SYM-AM-23-070

Excerpt from the Proceedings

of the
Twentieth Annual
Acquisition Research Symposium

**Acquisition Research:
Creating Synergy for Informed Change**

May 10–11, 2023

Published: April 30, 2023

ACQUISITION RESEARCH PROGRAM
DEPARTMENT OF DEFENSE MANAGEMENT
NAVAL POSTGRADUATE SCHOOL

ACQUISITION RESEARCH PROGRAM
DEPARTMENT OF DEFENSE MANAGEMENT
NAVAL POSTGRADUATE SCHOOL

Bridging the AI / Acquisition Divide: Why the Government Needs an Acquisition Revolution in the Coming Age of Artificial Intelligence

Tim Cooke, Ph.D.—is a veteran of the government professional services industry and an active contributor to better government organizations. As owner, President and CEO of ASI Government, LLC, he helps government clients leverage the marketplace through innovative outcome-focused acquisition to create mission essential capabilities. Prior to joining ASI, Cooke was a long-time executive at SRA International.

William (Will) Roe Roberts—is director of Acquisition Solutions for ASI, Government. He was most recently the Director of Acquisitions for the Joint Artificial Intelligence Center (JAIC), Office of the Department of Defense Chief Information Officer in Arlington, Virginia. As the first Acquisition Director for a DoD component focused on AI delivery, will was responsible for ensuring that acquisition pathways and methodologies are centered on the JAIC mission of providing AI capabilities to the warfighter and DoD workforce.

Michael Arendt, Ph.D.—is an executive-level consultant with proven technical expertise, domain experience, and business acumen to successfully transform commercial Artificial Intelligence (AI) and Machine Learning (ML) technologies into products, systems, and services for Government customers. Arendt has more than 18 years of experience in all facets of the business and program management life cycle, including identifying needs as well as defining program scope, resources, quality, costs, and ultimately the success/enablement of cutting-edge emerging technologies. Arendt received his PhD in Public Policy from the University of Maryland, College Park and has spent his career working across Government, Industry, and Academia.

Abstract

Artificial Intelligence technologies should be considered unique compared with the typical types of hardware and software solutions acquired by the Department of Defense (the Government). While at their heart, AI capabilities are indeed software, the journey required to build and deploy them successfully is very different. As a result, the Government must adapt its acquisition processes to support the AI development pipeline and include specific considerations for data acquisition, AI capability development, AI solution validation via test and evaluation, as well as ultimate deployment, adoption, and long-term refinement of the fielded AI capability. This research will seek to bridge the AI / acquisition divide by defining a detailed methodology to support execution of the AI acquisition life cycle.

Keywords: Artificial Intelligence, AI, Machine Learning, ML, Deep Learning, DL, acquisition, contracting, culture, technology, incentives, agile contracting

Introduction

The current revolution in conversational AI based on "foundation" or Large Language Models (LLM; e.g., GPT-3+) are capable of being adapted to a wide range of the Government missions and use cases. The disruptive change is rippling through the AI landscape like a hurricane tossing boats like toys. For many tasks, perhaps most, it reduces the need to build specialized systems and potentially replaces them with "prompt engineering" or the ability to interface with the LLMs based on knowledge of relevant domains. Such leverage of existing LLMs is an embryonic and rapidly developing endeavor which has made the need for developing new systems from scratch somewhat obsolete; replacing that work with the higher value, human work of aligning the technology to mission objectives and ensuring cultural values and trust are maintained. Many of the guiding lights and leaders of AI foresaw and started preparing for this day. In March 2023, they called for a

moratorium on developing more advanced models than GPT-4 out of caution for the potential to generate dysfunctional and dystopian possibilities.

This revolution in AI demands a revolutionary mindset and rethinking of the Government acquisition of AI capabilities and solutions. Considering the statements and policies of adversaries in China and Russia, the acquisition of AI is the new arms race essential to the next phase of ensuring American democracy. The enduring advantage of the American political experiment now lives in the people who lead our innovative technology industry guided by the aspirations of America's founders and its vibrant values. The acquisition workforce needs to create acquisition processes that attract industry professionals to bridge the gap between the Government and the AI technology community. The Government's ability to acquire AI in this brave new world is the highway to America's future.

