United States Government Accountability Office

Report to the Committee on Armed Services, U.S. Senate

ARTIFICIAL INTELLIGENCE

Status of Developing and Acquiring Capabilities for Weapon Systems

February 2022

ARTIFICIAL INTELLIGENCE

Status of Developing and Acquiring Capabilities for Weapon Systems

Why GAO Did This Study

DOD has reported that AI is poised to change future battlefields and the pace of threats the U.S. faces. AI capabilities could enable machines to perform tasks that usually require human intelligence, such as identifying potential threats or targets on the battlefield. DOD designated AI a top modernization area and is investing heavily in AI tools and capabilities. Other nations are making significant investments in this area that threaten to erode the U.S. military technological and operational advantage.

The National Security Commission on Artificial Intelligence concluded in its March 2021 report that the U.S. needs to act quickly to ensure AI readiness. AI experts from inside and outside DOD agree that ensuring the department has the necessary infrastructure in place will be essential to developing, acquiring, and scaling AI for weapon systems effectively.

Senate Report 116-236 includes a provision for GAO to review DOD's AI warfighting acquisition-related efforts. This report examines (1) the unique nature of AI and current status of AI capabilities that support weapon systems, and (2) how DOD is addressing challenges in developing, acquiring, and deploying AI capabilities for weapon systems.

To do this work, GAO interviewed officials from over 20 DOD entities and reviewed DOD documentation as well as reports and recommendations from the National Security Commission on Artificial Intelligence, among others.

What GAO Found

The Department of Defense (DOD) is actively pursuing artificial intelligence (AI) capabilities. AI refers to computer systems designed to replicate a range of human functions and continually get better at their assigned tasks. GAO previously identified three waves or types of AI, shown below.

Types of Artificial Intelligence (AI) and Associated DOD Examples
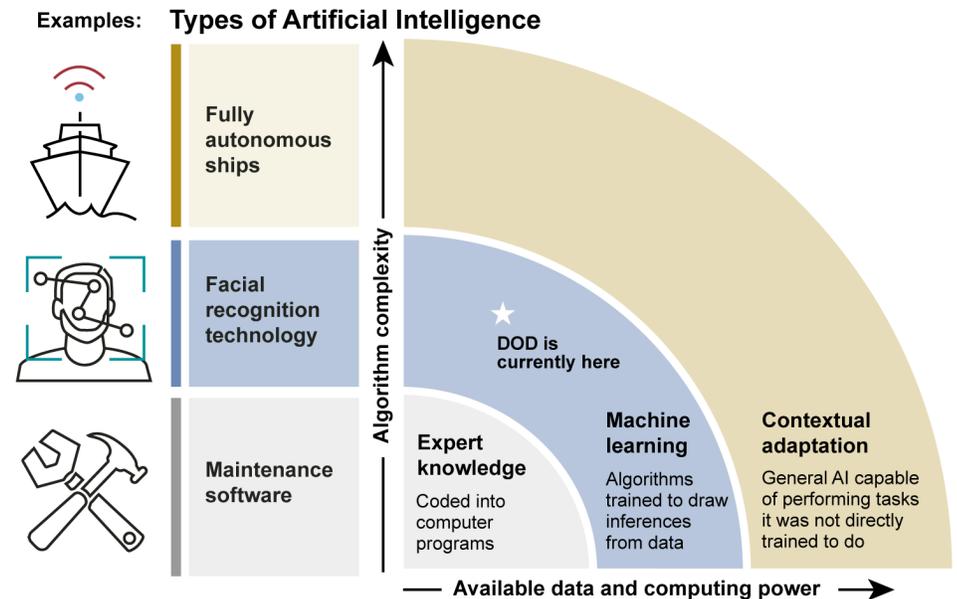
Source: GAO analysis of information from the Department of Defense (DOD) and academic experts. | GAO-22-104765

DOD recognizes that developing and using AI differs from traditional software. Traditional software is programmed to perform tasks based on static instructions, whereas AI is programmed to learn to improve at its given tasks. This requires large data sets, computing power, and continuous monitoring to ensure the capability performs as intended. The majority of AI capabilities that support DOD's warfighting mission are still in development. These capabilities largely focus on analyzing intelligence, enhancing weapon system platforms such as aircraft and ships that do not require human operators, and providing recommendations on the battlefield (such as where to move troops).

When acquiring new capabilities that depend on complex software, DOD has historically faced challenges, such as long acquisition processes and a shortage of skilled workers. GAO found that it continues to face these challenges along with others specific to AI, including having usable data available to train the AI. For example, AI for detecting an adversary's submarines requires gathering many images of various submarines and labeling them so the AI can learn to identify one on its own. DOD also faces difficulties integrating trained AI into existing weapon systems that were not designed for it and building trust in AI among its personnel. DOD initiated a variety of efforts—such as establishing a cross-service digital platform for AI and AI-specific trainings—to address these challenges and support its pursuit of AI, but it is too soon to assess effectiveness.

United States Government Accountability Office

Contents

Abbreviations

AI	artificial intelligence
DARPA	Defense Advanced Research Projects Agency
DOD	Department of Defense

February 17, 2022

The Honorable Jack Reed
Chairman
The Honorable James M. Inhofe
Ranking Member
Committee on Armed Services
United States Senate

Artificial intelligence (AI) is a transformative technology that the Department of Defense (DOD) has reported is poised to change future battlefields and the pace of threats the U.S. faces. AI capabilities will enable machines to perform tasks that normally require human intelligence, such as drawing conclusions and making predictions.[1] Moreover, AI-enabled machines can be expected to maneuver and change tactics at speeds that human operators cannot. Due in part to its potential to provide capabilities across a broad range of uses, DOD designated AI a top modernization area and is investing considerable effort and funds toward developing and acquiring AI tools and capabilities to support the warfighter. For fiscal year 2022, DOD is requesting $14.7 billion for science and technology programs as well as $874 million to directly support its AI efforts. According to DOD's 2018 Artificial Intelligence Strategy, failure to incorporate AI capabilities into weapon systems could hinder the ability of warfighters to defend our nation against near-peer adversaries.[2] Other nations are making significant investments in this area that threaten to erode the U.S. military technological and operational advantage.

The National Security Commission on Artificial Intelligence—established by Congress in 2018 to consider ways to advance the development of AI to address U.S. national security and defense needs—concluded in its March 2021 report that the U.S. is not prepared to defend itself in the AI

---

[1]For the purposes of this report, AI refers to machine learning capabilities unless otherwise noted.

[2]Department of Defense, *Summary of the 2018 Department of Defense Artificial Intelligence Strategy: Harnessing AI to Advance Our Security and Prosperity* (Feb. 12, 2019).

GAO-22-104765 Artificial Intelligence

era.[3] The commission noted that in order to ensure AI-readiness by 2025, the U.S. government needs to act quickly. The commission recommended that, among other things, DOD establish a digital ecosystem—including common development, testing, and data repository environments—to effectively scale AI across the department, restructure its workforce to better train and retain data scientists, and better use adaptive acquisition pathways and contracting flexibilities to procure AI capabilities.[4] AI experts from DOD and external organizations, such as RAND and the Congressional Research Service, agree that ensuring DOD has the necessary infrastructure, including tools and talent, in place will be essential to developing, acquiring, or scaling AI for weapon systems effectively.

Given the growing significance of AI to DOD's acquisition goals, the Senate Report accompanying a bill authorizing fiscal year 2021 appropriations included a provision for GAO to review DOD's AI warfighting acquisition related efforts.[5] This report examines (1) the unique nature of AI and the current status of DOD's AI capabilities that support weapon systems, and (2) DOD's efforts to address identified challenges in developing, acquiring, and deploying AI capabilities for weapon systems. AI capabilities supporting weapon systems are those that DOD can use in its warfighting operations that allow the U.S. to gain decision, speed, agility, and strategic advantages over its adversaries. To align with how DOD categorizes its AI capabilities, we refer to these as AI capabilities for warfighting operations throughout the report.

To understand the status of DOD's AI capabilities for warfighting operations and how DOD is addressing challenges in developing, acquiring, and deploying AI capabilities, we reviewed DOD documentation, such as its 2021 AI inventory and AI-related strategies. We also interviewed relevant officials from over 20 DOD organizations. These organizations include DOD's AI focal point—the Joint AI Center,

---

[3]John S. McCain National Defense Authorization Act for Fiscal Year 2019, Pub. L. No. 115-232, § 1051 (2018). The National Security Commission on Artificial Intelligence, Final Report (Mar. 1, 2021), available at https://www.nscai.gov/2021-final-report/.

[4]DOD established an Adaptive Acquisition Framework in January 2020 that provides six acquisition pathways—including urgent capability acquisition and software acquisition—that provide DOD entities with flexibilities in selecting an acquisition process that matches the characteristics of the capability being acquired. Department of Defense Instruction 5000.02, *Operation of the Adaptive Acquisition Framework* (Jan. 23, 2020).

[5]S. Rep. No. 116-236, at 131 (2020).

established in 2018—and entities within the military services primarily responsible for AI projects and related activities.[6] In addition, we reviewed recently published reports and recommendations by the National Security Commission on Artificial Intelligence and the RAND Corporation and interviewed the authors. We also interviewed officials from four selected private companies that work with DOD regarding the development and deployment of a variety of AI capabilities. For more information on our scope and methodology, see appendix I.

We conducted this performance audit from January 2021 to February 2022 in accordance with generally accepted government auditing standards. Those standards require that we plan and perform the audit to obtain sufficient, appropriate evidence to provide a reasonable basis for our findings and conclusions based on our audit objectives. We believe that the evidence obtained provides a reasonable basis for our findings and conclusions based on our audit objectives.

Background

Types and Complexity of Various AI Technologies

While there are various definitions of AI, in general, AI refers to computer systems that are able to solve problems and perform tasks that have traditionally required human intelligence and that continually get better at their assigned tasks. According to DOD and other AI experts, the field of AI has experienced technological breakthroughs for several reasons: recent advances in the tools and ability to develop and use large data sets in computer applications—often referred to as big data—and computing power, as well as the development of increasingly complex algorithms and the availability of open source-code libraries.[7] In prior

---

[6]In December 2021, the Deputy Secretary of Defense issued a memorandum announcing the establishment of a new position effective February 1, 2022 within the Office of the Secretary of Defense called the Chief Digital and AI Officer. According to this memo, the Office of the Chief Digital and AI Officer will serve as the successor to the Joint AI Center and is expected to be fully operational by June 1, 2022. Deputy Secretary of Defense Memorandum, *Establishment of the Chief Digital and Artificial Intelligence Officer* (Dec. 8, 2021).

[7]Big data, according to the National Institute of Standards and Technology, consists of extensive data sets that require scalable architecture for efficient storage, manipulation, and analysis. Algorithms generally refer to a sequence of instructions to solve a problem.

work, we identified three waves or types of AI, which are highlighted in figure 1.[8]

Figure 1: Types of Artificial Intelligence and Associated DOD Examples
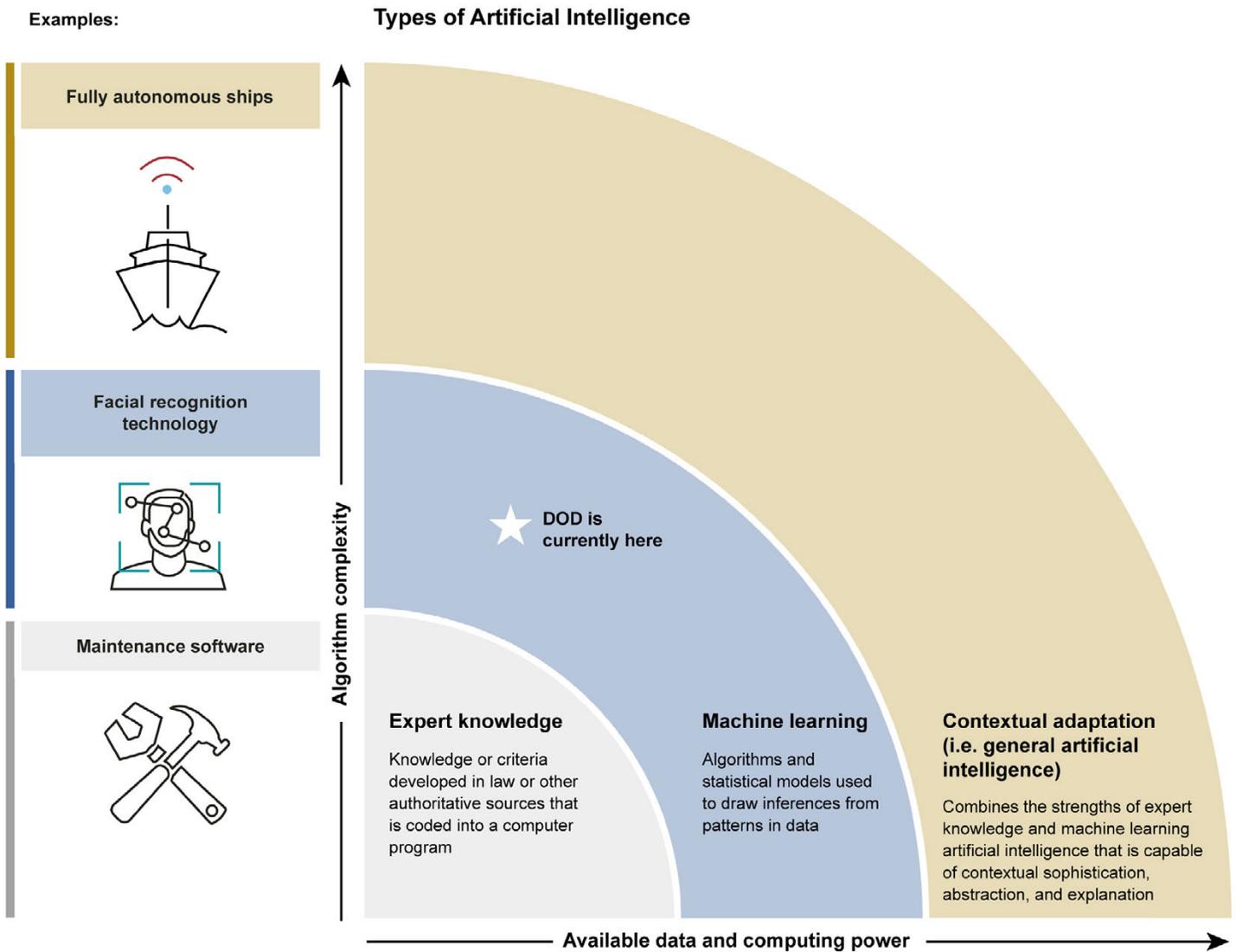
Source: GAO analysis of information from the Department of Defense (DOD) and academic experts. | GAO-22-104765

[8]GAO, *Artificial Intelligence: Emerging Opportunities, Challenges, and Implications,* GAO-18-142SP (Washington, D.C.: Mar. 28, 2018).

