Technology (NIST) Risk Management Framework (RMF) provides a process that integrates security and risk management activities into the system development life cycle. According to the NIST website, the risk-based approach to security control selection and specification considers effectiveness, efficiency, and constraints due to applicable laws, directives, executive orders, policies, standards, or regulations. In other words, their approach dictates how U.S. government security and computer systems must be designed, secured, and monitored. The six steps in the RMF are reproduced below:

1. ***Categorize*** *the security level of a project based on the potential impact on an organization if certain events occur which jeopardize the information and systems needed by the organization to accomplish its mission, protect its assets, fulfill its legal responsibilities, maintain functions, and protect individuals.*

2. ***Select*** *an initial set of baseline security controls for the system based on the security categorization; tailoring and supplementing the security control baseline as needed based on organization assessment of risk and local conditions.*

3. ***Implement*** *the security controls and document how the controls are deployed within the system and environment of operation.*

134 Ibid.

4. ***Assess*** *the security controls using appropriate procedures to determine the extent to which the controls are implemented correctly, operating as intended, and producing the desired outcome with respect to meeting the security requirements for the system.*

5. ***Authorize*** *system operation based upon a determination of the risk to organizational operations and assets, individuals, and other organizations and the Nation resulting from operation of the system and the decision that this risk is acceptable.*

6. ***Monitor*** *and assess selected security controls in the system on an ongoing basis including assessing security control effectiveness, documenting changes to the system or environment of operation, conducting security impact analyses of the associated changes, and reporting the security state of the system to appropriate organizational officials.*

A recent update to the RMF has several important additions, including connecting senior leaders to operations and incorporating supply chain risk management considerations. In the context of AI, the RMF provides a structure for safely developing the technology while also keeping more people accountable for the results. Ongoing monitoring ensures safety and privacy for all individuals involved and helps keep vulnerable information safe.[135]

135 "Risk Management Framework (RMF) Overview," National Institute of Standards and Technology (NIST) Computer Security Resource Center, https://csrc.nist.gov/projects/risk-management/risk-management-framework-(rmf)-overview.

ABOUT THE PROJECT DIRECTOR AND LEAD AUTHOR

Andrew P. Hunter is a senior fellow in the International Security Program and director of the Defense-Industrial Initiatives Group at CSIS. From 2011 to 2014, he served as a senior executive in the Department of Defense, serving first as chief of staff to undersecretaries of defense (AT&L) Ashton B. Carter and Frank Kendall, before directing the Joint Rapid Acquisition Cell. From 2005 to 2011, Mr. Hunter served as a professional staff member of the House Armed Services Committee. Mr. Hunter holds an M.A. degree in applied economics from the Johns Hopkins University and a B.A. in social studies from Harvard University.

Lindsey R. Sheppard is an associate fellow with the International Security Program at CSIS, where she supports various projects in emerging technology, including artificial intelligence and machine learning, and in security applications, ranging from strategic to tactical. Ms. Sheppard contributes expertise in modeling and simulation, system architecture, electronic warfare, and radar from several years of experience in defense research and development. Before joining CSIS, she was a member of the technical staff at the Charles Stark Draper Laboratory and the Georgia Tech Research Institute, during which time she served as the systems engineering lead on multi-year efforts building simulation capabilities to evaluate technology and deployment solutions to support military operations. She holds an M.S. and a B.S. in aerospace engineering from the Georgia Institute of Technology.

CSIS | CENTER FOR STRATEGIC & INTERNATIONAL STUDIES

1616 Rhode Island Avenue NW
Washington, DC 20036
202 887 0200 | www.csis.org