This research will seek to bridge the AI / acquisition divide by defining an overall methodology that includes the core processes necessary for acquiring data, developing an AI capability, validation (i.e., test and evaluation) of the AI capability, as well as ultimate deployment, adoption, and refinement of the AI capability within the Government. The methodology will organize the discussion of each of these core processes by presenting insights in four distinct areas that we believe are necessary for successful AI capability delivery. These four areas include organizational culture supportive of the AI acquisition life cycle; AI-centric technical considerations for data acquisition, application development, and capability refinement; flexible contracting approaches suitable for the AI acquisition life cycle; and well-designed incentives to effectively motivate all parties involved throughout the AI acquisition life cycle.

Data Acquisition—The Lynchpin for AI Capability Development

Cultural Keys for Success in Data Acquisition

Data is the lifeblood of AI and should be treated as an incredibly valuable shared asset across the Government. Nobody within the Government individually "owns" data. Instead, there is an incredibly complex hierarchy of "Data Stewards" who reside within various military, civilian, and contractor roles. These Data Stewards are the gatekeepers for access to the lifeblood of AI. For the Government to successfully begin its transformational journey developing and deploying AI, it must clear the way for access to data. The Government must make a concerted effort to change the culture of those who serve as Data Stewards from one that is restrictive and risk-averse, to one that recognizes the untapped potential that the coordination, collaboration, and sharing of data can yield.

Gaining access to this data and being able to share it freely across the Government could potentially lead to groundbreaking improvements vital to Government missions such as warfighting, cybersecurity, supply-chain/logistics and military healthcare and more. For example, from the health perspective, most diseases lack readily accessible, validated data sets in which the "truth" is defined relatively easily. While this tends to be true in the civilian world, within the Government and Department of Veterans Affairs (VA) there is a plethora of data that can be mined to help enable the development of AI tools. By opening access to this data, it could enable the Government to forecast force readiness, develop personalized training regimens, and even anticipate and intervene prior to potential service-member injuries or disabilities. However, the promise of these types of future AI solutions can only be realized if the Government culture supports it. More specifically, obtaining validated data sets for these highly complex problems will require greater flexibility by Data Stewards and the development of tools that can interrogate electronic health records to identify and annotate cases representing specific diagnoses. To achieve this, a shift in how the culture

views this information must occur while ensuring protection for patient privacy and personal information. Anonymized data in health records from across the Government, the VA, and even the civilian world for our warfighters and veterans might thus have to be treated as precious resources of potential benefit to warfighter and veteran health, in much the same way as public utilities such as drinking water are currently treated (Pisano, 2020).

The Government must build a "data culture" internally and cultivate one externally with industry and academia. A data culture is one that recognizes the importance of data within the Government's day-to-day business processes and works diligently to harness its value. The bullets below provide some insights into the fundamentals that are needed for the Government to build a world-class data culture, which recognizes the importance of "ours" vs. "mine" to support AI capability development, deployment, and adoption:

- Leadership prioritizes and invests in data collection, management, and analysis/knowledge production. Leadership prioritizes creative data literacy for the whole organization, not just IT personnel or technical staff.
- Staff are encouraged and supported to access, combine, and derive insight from the organization's data.
- Staff recognize data when they see it. They offer creative ways to use the organization's data to solve problems, make decisions and tell stories. (Rahulbot et al., 2017)

Technical Keys for Success in AI Data Acquisition

Along with cultural issues, removing technical barriers to data access is essential for AI because access to large amounts of diverse and high-quality data is critical for the development and training of machine learning models. Machine learning algorithms rely on vast amounts of data to learn and improve, and the availability of such data sets is crucial for the development of accurate and effective models. Unfortunately, many organizations, particularly private sector companies who have been around for decades and the federal government, are in a state of disarray when it comes to data storage and organization. Technical barriers to data access can include challenges such as data fragmentation, inadequate metadata, data privacy and security concerns, lack of data interoperability, and limited access to data storage and computing resources.

Overcoming these technical barriers requires a coordinated effort from a range of stakeholders, including data providers, software developers, policymakers, and researchers. Some strategies for removing these barriers include developing standardized data formats and metadata, implementing data-sharing policies and platforms, robust data quality/data labeling/data curation, and improving data security and privacy protections through both policy and practice. By removing technical barriers to data access, we can ensure that AI technologies are developed and trained on diverse and representative data sets, leading to more accurate and effective models that can be applied to a wide range of real-world problems.