GAO-22-104765 Artificial Intelligence

- **Expert knowledge.** The first and oldest type of AI whereby a computer is programmed with detailed rules based on human expertise or criteria and produces outputs consistent with its programming. An example of such rules-based AI capabilities for DOD is maintenance software for aircraft that requires users to input their information according to prespecified data formats and then processes that data according to rules programmed by human experts (i.e., maintenance professionals) to diagnose the cause of malfunctions.

- **Machine learning.** The second and current type of AI whereby a computer is given basic instructions and fed training data to learn how to predict specific outcomes. According to an academic publication, machine learning AI is an appropriate solution when writing a program for a machine to follow is too time-consuming or otherwise not possible.[9] Instead of explicit programming, this type of AI requires a developer to select an appropriate algorithm based on the desired result, feed it the appropriate training data, and watch to see if the algorithm learns what it is supposed to. If the AI model is not performing as expected, the developer can revise the training data, adjust the algorithm parameters, or chose a different algorithm. An example for DOD is facial recognition technology that uses a set of algorithms to identify individuals by instantaneously searching databases of faces and comparing them to those detected in a video or photograph.[10]

- **Contextual adaptation.** The third and potential future type of AI whereby a computer is capable of adapting to new situations without needing to be retrained while also being able to explain to users the reasoning behind its decisions or predictions. A potential example for DOD is a fully autonomous ship that uses algorithms to maneuver in situations it was not specifically trained for (such as inclement weather or contested waters) and is capable of planning, relaying, and carrying out military missions similar to the way a human would.

The field of AI moved into machine learning around the 1950s and, according to DOD documentation, has recently seen a massive increase in the number of real-world applications where machine learning AI is now

---

[9]Massachusetts Institute of Technology Sloan School of Management, *Machine learning, explained (*April 2021), available at https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained.

[10]In August 2021, GAO reported on facial recognition technology in the federal government. *GAO, Facial Recognition Technology: Current and Planned Uses by Federal Agencies*, GAO-21-526 (Washington, D.C.: Aug. 24, 2021).

practical and powerful. These AI applications include those for natural language processing (i.e., the ability to detect words and meaning from a human voice spoken in various languages), computer vision (i.e., image recognition), and robotic autonomy (i.e., ability for machines to automate decisions). Deep learning is a more complex subfield of machine learning AI that is comprised of neural networks, which are a specific category of algorithms that are loosely inspired by biological neurons in the human brain. The more layers or depth in the neural network, the more complex it is. Officials from the Office of the Under Secretary of Defense for Research and Engineering told us that DOD is just starting to explore deep learning neural networks, but does not currently have any in use. The private sector, in comparison, has begun to use deep learning. Apple's Siri and Amazon's Alexa virtual assistants are examples of deep learning AI that use complex neural networks to deliver results.

Academic and industry experts included in Stanford University's 2021 AI study panel agree that AI capabilities have not yet moved into the realm of contextual adaptation or general AI.[11] According to officials from the Army Research Laboratory, DOD's AI is not anywhere near being able to outthink a human, as current AI capabilities must be carefully programmed or trained for every situation they encounter. The Defense Advanced Research Projects Agency (DARPA) has ongoing work to develop machine common sense which, according to DARPA documentation, is the next step in moving toward general AI. Machine common sense aims to enable AI capabilities to understand new situations, monitor the reasonableness of their actions, communicate more effectively with people, and transfer learning to new domains.

Potential Uses and Ethical Principles for AI at DOD

DOD identified a variety of potential warfighting and non-warfighting uses for AI across the department. DOD's potential AI uses in warfighting operations include analyzing intelligence, surveillance, and reconnaissance sources; fusing data to provide a common operating picture on the battlefield; supporting semiautonomous and autonomous vehicles; and operating lethal autonomous weapon systems. Potential non-warfighting uses for AI (i.e., support and business operations) include resolving unmatched financial transactions, predicting maintenance needs, vetting security clearances, and analyzing warfighter health

---

[11]Michael L. Littman et al., "Gathering Strength, Gathering Storms: The One Hundred Year Study on Artificial Intelligence (AI100) 2021 Study Panel Report," Stanford University (September 2021), accessed September 16, 2021, http://ai100.stanford.edu/2021-report.

screenings. DOD's three categories of potential AI uses are shown in figure 2.

Figure 2: Uses of Artificial Intelligence (AI) at the Department of Defense
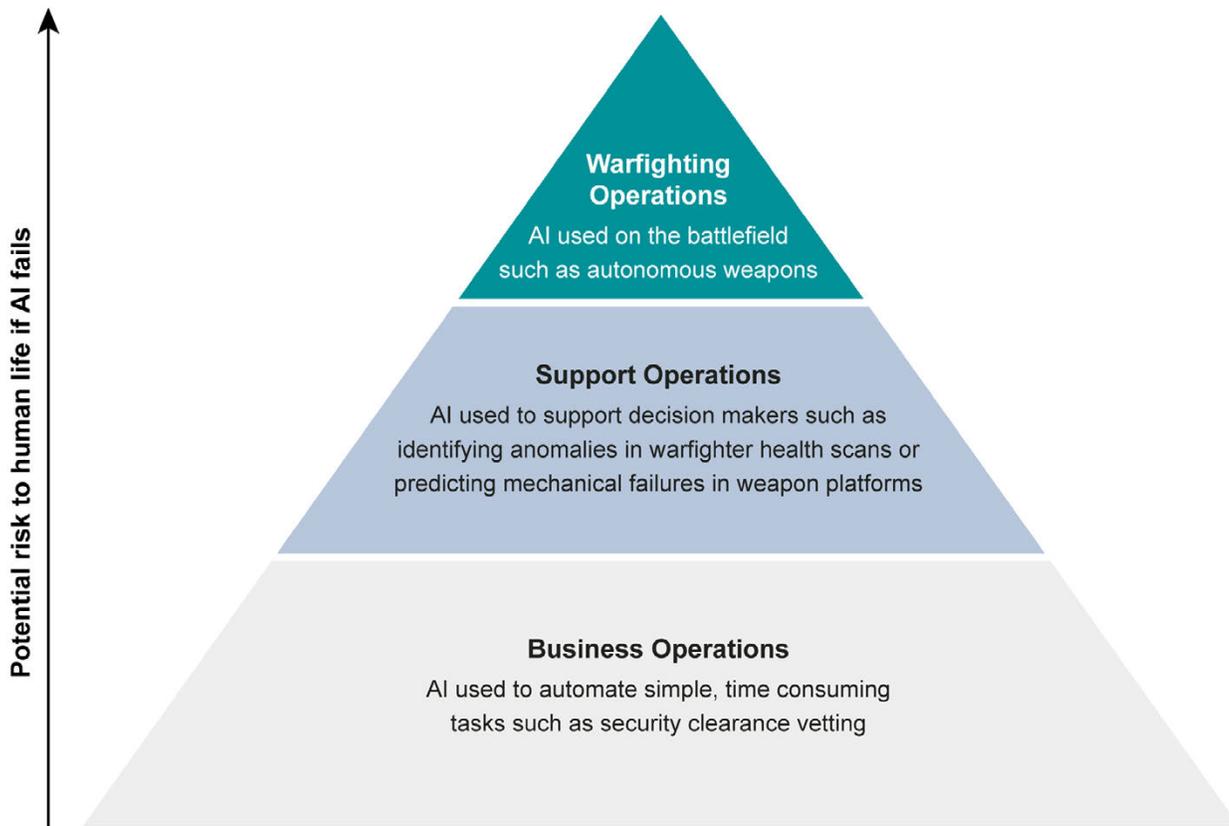
Source: GAO representation of Department of Defense information. | GAO-22-104765

According to a Joint AI Center official, AI for warfighting is an area where the department cannot rely on low performing AI because of its potential direct risk to human life. For example, if an AI model for precision strike fails during conflict, a civilian could be unintentionally targeted. To reduce the risk to human life, in February 2020, the department released its Ethical Principles for AI that will apply to both warfighting and non-

warfighting AI capabilities.[12] According to DOD, these principles affirm that the department is committed to AI that is:

- **Responsible.** Personnel will exercise appropriate levels of judgment and care, while remaining responsible for the development, deployment, and use of AI capabilities.

- **Equitable.** Personnel will take deliberate steps to help ensure AI capabilities do not unintentionally favor or harm a particular group.

- **Traceable.** Personnel will have an appropriate understanding of the AI capabilities' development processes and operational methods including transparent and auditable methodologies, data sources, and design procedures and documentation.

- **Reliable.** AI capabilities will have explicit, well-defined uses, and the safety, security, and effectiveness of such capabilities will be subject to testing and assurance within those defined uses across their entire life cycles.

- **Governable.** Personnel will design and engineer AI capabilities to fulfill their intended functions while possessing the ability to detect and avoid unintended consequences, and the ability to disengage or deactivate deployed systems that demonstrate unintended behavior.

In July 2020, the Joint AI Center completed a Responsible AI Champions Pilot to develop a better understanding of these principles, consider ways to operationalize them, and create a community of responsible AI ambassadors. According to Joint AI Center documentation, this pilot spanned over 10 weeks and leveraged the expertise of a multidisciplinary group of DOD officials involved with product design, test and evaluation, and acquisitions, among others. Moving forward, the Joint AI Center plans to build off the momentum of this pilot as it further builds out the

---

[12]DOD, *DOD Adopts Ethical Principles for Artificial Intelligence* (Feb. 24, 2020), available at https://www.defense.gov/News/Releases/Release/Article/2091996/dod-adopts-ethical-principles-for-artificial-intelligence/.

GAO-22-104765 Artificial Intelligence

curriculum and looks to scale the adoption of these AI ethical principles across the department.[13]

Entities Involved with AI for DOD

Numerous entities across private industry, academia, and DOD are involved in developing, transitioning, or deploying AI across the department. DOD's AI focal point is the Joint AI Center, which aims to provide a critical mass of expertise to help the department accelerate the delivery and adoption of AI. To do so, according to the center, it is providing that expertise by taking a holistic approach that includes:

- Establishing a common foundation that enables decentralized AI execution and experimentation;

- Evolving partnerships with industry, academia, international allies, and partners; and

- Cultivating a leading AI workforce.

The Joint AI Center's budget increased from $89 million in fiscal year 2019 to $278 million for fiscal year 2021. Initially, the Joint AI Center's activities focused on small AI technical demonstrations for the department, but the center announced a transition to more of a facilitator role in November 2020. While the Joint AI Center continues to build AI capabilities, following this transition, it also aims to assist DOD components in understanding the kinds of things they will need to do to support AI and providing technical support as needed. In our related report on AI published in February 2022, we report on DOD's efforts to collaborate on AI across the department.[14]

To develop AI capabilities for the department, DOD leverages private industry, academia, military service research labs, and DARPA. For example, in April 2020, the Army announced it awarded a contract to Carnegie Mellon University to pursue AI research and prototypes that will

---

[13]In addition to DOD's Ethical Principles for AI, the department issued a directive in November 2012 to govern autonomous and semi-autonomous weapon systems. This directive requires, among other things, that these weapon systems are designed to allow commanders and operators to exercise appropriate levels of human judgment over the use of force. Department of Defense Directive 3000.09, *Autonomy in Weapon Systems* (Nov. 2012), Incorporating Change 1 (May 2017). In April 2021, the Congressional Research Service reported that DOD does not have any lethal autonomous weapons in its inventory. Congressional Research Service, *International Discussions Concerning Lethal Autonomous Weapon Systems* (Apr. 19, 2021).

[14]GAO, *Artificial Intelligence: DOD Should Improve Strategies, Inventory Process, and Collaboration Guidance,* GAO-22-104516SU (Washington, D.C.: Feb. 16, 2022).

benefit the Army. Additionally, DARPA announced its Artificial Intelligence Exploration program in July 2018 to fund high-risk, high-payoff projects where researchers work to establish the feasibility of new AI concepts within 18 months of award. For example, in 2019 DARPA sought new approaches for autonomous teaming of various AI systems, like AI enabled drones or satellites, that can react to new or unexpected situations without access to centralized communication and human control. As of May 2021, DARPA officials told us that 26 projects have been awarded through this program.

Once developed, various DOD entities are involved in transitioning the AI capability to the end user, i.e., the military services, major commands, and other DOD components. For example, the Defense Innovation Unit helps the department adopt commercial technologies by reducing the time it takes to identify a problem, prototype a commercial solution, and implement it in the field. Additionally, military services such as the Army and Air Force established their own offices to facilitate the scaling of AI. The Army established its AI Integration Center—originally named the AI Task Force—in October 2018 to develop frameworks and methods to scale projects across the service, review policies that impede deployment of AI technologies to the Army, and establish an AI test bed, among other tasks. Additionally, the Air Force announced it signed a cooperative agreement with the Massachusetts Institute of Technology to conduct fundamental research that enables the rapid prototyping and scaling of AI to both the Air Force and Space Force. The relationship of the Joint AI Center to the various entities contributing to DOD's AI efforts is depicted in figure 3.

Figure 3: Entities Involved in Developing, Transitioning, and Using Artificial Intelligence (AI) for DOD, as of November 2021

Joint AI Center

Facilitates DOD's development and adoption of AI capabilities by providing resources and expertise to AI developers, users, and entities that assist in transitioning capabilities from research into operations.

