November 2021

# Federal Prize Competitions

## Using Competitions to Promote Innovation in Artificial Intelligence

CSET Policy Brief


AUTHORS
Ali Crawford
Ido Wulkan

## Executive Summary

From the marketplace to the sports arena, humans thrive on competition where they are often rewarded for their creativity and excellence. In the narrower field of technology innovation, competitions can advance the state of the art and reward breakthroughs in areas that have confounded researchers or slowed progress. Government, too, can play a role. It can provide powerful incentives for the private sector by organizing competitions that spur technological progress. If organized correctly, competitions can foster collaboration and innovation around unique national security challenges.

But choosing how to organize and host competitions involves far more than just offering monetary incentives. Competition organizers should consider the problem or issue that any given competition centers around, its reward system, and the number of participants or teams. They should also consider the best way to share the competition results, and the frequency of competitions. Over the last decade progress in artificial intelligence has occurred rapidly. AI-powered systems have already demonstrated great potential for solving some of the most difficult scientific challenges of the day. These include areas such as determining protein structure, helping drivers operate vehicles safely in urban environments, and identifying diseases. Given its transformational potential, governments should incentivize the development of AI applications for national security. U.S. federal agencies have begun to use competitions to do just this, but their use remains limited.

This policy brief examines 814 competitions conducted from 2010 to 2020 across the U.S. federal government as found in the Challenge.gov archive.[1] During this time period, the total prize purse reached roughly $243 million. In comparison, total federal research and development (R&D) outlays for the period were over $1.3 trillion.[2] While the prize purse does not represent all of the expenses involved in facilitating a competition—such as associated personnel costs or assembling the necessary infrastructure—the total costs still represent only a small fraction of yearly federal outlays and remain an underutilized tool to promote innovation. A review of federal competitions suggests that there are three factors

that contribute to the potential for a greater chance of satisfying organizer goals or useful implementation of a successfully-designed product.

These factors both increase the incentives for participants and increase the likelihood that promising entries are operationalized within the federal sector. First, prizes should reflect the time, effort, and resources involved for participants. Though larger prizes are more likely to attract high quality participants, it is not entirely necessary. Second, organizers should design competitions to ensure promising entries have a path towards rapid procurement or scaling versus the traditional acquisition and procurement process.[3] Finally, competitions can benefit from access to or the creation of professional networks among participants, which ultimately can improve public-private partnerships.

While the private sector also uses competitions and benchmarks to incentivize progress in AI, they have not addressed the full range of national security requirements. The federal government should consider additional competitions centered on the challenges of AI safety and security, AI trust and explainability, and AI applications for cybersecurity. Private sector work is occurring in each of these fields and can provide examples for how agencies may structure their own challenges. However, federal competitions have the potential to uplift the research and focus on the specific national security aspects of each of these topics.[4] With that, this brief makes three findings on the use of federal competitions to promote innovation in artificial intelligence:

- **Agencies have yet to fully leverage AI competitions for national security requirements:** Agencies could increase their use of prize competitions without significantly impacting their R&D budgets. Competitions have unique benefits that are less easily replicated through traditional R&D processes.[5] Disregarding additional expenditures, the sponsoring agency or department typically pays only for success, shifts risk of failure to participants, and is therefore able to establish ambitious goals. This form of open innovation attracts a broader range of participants by lessening bureaucratic impediments to working with the government. There are challenges happening within the private sector that the federal government could emulate.

- **Competitions provide a means to test operational effectiveness prior to potential prototyping, scaling, or procurement:** Running a competition prior to procurement allows the sponsoring agency to gain insight into the available tools and technologies and to test their operational performance by welcoming a diverse range of participants with various technological approaches. Because competitions transfer most of the risk of failure to participants, the agency could pursue more ambitious mission-oriented goals without risking great financial loss. A competition could also serve as an alternative means of acquisition which can speed the process for the government and allow opportunities for more private sector entities to participate.
- **The National AI R&D Strategic Plan and other federal strategies[6] provide a roadmap for the topics of future competitions:** Competitions can drive AI innovation for the government if they are designed around federal innovation strategies and priorities. Key topics include:
    - Safety, security, and trust in AI/machine learning (ML)-enabled systems: With increased emphasis on accelerating the adoption and integration of AI technologies into federal infrastructure and operations, the need for understanding AI-enabled system decision-making, trusting those decisions, and guaranteeing its continued performance is a critical area of strategic R&D focus.[7]
    - Deepening public-private partnerships: Deepening and strengthening public-private partnerships in AI R&D remains a key priority for most federal agencies and departments, because much of the talent, expertise, and current breakthroughs in ML are occurring in the private sector and academia. Competitions are an additional avenue for collaboration, partnership, and cooperation with these critical sectors. Furthermore, sponsors of federal competitions would benefit from working directly with technology incubators within their departments like the U.S. Department of Defense's Defense Innovation Unit or the U.S. Department of Health and Human Services' IDEA Lab to streamline the procurement of potential solutions.

## Table of Contents

## Introduction

In August 2016, seven computer towers stood quietly humming on stage at the Paris Hotel in Las Vegas. Underneath the stage, 300 kilowatts of electrical power and 180 tons of cold water powered the machines and kept them from overheating. Each of the towers represented one of seven teams competing for a $2 million prize. Despite the hushed buzz of excitement emanating from the audience, the teams could do nothing but watch as the computers competed. This was the Cyber Grand Challenge, the world's first machine-only cyber defense competition, organized by the Defense Advanced Research Projects Agency to advance the field of automated cybersecurity and vulnerability discovery.

The challenge spanned several years. From the first call for applications in 2013, DARPA received over one hundred applications. Three qualifying events narrowed them down to seven teams,[8] with each receiving $750,000 to build a capability to automatically discover vulnerabilities and defend a simulated network. Each team designed their entry, termed a Cyber Reasoning System, to be a fully autonomous computer system that defended a set of software services and demonstrated the autonomous application of firewall rules, vulnerability detection, and patching.[9] Essentially, the CRS attempted to emulate many human cyber defense skills and apply them at machine speed against other machines. On the day of the competition, the participants watched as their programs played a virtual version of a capture-the-flag competition in a digital area that DARPA had riddled with software bugs.

This competition was notable for its complexity and cost. For this grand challenge, DARPA built a custom air-gapped network not accessible to the internet and an associated operating system.[10] Built into this custom network were numerous vulnerabilities. Gameplay consisted of each team's CRS discovering vulnerabilities, creating patches, and submitting the patches to DARPA referees. The referees performed tests to check the patches and exploits for functionality and produced scores for each team based on proof of vulnerability demonstration, system security, and system availability. An important objective for the competition's organizers

was the ability of the CRS to ensure the continuous operation of services, and they penalized any deployed firewall rules or patches that disrupted service availability.[11] After nearly twelve hours of play, a machine called "Mayhem," building on years of earlier research, was declared the winner. After the competition, the U.S. Department of Defense's technology accelerator organization, the Defense Innovation Unit (DIU), awarded the team an $8 million contract to take Mayhem to production, and in 2020, the DOD awarded a $45 million contract to deploy Mayhem across multiple DOD networks.[12]

The Cyber Grand Challenge demonstrates some of the promises and the challenges of federal prize competitions. The competition exceeded the organizer's expectations as teams demonstrated the ability to discover and patch vulnerabilities quickly and effectively within the constraints of the gameplay architecture.[13] In addition, the winning system was valuable enough to warrant acquisition into DOD network architecture. Competitions like the Cyber Grand Challenge can catalyze research and prototype development. However, they can be expensive, with no guaranteed return on investment. Between the qualification rounds and the grand prize awards, DARPA spent about $9 million on prize money alone. This figure does not include the cost of the venue, or the computing infrastructure required for both participants and gameplay. However, these are also modest in comparison to the contracts awarded to Mayhem following the competition.