Contracting Considerations to Ensure Successful Data Acquisition

The backbone of any good AI data acquisition strategy will be found in the contracts that transform that vision into reality. While current contracts may be not ideally suited to support the level of data sharing and collaboration that is necessary to achieve wide-spread AI successes in the short run, the structure of future contracts for data are where the Government can make the largest contractual improvements to enable the deployment of AI across the Department. Below are two specific recommendations that the Government should consider employing to ensure contracts are structured to permit successful acquisition of data for AI.

Develop and Leverage Contractual Vehicles Specifically for AI Data Services.

The function of acquiring and accessing data has many tentacles in the life cycle process, but one of the key ingredients to being successful in this effort is to select a well-designed acquisition strategy that is tailored to facilitate data access for AI from the very beginning. To do this, the Government should develop multiple award contracts specifically for AI-relevant data services that give special consideration to:

- The composition of their competitive pool of vendors;
- The methods of defining the data acquisition requirements;
- Development of targeted evaluation criteria that attract strong AI Data Service provider solutions;
- Incentives that reward vendors for efficiently generating quality datasets that lead to impactful AI solutions.

The most important consideration is ensuring that the Government agency has a strong pool of vendors to solicit in their acquisition vehicle. These vendors must have the requisite technical expertise specifically related to data services for AI. In many cases, vendors who excel in this space may not be vendors who have experience in Federal contracting and are thus considered "non-traditional." This pool of vendors should be able to demonstrate their existing capabilities through more than a paper proposal alone. Such an approach levels the playing field for those without extensive past performance with the Federal Government. Moreover, as this technology is ever changing, the vendors within this pool should also be able to demonstrate that they have the capacity to collaborate with a wide variety of partners including other big businesses, small businesses, start-ups, academia, laboratories, or non-profit organizations. The next most important consideration for contract vehicles for AI Data Services is tied to how the Agency defines its requirements and structures its evaluation criteria to best solicit innovative solutions. It is not uncommon that industry and other acquisition partners often understand the Government's AI requirements better than the Government does. Thus, this approach should be employed whereby the requirements are presented in terms of a problem statement or a series of objectives for providing AI related data services to the Government. Such an approach maximizes vendor flexibility in proposed solutions and permits the leveraging of cutting-edge commercial technologies that may be used off the shelf and could be outside the norm of what the Government may typically use. To ensure these considerations are built into the contract, corresponding evaluation criteria must be used that focus on the matureness and those proposed industry solutions and how they may be leveraged specifically in the Government environment with minimal customization.

Instill Best Practices for AI Data Considerations into Other Contracts

Given the importance for data as part of building AI capabilities across the Department, the Government should ensure that best practices for data collection, curation, and sharing across stovepipes are built into future program requirements—even for those programs that don't think they have anything to do with AI. The reason for the breadth of this assertion is because the contracts that are created for programs across the Government ultimately become the brick and mortar of the data stovepipes that desperately need to be torn down. Terms and conditions are sometimes narrowly focused, over-prescribed, and may not have flexibility built into them for an ever-changing technological world. By building AI data considerations into the contracting process from the very beginning for all programs, it ensures that if/when data sharing needs for a future AI project are necessary, access to quality data can be quickly realized.

Data specific contractual requirements to support development of AI solutions can be treated as standard terms and conditions to enable AI much like existing requirements for cyber security or environmental considerations. These considerations must be included within every step of the contracting process starting with the creation of an IP/data right strategy at the beginning of any program. This strategy must then become enacted through the contractual process and any other agreements between parties. In addition, a data rights/IP risk assessment and cost assessment are recommended. Moreover, data for AI should be considered and discussed as part of any RFIs, RFPs, evaluation factors, source selections, and should ultimately work their way directly into the contract line items, work statements, contract clauses, contract data requirements lists, data rights assertion provisions, instructions, conditions, and notices to offerors as well as any quality assurance surveillance plans or related evaluations by Contracting Officers Representatives (CORs) during the contract period of performance.

AI Capability Development—Where the Rubber Meets the Road

Cultural Keys for Success in AI Capability Development

Everyone Can Develop AI Capabilities—No Degree Required!