AI developers

Military service labs:
- Naval Research Laboratory
- Air Force Research Laboratory
- Marine Corps Warfighting Laboratory
- Army Research Laboratory

Defense Advanced Research Projects Agency

Academia (such as Carnegie Mellon University, Massachusetts Institute of Technology, and Johns Hopkins University, among others)

Private industry (i.e. artificial intelligence innovators, service providers, etc.)

AI capability transition

Some of the DOD entities involved include:
- Defense Innovation Unit
- Army AI Integration Center
- Air Force AI Accelerator

AI users
- U.S. Marines
- U.S. Navy
- DOD Major Commands
- U.S. Army
- U.S. Air Force
- DOD Joint Program Offices

Source: GAO depiction of Department of Defense (DOD) information. I GAO-22-104765

Note: Information in this figure is accurate as of November 2021. In December 2021, the Deputy Secretary of Defense issued a memorandum announcing the establishment of the Office of the Chief Digital and AI Officer that will serve as successor to the Joint AI Center in 2022.

AI Is Different from Traditional Software and Capabilities for Warfighting Are in Development

Developing and using AI differs from traditional software, in that AI requires vast amounts of data for development and continuous monitoring once deployed. There is also a need for speedy development for AI to keep pace with changing technology. We found that the majority of AI activities supporting DOD's warfighting mission are in research and development, with a focus on developing autonomy for uncrewed systems, recognizing targets, and providing recommendations to commanders on the battlefield. Additionally, various DOD officials told us that, as with other emerging technologies, AI capabilities are subject to difficulties in transitioning from research and development to the end user.

GAO-22-104765 Artificial Intelligence

Developing and Using AI Is Different than Traditional Software

AI is a software-centric capability, but DOD recognizes that developing and using it is different from traditional software in key ways.[15] At its core, software is programmed to perform tasks based on static instructions written as code, producing the same result based on the instructions coded into the system. For example, software designed to identify late invoices by changing the color of the text on screen to red if the due date is after today's date will always change the text color to red so long as the date of the invoice is after the designated date. AI, by contrast, is software that is programmed with general parameters for performing the applicable task (referred to as algorithms) so that the AI model can learn to perform the task and improve over time. The AI model learning process is achieved by providing the AI algorithm with large data sets that identify the desired outcome, with the AI developer validating that the model is producing the desired results. For example, training an AI model to recognize a submarine from a video feed requires a large data set of images of various types of submarines that are identified as submarines.[16] During training, the system will be exposed to images of submarines and staff involved in the training will validate when the AI model identifies a submarine correctly and when it does not. Erroneous training outcomes are evaluated by staff to determine whether the AI algorithm should be modified or if more training is required.

According to the Defense Innovation Board's 2019 Software Acquisition and Practices Study, training a computer to learn is inherently different from programming instructions.[17] These differences include the importance of data, computing power needs, and methods to identify errors. See table 1 for a more detailed discussion on the differences between traditional software and AI.

---

[15]For the purposes of this report, we use the term traditional software to mean computer programs, procedures, rules, and possibly documentation and data pertaining to the operation of the computer system. Some examples of traditional software include enterprise systems, business systems, and combat systems.

[16]There are ways to train an AI model that do not require data that has the desired result already identified (i.e., labeled). This includes a method known as unsupervised learning, where unlabeled data are fed to an AI capability to identify clusters and associations without the need for human intervention. Another method is reinforcement learning, where the AI model collects and explores data on its own in order to receive human-defined rewards for correct actions. According to DOD documentation, these types of learning can be less predictable and more challenging.

[17]Defense Innovation Board, *Software Is Never Done: Refactoring the Acquisition Code for Competitive Advantage* (May 3, 2019).

GAO-22-104765 Artificial Intelligence

Table 1: Key Differences between Traditional Software and Artificial Intelligence

	Traditional software	Artificial Intelligence (AI)
Developing	• Static code is the primary input for traditional software. These are instructions written by a human for machines to execute. These instructions may include algorithms—a series of specific instructions to execute under specific conditions. • Data are not necessary to develop traditional software. • Running traditional software requires variable computing power. For example, a simple software program like a word processor does not require significant computing power, but a command-and-control tool that has performance requirements to render high-quality images and simulate real-world conditions would.	• AI-specific algorithms are one of the two primary inputs for an AI model. Available AI algorithms can produce a range of desired outcomes such as classifying images or providing a prediction based on patterns in data. Rather than codify explicit instructions for a computer to process, an AI developer creates or selects one or more AI algorithms that outline how the computer will learn on its own from the data. • Data are the other primary input for an AI model. Large sets of data are required to train the AI model. For example, to train an AI model to recognize a target such as a car, an AI developer must feed it thousands of car images so that the AI model can learn all the aspects of what a car looks like to identify one on its own. • Significant computing power is necessary to train and store data for an AI model. Specifically, an AI model typically requires access to high performance computing or cloud computing to conduct complex calculations that are otherwise difficult to do on a single computer. Cloud computing harnesses the ability to temporarily access servers to support such calculations while also moving and storing vast amounts of data necessary to train an AI model quickly.[a]
Using	• Traditional software is deterministic in that it behaves the same way every time until updated by a human. This means that the same inputs would result in the same outputs regardless of how many times the software program is run. For example, software that generates a location status report from a personnel database will generate the same report each time unless a human enters new data into the database or changes the code that creates the report. • Traditional software is periodically tested for errors—known as debugging—once deployed. The debugging process detects anomalies or unintended inputs in the code that may cause software to crash or not run as intended. The deterministic nature of traditional software makes it easier to identify and correct the source of errors.	• AI is dynamic and adaptive, not always deterministic. In some cases, AI algorithms have an element of randomness such that the same or very similar inputs may generate different outputs. AI algorithms produce probabilistic outcomes, meaning the output of the AI model is a prediction, not a certainty. Moreover, retraining or adapting an AI model based on new data leads to different outputs for given inputs—the desired result of such adaptation being improved performance. • Entities—such as the AI developer and end users—should continuously monitor the AI model to ensure it is working as intended after it is deployed. If the data used to train an AI model do not match the data being used to run the model, it can result in model drift and performance degradation. For example, if a predictive maintenance model is trained on data from vehicles in desert conditions and then applied to vehicles in non-desert conditions, model performance may suffer. The non-deterministic nature of AI can make it difficult to identify a point of failure or pinpoint why an AI model produced the wrong result.

Source: GAO representation of Department of Defense and other subject matter expert provided information. I GAO-22-104765

Note: AI refers to machine learning capabilities—the predominant type of AI being developed today.

[a]As defined by the National Institute of Standards and Technology, cloud computing is a means for enabling on-demand access to shared pools of configurable computing resources.

The AI development process should leverage Agile software development principles to keep pace with rapidly changing technologies. Agile software development focuses on iterative product development that encourages collaboration across an organization and allows requirements to evolve as a program progresses.[18] GAO's recently published AI accountability framework echoes this idea, noting that the four phases of the AI life cycle—design, development, deployment, and continuous monitoring—are often iterative and not necessarily sequential.[19] An official from the Office of the Under Secretary of Defense for Research and Engineering told us that in her experience developing AI capabilities, it is important to have the end user representatives working with AI developers. This official added that user representatives can provide sample data as available and feedback throughout the development process. Additionally, we reported in 2019 that user involvement is critical to the success of any software development effort by decreasing the risk of delivering systems that do not meet user needs.[20] Officials from the Air Force U-2 federal laboratory highlighted this point, telling us that involving stakeholders, particularly end users (e.g., the warfighter), during AI development helps to understand their needs and expedites delivery of the model into operation.[21]

While various DOD officials generally agree that AI development should leverage Agile principles, there are likely to be aspects of Agile that may

---

[18]GAO, *Science & Tech Spotlight: Agile Software Development*, GAO-20-713SP (Washington, D.C.: Sept. 29, 2020).

[19]GAO, *Artificial Intelligence: An Accountability Framework for Federal Agencies and Other Entities*, GAO-21-519SP (Washington, D.C.: June 30, 2021). For more on this framework, see appendix III.

[20]GAO, *DOD Space Acquisitions: Including Users Early and Often in Software Development Could Benefit Programs*, GAO-19-136 (Washington, D.C.: Mar. 18, 2019). In November 2020, GAO reported similar concerns related to stakeholder—such as the end user—engagement in the Defense Intelligence Agency's development of its new military intelligence system. GAO, *Defense Intelligence: Comprehensive Plan Needed to Improve Stakeholder Engagement in the Development of New Military Intelligence System*, GAO-21-57 (Washington, D.C.: Nov. 19, 2020).

[21]According to Air Force officials, the U-2 federal laboratory was established to fast field advanced technologies to the warfighter and does so by embedding developers with the warfighters who operate the weapon systems. In 2019, the U-2 federal laboratory requested that the National Institute of Standards and Technology establish an accreditation program for these type of labs, which was approved in July 2021. In November 2021, an Air Force official stated that the lab has since undergone a formal name change to "ACC Federal Laboratory_Beale."

look different for AI. For instance, fielding the most basic version of a program that adds value quickly is a key tenet of Agile for traditional software development. Depending on the complexity of the planned AI capability, developing a simpler AI capability that starts adding value quicker can be useful while developers continue to add functionality over time. However, since AI typically requires vast amounts of data to learn a function and can behave differently or unexpectedly once deployed, officials from the Joint AI Center and DOD's Defense Digital Service cautioned against rushing to a proof of concept or minimum viable product for AI capabilities that support warfighting operations.

According to an Army publication, to be minimally viable, the AI model should be performing its designated function (solving the problem it was designed to) at the time it is deployed at an established performance metric and exceed what a human is capable of (i.e. demonstrating a positive effect on the operating environment).[22] For example, according to this publication, an AI model intended to translate intelligence-related documents and data would be deemed minimally viable if the trained model accurately translated the text at the agreed upon threshold and outperformed the human analysts typically responsible for the work. Given the need for data, training, and testing for operational usefulness, it may take longer to get to a minimally viable product for AI than for traditional software Agile projects. Once deployed, the AI capability should continue to get better at its task as it is exposed to more data, but needs to be continuously monitored to ensure the new data are not negatively affecting performance.

The Majority of AI Capabilities in DOD's Initial AI Inventory Are in Research and Development

DOD is actively pursuing AI capabilities, but according to officials from the Joint AI Center and military services, the majority of its advanced AI capabilities for warfighting are still in development as DOD grapples with their differences from traditional software. However, a few AI capabilities for warfighting are in operation, including those produced by Project Maven, which was launched by the Deputy Secretary of Defense in April 2017.[23] In 2019, the RAND Corporation reported that scaling DOD's operational AI (such as that for warfighting) is not likely to occur in the

---

[22]Courtney Crosby, *Operationalizing Artificial Intelligence for Algorithmic Warfare*, Military Review (July-August 2020).

[23]Project Maven—also known as the Algorithmic Warfare Cross-Functional Team—develops computer vision machine learning AI capabilities that analyze massive amounts of full-motion video collected by intelligence, surveillance, and reconnaissance assets to identify objects of interest.

near term.[24] Also in this report, RAND found that DOD had not established metrics to assess progress toward its AI goals. In our related AI report, we recommended, and DOD concurred, that the department establish guidance that defines outcomes and includes AI key performance indicators.[25]

DOD has begun efforts to formally track its AI activities in response to congressional concern over the department's visibility into the activities, and published its initial AI inventory in April 2021.[26] In this initial inventory, DOD identified 685 AI projects—including, but not limited to those supporting its warfighting mission—by analyzing its unclassified research and development and procurement budget documents, but it was unable to provide the estimated funding associated with these projects.[27] According to a DOD official leading the effort, it is difficult to identify the discrete funding for AI efforts because AI is typically a piece of a program, rather than a program itself, and DOD's budget documents do not separate AI project funding from the rest of the program. DOD's initial inventory has other limitations—such as omitting classified AI activities and those funded through operations and maintenance—that DOD officials said they are planning to address in future iterations of the inventory. In our related AI report, we recommended that the department develop a high-level plan that captures all requirements, activities, and milestones supporting the AI inventory.[28] Of the AI projects included in the

---

[24]RAND Corporation, *The Department of Defense Posture for Artificial Intelligence: Assessment and Recommendations* (Santa Monica, CA: 2019).

[25]GAO-22-104516SU.

[26]166 Cong. Rec. H8251 (Dec. 21, 2020). This Joint Explanatory Statement that accompanies the Consolidated Appropriations Act, 2021, Pub. L. No. 116-260 (2020) supports DOD's AI activities that are intended to improve the affordability and effectiveness of military operations. However, the statement reflects a concern about a lack of coordination among the myriad of AI programs within the department and the military services. Therefore, Congress directed the Director of the Joint AI Center to provide the congressional defense committees, not later than 120 days after the enactment of the act, an inventory of all AI activities, to include each program's appropriation, project, and line number; the current and future years' defense program funding; the identification of academic or industry mission partners, if applicable; and any planned transition partners

[27]Joint AI Center officials told us that they expect more AI projects to be identified when they are able to analyze classified as well as operations and maintenance budget documentation, but could not estimate how many. For additional information on this initial inventory, see appendix II.

[28]GAO-22-104516SU.

initial inventory, 88 percent were identified from research and development budget documentation, which aligns with various DOD officials' assertion that most of the department's AI capabilities, especially those supporting the warfighting mission, are still in development.[29]

DOD's identified AI projects include efforts specifically aligned to individual programs and others that are potentially broadly applicable. We found that 17 of DOD's 88 reported major weapon systems in fiscal year 2021—such as the Joint Light Tactical Vehicle and MQ-9 Unmanned Aerial Vehicle—had associated AI projects clearly identified in DOD's inventory.[30] However, most of the 685 identified projects are not yet aligned to specific systems but have potentially broad applicability to multiple systems. DOD is currently pursuing AI capabilities for warfighting that largely focus on (1) recognizing targets through intelligence and surveillance analysis, (2) providing recommendations to operators on the battlefield (such as where to move troops or which weapon is best positioned to respond to a threat), and (3) increasing the autonomy of uncrewed systems. Several potential applications of AI—including those listed above—can be trained for integration into multiple weapon platforms, such as aerial drones that are used across all the military services. See figure 4 for a depiction of how these types of AI capabilities can work.