This report begins by reviewing the history of federal prize competitions and the purposes they serve in the larger innovation ecosystem. We explore more recent federal competitions by examining data in the General Services Administration's (GSA) Challenge.gov archives to identify factors in successful competitions and applicable lessons learned. We conclude by exploring how federal departments and agencies could expand their use before offering conclusions and findings.

## Federal Prize Competition Policy and Practice

In the drive to sustain U.S. technological leadership following World War II, the federal, private, and academic sectors collaborated to deliver incredible innovations such as the semiconductor, the early internet, and global positioning satellites. Many of these innovations were supported by generous government funding and this system of innovation thrived under the leadership of the federal government for many years. Commercialization and globalization fueled private sector growth over time, and beginning in the 1980s, the private sector emerged as the dominant force in the U.S. innovation ecosystem.

Presently the private sector funds and conducts 70 percent of U.S. research and development.[14] While the federal government is the second-largest funder of total U.S. R&D, it is the largest funder of academic and basic research.[15] Historically, geostrategic competition drove significant federal government investment in R&D. It used the funding to spur innovation in areas of critical importance, but in more recent years the private sector has been the primary source of innovative technologies. While the government often benefits from these advances, the private sector is unlikely to meet unique government requirements without greater incentives.[16] In these cases, federal investment in the form of targeted R&D may be necessary, one form of which is federal prize competitions.

These competitions are contests sponsored by one or more government agencies, sometimes in partnership with private companies, in which the government uses monetary or nonmonetary incentives to advance knowledge within a particular field or to solicit tools and solutions for specific problems.[17] Sometimes, a competition will serve both purposes. Competitions provide a range of unique benefits that differ from other R&D and acquisition efforts. Hosting a competition allows a diverse group of participants to submit ideas with fewer bureaucratic impediments to participation. The sponsor also shifts some of the risk of failure to the participants.[18] While this might dissuade some potential participants, it generally benefits both the sponsor and the competitors. The sponsoring department can pursue more

ambitious goals without significant upfront costs. Likewise, participants benefit from incentives such as prize purses, exposure, access to professional networks, and commercialization opportunities.

However, competitions are not a replacement for long-term and sustained investment in basic R&D or other policy tools.[19] Instead, they are meant to augment them. Ideally, competitions will lead to breakthrough innovations, but progress can still occur even when competitions are unsuccessful. For example, the competition may identify weaknesses or areas in which additional R&D is required. This suggests that competitions are best suited for the applied R&D of experimental, exploratory, or early-stage technologies allowing competition organizers to measure the operational effectiveness of entries prior to scaling, integration, or deployment.

The use of competitions received a boost in 2010 with the America COMPETES Reauthorization Act,[20] which extended broad authorities to all federal agencies and departments to conduct competitions to further agency missions or to pursue fields of interest for which the private sector lacked incentive. Shortly after this, the GSA created the Challenge platform to host information for federal prize challenges, agency tool kits, case studies, and both past and present federal competitions.[21] To assess the use of competitions, we examined a dataset of archived competitions on the Challenge.gov platform. While the Challenge archive is a comprehensive and useful resource, it is not complete. There are known instances of federal competitions that agencies did not add onto the platform and may be found or hosted elsewhere. Also, the archive does not maintain information for competitions held prior to 2010. Therefore, the Challenge archive data is intended to act as a sample of the federal prize competition ecosystem.

We identified 814 federal competitions between 2010 and 2020, with total prize money offered totaling roughly $243 million. It is important to note that total prize money awarded does not reflect associated costs with facilitating a competition. However, even taking these additional costs into account, the prize money awarded is small (.02 percent) compared to total federal R&D outlays, which was nearly $1.3 trillion over the same period.[22] Of

the 814 competitions, 71 specifically had an AI focus, purpose, or goal with total prizes of $21 million. The DOD, NASA, and DHHS conducted the most competitions. For AI competitions, the DOD, the Intelligence Advanced Research Projects Agency, and NASA conducted the most. The National Artificial Intelligence Initiative Office has also started tracking AI-specific competitions, suggesting interest in the continued analysis of the outcomes of federal AI competitions.[23]

As shown in Table 1, "Software and Applications" is the most commonly reported challenge goal type followed by "Ideas" for both federal competitions in general and AI competitions specifically. Though some competitions may list more than one goal type, the data shows that agencies are seeking tangible solutions to practical problems as well as novel ways to make use of federal datasets.

Table 1: Challenge Goal Types

| Type | Description | Number of Competitions with this Goal | Number of AI Competitions with this Goal |
|---|---|---|---|
| **Software and Applications** | *Creating a software application to solve an existing problem or to draw attention to potential uses of available datasets* | 235 | 33 |
| **Ideas** | *Seeking new ways of understanding and framing problems, new processes to solve existing problems, and innovating implementations as solutions* | 202 | 11 |
| **Creative, Design, and Multimedia** | *Capturing and communicating a concept or aesthetic that would be difficult to achieve with a grant or contract* | 182 | 1 |

| | | | |
|---|---|---|---|
| **Scientific** | *Understanding the problem, solution, or outcome using empirical or measurable evidence-based practices* | 178 | 18 |
| **Technology Demonstration and Hardware** | *Seeking prototypes or fully developed solutions to catalyze and demonstrate breakthrough technical innovations* | 77 | 7 |
| **Analytics, Visualizations, and Algorithms** | *Improvements to interpreting or communicating data* | 40 | 21 |
| **Entrepreneurship and Business Plans** | *Training and equipping entrepreneurs or launching ventures* | 8 | 1 |

Source: CSET analysis of Challenge.gov archive data and Challenge.gov.