One way to ensure that these embedded end users can make an immediate impact is by recruiting them to participate in the AI capability development process itself versus simply advising the development team as a subject matter expert. There are several existing examples where this is already being done across the Federal Government. In fact, in the recent report entitled *Government by Algorithm*, a widespread survey was conducted across Federal Civilian agencies to assess the current use of AI solutions. Of the 157 identified instances of civilian agencies using AI to augment their operations, some 84 or 53% of those solutions were developed in-house by current civilian employees in lieu being outsourced to contractors as shown in the figure below (Engstrom et al., 2020).
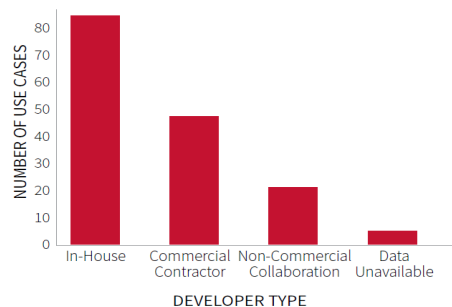

Figure 1. AI Use Cases by Developer Type.
(Engstrom et al., 2020).

One shining example of this approach is at the Social Security Administration's Office of Appellate Operations (OAO; Engstrom et al., 2020). Gerald Ray, who spent most of his career at SSA, and served as Administrative Appeals Judge and then deputy executive director of the OAO. During his time at OAO, he led the successful development of prototype AI tools by civilian personnel within his office. Described by one co-worker as the "Steve Jobs of the SSA," Ray realized that AI tools could be used to automate business processes which could be incredibly helpful to support personnel engaged in the SSA adjudication process. Because the OAO was unable to hire outside contractors and was only authorized to employ attorneys, Ray identified attorneys within the organization who had an existing background data analysis and software engineering. These lawyers

ultimately became the became the core team that built out SSA's early AI prototypes that were used to improve the SSA adjudication processes within the OAO. (Engstrom et al., 2020)

Ultimately, AI will have to evolve into a low-code, no-code form that is like most common business applications. This would eliminate the need to have specifically skilled AI experts involved in development (Petrocelli, 2018). Further, the emerging ecosystem, consisting of marketplaces for data, algorithms and computing infrastructure, will also make it easier for developers to pick up relevant AI skills. The net result will be lower costs to train and hire talent. The above two factors will be particularly powerful in vertical (industry-specific) use cases such as weather forecasting, healthcare/disease diagnostics, drug discovery, and financial risk assessment that have been traditionally cost prohibitive (Hosanager, 2017).

As a result of these factors, the Government should take advantage of AI development platforms that minimize the need for highly technical training and minimize the gap between warfighter and AI capability developer. Lastly, instead of re-creating a new training program, the Government should leverage as many existing training programs as possible via a series of partnerships with these organizations to streamline access for warfighters and civilians.

Collapse the AI Capability Development Pipeline

When an AI solution is developed in a stovepipe or vacuum without direct end-user involvement, the likelihood that it will be adopted successfully in the field is significantly diminished (Fountaine et al, 2020). Thus, one key to development of meaningful AI capabilities is ensuring that end-users are fully integrated as an active part of the product development team for an AI capability. For some organizations, this could be a potential radical shift in the technology development process where historically there may have been several bureaucratic layers between those who write requirements, develop technological solutions, those who test and evaluate solutions, and those who finally get an opportunity to field test them prior to limited or full rate production.

Thus, for successful AI capability development, the Government should seek to collapse the capability development chain by integrating all core players into an AI capability development product team. Collapsing the capability development chain means that the Government should reduce the distance between those ultimate end users who are in the field and those who develop AI/ML capabilities that can become impactful across a range of use cases. The impact of this approach will be ensuring AI/ML solutions are applicable to the real-world, usable by operators, who can provide meaningful insights prior to initial field testing.

Technical Keys for Success in AI Capability Development

Developing strong AI capabilities requires a marriage of technical expertise and thoughtful strategic planning. Here are some key technical factors that can contribute to success in AI capability development:

- **Understand the Problem You are Solving**. Before pitching a solution, make sure you really understand the use case and the challenges. Talk to experts and the potential customers you want to serve. Understand what it is like to go through their day. Until you get this right, the next steps will not matter because you will get a faulty solution that does not solve the problem.

- **Choose the Right Algorithms**. Choosing the right algorithms is essential for building effective AI models. You need to have a good understanding of the

strengths and weaknesses of different algorithms and be able to select the most appropriate ones for your specific use case. AI algorithms are used to take this data and turn it into something useful that can serve to automate processes, personalize experiences, and make complex predictions. Don't be afraid to use a commercial off the shelf solution (COTS). There is no need to reinvent the wheel when a solution already exists. Incorporating a hybrid approach also works well by taking a COTS solution and customizing it for your needs.