[29]DOD's appropriations are divided into multiple categories and accounts, including Operations and Maintenance; Military Personnel; Research, Development, Test, and Evaluation; Procurement; and Military Construction, among others. The Research, Development, Test, and Evaluation account (which we refer to as research and development) generally funds the scientific research and military development of new technologies. The procurement account generally funds the procurement, manufacturing, and modification of DOD weapon systems and associated parts. The Operations and Maintenance account generally funds current operations, equipment maintenance, and civilian salaries.

[30]Major weapon systems are weapon systems that are acquired pursuant to a major defense acquisition program. Defense Federal Acquisition Regulation Supplement § 234.7001. Major defense acquisition programs are acquisition programs that are designated by the Secretary of Defense, as delegated, or are estimated to require an eventual total expenditure for research, development, test, and evaluation, including all planned increments, of more than $525 million in fiscal year 2020 constant dollars or, for procurement, of more than $3.065 billion in fiscal year 2020 constant dollars. 10 U.S.C. § 2430; Department of Defense Instruction 5000.85, *Major Capability Acquisition* (Aug. 6, 2020), incorporating Change 1 (Nov. 4, 2021).

GAO-22-104765 Artificial Intelligence

Figure 4: Notional Example of Artificial Intelligence on the Battlefield
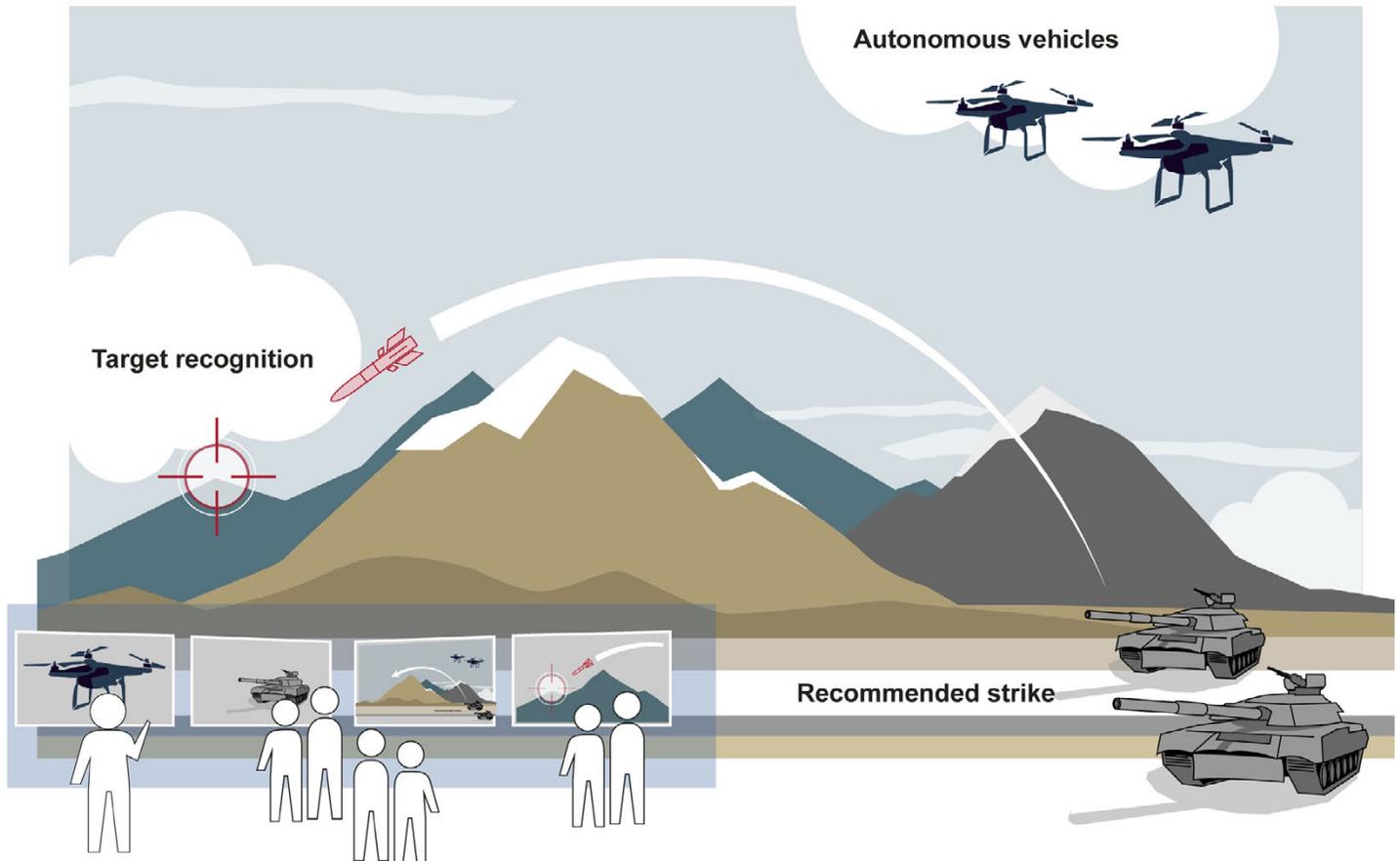
Source: GAO depiction of Department of Defense (DOD) provided information. | GAO-22-104765

Various DOD entities are developing AI capabilities in these areas. Examples include:

- **Joint AI Center.** The Joint AI Center is working with the military services on a targeting AI capability known as project "Smart Sensor." Smart Sensor is a video processing AI prototype that rides on uncrewed aerial vehicles and is trained to identify threats and immediately transmit the video of those threats back to analysts for real-time monitoring and evaluation. According to a Joint AI Center official, the center is developing this AI capability to be able to operate on other airborne platforms and committed roughly $50 million of its appropriations toward this capability in fiscal year 2021.

GAO-22-104765 Artificial Intelligence

- **Air Force and Space Force.** In December 2020, the Air Force demonstrated an AI capability—known as Artuu—that was able to pilot the U-2 platform assuming responsibility for finding enemy launchers during a simulated mission strike.[31] According to Air Force officials, the U-2 federal laboratory developed Artuu in 35 days and, while Artuu was successful in its initial demonstration, it would require significantly more training to be operational in a real-world environment. The Air Force is also developing the Air Force Advanced Battle Management System, which is designed to provide a real-time operational picture of threats across all domains in part by leveraging AI.[32] According to Space Force officials, the Space Force is working with the Air Force to develop an AI domain awareness capability specific to space.

- **Army.** The Army is pursuing a target recognition AI capability—known as Scarlet Dragon—that uses data from Project Maven to support airborne combat operations. According to an Army official working on the capability, development of this AI capability is primarily funded through Project Maven and is being used in live fire drills every 90 days.[33] The latest major demonstration, held in October 2021, used the AI capability across various Army, Air Force, and Navy weapon platforms to identify and eliminate targets. The Army is also developing a similar AI capability, known as Prometheus, to sense and identify targets using space-based capabilities (i.e., satellite imagery).

- **Navy and Marine Corps.** The Navy is developing a decision support AI capability—known as the Undersea Warfare Decision Support System—that will be able to provide a common operating picture across Navy platforms to help operators plan and execute Navy undersea warfare missions. According to a Navy official, the Navy is also investing in AI capabilities to improve targeting for the Naval

---

[31]The U-2 is a single-jet engine, high-altitude reconnaissance aircraft developed by Lockheed Martin for the U.S. Air Force. The aircraft first flew in 1955 and is capable of gathering surveillance and signals intelligence data in real-time and can be deployed anywhere in the world.

[32]GAO previously reported on the Air Force's efforts to develop its Advanced Battle Management System. GAO, *Defense Acquisitions: Action Is Needed to Provide Clarity and Mitigate Risks of the Air Force's Planned Advanced Battle Management System*, GAO-20-389 (Washington, D.C.: Apr. 16, 2020).

[33]According to an Army official working on Scarlet Dragon, roughly 70 percent of this capability is funded by Project Maven and the remaining 30 percent is funded through the Army's Tactical Radio Application Extension program. The Fiscal Year 2021 National Defense Authorization Act provided $250 million for Project Maven.

GAO-22-104765 Artificial Intelligence

Strike Missile and the Mark 48 torpedo. According to Marine Corps officials, the service is working to incorporate algorithms developed as part of Project Maven into their capabilities and to modernize legacy weapon systems, such as integrating AI target sensors onto uncrewed aerial vehicles.

DOD Recognizes It Faces Difficulties Transitioning AI Capabilities

As DOD's AI capabilities mature, officials from the military labs told us that the department is likely to face difficulties with transitioning these capabilities to the end user that are similar to those experienced with other emerging technologies. DOD has long recognized these difficulties, particularly in moving technologies from research and development to further maturity and production within the acquisition community for eventual delivery to the warfighter. See figure 5 for a depiction of DOD's technology development process and the potential gap in transitioning— often referred to as the valley of death.

Figure 5: Representation of DOD Technology Development Process and Potential Transition Challenge
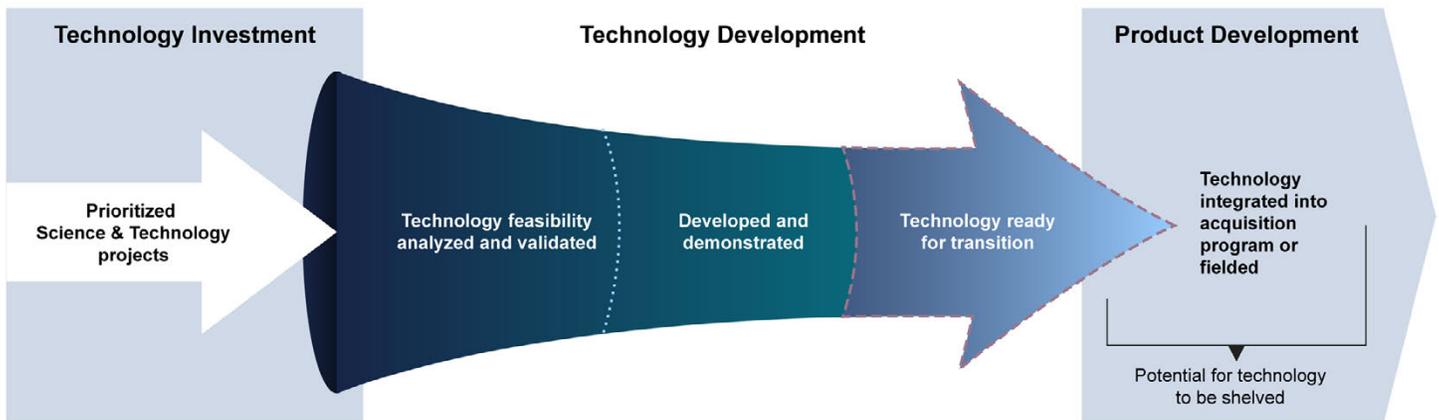
Source: GAO analysis of Department of Defense (DOD) information. | GAO-22-104765

Our prior reports note that this gap exists because the acquisition community often requires a higher level of technology maturity than the science and technology community is willing to fund and develop.[34] As a result, DOD components may shelve technologies that are developed and demonstrated within the research labs.

---

[34]See, for example, GAO, *Defense Advanced Research Projects Agency: Key Factors Drive Transition of Technologies, but Better Training and Data Dissemination Can Increase Success*, GAO-16-5 (Washington, D.C.: Nov. 18, 2015).

According to officials from DARPA and the military service labs, they are considering and starting to employ various strategies to facilitate the transition of AI capabilities to program offices across the military services. These strategies tend to center around engaging the end user or program office as early as possible. Specifically, officials from DARPA told us they are trying to identify and meet with end users prior to starting work on a project to discuss the needs or objectives that program offices are trying to accomplish with AI to ensure the eventual capability will be usable. Additionally, an official from the Dahlgren Division of the Naval Surface Warfare Center told us that before starting work on an AI capability, center officials assess the feasibility of transition to an acquisition program, including the data and computing power available, with the intention of making the process easier.

AI Presents Challenges That DOD Is Beginning to Address

AI exacerbates existing challenges—such as cybersecurity and workforce issues—that DOD faces in developing, acquiring, and scaling software products. Outside of these existing challenges, AI also introduces its own challenges related to data, integration, and trustworthiness, for example. DOD is fielding several initiatives to support its pursuit of AI and address these challenges. These include efforts by the Joint AI Center, as well as by the military services. Most of these initiatives are expected to be rolled out across the department by 2023.

AI Exacerbates DOD's Existing Software Development Challenges and Introduces New Ones

DOD has historically faced challenges in developing, procuring, and deploying software-centric capabilities. In 2019, the Defense Innovation Board's Software Acquisition and Practices study highlighted DOD's lack of fast, Agile acquisition processes that incorporate cybersecurity, cross-service digital infrastructure, and pathways for cultivating digital talent as impediments to modern software.[35] Our prior work identified similar challenges. Specifically, in July 2017, we reported that the complexity and length of DOD's acquisition process was an impediment for industry to working with DOD on software development, among other things.[36] Additionally, our October 2018 report on weapon systems cybersecurity found that DOD faces mounting challenges in protecting its weapon systems from cyber threats due, in part, to their increasing dependence

---

[35]Defense Innovation Board, *Software Is Never Done: Refactoring the Acquisition Code for Competitive Advantage* (May 2019).

[36]GAO, *Military Acquisitions: DOD Is Taking Steps to Address Challenges Faced by Certain Companies*, GAO-17-644 (Washington, D.C.: July 20, 2017).