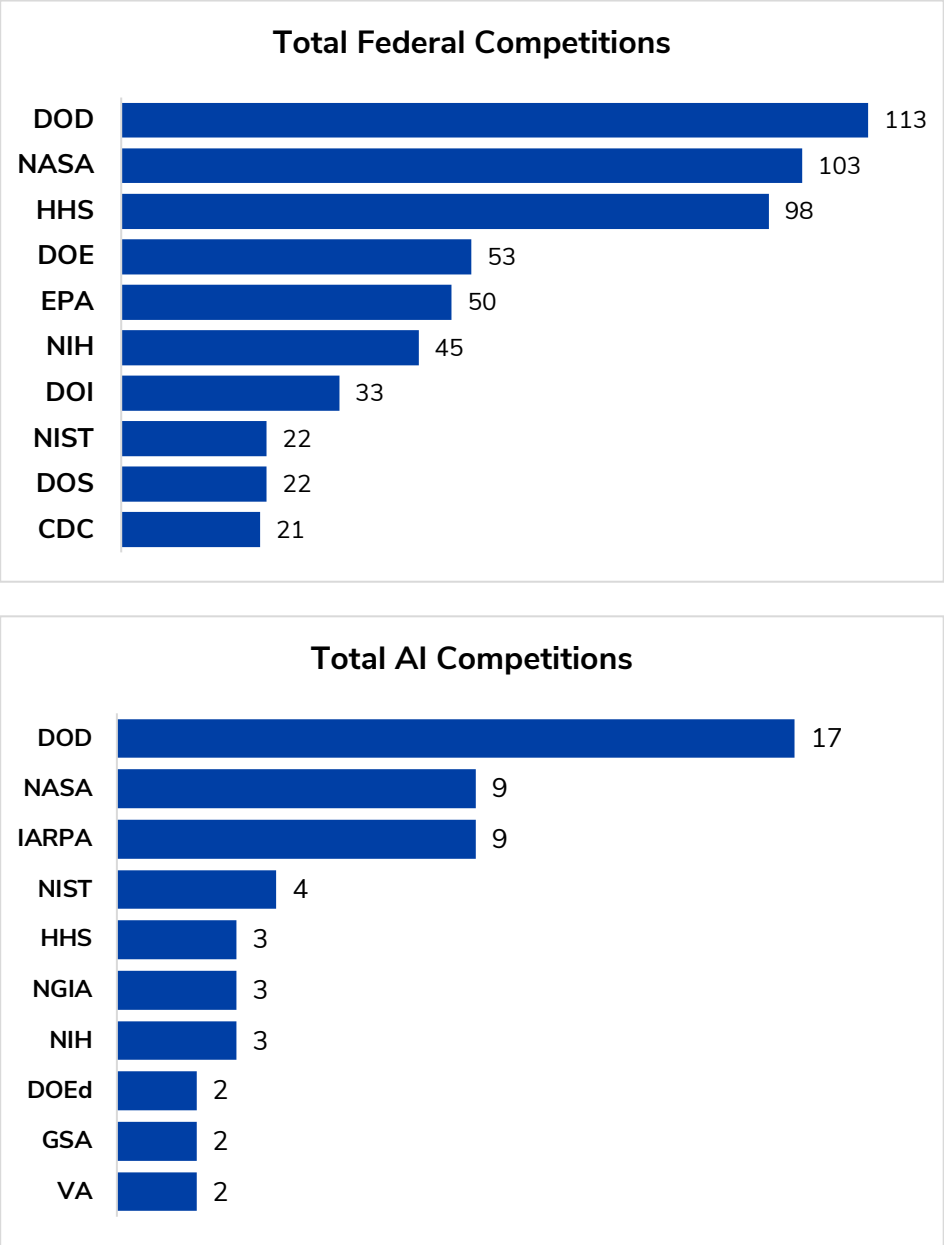
Table 2: Breakdown of Federal Competition Data, 2010–2020

| Year | Number of Competitions | Total Prize Purse *(in millions)* |
|---|---|---|
| 2010 | 26 | $11.2 |
| 2011 | 81 | $6.6 |
| 2012 | 88 | $8.0 |
| 2013 | 68 | $14.8 |
| 2014 | 75 | $13.0 |
| 2015 | 86 | $28.2 |
| 2016 | 82 | $11.5 |
| 2017 | 60 | $6.9 |
| 2018 | 64 | $55.5 |
| 2019 | 90 | $41.7 |
| 2020 | 94 | $46.0 |
| **Totals:** | **814** | **$243.9 million** |

| Year | Number of AI-Specific Competitions | Total AI Prize Purse *(in thousands)* |
|---|---|---|
| 2010 | 2 | $50 |
| 2011 | 2 | $60 |
| 2012 | 2 | $100 |
| 2013 | 3 | $1,057.5 |
| 2014 | 5 | $152 |
| 2015 | 4 | $160 |
| 2016 | 4 | $133 |
| 2017 | 8 | $1,589.5 |
| 2018 | 9 | $4,900.5 |
| 2019 | 15 | $6,940 |
| 2020 | 17 | $5,874.8 |
| **Totals:** | **71** | **$21 million** |

Source: CSET analysis of Challenge archive data.

Figure 1: Lead Sponsors of Federal Competitions and AI-Specific Competitions, 2010–2020*




Source: CSET analysis of Challenge archive data.

* For length, only the top 10 agencies or departments are represented in these graphs.

### *Additional Factors in Competitions*

Competitions with defined problems, transparent rules for contestants, and a clear means to identify the winners have the potential for a higher payoff.[24] Additionally, there are other contributing factors. Incentives, either monetary or nonmonetary, are important components in attracting high-quality participants. Ideally, competitions are also tied to an agency's acquisition and procurement processes, as they can be time-consuming and difficult to navigate. Competitions can avoid acquisition delays to increase speed of adoption of solutions. Other factors, like engagement with professional networks and outside partnerships, benefit all participants by expanding the competitor pool, increasing engagement and exposure, and deepening federal engagement with critical private sector industries.

Prizes and nonmonetary incentives

Agencies have used prizes and nonmonetary incentives of varying amounts and types in federal competitions to both attract participants and source viable solutions. Well-advertised and high-profile competitions are more likely to attract a wider field of competitors who see reputational benefits from participating and potentially winning. If the investment required from participants is not reflected in the prize purse, both the quantity and quality of the participant pool may suffer. However, this does not mean that only large prize purses will attract participants. For the GSA's End-User License Agreement challenge, the prize purse totaled just $20,000 and yet the winning team directly solved what the administration hoped to achieve. This stands in contrast to the Cyber Grand Challenge's $2 million grand prize, but the outcomes of each competition were both positive and met established criteria.

Nonmonetary incentives such as access to professional networks, developer tools, commercialization opportunities, and publicity can also play a role.[25] For example, an Air Force Research Laboratory competition sought innovations to improve the features of specialized chips or Application Specific Integrated Circuits but did not offer monetary prize packages.[26] Instead, AFRL offered

selected participants access to developmental and design tools valued at $10 million per license.

While cash awards may seem like a large outlay for a federal agency, the reality is more nuanced. Aside from the costs of running the competition, the sponsoring agency generally pays only for success. This is a key distinction between competitions and traditional contract awarding, where payment is administered incrementally for work performed or a service as agreed upon by the signing parties.[27] If the participants do not meet the requirements of the competition and no winner is declared, the prize purse may not be awarded. Originally, DARPA's Autonomous Vehicles Grand Challenge did not award its $1 million prize because no contestant finished the predetermined course.[28] A subsequent round was held a year later, with an updated grand prize of $2 million awarded to the first of the five teams to complete the course.

There are other ways organizers can award prizes. For example, a challenge executed by the research arm of HHS in 2018 required teams to build an app to simplify the process of collecting, interpreting, and sharing certain types of patient data called Patient Reported Outcomes. PRO are self-assessment health reports given by the patient without interpretation by a healthcare provider.[29] The agency divided the grand prize of $75,000 into a $35,000 up-front payment and an additional $40,000 payment contingent on the successful completion of a pilot period.[30] In the Environmental Protection Agency's Smart City Air Challenge, the agency sought new ways to use data from air quality sensors in a community. The organizers awarded two grand prizes of $40,000. Winners initially received $20,000 with the remainder paid a year later following successful implementation.[31] Seed money can also motivate startups or individuals to participate in a challenge who otherwise lack the resources to fully develop their solutions or are unable to navigate the federal acquisition bureaucracy.

Procurement

Ideally, organizers will design competitions to enable the rapid transfer of winning submissions into the procurement process. The

solicitation and evaluation stage of the federal acquisition process requires agencies to publicize the proposed acquisition and evaluate submissions based on determining criteria before awarding contracts. Agencies can speed the acquisition and deployment of winning submissions by structuring the competition to satisfy procurement authority evaluation requirements in advance using Challenge-based Acquisition, a term coined by MITRE.[32] This technique effectively streamlines the acquisition process and leverages the benefits of real-world competitions to supplement proposals.

Furthermore, sponsors of federal competitions can benefit from working directly with technology incubators within their departments like DOD's DIU or the HHS IDEA Lab to streamline the procurement of potential solutions. For example, the DIU is responsible for accelerating and facilitating the military's adoption of emerging commercial technologies by reducing the length of the average project life cycle and facilitating the transition of commercial products to large-volume DOD procurement. The DOD's Joint Artificial Intelligence Center has launched its own accelerator to more rapidly acquire AI systems. In both the DIU and JAIC examples, tying competition winners to these existing rapid procurement processes can speed promising new capabilities into the pipeline and give participants a clear path to government procurement.