- **Don't Forget Infrastructure and Tools.** Building and training AI models requires significant computing power and specialized tools. You need to have a robust infrastructure in place, including high-performance computing resources and the right software tools. While there are many different types of software tools available for use, not all tools provide the same ease of use and accessibility. Tensorflow and Pytorch are two examples of very popular tools which have a plethora of online training resources available to support AI algorithm development.

- **Don't Be Afraid to Experiment and Iterate.** Developing AI capabilities is an iterative process that involves continuous experimentation and refinement. You need to have a culture of experimentation and be willing to try out different approaches, learn from your mistakes, and make improvements. Experimentation here means augmenting the insights from the data you have and planning business processes on a test and learn a basis to see how they respond. The process of experimentation can reap multiple rewards for businesses, considering how they will find themselves in a better position to continue with a given strategy if it proves to be successful.

- **Ensure You Have Domain Expertise.** To build effective AI models, you need to have a deep understanding of the domain you're working in. This includes not only technical knowledge of machine learning and data science, but also a strong understanding of the business context and the needs of your stakeholders.

Contracting Considerations to Ensure Successful AI Capability Development

The innovative and evolving nature of AI development requires an acquisition approach that is similarly innovative and evolving. The use of traditional contracting structures doesn't work well for AI development efforts for several reasons. First, traditional FAR-based acquisition approaches tend to encourage participation from traditional government contractors that specialize in the government over the technology. Second, traditional application of the FAR does not adequately address the need for agile contracting teams to instill fluidity and agility in the process. This kind of flexibility is essential for procuring and delivering emerging technology, and particularly for buying and adopting AI.

It is worth noting that "being traditional" is not equivalent to "using the FAR." For those with knowledge and skill, the FAR provides enough flexibility to engage in meaningful agile contracting. In the book and website *Agile Government Contracting,* the authors emphasize this point, and further show that *mature* agile teams will work within the regulatory parameters of federal procurement. Agile teams that ignore these parameters are just as ineffective as traditional contracting professionals that rely on outdated processes for rule compliance. There must be a balance to achieve true success and successful AI adoption (*Agile Government Contracting*, n.d.).

There are three major contracting areas that are unique to AI and that, in our experience, greatly increase the chances of successful AI adoption: (1) a meaningful Intellectual Property (IP) Strategy; (2) effective language on responsible use of AI (RAI); and (3) Agile contract structure and performance metrics centered on the Way of Working

(WoW) during contract performance and end-user value. This paper will cover these three areas in more detail in the following sections.

In some cases, it may make sense to consider the use of non-FAR-based acquisition authorities such as Other Transactions Authority (OTA). This approach offers a variety of methods to collaborate with innovative and emerging leaders in the AI arena. An example of this method is found in CDAO's Tradewind OTA Consortium vehicle. Tradewind is available for all DoD customers and the goal of the vehicle is two-fold: (1) provide a means to effectively procure and deliver AI from the best companies; and (2) help DoD customers learn the benefits of agile contracting so that they can start to practice this technique for future AI buys (whether they continue to use Tradewind or not). This second objective from CDAO emphasizes the realization that, right now, it is most beneficial for DoD agencies to help each other "learn" these new processes and take on a servant-type role of enabling the success of other agencies, rather than promoting oneself as the procurement expert organization for AI. More emerging tech-focused agencies should follow this example.

AI Validation (T&E)—Fingers Crossed, This Thing Actually Works!

Deploying AI capabilities requires rigorous testing and evaluation to ensure that they are performing as expected. Therefore, a comprehensive testing and evaluation plan must be in place to assess the effectiveness of the AI capabilities. The testing should cover a wide range of scenarios, including adverse scenarios, to identify any vulnerabilities or weaknesses in the system. Moreover, testing should be done continuously to identify any issues and address them promptly. The tests should focus on data and model quality, accuracy of predicted results, usability, and identify any future datasets that would improve the quality of the AI generated results.