GAO-22-104765 Artificial Intelligence

on software.[37] In our 2021 Weapon Systems Annual Assessment, we found that major defense acquisition programs and programs using the middle tier acquisition pathway reported challenges related to their software development workforce. For example, over half of these programs reported staffing challenges, including hiring contractor and government staff in time to perform planned work and identifying contractor and government staff with expertise in software development.[38]

DOD has taken some steps to address these challenges for traditional software acquisitions. For example, it established the Software Acquisition Pathway in October 2020 to enable the timely acquisition of software capabilities.[39] DOD also established software factories, which act as in-house innovation hubs within some of the military services that build, test, and deliver warfighting software applications. Yet, the National Security Commission on Artificial Intelligence and various DOD entities maintain that the department continues to face these same software-related challenges as they pursue increasingly complex AI capabilities.[40] Through our analysis of interview responses from various DOD entities—such as the military services and Joint AI Center, among others—as well as external subject matter experts and private industry officials, we identified the following AI-related challenges:

---

[37]GAO, *Weapon Systems Cybersecurity: DOD Just Beginning to Grapple with Scale of Vulnerabilities*, GAO-19-128 (Washington, D.C.: Oct. 9, 2018).

[38]GAO, *Weapon Systems Annual Assessment: Updated Program Oversight Approach Needed*, GAO-21-222 (Washington, D.C.: June 8, 2021).

[39]Department of Defense Instruction 5000.87, *Operation of the Software Acquisition Pathway* (Oct. 2, 2020). This pathway is designed for software-intensive systems to facilitate the rapid and iterative delivery of software capability to the user. It integrates modern software development practices with a focus on active user engagement. According to officials from the Office of the Under Secretary of Defense for Acquisition and Sustainment, 21 programs were using the software acquisition pathway as of July 2021.

[40]The National Security Commission on Artificial Intelligence made recommendations for DOD to confront these challenges. For example, the commission found that DOD's acquisition policies are inadequate to prepare it for the future of AI, and recommended that the department expand the use of specialized acquisition pathways and contracting approaches for AI. In addition, it encouraged DOD to invest in its digital infrastructure, to include shared cloud computing access that is accessible across DOD to internal AI developers and industry partners. It also encouraged DOD to create digital career pathways in the services and to bolster cybersecurity support for partner research institutions.

- **Linear and time-consuming acquisition processes.** DOD's traditional acquisition processes were designed for hardware-intensive systems and are typically linear and time-consuming. The National Security Commission on Artificial Intelligence reported and industry group officials told us that this process is not well-suited to AI. Specifically, they cited the long amount of time it typically takes DOD to acquire a new capability as incompatible with the fast speed at which AI technology is developed. DOD's acquisition processes, according to industry officials, can also be a potential barrier to entry for small and nontraditional companies who may be less familiar with DOD-specific requirements than large companies.[41]

- **Lack of cross-service digital infrastructure to support AI.** DOD does not have the necessary digital infrastructure in place to develop and scale AI across the department, a point echoed by the National Security Commission on Artificial Intelligence and various DOD officials. According to Joint AI Center documentation, AI development requires tools, technologies, and computing infrastructure and lack of access can be a deterrent to adopting or experimenting with AI capabilities. Additionally, officials across DOD headquarters entities and the military services told us that the department's current approach to AI development—in which each component works in isolation—is a hindrance as it inhibits the sharing of data and development tools.

- **Digital talent deficit.** DOD's workforce often lacks the skills necessary to successfully develop, buy, or use AI capabilities. The National Security Commission on Artificial Intelligence designated the digital talent deficit as the government's greatest inhibitor to buying, building, and fielding AI-enabled technologies. DOD's 2020 AI Education Strategy also states that AI talent is in short supply and highlights that commercial companies have a distinct advantage over DOD when it comes to recruiting and retaining top AI talent.[42]

- **AI vulnerability to traditional and new cyberattacks.** DOD's AI capabilities are vulnerable to traditional and new forms of cyberattack. The National Security Commission on Artificial Intelligence found that commercial firms and researchers have documented attacks that involve evasion, data poisoning, model replication, and exploiting traditional software flaws to deceive, manipulate, compromise, and render AI systems ineffective. An official from DOD's Defense Digital

---

[41]For additional information on nontraditional companies and DOD, see GAO-17-644.

[42]Department of Defense, *2020 DOD Artificial Intelligence Education Strategy* (Arlington, VA: 2020).

Service told us that the data needed to train and run an AI capability introduce a new means for cyberattack through data poisoning. In the case of imagery data, this means that an adversary could alter a portion of an image in the training data to trick the model into misidentifying the image, which may still look the same to the human eye. Cybersecurity of emerging technologies including AI and cyber workforce management challenges are two of the 10 critical actions we have recommended that agencies need to address to improve the nation's cybersecurity.[43]

In addition to those above, we identified other challenges DOD faces unique to AI related to the availability of usable data, integrating AI into existing weapon systems, and increasing trust in the technology.

- **Availability of usable data to develop and train AI.** Data are critical to the development and operation of AI-enabled capabilities, but according to various DOD officials, the department often lacks data that are usable for AI. High performing AI typically requires accurately labeled historical data to train the system.[44] Labeled data refers to raw data (images, text files, videos, etc.) that have been tagged with one or more identifiers to provide context so that the AI algorithm can learn from it. For example, an intelligence, surveillance, and reconnaissance AI capability trained using labeled data to identify tanks would require pictures of various tanks tagged as such. Figure 6 provides another example of such labeling.

---

[43]GAO, *High-Risk Series: Federal Government Needs to Urgently Pursue Critical Actions to Address Major Cybersecurity Challenges*, GAO-21-288 (Washington, D.C.: Mar. 24, 2021).

[44]There are ways to train an AI model that do not require labeled data. This includes a method known as unsupervised learning where unlabeled data are fed to an AI capability to identify clusters and associations without the need for human intervention or reinforcement learning where the AI model collects and explores data on its own to receive human defined rewards for correct actions. According to DOD documentation, these types of learning can be less predictable and more challenging.

GAO-22-104765 Artificial Intelligence

Figure 6: Example of Labeled Imagery Data
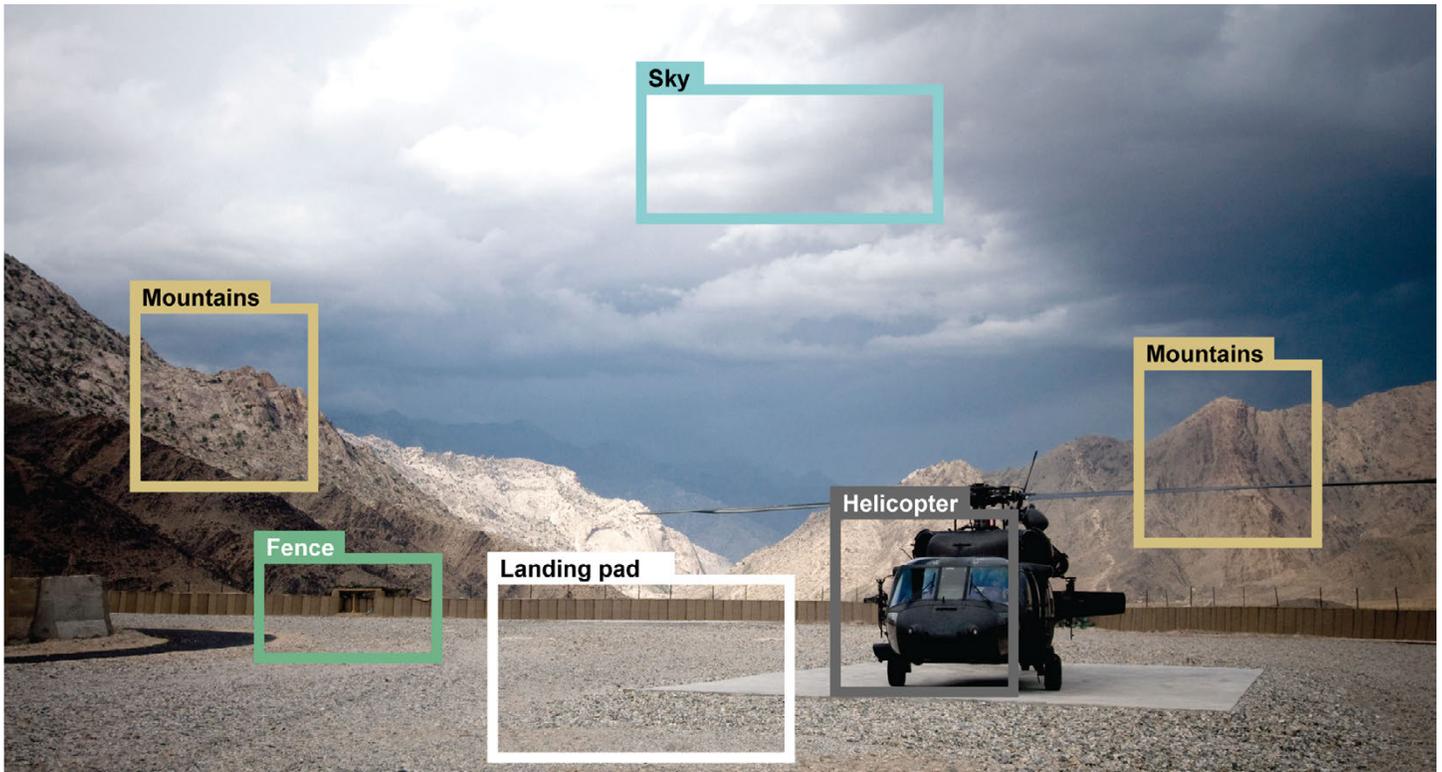
Source: U.S. Army/Staff Sergeant S. Morse/Defense Visual Information Distribution Service. | GAO-22-104765

The 2020 DOD Data Strategy states that data sets for AI training and algorithmic models will increasingly become DOD's most valuable digital assets.[45] However, making sure the right data are available and in a usable format poses unique challenges for DOD. Officials from the Office of the Under Secretary of Defense for Research and Engineering noted that, even when a DOD component or weapon program office says that data are available, they are typically unusable for AI. Specifically, they told us that these entities often have raw, unlabeled data that cannot be used to develop or train an AI model. DARPA officials told us that that labeling all previously gathered data would be too challenging and DOD should focus on incentivizing programs to collect and store data in a standardized format usable by AI systems. In addition, according to an official from the Office of the Under Secretary of Defense for Acquisition and

---

[45]Department of Defense, *DOD Data Strategy* (September 2020).

GAO-22-104765 Artificial Intelligence

Sustainment, a weapon platform should be built with a variety of sensors to collect data that can be used to train and support future AI capabilities.

- **Integration of AI into existing weapon platforms.** Integration of AI into weapon platforms involves synchronizing development and testing of capabilities with the training of operators who will use them in the field. Industry officials told us that one of the benefits and challenges of successfully implementing AI capabilities is extending them out to military operations in the field. DOD documentation acknowledges that future confrontations with adversaries are likely to be fought in contested areas with denied access to digital infrastructure like the cloud.[46] As such, it is important for AI capabilities embedded in weapon platforms to be able to function without this type of access. Integrating AI capabilities into a weapon platform requires computing equipment, which can take up space and add weight, as well as electrical power that may not be available in some existing weapon platforms. For example, officials from the Air Force's U-2 laboratory told us that when they wanted to install an AI capability onto the U-2 aircraft, the program office initially told them that there was no available processing space. To verify, these officials physically inspected the aircraft and found one obsolete computing area they could use. This is not unique to the U-2 aircraft; officials from each of the military services told us that understanding and creating the capacity needed aboard existing weapon platforms to integrate an AI capability will be difficult.

- **Hesitancy to trust AI capabilities.** According to the National Security Commission on AI, AI systems challenge DOD's existing evaluation strategies and ethical standards for capabilities, which can result in hesitancy to use them. Figure 7 provides a notional example of AI model complexity and the questions a user may need to be able to answer to trust the AI's decision or recommendation.

---

[46]As defined by the National Institute of Standards and Technology, cloud computing is a means for enabling on-demand access to shared pools of configurable computing resources (e.g., networks, servers, storage applications, and services) that can be rapidly provisioned and released.

Figure 7: Notional Example of Artificial Intelligence Model Complexity
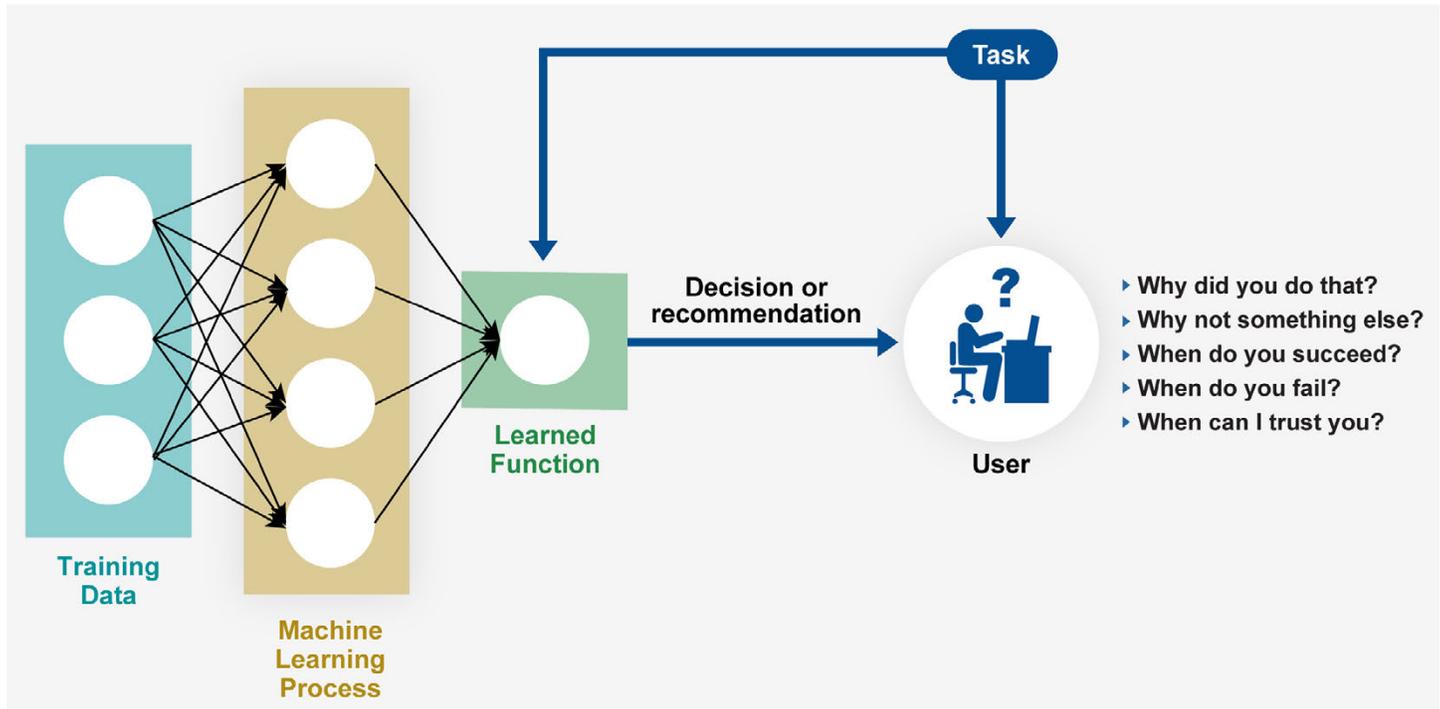
Source: GAO representation of Defense Advanced Research Projects Agency and other information. | GAO-22-104765