Similarly, the Small Business Innovation Research and Small Business Technology Transfer programs provide additional opportunities for small U.S. businesses to compete for federal investment with the potential for commercialization, and can also be used as additional incentives for participants. For example, the U.S. Army's xTECH SBIR competition series welcomed submissions in six topic areas to address crucial capability gaps and to accelerate prototype development.[33] Whereas the typical SBIR process has three phases, the first of which requires applicants to demonstrate technical merit and feasibility, winners of the xTECH series were able to bypass the first phase and proceed directly to second phase contracts.

Professional, academic, and private sector networks

Competitions can leverage existing networks of researchers, academic institutions, entrepreneurs, and other professionals, as well as foster the creation of new ones. These networks can be a valuable resource for planning a competition, increasing the reach to target audiences, and enhancing collaboration before and after a competition. For the 2016 Army Cyber Innovation Challenge, the army solicited and tested defensive software solutions for its network infrastructure through an organization called the Consortium for Command, Control, and Communications in Cyberspace, or C5. This is a network of professional institutions and companies in command, control, communications, computers, intelligence, surveillance and reconnaissance, also known as C4ISR, with which the army maintains a standing Other Transaction Authority.[34] Competitions can also help build a network of former grantees who can identify potential competitors and provide feedback to assist in future competition design. Former grantees can serve as mentors for current competitors and offer participants advice and guidance on innovation and entrepreneurial skills.

## Examining Previous Federal AI Competitions

A number of previous federal competitions demonstrate the benefits of prize challenges while also illuminating the factors described in the previous section. The following examples provide insight into how various agencies used competitions to source new technologies or to discover novel methods of conducting mission-oriented tasks.

### The Food and Drug Administration's Open Data Challenge Detecting Adverse Event Anomalies

In 2020, the FDA asked participants to develop an algorithm to automatically detect adverse event anomalies using publicly-available data and without the use of labeled training data.[35] Currently, the FDA uses a voluntary reporting system for consumers, health care professionals, product manufacturers, and others to report adverse reactions to food, drug, cosmetics, and more. Human analysts monitor these reports and track for signals that may indicate risk. Adverse event anomalies arise when data is illogical or otherwise incorrectly reported.[36] For example, a drug which receives adverse event reports before it was legally marketed or in clinical trials would be considered an adverse event anomaly.

The FDA competition was part of their data modernizing strategy, which included efforts to develop novel methods such as AI/ML-enabled tools to process and analyze FDA data. Two top performers were recognized for their innovative submissions, and the FDA said that it had three major outcomes.[37] First, it improved the agency's understanding of the possible AI/ML-enabled approaches to monitoring key datasets. The competition also served as a valuable first step in the FDA's use of AI techniques to facilitate monitoring and surveillance of adverse event reporting. It also demonstrated the limitations of AI techniques, at least at the present, and suggested that human experience and intuition is still necessary.[38] While AI reduced the burden on human analysts, the background knowledge and experience of a human event reviewer proved invaluable. In this example, the competition helped the agency better understand both the promise and limitations of

existing AI techniques when applied to FDA data. It also helped the agency begin to see where AI capabilities might best be applied within the organization.

The U.S. Navy's AI Applications to Autonomous Cybersecurity Series (AI ATAC)

Beginning in 2019, the DOD's Naval Information Warfare Systems Command launched three competitions aimed at improving cybersecurity by automating the detection of and response to adversarial campaigns.[39] In the first competition, the navy solicited white papers and corresponding AI/ML-enabled tools that could automate end-point security detection to better protect individual computers on a network. The second competition sought tools for automated detection of advanced persistent threat activity at the network level. The final and most ambitious competition sought AI/ML-enabled tools to automate security orchestration processes to enhance the detection and prevention of APT activity. The total prize package for all three competitions totaled $1.3 million, and the navy was authorized to award follow-on production contracts. The competition demonstrated the need to reduce the burden on the navy's human analysts as a large amount of time is spent triaging security alerts.[40] The competitions provided a realistic environment to examine new capabilities due to the navy's partnership with Oak Ridge National Laboratory, and they allowed the navy to gain insight into how the tools would perform in operational settings.

National Institute of Standards and Technology Differential Privacy Challenge Series

Starting in 2018, NIST began sponsoring a series of competitions to advance research into differential privacy, a technique for preserving individual privacy in large datasets while maintaining the dataset's utility in training ML systems. Public agencies collect vast amounts of data containing information that is useful for policymakers, researchers, and the general public. However, current privacy-protecting techniques either provide insufficient protection, or the resulting synthetic data does not accurately represent the original data.[41] The first NIST competition solicited

concept papers from participants describing their methods and techniques.

The second iteration built upon the first phase by challenging the participants to apply their concepts in a sequence of short contests where their solutions were tested. The current iteration builds upon the successes of the previous competitions by asking participants to design algorithms that are capable of anonymizing data containing time and spatial information. The total potential prize is roughly $456,000. From these challenges, NIST was able to clearly benchmark competing differential privacy approaches against each other for the first time and established a measurement-based approach to fostering data-driven R&D in this area. This series highlighted approaches that will become the basis for future growth and innovation in this area.[42]

General Services Administration's End User License Agreement Challenge

In 2020, GSA's End User License Agreement Challenge sought AI/ML-enabled solutions to automatically review EULAs for terms and conditions that are unacceptable to the federal government and to shorten the average seven to 14 days it took for its contracting officers and parties to review, negotiate, and accept terms and conditions for contracts.[43] With a total prize purse of $20,000, the competition produced three winning solutions that GSA was able to license. GSA noted how the challenge illuminated ways to get AI/ML into the acquisition process to improve efficiency and effectiveness across business processes, and that the challenge offered insight into the types of commercial AI/ML tools available.

Though each of these four federal competitions had very different goals, their outcomes were similar. Each sponsoring agency reported gaining valuable insight into the tools and solutions available in the commercial sector and how to apply them to their specific missions. They were able to see how these tools performed in operational settings prior to any major procurement and acquisition decisions. These competitions also demonstrated that both large and small cash prizes, as well as nonmonetary

incentives, motivated participants. The prizes varied significantly, but all were able to solicit solutions, demonstrating utility or novelty, for their specific problems. The FDA, GSA, and navy all gained valuable insight and experience on the interplay between human operators and AI/ML-enabled tools that likely have implications for other AI/ML deployments in their agencies. The tools also reduced analytical burdens on human operators with potential increased productivity in these settings. For its part, NIST's differential privacy challenges helped guide the agency in establishing R&D benchmarks and measuring progress on a topic central to the application of AI in government settings. The benefits also would likely extend beyond NIST as federal agencies seek to develop and deploy ML systems while seeking to protect the privacy of citizen data.

## Using Competitions to Further National Artificial Intelligence R&D Goals

Competitions have demonstrated that progress and innovation in AI can be achieved through iterative challenges. Private sector competitions offer several useful examples that demonstrate how competitions have progressed innovation in AI. Beginning in 1994, the Critical Assessment of Techniques for Protein Structure Prediction (CASP) sought to advance and accelerate methods for identifying protein structures from amino acid sequences and is considered the "Olympics" of protein folding research.[44] A protein's function largely depends on its unique structure. Understanding its shape is important for understanding its function, which can lead to major advancements in science and medicine. Called the "protein folding problem," replicating this process in a laboratory is time-consuming and costly.