Cultural Keys for Success in Validation of an AI Capability

"Field to Learn"—Don't Make Perfect the Enemy of the Good in AI/ML

The Government should not allow the perfect be the enemy of the good when it comes to development and fielding of AI/ML capabilities in the near-term. The culture must shift from one that puts speed to fielding first, even if the current iteration of a solution is less than ideal. In short, for AI and ML technologies to be adopted and improved, they must get into the hands of users as quickly as possible with the understanding that there will be a pipeline in place to push feedback for real-time or near-real-time refinements to improve functional capabilities and usability as additional data is collected. While on one hand, AI has shown that it can translate speech, diagnose cancer, and beat humans at poker. On the other hand, there have been missteps along the way that imply the technology may still be a long way off from full maturity. "For example, image recognition algorithms can now distinguish dog breeds better than you can, yet they sometimes mistake a chihuahua for a blueberry muffin. AIs can play classic Atari video games such as *Space Invaders* with superhuman skill, but when you remove all the aliens but one, the AI falters inexplicably" (Hutson, 2018). Such mistakes are indeed inevitable as nothing can fully replicate the fielding of a solution in a real-world environment. The important factor here is to ensure that the early fielding is done in a safe and responsible manner to ensure there are no serious impacts to the mission.

Technical Keys for Success in Validation of an AI Capability

As AI becomes more sophisticated and ubiquitous, it is critical that we ensure that it is used ethically, securely, and with trust. In this section, we will explore the technical keys for success in validating an AI capability that incorporates security, ethics, and trust. Below are four key technical considerations for the validation of an AI capability.

Understand the Data

The first technical key to success in validating an AI capability is to understand the data that will be used to train the AI algorithm. Data is the foundation of any AI system, and it is essential to ensure that the data is accurate, relevant, and representative of the real-world scenarios that the AI system will be deployed in.

In addition, it is crucial to understand the biases that may exist in the data. Biases in the data can lead to biased AI algorithms, which can have serious consequences. Therefore, it is essential to analyze the data for any biases and take steps to mitigate them. One example highlighting this problem is facial recognition technology. Facial recognition has been shown to have discrepancies in accuracy with females and minorities (SITNFlash, 2020). One of the problems is that early facial recognition programs were trained on friends and family members of the founders, and the dataset was mostly faces of white males. Fixing the training data by using more diverse groups of people to train the model, goes a long way to solving this problem, but it remains to be seen as to whether this action by itself resolves the problem. If a company or organization uses facial recognition for security, it is imperative to keep in mind the flaws and that any action taken as a result of a facial recognition process should be done with the utmost professionalism and care and should rely on other analytic methods.

Understand the Algorithm

The second technical key to success is to choose the right algorithm for the AI system. There are numerous AI algorithms available, and each has its strengths and weaknesses. Some algorithms may be better suited for certain tasks than others. It is also essential to understand the limitations of the chosen algorithm. No AI algorithm is perfect, and it is important to understand what the algorithm is capable of and what its limitations are. Understanding the limitations of the algorithm will help to manage expectations and ensure that the AI system is deployed in a manner that is safe, ethical, and trustworthy. This kind of understanding is also important for making sure your models are relevant. The world is constantly changing. If you understand the capabilities and limitations of your model, you are quicker to react and modify the model accordingly if a new dynamic occurs forcing you to change. As you validate the AI model, you should have a team separate from the development team to validate the model.

Ensure Security

The third technical key to success is to ensure the security of the AI system. Security is critical in any system, but it is especially important in an AI system that is handling sensitive data. Ensuring the security of the AI system involves taking steps to prevent unauthorized access, safeguarding against cyber-attacks, and ensuring that the data is stored securely.

Incorporate Ethical Considerations

The fourth technical key to success is to incorporate ethical considerations into the design and development of the AI system. Ethical considerations are critical because AI systems can have far-reaching consequences, both positive and negative. Incorporating ethical considerations involves ensuring that the AI system is designed to promote fairness, transparency, and accountability. AI development also requires an awareness of ethical considerations, such as privacy, bias, and fairness. You need to design and develop AI models that are ethical and responsible, and that respect the rights and interests of all stakeholders. Organizational leaders should constantly assess the following criteria:

1. Safe and secure

2. Private
3. Responsible
4. Robust and reliable
5. Transparent and explainable
6. Accountable
7. Fair and impartial

By focusing on these technical factors, you can build strong AI capabilities that deliver value to your organization and stakeholders.

Build Trust

The final technical key to success is to build trust in the AI system. Trust is critical because it determines how users perceive and interact with the AI system. Building trust involves designing the AI system to be transparent, explainable, and accountable. It also involves ensuring that the AI system is used in a responsible and ethical manner. Education is another key aspect to building trust.

Validating an AI capability that incorporates security, ethics, and trust requires technical expertise and a deep understanding of the data, algorithms, and ethical considerations involved. Ensuring the security of the AI system, choosing the right algorithm, incorporating ethical considerations, and building trust are all critical technical keys to success. By incorporating these principals, we can ensure that AI is used in a responsible, ethical, and trustworthy manner, ultimately benefiting society as a whole. Finally, you must always assess everything from your data, modeling, use cases, and even ensuring your workforce is cognizant of not just how the AI process works, but understand the ethics and trust issues which often lag behind technical development.