Trustworthy AI refers to AI capabilities that exhibit characteristics such as resilience, security, and privacy so that people can adopt them without fear.[47] To achieve this, an AI capability must be traceable, meaning that it is developed and deployed such that relevant personnel possess an appropriate understanding of the technology, development processes, and operational methods applicable to AI capabilities in line with the Joint AI Center's Ethical Principles for AI. According to DOD officials we spoke with, the more advanced the AI capability, the harder it is to understand and explain why the model is producing a certain output. For example, according to officials from the Office of the Director, Operational Test & Evaluation, AI has internal complexities and external vulnerabilities that pose obstacles to testers in understanding the decisions being made by the system. This inability to explain in simple terms what the AI capability does, how it does it, and how to use the output can lead to distrust and hesitation on the part of the end user, according to DOD documentation.

---

[47]National Institute of Standards and Technology, *Draft- Taxonomy of AI Risk* (October 2021).

GAO-22-104765 Artificial Intelligence

Further, Navy officials told us that if operators or decision makers are still required to perform all functions of an AI capability manually or they do not build trust with the capability, then the benefit of AI is limited.

Ethical concerns and notions of trustworthiness are important for DOD as it seeks to work with private companies to develop capabilities. For example, RAND reported that in 2018, Google employees protested the company's involvement with Project Maven for ethical reasons, resulting in Google pulling out of the project completely.[48] DOD found other vendors to continue the work. Since that time, Google has been working with DOD on some AI projects such as AI capabilities benefiting natural disaster first responders and increasing the speed and accuracy of service members' cancer diagnoses.

DOD Initiated Efforts to Support Its Pursuit of AI and Address Challenges

In part to address the challenges discussed above, the Joint AI Center initiated a number of recent efforts, investing approximately $610 million from fiscal years 2019 to 2021.[49] We found that for each challenge identified above, DOD has at least one ongoing effort to help address it and, in some cases, specific efforts are aimed at addressing multiple challenges, as shown in figure 8. DOD either fielded these efforts in calendar year 2021 or expects them to be available within the next 2 years. For this reason, it is too soon to assess their effectiveness, and we plan to continue monitoring the department's efforts in mitigating these challenges.

---

[48]RAND Corporation, *Military Applications of Artificial Intelligence* (Santa Monica, CA: 2020).

[49]Total investment for fiscal year 2021 includes efforts related to the Joint AI Center's mission initiatives, testing and evaluation, strategy and policy, infrastructure and platform, and program support.

Figure 8: Selected Joint AI Center Initiatives and the Challenges They May Address
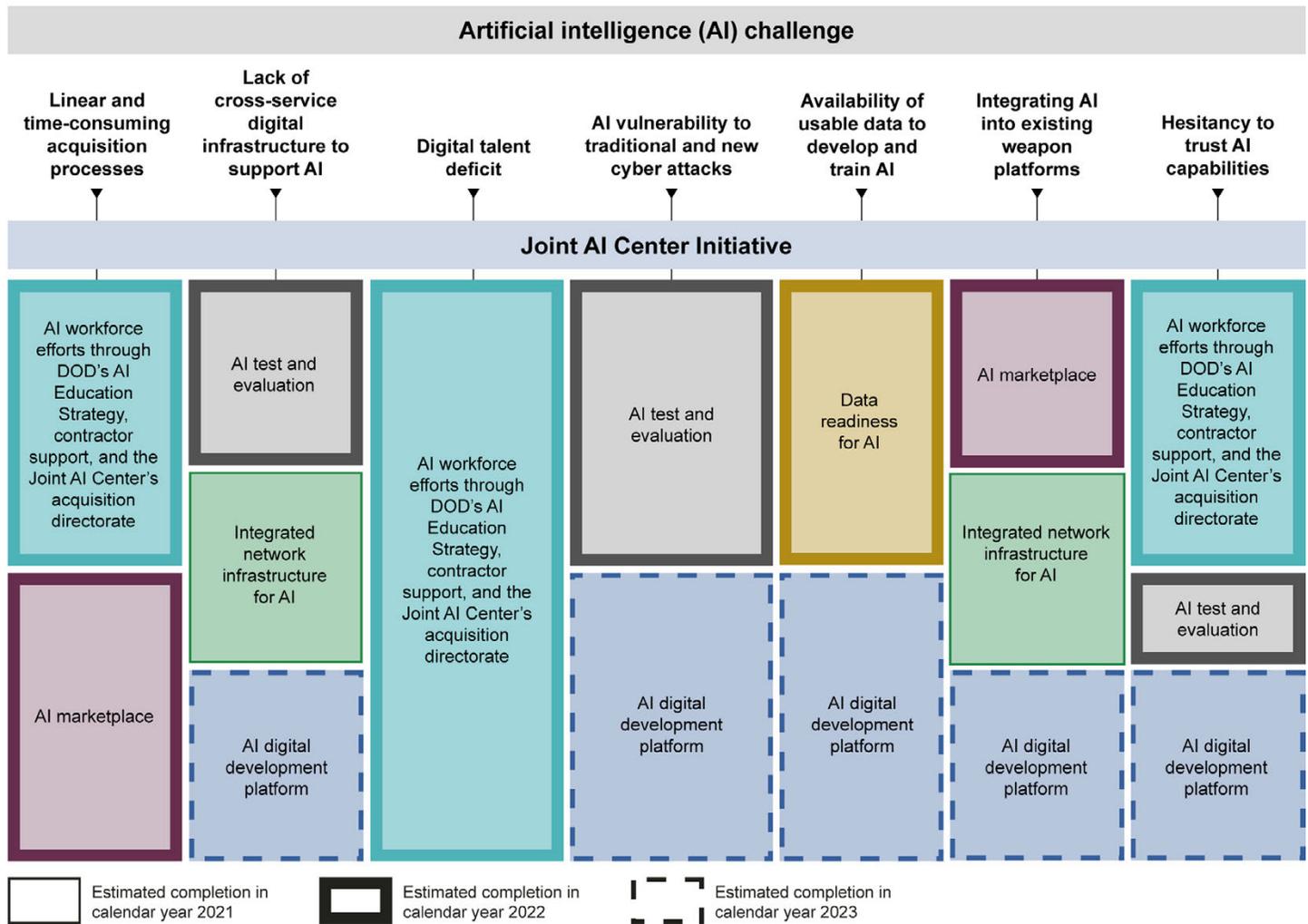
Source: GAO representation of Department of Defense information. | GAO-22-104765

Note: Completion can refer to contract award, full operational capability, or conclusion of the pilot phase.

Further information on the planned initiatives follows.

AI workforce efforts. The Joint AI Center has taken steps to address concerns with acquisition, DOD's AI talent pool, and user trust in AI through efforts aimed at enhancing its AI workforce. It is developing training to provide AI-related skills to current employees, including senior leadership. Details of these efforts are captured in DOD's September

2020 AI Education Strategy, led by the Joint AI Center.[50] The Education Strategy categorizes six AI functional groups across the department based on the necessary AI knowledge or skills each need to be successful in their respective role. These functional groups are:

- Senior leaders who lead DOD's use of AI;
- Managers of AI personnel and AI-enabled tool development, including acquisition officials, who drive DOD toward AI usage;
- Technical experts who help deliver AI to end users;
- Personnel embedded with end users to ensure AI is used effectively and in support of the mission;
- Personnel responsible for bridging the gap between technical experts and end users; and
- End users within, for example, combatant commands.

The Education Strategy states that DOD is initially prioritizing raising AI awareness for DOD leadership to build support and accelerate adoption of these capabilities. In addition, it noted that DOD was either planning or conducting pilot training programs for the skills needed in each of the AI functional groups in fiscal year 2021. For example, DOD planned to focus efforts on training acquisition managers to understand AI-specific acquisition models and testing personnel to identify AI failure modes. DOD planned to pilot and scale these trainings across the department through October 2022.

Outside of DOD's AI Education Strategy, the Joint AI Center is establishing its own Acquisition Directorate to provide expertise that will assist the department in rapidly procuring AI capabilities using the most appropriate acquisition pathway and contracting mechanisms. This directorate will act as the workforce education advisor for the Joint AI Center and DOD on acquisition matters; develop AI-centered training for acquisition personnel; and build tools, contract language, and templates that can be leveraged across the department. The Joint AI Center expects the Acquisition Directorate to be in operation by March 2022.

While DOD builds its AI expertise internally, from 2020 to 2021, the Joint AI Center awarded blanket purchase agreements—known as AI Talent—

---

[50]*2020 DOD Artificial Intelligence Education Strategy.*

to six vendors to bring in contractor support as necessary, according to DOD officials.[51] According to Joint AI Center documentation, these agreements are intended to ensure that DOD grows its talent pool to keep pace with AI technologies and allows for the rapid onboarding of experts from 16 labor categories, including systems engineers, data scientists, and operational test engineers. As of October 2021, the department had obligated approximately $8.1 million dollars across the six vendors.

AI marketplace. The Joint AI Center has taken steps to address concerns over acquiring AI and pursuing integration of AI into existing weapon system platforms by developing an AI marketplace known as Tradewind. Tradewind provides a forum that brings together DOD end users, private industry, and academia to expedite the procurement of AI capabilities. This marketplace serves as a new acquisition business model for AI. It is intended to break down barriers for nontraditional companies to develop AI for DOD and enable the department to quickly award prototype agreements with language suited to the unique aspects of AI, such as addressing intellectual property concerns and training data needs.[52] Tradewind's processes are intended to enable DOD to move from identifying AI needs to awarding a prototype other transaction agreement in 30-60 days and will serve as the focal point for templates, processes, and assistance for AI-unique contracting.[53]

A Joint AI Center official told us that Tradewind will also help with integrating AI capabilities within weapon systems by requiring the end user and AI developer to think through the logistics for integration and ensure that the end user and the AI developer can provide instantaneous feedback throughout the process. The Joint AI Center is also focused on incorporating ethics into this effort by including a responsible AI roadmap

---

[51]Blanket purchase agreements are a simplified method of fulfilling repetitive needs for supplies and services by establishing charge accounts with qualified sources of supply. Agencies may award these agreements to one vendor or to more than one vendor, and then issue individual orders to fulfill requirements for goods and services as they arise. See Federal Acquisition Regulation §§ 8.405-3, 13.303.

[52]Nontraditional companies refers to companies that do not typically sell or develop products for DOD's use.

[53]Other transaction agreements can enable DOD and companies to negotiate terms and conditions specific to a project without requiring them to comply with certain federal regulations that apply to government procurement contracts under the Federal Acquisition Regulation. See 10 U.S.C. § 4003. See also GAO, *Defense Acquisitions: DOD's Use of Other Transactions for Prototype Projects Has Increased*, GAO-20-84 (Washington, D.C.: Nov. 22, 2019).

for Tradewind. An industry official said smaller commercial entities encounter challenges when they compete against larger companies that have more diverse AI portfolios and a longer working history with DOD. As of November 2021, Tradewind was expected to be functional by January 2022, according to a Joint AI Center official.

Data readiness for AI. To address concerns regarding the availability of usable data to develop and train AI, the Joint AI Center plans to establish a framework agreement to ensure data readiness for AI. In particular, it is developing a basic ordering agreement—which generally is a written agreement with one or more vendors to provide specified services at a future point in time—that DOD components can use to prepare their data to develop and train AI capabilities.[54] This agreement, which DOD intends to award to multiple vendors, will allow DOD components and other government agencies to access commercial data preparation services through an approved list of private vendors to meet their AI data needs. These needs may include data curation, data labeling, securing and encryption, and packaging. Data services acquired or developed by vendors under this agreement will be required to integrate and operate with new or existing DOD cloud platforms and leverage standardized intellectual property terms to ensure government ownership of prepared data, as well as other intellectual property protections for AI. These agreements are expected to be awarded and available in February 2022, according to Joint AI Center officials.

AI digital development platform. To address several challenges (the lack of cross-service digital infrastructure, AI vulnerability to cyberattacks, the availability of usable data to develop and train AI, integrating AI into existing weapon platforms, and hesitancy to trust AI capabilities), the Joint AI Center is establishing an AI digital development platform known as the Joint Common Foundation. The Joint Common Foundation is intended to be the department's digital platform to design, develop, and test AI capabilities, and to provide project developers and individual users with access to the cloud, open-source, and commercially available AI development tools (such as open source algorithms), and shared data sets. According to a DOD official, the infrastructure contains various tools and is available to a variety of users, as shown in figure 9.

---

[54]In general, a basic ordering agreement contains the terms and clauses applying to future orders, describes the types of supplies and services to be provided, and contains the methods for pricing, issuing, and delivering future orders. A basic ordering agreement is not a contract. Federal Acquisition Regulation § 16.703.