For CASP13, DeepMind's AlphaFold produced unprecedented progress in protein shape prediction, marking the first time that CASP organizers witnessed the effective application of AI.[45] AlphaFold applied machine learning techniques using neural networks to determine characteristics such as the distance between individual amino acids to predict protein shape. In CASP14, an improved AlphaFold2 showed even more progress, resulting in a 92.4 percent modeling accuracy.[46] Because of this scientific advancement, and arguably because of the existence the CASP series, AlphaFold is capable of producing high-quality predictions for every protein in the human body and for proteins in 20 additional organisms.[47] This significant AI-assisted contribution to scientific discovery is an example of the promise of AI competitions.

The discussion that follows considers how competitions can be used to advance national AI priorities as specified in a variety of federal documents.[48] It examines private sector competitions, and in tandem analyzes the innovation strategies and priorities of various federal organizations to identify areas for future competitions in the national security sector.

Safety and security of AI/ML-enabled systems

The need to ensure the safety and security of AI/ML-enabled systems is a top priority set out in many federal documents.[49] Machine learning systems are susceptible to a variety of relatively simple but significant attacks known as adversarial examples. In these examples, an attacker slightly modifies the input data to cause a misclassification. These types of attacks are typically extremely difficult to detect because, in many instances, the changes are imperceptible to the human eye.[50] Cognizant of these and related vulnerabilities, the National Security AI Commission recently found that the current testing, evaluation, verification, and validation processes are not sufficient at providing the necessary assurances for ML systems.[51]

The federal government could emulate CSAW's 2020 HackML challenge, which asked participants to develop innovative backdoor defenses for ML models and detections of adversarial attacks.[52] Although participation in this competition was limited to undergraduate and graduate students and awarded a relatively small prize purse of $2,250, the goal of the competition was directly applicable to the shared federal priority of robust and secure AI systems. HackML organizers provided a "backdoored" ML model, a model designed to purposefully misbehave when fed certain data, in which participants were required to mitigate unwanted behaviors. Similarly in 2017, Google Brain organized the Non-targeted Adversarial Attack challenge for the annual Neural Information Processing Systems conference for the purpose of accelerating research on adversarial examples by building robust systems.[53] Google Brain's challenge did not award a monetary prize. The main purpose of the two competitions was to accelerate innovation and understanding of critical security issues in ML.

A federal competition could take a similar approach but through a national security lens. Participants could be asked to find flaws or uncover vulnerabilities in a segment or copy of a federal AI/ML system, like the U.S. Air Force's 2020 Hack-a-Sat competition, in which participants who qualified for the final round were granted access to attempt to hack into a live on-orbit U.S. satellite. Although not an AI competition, it was seen as a way to bridge the

DOD and the security researcher community, to spur interest in the field of aerospace cybersecurity, and to gain a sense of satellite vulnerabilities.

Trust and Explainability

Robust safety and security measures are related to trust and explainability. With increased emphasis on accelerating the adoption and integration of AI technologies into federal infrastructure and operations, the need for understanding AI-enabled system decision-making, trusting those decisions, and guaranteeing its continued performance is a critical area of strategic R&D focus.[54] The development of methods for human-machine collaboration is a key priority for organizations intent on augmenting human capabilities.[55] With the exception of NIST's Differential Privacy Series, the previous examples of federal competitions demonstrate a continued need for understanding how human-machine collaboration would function in an operational capacity. The after-action analysis of each competition shows that human experience and intuition remains invaluable, while AI can help reduce the analytical burden on human operators freeing them to perform more cognitive intensive tasks. Likewise, as AI-enabled tools are incorporated into military and intelligence operations, decision makers need to understand how the system reached its conclusions and need to trust that process and the underlying system. Current research focuses largely on technological solutions and less on research and experimentation under operational conditions.[56]

The Innovare Advancement Center, in partnership with the Air Force Research Laboratory, the National Security Innovation Network, and other major private entities, facilitated the Trusted AI Challenge Series to advance four key areas of AI: Verification of Autonomous Systems, Trust and Joint Action for Digital Data Analysis, Dynamic Bi-Directional Trust in Human-AI Collaborative Systems, and Trustworthy AI Certification.[57] For its part, the AFRL helped to engage partners and tech startups and to initiate entrepreneurship in these strategic AI areas. The challenges were designed to gain insight into critical path requirements for building reliable and robust AI to safely operate in society.[58] In particular,

the Dynamic Bi-Directional Trust challenge sought novel ideas to build trust in human-AI collaboration. Successful white paper proposals received $10,000 at the time of award, $15,000 after justification of effort, and $50,000 at completion of an evaluation period.[59] This $75,000 prize package was a preliminary grant to conduct further development and prototyping of promising proposals.

Trusting AI/ML-enabled systems is not just limited to a human operator's confidence in the machine. Ultimately, trust and explainability is informed by the overall safety and security of these systems, which starts at the beginning of the development life cycle. ML models in particular are vulnerable at the start of their development where attacks on shared resources like ML libraries, pretrained models, and training datasets can be extremely difficult to detect.[60] Maintaining confidentiality, integrity, and accessibility has long been the gold standard in cybersecurity and the same applies for AI/ML systems. Competitions built around one or more of these objectives can serve as a means to assess current progress and uncover promising solutions.

It is not uncommon for prize challenges to first solicit white papers before challenging participants to actually develop the tool or a prototype. For AI fields like trust and explainability where measurements and concepts are somewhat abstract, white paper competitions may be just as useful as promoting innovation by simply encouraging researchers and practitioners to think about new ways to solve an agency's well-defined problem to later fund development or scaling or to determine how levels of trust can be built into a system. Depending on the goal, the agency would pay only for success, maturity, or feasibility. For example, in the navy's AI ATAC challenge series, teams were required to submit white papers which researchers at Oak Ridge National Laboratory's Cybersecurity Research Group and actual Navy Security Operations Center operators then evaluated. This style of competition allows sponsoring agencies to gain a sense of available technologies and to actually test and workshop the ideas or submissions.

Developing and Testing Reinforcement Learning Agents for Cyber Defense

As a final example, the DOD or DHS could sponsor a follow-up competition to the Cyber Grand Challenge by examining other automated cyber defense technologies. Machine learning already contributes much to cybersecurity, but some applications for cyber defense remain theoretical. In recent years, researchers have considered reinforcement learning agents (RL agents) as a potential game changer in cybersecurity. After an attacker makes their way into their intended victim's network, they often must spend considerable effort to understand the network's design, orient themselves, and then move laterally across the network to access desired files. Because the defender controls the environment that the attacker must work in, this is an advantage for the defense. However, using the landscape to the defender's advantage has proven difficult. Only organizations dedicated to extreme cybersecurity, like intelligence agencies, are able to build their networks with defense in mind.

RL agents for cyber defense would change, or perhaps level, the playing field. Researchers believe that RL agents, empowered to change network design and security policies, could disorient attackers and prevent them from reaching their objectives by changing the network landscape. So far, some researchers have designed algorithms capable of changing networks to prevent successful cyberattacks, but their research has proven that it is difficult to transfer to real networks.[61] The RL agents performed well in their training environments but poorly on other live networks. Moreover, like other AI systems, RL agents are vulnerable to adversarial attacks.

Researchers at MITRE and the National Security Agency published the *Framework for Advanced Reinforcement Learning for Autonomous Network Defense* in early 2021.[62] FARLAND is a training ground, canonically called a gym, for RL agents designed to defend computer networks. Researchers can train their algorithms in FARLAND while modifying the gym's settings, allowing the environment to simulate network attacks with common tactics, techniques, and procedures used by real

attackers. The federal government funded FARLAND's development, and could use the environment, either in its totality or as inspiration, for a series of competitions. Not only would this enable private sector entities with promising AI/ML solutions to train in an adaptable environment designed to simulate live networks, the shared testing space would undoubtedly strengthen private partnerships or professional networks.