Contracting Considerations to Ensure Successful Validation of an AI Capability

The Importance of Third Party Neutrality to Establish "Trust"

In many cases, the developer of the AI model (or the large prime contractor of the project if the Government set up their contract to rely on a prime integrator) will offer services to test the reliability of their own models. In some cases, such as automated testing, this will be the most logical option. However, in most cases it will be essential for the Government to arrange a separate contract for third party testing. This establishes neutrality which furthers overall trust in the model. In such cases, it may be beneficial to arrange for a small multiple award ordering vehicle among a selected pool of trusted third-party AI testers. The vehicle will not only allow for faster repetitive orders but establish the basic parameters of trustworthy AI at the base contract level. If performed with skill, the CO can include terms and conditions in the base contracts which establish predictability in performance in all vendors. This will, in turn, begin to solidify a reliable "way of work" among the agile teams as they bring the testing companies into the overall process.

Testing for "Trust" in Responsible Use of AI (RAI)

There are two primary areas of "trust" needed in order for end users to become early adopters of AI technology for the benefit of government missions: (1) Trust in the working functionality of the technology to enhance the user's job and mission; and (2) Trust in the responsible use of the technology to prevent harm to innocent life and bias in outcomes. Although the first area of trust has been the primary focus of Test and Evaluation (T&E) activities, testers must address RAI with the same level of scrutiny. In this sense, AI provides a unique challenge, and COs must be cognizant of the need to imbed RAI considerations in the performance metrics. As with almost every element in AI acquisition, the contract is the mechanism in which the parties mutually agree on how RAI is addressed,

and this places importance on the role of the CO. The two most challenging but equally impactful areas to address RAI in the contracting process are (1) at the source selection stage as a discriminator for award selection; and (2) at the T&E stage as a metric for contractor performance. In both areas, the CO must navigate through the dangers of ambiguity in defining what "responsible" and "trustworthy" look like. Challenging as it may be, it is essential for the success of the end users' AI adoption. The Defense Innovation Unit (DIU), the Air Force MIT Accelerator, and CDAO all provide guidance that can help the acquisition team select the appropriate contract language that allows for testing AI trustworthiness.

Deployment, Adoption, and Refinement—All Aboard the AI Bullet Train

Cultural Keys for Success in Deployment, Adoption, and Refinement of an AI Capability Across the Government

IT systems have made it possible to give senior commanders instantaneous access to information across the globe. While that can provide a strategic advantage, it also leads to micromanagement and tighter measures of control. When it comes to AI, having tools at the command center is crucial, but there must be a deliberate effort to push technology to the edge. Simply put, giving each individual rifleman the AI tools needed to plan and execute missions, anticipate the enemy, and provide force protection is essential to success on future battlefields. It is also critical not to overlook critical support functions such as predicting maintenance issues, assisting with logistics, or using AI to assess the wellbeing of our military personnel.

Build a Sense of Urgency

As noted in many articles over the last few years, the Government has to take concrete steps to improve the acquisition process so that military personnel have the best resources at their disposal. The entire process from requirement identification to signing the contract must move with urgency. Far too often, companies that try to help the government end up shutting their doors because they have run out of cash and their solutions were so unique that there was not a private sector counterpart. Without fixing this key aspect, many companies may shy away from working with the government. Contrast that with the war in Ukraine and how the Ukrainians have developed unique technical solutions at scale. After all, as noted by former Google CEO Eric Schmidt innovation power is the decisive form of power for world dominance. If the government cannot speed up acquisition, it will not be able to innovate effectively, causing the United States to fall behind as an innovative power (Schmidt, 2023).

Ensure Effective Change Management

Deploying AI capabilities across the Department of Defense requires effective change management processes to ensure that all stakeholders are on board with the deployment plan. This includes communicating the benefits of AI capabilities, training personnel, and addressing any concerns or objections that may arise. Additionally, it is essential to have a clear plan for managing the transition to AI capabilities, including the necessary infrastructure changes and modifications to existing processes.

Technical Keys for Success in Deployment, Adoption, and Refinement of an AI Capability Across the Government

AI is transforming the Government by improving decision-making, increasing efficiency, and enhancing operational effectiveness. However, deploying, adopting, and refining an AI capability across the Government requires a unique set of technical skills and expertise.