Figure 9: Notional Depiction of DOD's Joint Common Foundation for Artificial Intelligence (AI)
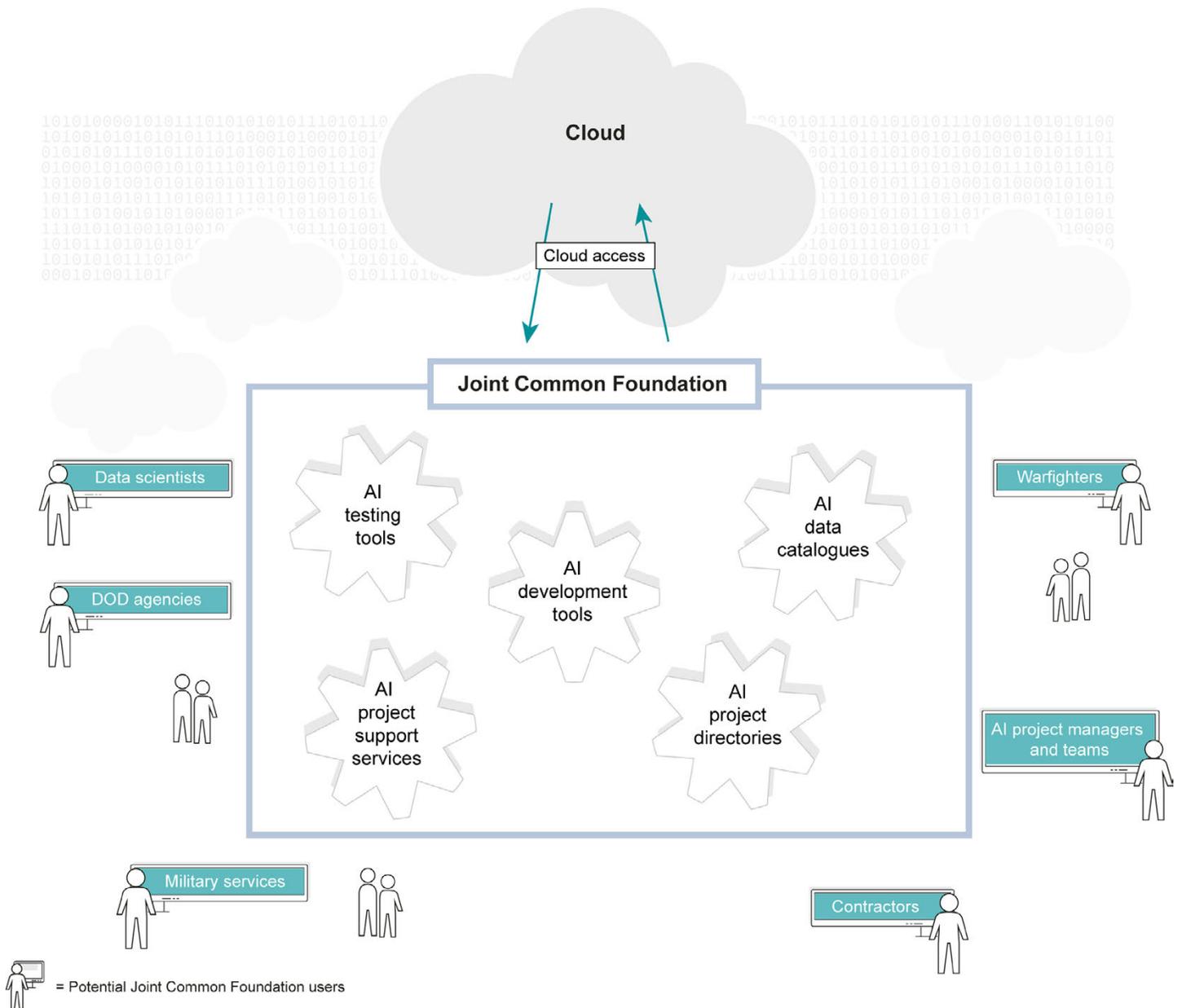
Source: GAO depiction of Department of Defense (DOD) provided information. | GAO-22-104765

According to Joint AI Center documentation, the Joint Common Foundation is being developed in line with DOD's zero trust cybersecurity

architecture.[55] The platform is expected to contain shared elements for its users to develop AI, which includes the data catalog, various open-source tools, and project directories and support services. Additionally, a Joint AI Center official told us that the platform also supports projects that have been approved for initiation by providing the project team with cloud access and the additional tools needed to package, secure, and deploy an AI capability. This official said that final capabilities to be offered in the Joint Common Foundation are evolving. The official added that, for fiscal year 2022, the Joint AI Center is focused on making key AI development tools available while also identifying additional open source and commercially-provided tools that could be useful in the future. According to Joint AI Center officials, this capability became initially operational in March 2021, and as of July 2021, there were 100 unique users and seven projects being developed in the Joint Common Foundation. Even though the final capabilities to be offered in the platform are in flux, the Joint Common Foundation is expected to incorporate classified capabilities by fiscal year 2023. In our related AI report, we recommended that DOD issue a roadmap or a high-level plan that captures all requirements and milestones for developing and onboarding users to the Joint Common Foundation.[56]

Integrated network infrastructure for AI. To address the lack of cross-service digital infrastructure and to integrate AI into existing weapon platforms department-wide, the Joint AI Center is working to establish an integrated network infrastructure and an AI-enabled operating system. Known as Project AI Data Accelerator, the infrastructure is intended to allow the use of AI capabilities across the combatant commands to link weapon platforms and support complex decision-making in battle.[57] According to an official from the Office of the Under Secretary of Defense for Acquisition and Sustainment, this project is intended to ultimately enable the integration of AI capabilities into all service weapon platforms. Additionally, the Joint AI Center is developing an AI needs form that

---

[55]The National Institute of Standards and Technology defines zero trust as an evolving set of cybersecurity paradigms that move defenses from static, network-based perimeters to focus on users, assets, and resources. Zero trust assumes there is no implicit trust granted to assets or user accounts based solely on their physical or network location (i.e., local area networks versus the internet) or based on asset ownership (enterprise or personally owned).

[56]GAO-22-104516SU.

[57]According to DOD officials, Project AI Data Accelerator is the operational portion of DOD's broader AI Data Accelerator initiative that was announced by the Deputy Secretary of Defense in June 2021.

prompts DOD components to consider the technical specifications of the weapon platform the AI will be employed on, among other considerations, before pursuing an AI capability through Tradewind. As of November 2021, DOD was in the process of evaluating vendor proposals with the expectation of a December 2021 award, according to an official from the Office of the Under Secretary of Defense for Acquisition and Sustainment.

AI test and evaluation. To address the lack of cross-service digital infrastructure, vulnerability to traditional and new cyberattacks, and hesitancy to trust AI capabilities, the Joint AI Center has taken several steps to address testing and evaluation of AI. For example, the center drafted a Test and Evaluation framework specific to AI in July 2020. This framework is intended to provide a sequential process for verifying and validating an AI capability in line with DOD's ethical principles and focuses on ensuring security, resilience, and robustness, among other things. According to a Joint AI Center official, the center recently began an initiative to translate DOD's ethical principles into testable requirements, but does not have an estimated time frame for completion.

Additionally, in February 2021, the Joint AI Center released a request for proposals for test and evaluation services blanket purchase agreements that are expected to provide department-wide access to test technology and tools for a variety of AI, automation, and autonomy applications. Specifically, the agreements are intended to streamline the procurement process and increase safety and security of AI capabilities by standardizing the testing and evaluation process while ensuring an independent and unbiased assessment of the quality and readiness of AI-enabled systems to increase confidence by end users, according to Joint AI Center documentation. Eventually, the Joint AI Center anticipates integrating AI-specific test and evaluation processes into the Joint Common Foundation and synchronizing testing and evaluation for AI across the department. Services include analysis of decisions or recommendation made by the AI capability, testing services, and identifying new technologies and development efforts. The AI Test and Evaluation blanket purchase agreements are expected to be awarded and made available to DOD components in March 2022, according to Joint AI Center officials.

According to officials from the Joint AI Center and Office of the Under Secretary of Defense for Acquisition and Sustainment, DOD is not planning to mandate use of these tools. Instead, Joint AI Center officials told us, they are striving to develop tools that are simple and widely available to entice broad use across DOD components. Center officials

noted that they do not want to stifle the efforts and expertise that may already exist throughout the department, particularly within the military services. Instead, DOD wants these tools to complement ongoing efforts and be a valuable resource to DOD components that may not have the expertise or tools developed internally. The National Security Commission on Artificial Intelligence shared similar views, reporting that the Joint AI Center cannot develop and proliferate AI capabilities for every user group or mission area within DOD. The commission recommended that DOD create an organizational structure that pairs top-down strategy with bottom-up development. According to Joint AI Center documentation, the center is developing internal key performance indicators that will assess how effectively it is achieving its mission to transform the department through AI and expects to report quantitative measures on its metrics in the first quarter of fiscal year 2022.

According to various DOD officials, in addition to the initiatives listed above, other DOD entities, including the Chief Data Office and Office of the Under Secretary of Defense for Acquisition and Sustainment—in conjunction with the Joint AI Center—are working on other AI department-wide efforts. Specifically, officials from DOD's Chief Data Office told us they are establishing data teams to deploy to each of the combatant commands as part of the Deputy Secretary of Defense's AI Data Accelerator initiative.[58] These data teams will help the combatant commands to catalog, manage, and automate the collection of data. Building on the work of these data teams, DOD plans to deploy teams of technical experts to help the combatant commands integrate AI into their workflows. According to officials from DOD's Chief Data Office, these data teams will be deployed to each of the combatant commands by the beginning of calendar year 2022. Additionally, the Office of the Under Secretary of Defense for Acquisition and Sustainment is working to adjust the software acquisition pathway for AI and issue updated guidance to DOD components as it becomes available, according to officials from that office.[59] These officials also told us that these adjustments include adding in automated testing and requiring a data strategy for AI projects.

---

[58]According to the Deputy Defense Secretary, the AI Data Accelerator's goal is to rapidly advance data and AI dependent concepts, like joint all-domain command and control, to generate foundational capabilities through a series of implementation experiments or exercises, each one purposefully building understanding through successive and incremental learning.

[59]Department of Defense Instruction 5000.87, *Operation of the Software Acquisition Pathway* (Oct. 2, 2020).

Additionally, we found that each of the military services has ongoing efforts to enable the development and scaling of AI internally. These include:

- **Army.** The Army AI Integration Center is developing its own digital AI platform, called Coeus. According to Army officials, Coeus is the Army's AI development system, comprising both hardware and software infrastructure for algorithm development, that provides a suite of data science and coding tools. DOD officials told us that the Army initiated the development of this effort prior to the Joint AI Center's Joint Common Foundation and that the Army intends for the two systems to link in the future. In addition, the Army is piloting an education program for AI across the Army with the goal of having 500,000 soldiers become AI literate—understanding how to use data and incorporate AI capabilities into their respective units and organizations—by 2024, according to officials. They told us that the program is tailored to meet the needs of three various groups within the Army: senior leadership; data scientists and engineers; and prospective AI end users.

- **Air Force and Space Force.** The Air Force Chief Data Office's Visible, Accessible, Understandable, Linked, and Trusted Data Platform provides a set of tools to support data exploitation activities such as gathering, management, and cleaning. Air Force officials told us this is similar to the Joint AI Center's Data Readiness for AI Development initiative, discussed above. Additionally, officials from the Joint AI Center told us that they are leveraging other tools already in use by the Air Force, including Cloud One and Platform One, to incorporate into the Joint Common Foundation. Cloud One is the Air Force's secure cloud computing platform for DOD, which Joint AI Center officials said provides the Joint Common Foundation with secure access to different cloud services. The Air Force's Platform One manages software factories for development teams and, according to Joint AI Center officials, provides the ability to package AI to Joint Common Foundation users. In addition, Space Force officials told us that they are working to add a project specific to space domain awareness with the Air Force AI Accelerator, the Air Force's main AI hub.

- **Navy and Marine Corps.** According to a Navy AI official, the Navy Information Warfare Center Atlantic has been tracking AI-related skills like data science and engineering possessed by its civilian workforce and has done well at making sure these skills are applied where needed. However, this Navy AI official noted that there is currently no occupational series for AI in the military workforce and the Navy is

considering creating one in line with the National Security Commission on AI's recommendation. In addition, Marine Corps officials said that they are working on a draft strategy document for AI. They are also planning to update a previous use-case checklist that they provided to Marine Corps Systems Command. This checklist is intended to guide potential users—such as program offices—through the process of suggesting potential AI technologies and considerations needed to translate these ideas into development, such as data available.

Agency Comments and Our Evaluation

We provided a draft of this report to DOD for review and comment. DOD provided technical comments, which we incorporated as appropriate.

We are sending copies of this report to the appropriate congressional committees and the Secretary of Defense. In addition, the report will be available at no charge on GAO's website at http://www.gao.gov.

If you or your staff have any questions about this report, please contact Jon Ludwigson at (202) 512-4841 or ludwigsonj@gao.gov or Candice N. Wright at (202) 512-6888 or wrightc@gao.gov. Contact points for our Offices of Congressional Relations and Public Affairs may be found on the last page of this report. GAO staff who made key contributions to this report are listed in appendix IV.

Jon Ludwigson
Director, Contracting and National Security Acquisitions

Candice N. Wright
Director, Science, Technology Assessment, and Analytics

Appendix I: Objectives, Scope, and Methodology

The Senate Report accompanying a bill authorizing fiscal year 2021 appropriations included a provision for GAO to review the Department of Defense's (DOD) artificial intelligence (AI) warfighting acquisition related efforts.[1] This report examines (1) the unique nature of AI and the current status of DOD's AI capabilities that support weapon systems, and (2) DOD's efforts to address identified challenges in developing, acquiring, and deploying AI capabilities for weapon systems. AI capabilities supporting weapon systems are those that DOD can use in its warfighting operations that allow the U.S. to gain decision, speed, agility, and strategic advantages over its adversaries. To align with how DOD categorizes its AI capabilities, we refer to these as AI capabilities for warfighting operations throughout the report.