A FARLAND Grand Challenge could spur innovation, new businesses, and advanced research that is required to advance RL agents for cyber defense. Besides proving capable of defending networks, the RL agents must be able to withstand adversarial attacks. The Cyber Grand Challenge moved automated vulnerability discovery from the lab to the field, the FARLAND Grand Challenge may do something similar for another promising technology. Like the Cyber Grand Challenge, the FARLAND series should end with a round of human versus machine competition. If RL agents pass muster for cyber defense, they could be deployed on networks across society, including military networks that have already been testing automated cyber defenses.

Public-private partnerships

Deepening and strengthening public-private partnerships in AI R&D remains a key priority for most federal agencies and departments, because much of the talent, expertise and current breakthroughs in ML are occurring in the private sector and academia. The national AI R&D strategic plan and NIST both call for shared public datasets, access to high-performance and cloud-computing resources, and training environments for technologies.[63] These resources are essential for researchers seeking to use actual operational data for modeling and experimentation to train real-world systems. If test beds and access to high-performance computing are limited, progress could suffer as AI development would be limited to a few well-resourced private sector entities.[64]

Data-sharing is also a priority. Machine learning models require significant amounts of data during training. Ideally large, representative datasets would be available for this training. The more robust the dataset, generally the more accurate the model.

However, the collection of curated datasets can be one of the most time-consuming parts of the ML production process. Due to the centrality of data for training AI models, any competition that advances the generation, collection, application and analysis of data can advance AI research and development.

If developed, agencies could utilize the test bed infrastructure in future competitions and reduce costs and avoid duplication in the process. Initial steps to develop these professional and academic partnerships for testing and training could include establishing networks with the recently created National Science Foundation AI Research Institutes.[65] These institutes cover a variety of topics intersecting with AI such as advanced cyber infrastructure and environmental science. These institutes are designed to act as connections in a broader nationwide network to pursue transformational advances in their respective fields through the application of AI.[66] Experimental methods, especially those being developed for national security purposes, could be continuously tested and validated at the corresponding NSF's AI Institute. In short, these institutes provide an ideal location for future AI competitions across many domains as well centers for testing and verification.

## Conclusion

The federal government can and should continue to leverage competitions to advance innovation in emerging technologies. To keep pace with rapid advances in AI, the scope and scale of these competitions should increase. Leveraging the promise of AI requires a recalibration of the innovation system and deeper relationships between the federal, private, and academic sectors. This will also require identifying and adopting the best commercial products and solutions at a speed that matches the pace of technological innovation. This report's findings, which follow below, suggest that prize competitions are an important and underused tool that the federal government can apply to facilitate innovation.

**Competitions offer a valuable tool to further agency missions.** When structured around a defined problem with transparent rules, competitions can help drive innovation. Ideally, departments and agencies should use competitions to drive progress on specific national security issues that are not already well-covered in the private sector. Competitions have several advantages over traditional processes. First, the federal sponsor typically pays only when participants successfully meet the objectives of the competition. Traditional acquisitions are more uncertain and typically require an upfront contract agreement for incremental financial installments. Second, removing bureaucratic impediments may incentivize a more diverse pool of participants, particularly small companies and academic institutions, who may be reluctant, unfamiliar with, or unsuccessful in entering into traditional government acquisition processes. Third, competitions can facilitate broader collaboration around a specific problem to encourage rapid or iterative innovation, as was the case with DARPA's autonomous vehicles challenge and NIST's facial recognition testing series. In addition, because competitions shift much of the risk of failure onto participants, this allows the sponsor to pursue more ambitious goals. Finally, competitions can boost promising research and prototypes.

Competitions do have limitations and are not meant to replace traditional R&D. They require personnel to organize and execute

the competition, which depending upon their size, can be non-trivial. Likewise, participation may be self-limiting, with more well-resourced organizations choosing not to compete. If no participant meets the competition's objectives, the return on this investment is limited. However, even if no participants meet the contest's goals, failures can still lead to new insights and eliminate impractical solutions. This may be especially true for more complex areas like trust in human-machine teaming and real-time operational performance.

**Competitions provide a means to test operational effectiveness prior to potential prototyping, scaling, or procurement.** When an agency runs a competition to solicit potential tools or applications, it can speed the acquisition and deployment of winning submissions by structuring the competition to satisfy procurement authority requirements in advance. Furthermore, challenges do not need to be complex or have significant prizes to achieve agency goals. In the GSA's EULA Challenge, which had a total prize purse of only $20,000, the administration structured the terms and conditions to receive a royalty-free license to use the winning team's tool. In competitions, awards are typically only administered for mature and functional solutions that meet the requirements of the challenge goals. This is the greatest difference between competitions and contracts. In contract awarding, payment is administered incrementally for work performed or a service as agreed upon by the signing parties.

Technologies developed in the private sector may not have considered the same requirements, such as safety and privacy, as those used in a federal context. Therefore, running a competition prior to procurement allows the sponsoring agency to gain insight into the available tools and technologies and to test their performance. Following its Open Data Challenge, the FDA noted it was a significant first step in its strategic implementation of AI because the agency was able to see how AI/ML-enabled tools perform. Furthermore, competitions allow agencies to shift the risk of failure onto the participants, meaning the agency may pursue more ambitious goals. Competitions could fit into this innovation ecosystem by testing private sector solutions against agency-defined operational parameters and acceptable metrics without

major financial loss. Competitions are best suited for the applied R&D of experimental, exploratory, or early-stage technologies, allowing their operational effectiveness to be measured or tested prior to integration and deployment.

**AI competitions are increasing and drive innovation in AI, but are still underutilized at the federal level.** Given federal priorities for AI R&D, competitions are an advantageous and additional method for sourcing and testing new technologies, as well as advancing innovation and understanding in various subfields of AI. Previous federal AI competitions demonstrate how the competitions helped sponsoring agencies gain valuable insight into how AI-enabled technologies could fit into their existing infrastructure and augment operations. Much like AlphaFold's unprecedented success in CASP13 and 14, a competition may highlight areas in which further R&D is required or areas in which unforeseen weaknesses are present. This is especially important for testing AI-enabled technologies for national security.

**The national R&D priorities for AI provides a roadmap for future competitions.** Various federal AI innovation strategies and priorities are focused on areas such as safety and security, as well as trust and explainability. There is an enhanced focus on developing and sustaining private partnerships because most of the talent, expertise, and current breakthroughs are occurring in the private and academic sectors. These categories are not static, but build upon one another to create a robust system. Sponsors of federal competitions would benefit from working directly with technology incubators within their departments like DOD's DIU and the DHHS IDEA Lab to streamline the procurement of potential solutions. In addition, the National Science Foundation's new AI Research Institutes offer a new locus for innovation across many domains and are well-positioned to leverage competitions to inform and progress their important work.

In conclusion, competitions are valuable yet underutilized methods for catalyzing innovation, sourcing technologies, and incentivizing the private sector to develop novel solutions for national security problems and adopting these solutions at the speed of relevance.

## Authors

Ali Crawford is a research analyst at CSET, where she works on the CyberAI Project.

Ido Wulkan is a former CSET research assistant who worked on the CyberAI Project.

## Acknowledgments

The authors are indebted to John Bansemer, Micah Musser, and Drew Lohn for thoughtful and constructive feedback on earlier versions of this paper, and to Dakota Cary for contributing to a section to this paper. We would also like to thank Joshua Powers, Jarah Meador, and Luciano Kay for providing additional comments and valuable insight.




Document Identifier: doi: 10.51593/2021CA002

## Endnotes

[1] General Services Administration, "Challenge.gov," accessed August 2, 2021, https://www.challenge.gov/.

[2] Office of Management and Budget, "Table 9.7: Summary of Outlays for the Conduct of Research and Development: 1949–2022 (In Current Dollars, as Percentages of Total Outlays, as Percentages of GDP, and in Constant (FY 2012) Dollars)," The White House, accessed June 8, 2021, https://www.whitehouse.gov/omb/historical-tables/.

[3] Stephen Roe et al., "Challenge-Based Acquisition: 5th Edition" (MITRE, March 2020), https://www.mitre.org/publications/technical-papers/challenge-based-acquisition-5th-edition.

[4] The White House Archives, "Artificial Intelligence for the American People."

[5] Marcy E. Gallo, "Federal Prize Competitions," Congressional Research Service, April 6, 2020, https://fas.org/sgp/crs/misc/R45271.pdf.

[6] Select Committee on Artificial Intelligence of the National Science and Technology Council, *The National Artificial Intelligence Research and Development Strategic Plan: 2019 Update* (Washington, DC: Executive Office of the President, June 2019), https://www.nitrd.gov/pubs/National-AI-RD-Strategy-2019.pdf; U.S. Department of Defense, *Summary of the 2018 Department of Defense Artificial Intelligence Strategy: Harnessing AI to Advance Our Security and Prosperity* (Washington, DC: Department of Defense, 2018), https://media.defense.gov/2019/Feb/12/2002088963/-1/-1/1/SUMMARY-OF-DOD-AI-STRATEGY.PDF; National Institute of Standards and Technology, *Plan for U.S. Leadership in AI: A Plan for Federal Engagement in Developing Technical Standards and Related Tools* (Washington, DC: Department of Commerce, 2019), https://www.nist.gov/system/files/documents/2019/08/10/ai_standards_fedengagement_plan_9aug2019.pdf; and Office of the Director of National Intelligence, *The AIM Initiative: A Strategy for Augmenting Intelligence Using Machines* (Washington, DC: Office of the DNI, 2019), https://www.dni.gov/files/ODNI/documents/AIM-Strategy.pdf.

[7] Office of the Director of National Intelligence, *The AIM Initiative*; National Institute of Standards and Technology, *Plan for U.S. Leadership in AI*; and Select Committee on Artificial Intelligence of the National Science and Technology Council, *The National Artificial Intelligence Research and Development Strategic Plan*, 4.

[8] ForAllSecure (*Pittsburgh, PA*), Deep Red (*Arlington, VA*), Shellphish (*Santa Barabara, CA*), TechX (*Ithaca, NY; Charlottesville, VA*), disekt (*Athens, GA*),

CSDS (*Moscow, ID*), and Code Jitsu (*Berkley, CA; Syracuse, NY; Lausanne, Switzerland*).

[9] Thanassis Avgerinos et al., "The Mayhem Cyber Reasoning System," in *IEEE Security & Privacy* 16, no. 2 (March/April 2018): 52-60, https://ieeexplore.ieee.org/abstract/document/8328972/.

[10] "Cyber Grand Challenge Qualifying Event," Lincoln Laboratory at the Massachusetts Institute of Technology, accessed February 23, 2021), https://www.ll.mit.edu/research-and-development/cyber-security-and-information-sciences/cyber-grand-challenge/qualifying.

[11] Avgerinos et al., "The Mayhem Cyber Reasoning System."

[12] "About Us," ForAllSecure, accessed February 19, 2021, https://forallsecure.com/about-us.

[13] DARPAtv, "DARPA's Cyber Grand Challenge: Early Highlights from the Competition," YouTube, August 5, 2016, https://youtu.be/WEDO2GgL20Q.

[14] Beethika Khan, Carol Robbins, and Abigail Okrent, "The State of U.S. Science and Engineering: U.S. R&D Performance and Funding," National Science Foundation | National Science Board, January 15, 2020, https://ncses.nsf.gov/pubs/nsb20201/u-s-r-d-performance-and-funding.

[15] Khan, Robbins, and Okrent, "The State of U.S. Science and Engineering: U.S. R&D Performance and Funding."

[16] Select Committee on Artificial Intelligence of the National Science and Technology Council, *The National Artificial Intelligence Research and Development Strategic Plan*; and National Security Commission on Artificial Intelligence, *Final Report* (Washington, DC: NSCAI, March 2021), https://www.nscai.gov/wp-content/uploads/2021/03/Full-Report-Digital-1.pdf.

[17] Gallo, "Federal Prize Competitions"; and Luciano Kay, "Managing Innovation Prizes in Government" (IBM Center for The Business of Government, 2011), http://www.businessofgovernment.org/report/managing-innovation-prizes-government; and Office of Science and Technology Policy, *Implementation of Federal Prize and Citizen Science Authority Fiscal Years 2017-18* (Washington, DC: Executive Office of the President of the United States, June 2019), https://trumpwhitehouse.archives.gov/wp-content/uploads/2019/06/Federal-Prize-and-Citizen-Science-Implementation-FY17-18-Report-June-2019.pdf.

[18] Gallo, "Federal Prize Competitions."

[19] Gallo, "Federal Prize Competitions"; Kay, "Managing Innovation Prizes in Government"; and Office of Science and Technology Policy, *Implementation of Federal Prize and Citizen Science Authority*.

[20] The COMPETES Act is complementary to existing authorities available to several federal departments and agencies including the Department of Defense (DOD), the Department of Energy (DOE), the National Aeronautics and Space Administration (NASA), the Department of Health and Human Services (DHHS), and the National Science Foundation (NSF).

[21] See https://www.challenge.gov/.

[22] Office of Management and Budget, "Table 9.7: Summary of Outlays for the Conduct of Research and Development: 1949–2022 (In Current Dollars, as Percentages of Total Outlays, as Percentages of GDP, and in Constant (FY 2012) Dollars)," The White House.

[23] National Artificial Intelligence Initiative Office, "Challenge Competitions," accessed September 27, 2021, https://www.ai.gov/category/challenge-competition/.

[24] Gallo, "Federal Prize Competitions."

[25] Sarah Z. Tang, Steven N. Rader, Peter Phillips, *Surprising Results from Large Crowds Using Micro-Purchase Challenges - Using Contests on Freelancing Communities to Source Innovative, Impactful and Cost-Effective Solutions* (Washington, DC: NASA, August 8, 2018), 5, https://www.nasa.gov/sites/default/files/atoms/files/surprising_results_from_large_crowds_using_micro-purchase_challenges.pdf.

[26] General Services Administration, "Advanced Microelectronics Design & Prototype Challenge," Challenge.gov, https://www.challenge.gov/challenge/advanced-microelectronics-design-prototype-challenge/.

[27] General Services Administration, "Challenge Toolkit: Frequently Asked Questions," Challenge.gov, accessed March 23, 2021), https://www.challenge.gov/toolkit/faq/.

[28] Defense Advanced Research Projects Agency, "The Grand Challenge," accessed March 24, 2021), https://www.darpa.mil/about-us/timeline/-grand-challenge-for-autonomous-vehicles.

[29] General Services Administration, "AHRQ Step Up App Challenge," Challenge.gov, https://www.challenge.gov/challenge/ahrq-step-up-app-challenge/.

[30] General Services Administration, "AHRQ Step Up App Challenge."

[31] General Services Administration, "Smart City Air Challenge," Challenge.gov, accessed February 16, 2021), https://www.challenge.gov/challenge/smart-city-air-challenge/#prize.

[32] Roe et al., "Challenge-based Acquisition."

[33] General Services Administration, "Army xTechSBIR," Challenge.gov, https://www.challenge.gov/challenge/army-xtech-sbir/.

[34] Roe et al., "Challenge-Based Acquisition."