To inform our views across both objectives, we conducted interviews with or obtained written responses from the following organizations:

- Office of the Secretary of Defense organizations and other key DOD entities: Office of the Director, Operational Test and Evaluation; Office of the Chief Information Officer, specifically the Joint AI Center and the Chief Data Officer; Office of the Under Secretary of Defense for Research and Engineering; Office of the Under Secretary of Defense for Acquisition and Sustainment; Defense Technical Information Center; Defense Digital Service; and Defense Innovation Unit.

- Military services including program offices pursuing AI capabilities: Army AI Integration Center; Army XVIII Airborne Corps; Air Force AI Accelerator; Air Force U-2 Federal Laboratory; Office of the Chief of Navy Operations; Marine Corps Headquarters Office of the Commandant for Information; Space Force Chief Technology Innovation Office; and the Naval Undersea Warfare Center.

- DOD research labs and other research entities: Defense Advanced Research Projects Agency; Naval Surface Warfare Center; Army Combat Capabilities and Development Command including Army Research Lab; and Air Force Research Lab.

- Selected organizations with subject matter expertise: National Security Commission on AI; the National Institute of Standards and Technology; National Defense Industrial Association; and the RAND Corporation.

---

[1]S. Rep. No. 116-236, at 131 (2020).

GAO-22-104765 Artificial Intelligence

- Selected private industry AI vendors working with DOD: Ball Corporation; C3 AI; Zel Technologies; and TRAX International Corporation.

To understand the current state of DOD's AI capabilities for weapon systems, we analyzed DOD documentation and interviewed relevant officials from across the department. Specifically, we analyzed DOD Joint AI Center's fiscal year 2021 AI Inventory and determined the data included were sufficiently reliable for the purposes of illustrating the distribution of identified AI projects across the research and development and procurement funding lines as well as ties to particular major weapon systems. We compared the AI project titles included in the inventory to DOD identified major weapon systems for fiscal year 2021.[2] We also reviewed DOD's 2021 Biannual Reports to Congress. Additionally, we interviewed and reviewed associated documentation—such as the 2019 Defense Innovation Board's Software Acquisition and Practices Study— from a variety of DOD entities involved in developing or procuring AI to understand how AI is different from traditional software, the types of capabilities DOD is pursuing, and strategies for transitioning AI capabilities from research into operations.

To identify the types of challenges DOD faces in developing, acquiring, and deploying AI capabilities for weapon systems, we reviewed recently published reports and recommendations by the National Security Commission on AI and the RAND Corporation and interviewed the authors. We also reviewed DOD documentation and interviewed relevant officials from the Joint AI Center, Office of the Under Secretary of Defense for Research and Engineering, and the military services. This included DOD's AI Strategy, documentation from the Joint AI Center's 2021 AI Symposium, and the military services' AI strategy annexes, among others. Additionally, we selected three private companies working with DOD on AI activities using a non-generalizable random sample stratified into three groups based on the types of projects they work on (AI-enabled, AI-enabling, and core AI) and the military service responsible for these projects (Army, Navy, and Air Force). We interviewed officials from each of the randomly selected, and one additionally selected, private companies to further understand the challenges. To characterize existing and new challenges posed by AI, we analyzed interview responses from

---

[2]Office of the Under Secretary of Defense (Comptroller)/Chief Financial Officer, *Program Acquisition Cost by Weapon System* (February 2020).

each of the entities described above and reviewed challenges described in previously published GAO and DOD reports.

To examine DOD's efforts to address these challenges, we reviewed the Joint AI Center's documentation such as the performance work statements for its data readiness and test and evaluation contracting vehicles as well as its AI digital development platform (known as the Joint Common Foundation) and AI acquisition marketplace (known as Tradewind). We also examined DOD's AI Education Strategy and associated workforce initiatives. To understand similar initiatives undertaken by the military services, we reviewed additional documentation such as the Army's information paper on Workforce Tiers, Education, and Roles as well as the Air Force's AI Strategy Annex and the Navy's AI strategy. To further understand these initiatives and how they will address AI development, acquisition, and integration challenges, we interviewed relevant officials from DOD's Joint AI Center, Chief Data Officer, the Director, Operational Test and Evaluation's Office, and each of the military services such as the Office of the Chief of Naval Operations and Space Force Chief Technology Innovation Office, among others.

We conducted this performance audit from January 2021 to February 2022 in accordance with generally accepted government auditing standards. Those standards require that we plan and perform the audit to obtain sufficient, appropriate evidence to provide a reasonable basis for our findings and conclusions based on our audit objectives. We believe that the evidence obtained provides a reasonable basis for our findings and conclusions based on our audit objectives.

Appendix II: Overview of the Department of Defense's Initial Artificial Intelligence Inventory

To fulfill a congressional mandate, the Department of Defense (DOD) recently began efforts to catalog all known artificial intelligence (AI) projects—including, but not limited to, those for warfighting—across the department that will eventually provide visibility into DOD's AI spending as well as the types of AI capabilities being pursued and what they hope to accomplish.[1] DOD's initial AI inventory, which only included those AI projects funded through DOD's research and development and procurement accounts, identified 685 AI projects as of April 2021. DOD's initial inventory does not include either classified AI projects or those funded through operations and maintenance. According to Joint AI Center documentation, budget justifications for classified programs and operations and maintenance are not machine readable and therefore were not analyzed for this year's inventory. See table 2 for a detailed breakdown of AI projects identified by DOD components in the initial inventory.

Table 2: Artificial Intelligence (AI) Projects by Identified DOD Components, as of April 2021

DOD component	Number of AI projects funded through research and development	Number of AI projects funded through procurement	Total number of AI projects[c]
Air Force (including Space Force)[a]	74	6	80
Army	209	23	232
Marine Corps	26	7	33
Navy	176	39	215
Other DOD entities[b]	117	8	125

Source: GAO analysis of Department of Defense (DOD) information. I GAO-22-104765

[a]DOD's methodology combined AI projects from the Air Force and Space Force.

[b]Other DOD entities include combatant commands and other unspecified DOD components included in the Joint AI Center's methodology.

[c]DOD's initial inventory does not include classified AI projects or those funded through operations and maintenance.

DOD's initial inventory does not capture all AI projects nor does it provide insight into the capability area or cost associated with each project. Officials from the Joint AI Center stated that they expect to identify additional AI projects funded through operations and maintenance, but

---

[1]166 Cong. Rec. H8251 (Dec. 21, 2020). This Joint Explanatory Statement accompanies the Consolidated Appropriations Act, 2021, Pub. L. No. 116-260 (2020).

cannot estimate how many. As a result, the reported number does not fully reflect the number or range of projects currently underway. Additionally, the budgetary documentation that was able to be analyzed to identify AI projects does not contain sufficient information such as a description of AI capability (i.e. for warfighting or weapon systems) or the portion of the project cost that is allocated to develop the AI capability. As a result, the department is unable to identify the overall funding directed toward AI or the eventual use for these projects (i.e. warfighting, logistics, etc.)

DOD officials told us they are working to refine their methodology to address these limitations that will allow the department to systematically track the warfighting or business requirements being addressed. DOD expects this new system—known as the DOD AI Inventory Portfolio Analytics Tool—to support future iterations of the Joint AI Center's inventory of AI activities provided to Congress. In our related AI report issued in February 2022, we recommended that the department develop a high-level plan or roadmap documenting requirements, activities, and milestones that support the preparation of the department's AI portfolio inventory and budget data.[2]

---

[2]GAO, *Artificial Intelligence: DOD Should Improve Strategies, Inventory Process, and Collaboration Guidance,* GAO-22-104516SU (Washington, D.C.: Feb. 16, 2022).

Appendix III: Overview of GAO's Artificial Intelligence Accountability Framework

In June 2021, we published an artificial intelligence (AI) accountability framework that identifies key practices to help ensure accountability and responsible use of AI by federal agencies and other entities involved in the design, development, deployment, and continuous monitoring of AI systems.[1] The AI accountability framework is organized around four complementary principles addressing governance, data, performance, and monitoring. For each principle, the framework describes key practices for federal agencies and other entities that are considering, selecting, and implementing AI systems. Each practice includes a set of questions for entities, auditors, and third-party assessors to consider, as well as procedures for auditors and third-party assessors. Figure 10 illustrates the key points of each principle.

---

[1]GAO, *Artificial Intelligence: An Accountability Framework for Federal Agencies and Other Entities,* GAO-21-519SP (Washington, D.C.: June 2021). To develop this framework, we convened a Comptroller General Forum, in September of 2020, of experts in industry, government, nonprofits, and academia to discuss factors affecting oversight of AI, including AI governance, sources of evidence, methods to assess implementation of AI systems, and identifying and mitigating potential bias and inequities. We selected 23 experts representing 20 organizations to participate in the forum. These individuals presented a variety of perspectives, including those of software developers, data scientists, privacy/security experts, risk management professionals, legal counsel, civil liberties advocates, users, and individuals affected by AI systems. We also conducted an extensive literature review and obtained independent validation of key practices from program officials and subject matter experts. In addition, we interviewed AI subject matter experts representing industry, state audit associations, nonprofit entities, and other organizations, as well as officials from federal agencies and Offices of Inspector General.

Figure 10: Artificial Intelligence (AI) Accountability Framework

Data
Ensure quality, reliability, and representativeness of data sources and processing.

Data Used to Develop an AI Model
Entities should document sources and origins of data, ensure the reliability of data, and assess data attributes, variables, and augmentation/enhancement for appropriateness.

Data Used to Operate an AI System
Entities should assess the interconnectivities and dependencies of data streams that operationalize an AI system, identify potential biases, and assess data security and privacy.

Monitoring
Ensure reliability and relevance over time.

Continuous Monitoring of Performance
Entities should develop plans for continuous or routine monitoring of the AI system and document results and corrective actions taken to ensure the system produces desired results.

Assessing Sustainment and Expanded Use
Entities should assess the utility of the AI system to ensure its relevance and identify conditions under which the AI system may or may not be scaled or expanded beyond its current use.

Governance
Promote accountability by establishing processes to manage, operate, and oversee implementation.

Governance at the Organizational Level
Entities should define clear goals, roles, and responsibilities, demonstrate values and principles to foster trust, develop a competent workforce, engage stakeholders with diverse perspectives to mitigate risks, and implement an AI-specific risk management plan.

Governance at the System Level
Entities should establish technical specifications to ensure the AI system meets its intended purpose and complies with relevant laws, regulations, standards, and guidance. Entities should promote transparency by enabling external stakeholders to access information on the AI system.

Performance
Produce results that are consistent with program objectives.

Performance at the Component Level
Entities should catalog model and non-model components that make up the AI system, define metrics, and assess performance and outputs of each component.

Performance at the System Level
Entities should define metrics and assess performance of the AI system. In addition, entities should document methods for assessment, performance metrics, and outcomes; identify potential biases; and define and develop procedures for human supervision of the AI system.

Source: GAO. | GAO-22-104765

Appendix IV: GAO Contacts and Staff Acknowledgments

GAO Contacts	Jon Ludwigson, (202) 512-4841 or ludwigsonj@gao.gov Candice N. Wright, (202) 512-6888 or wrightc@gao.gov
Staff Acknowledgments	In addition to the contacts named above, Raj Chitikila (Assistant Director), Erin Butkowski (Analyst in Charge), Lori Fields, Laura Greifner, Sean Manzano, and Megan Stewart were principal contributors. In addition, the following people made contributions to this report: Jennifer Andreone, Penney Harwell Caramia, Hannah Hubbard, Joshua Leiling, Gabrielle Matuzsan, Miranda Riemer, Andrew Stavisky, and Alyssa Weir.

Related GAO Products

Artificial Intelligence: DOD Should Improve Strategies, Inventory Process, and Collaboration Guidance. GAO-22-104516SU. Washington, D.C.: February 16, 2022.

Artificial Intelligence: An Accountability Framework for Federal Agencies and Other Entities. GAO-21-519SP. Washington, D.C.: June 30, 2021.

Weapon Systems Annual Assessment: Updated Program Oversight Approach Needed. GAO-21-222. Washington, D.C.: June 8, 2021.

High-Risk Series: Federal Government Needs to Urgently Pursue Critical Actions to Address Major Cybersecurity Challenges. GAO-21-288. Washington, D.C.: March 24, 2021.

Artificial Intelligence in Health Care: Benefits and Challenges of Technologies to Augment Patient Care. GAO-21-7SP. Washington, D.C.: November 30, 2020.

Science & Tech Spotlight: Agile Software Development. GAO-20-713SP. Washington, D.C.: September 29, 2020.

Defense Science and Technology: Opportunities to Better Integrate Industry Independent Research and Development into DOD Planning. GAO-20-578. Washington, D.C.: September 3, 2020.

Science & Tech Spotlight: Deepfakes. GAO-20-379SP. Washington, D.C.: February 20, 2020.

Weapon Systems Cybersecurity: DOD Just Beginning to Grapple with Scale of Vulnerabilities. GAO-19-128. Washington, D.C.: October 9, 2018.

Artificial Intelligence: Emerging Opportunities, Challenges, and Implications for Policy and Research. GAO-18-644T. Washington, D.C.: June 26, 2018.

Technology Assessment: Artificial Intelligence: Emerging Opportunities, Challenges, and Implications. GAO-18-142SP. Washington, D.C.: March 28, 2018.

Military Acquisitions: DOD Is Taking Steps to Address Challenges Faced by Certain Companies. GAO-17-644. Washington, D.C.: July 20, 2017.

Defense Science and Technology: Adopting Best Practices Can Improve Innovation Investments and Management. GAO-17-499. Washington, D.C.: June 29, 2017.

Weapon Systems: Prototyping Has Benefited Acquisition Programs, but More Can Be Done to Support Innovation Initiatives. GAO-17-309. Washington, D.C.: June 27, 2017.

Defense Advanced Research Projects Agency: Key Factors Drive Transition of Technologies, but Better Training and Data Dissemination Can Increase Success. GAO-16-5. Washington, D.C.: November 18, 2015.