[35] precisionFDA, "Gaining New Insights by Detecting Adverse Event Anomalies Using FDA Open Data," Food and Drug Administration, accessed June 22, 2021, https://precision.fda.gov/challenges/9.

[36] precisionFDA, "Gaining New Insights by Detecting Adverse Event Anomalies Using FDA Open Data."

[37] precisionFDA, "Gaining New Insights by Detecting Adverse Event Anomalies Using FDA Open Data."

[38] precisionFDA, "Gaining New Insights by Detecting Adverse Event Anomalies Using FDA Open Data."

[39] General Services Administration, "Artificial Intelligence Applications to Autonomous Cybersecurity (AI ATAC) Challenge," Challenge.gov, https://www.challenge.gov/challenge/artificial-intelligence-applications-to-autonomous-cybersecurity-challenge/; General Services Administration, "Network Detection of Adversarial Campaigns using Artificial Intelligence and Machine Learning," Challenge.gov, https://www.challenge.gov/challenge/network-detection-of-adversarial-campaigns/; and General Services Administration, "AI ATAC 3 Challenge: Efficiency & Effectiveness Afforded by Security Orchestration & Automated Response (SOAR) Capabilities," https://www.challenge.gov/challenge/AI-ATAC-3-challenge/.

[40] Kara McDermott, "Winners of Artificial Intelligence Challenge Announced," U.S. Navy, December 7, 2020, https://www.navy.mil/Press-Office/News-Stories/Article/2436651/winners-of-artificial-intelligence-challenge-announced/; and Kara McDermott, "NAVWAR Enterprise Launches Third Artificial Intelligence Prize Challenge in Series; Increases Award to $750K," Defense Visual Information Distribution Service, December 17, 2020), https://www.dvidshub.net/news/385243/navwar-enterprise-launches-third-artificial-intelligence-prize-challenge-series-increases-award-750k.

[41] National Institute of Standards and Technology, "2020 Differential Privacy Temporal Map Challenge," https://www.nist.gov/ctl/pscr/open-innovation-prize-challenges/current-and-upcoming-prize-challenges/2020-differential; and National Institute of Standards and Technology, "2018 Differential Privacy Synthetic Data Challenge," https://www.nist.gov/ctl/pscr/open-innovation-prize-challenges/past-prize-challenges/2018-differential-privacy-synthetic.

[42] "NIST Differential Privacy Synthetic Data Challenge," HeroX, accessed June 23, 2021), https://www.herox.com/Differential-Privacy-Synthetic-Data-Challenge/updates; and National Institute of Standards and Technology, "2020 Differential Privacy Temporal Map Challenge"; and National Institute of Standards and Technology, "2018 Differential Privacy Synthetic Data Challenge."

[43] General Services Administration, "GSA Artificial Intelligence and Machine Learning End-User License Agreement Challenge 2020," accessed June 23, 2021), https://www.challenge.gov/challenge/GSA-artificial-intelligence-AI-machine-learning-ML-challenge/.

[44] Andrew Senior et al., "AlphaFold: Using AI for scientific discovery," *DeepMind*, January 15, 2020, https://deepmind.com/blog/article/AlphaFold-Using-AI-for-scientific-discovery; and DeepMind, "AlphaFold: The making of a scientific breakthrough," YouTube, November 30, 2020, https://www.youtube.com/watch?v=gg7WjuFs8F4.

[45] YouTube. AlphaFold: The making of a scientific breakthrough. (30 November 2020). Retrieved from: https://www.youtube.com/watch?v=gg7WjuFs8F4

[46] DeepMind. AlphaFold: a solution to a 50-year-old grand challenge in biology. (30 November 2020). Retrieved from: https://deepmind.com/blog/article/alphafold-a-solution-to-a-50-year-old-grand-challenge-in-biology

[47] DeepMind. Putting the Power of AlphaFold into the World's Hands. (22 July 2021). Retrieved from: https://deepmind.com/blog/article/putting-the-power-of-alphafold-into-the-worlds-hands

[48] Select Committee on Artificial Intelligence of the National Science and Technology Council, *The National Artificial Intelligence Research and Development Strategic Plan*; U.S. Department of Defense, *Summary of the 2018 Department of Defense Artificial Intelligence Strategy*; National Institute of Standards and Technology, *Plan for U.S. Leadership in AI*; and Office of the Director of National Intelligence, *The AIM Initiative*.

[49] Select Committee on Artificial Intelligence of the National Science and Technology Council, *The National Artificial Intelligence Research and Development Strategic Plan*; U.S. Department of Defense, *Summary of the 2018 Department of Defense Artificial Intelligence Strategy*; and National Institute of Standards and Technology, *Plan for U.S. Leadership in AI*.

[50] Andrew Lohn, "Hacking AI: A Primer for Policymakers on Machine Learning Cybersecurity" (Center for Security and Emerging Technology, December 2020), https://doi.org/10.51593/2020CA006.

[51] National Commission on Artificial Intelligence, *Final Report*, Chapter 7, pp. 137.

[52] "CSAW'20 HackML," New York University Center for Cybersecurity, accessed July 15, 2021, https://wp.nyu.edu/csaw_hackml_2020/rules-and-regulations/.

[53] "NIPS 2017: Non-targeted Adversarial Attack: Imperceptibly transform images in ways that fool classification models," Kaggle, https://www.kaggle.com/c/nips-2017-non-targeted-adversarial-attack.

[54] Office of the Director of National Intelligence, *The AIM Initiative*; National Institute of Standards and Technology, *Plan for U.S. Leadership in AI*; Select Committee on Artificial Intelligence of the National Science and Technology Council, *The National Artificial Intelligence Research and Development Strategic Plan*, 4.

[55] Office of the Director of National Intelligence, *The AIM Initiative*; and National Institute of Standards and Technology, *Plan for U.S. Leadership in AI*.

[56] Margarita Konaev, Tina Huang, and Husanjot Chahal, "Trusted Partners: Human-Machine Teaming and the Future of Military AI" (Center for Security and Emerging Technology, February 2021), https://cset.georgetown.edu/publication/trusted-partners/.

[57] "Trusted AI Challenge Series," Innovare Advancement Center, accessed July 13, 2021), https://assets.website-files.com/5f47f05cf743023a854e9982/60885a15a1cfec5818fee5f4_TAI%20Topic%201%20RFP%20AFRL_AFOSR%20final%2026Apr.pdf.

[58] "Trusted AI Challenge Series."

[59] "Trusted AI Series: Dynamic Bi-Directional Trust in Human-Artificial Intelligence (AI) Collaborative Systems," Innovare Advancement Center, accessed July 14, 2021), https://assets.website-files.com/5f47f05cf743023a854e9982/60885a1587a12f5393a33f67_TAI%20Topic%203%20RFP%20NSIN_NYSTEC%20final%2026Apr.pdf.

[60] Andrew Lohn, "Poison in the Well: Securing the Shared Resources of Machine Learning" (Center for Security and Emerging Technology, June 2021), https://doi.org/10.51593/2020CA013.

[61] Lohn, "Poison in the Well."

[62] Lohn, "Poison in the Well."

[63] National Institute of Standards and Technology, *Plan for U.S. Leadership in AI*; and Select Committee on Artificial Intelligence of the National Science and

Technology Council, *The National Artificial Intelligence Research and Development Strategic Plan*.

[64] Select Committee on Artificial Intelligence of the National Science and Technology Council, *The National Artificial Intelligence Research and Development Strategic Plan*.

[65] National Science Foundation, "NSF partnerships expand National AI Research Institutes to 40 states," July 27, 2021, https://www.nsf.gov/news/news_summ.jsp?cntn_id=303176.

[66] National Science Foundation, "NSF partnerships expand National AI Research Institutes to 40 states."