Develop a Comprehensive Plan

The first technical key to success is to develop a comprehensive plan for the deployment, adoption, and refinement of the AI capability based on mission need. This plan should include a clear and concise description of the AI capability, the benefits it provides, and how it will be deployed and adopted across the Government. It is essential to identify key tasks and the person assigned to those tasks. The plan should also include a roadmap for the refinement and improvement of the AI capability over time. This roadmap should identify the technical challenges and risks associated with the deployment and adoption of the AI capability, as well as the strategies that will be used to mitigate these risks. Once implemented, the Government will have to constantly assess and modify technology and resources to adapt to new trends. For example, the arrival of ChatGPT already threatens to place previous AI solutions into obsolescence. ChatGPT is so powerful that recently, several tech insiders have written a letter calling for a pause of the development of such models (Blake, 2023).

Robust Data Management

The quality of data is a critical factor in the development of AI capabilities. Therefore, it is essential to have a robust data management system in place. The system should be able to store, manage, and protect sensitive data from unauthorized access. Additionally, it should be able to handle large volumes of data and provide easy access to relevant data for analysis. A robust data management system ensures that AI capabilities are developed based on accurate and relevant data, leading to better outcomes.

Addressing Technical Challenges

The second technical key to success is to address the technical challenges associated with the deployment and adoption of the AI capability. The deployment and adoption of AI across the Government involves a complex set of technical challenges, including data integration, algorithm development, and system interoperability. It is important to identify and address these technical challenges early and ensure that you not only have the necessary technical infrastructure and capabilities, but also the key technical experts across the entire enterprise. Military personnel, ranging from the most senior general to the lowest ranking private have a role. After all, they make the decisions and execute military orders. They need to articulate what they need, address any modifications that need to be made to any AI solution, and validate that the solutions do indeed make their job easier and that they are rapidly deployable.

Strong AI Infrastructure

Deploying and operating AI capabilities requires a robust infrastructure that includes hardware, software, and networks. This infrastructure must be scalable, reliable, and secure. Furthermore, it must be designed to handle the complex computations required for AI applications. Therefore, investing in modern hardware and software technologies is critical to ensuring a successful deployment, adoption, and refinement of AI capabilities. The infrastructure must allow data to flow up and down the chain of command to not only provide the best AI modeling, but to allow full spectrum analysis of the situation in order for leaders to make the best battlefield. decisions.

Strong Cybersecurity Measures

Deploying AI capabilities across the Department of Defense requires robust cybersecurity measures to protect sensitive data and systems from cyber threats. This includes implementing strong access controls, encryption, and network segmentation to prevent unauthorized access. Additionally, it is essential to conduct regular security audits

and implement continuous monitoring to detect and respond to any potential security threats.

Documentation is Key!

With a constant turnover, documentation is key. Oftentimes, this is one of the most overlooked aspects of an AI program, and indeed, just about any IT program. One of the reasons data storage and model building is somewhat disjointed is because of turnover and the lack of documentation.

In conclusion, deploying, adopting, and refining AI capabilities across the Department of Defense requires a robust technical foundation that includes robust data management, strong AI infrastructure, robust testing and evaluation, strong cybersecurity measures, and effective change management. With these key technical factors in place, the Department of Defense can develop and deploy AI capabilities that are effective, efficient, and secure.

Contracting Considerations to Ensure Successful Deployment, Adoption, and Refinement of AI Capabilities Across the Government

There is much to address in contracting for successful AI adoption, and this paper is not able to delve into all of these concerns in detail. As a foundational statement, it should be noted that successful deployment, adoption, and refinement are the most neglected areas in AI procurement. Consequently, prototypes die in the "valley of death," end users never experience or enjoy the impact that AI can have in their jobs, and the technology adoption by foreign adversaries continues to be a very real and disconcerting threat. The acquisition professionals hold the key to changing this reality, and contracting officers must adjust their practices in order to reverse the circumstances. The good news is that such changes are realistic and achievable, they just need to be accepted and practiced. It should also be noted that this is not a "top-down" problem, and real change can and should occur at the operational and tactical grassroots level (a.k.a., the acquisition team).

Strategizing Intellectual Property to Maximize Market Participation

An area in contracting that directly impacts the deployment and sustainment of an AI model is the topic of intellectual property rights. When it comes to AI, IP strategies take a slightly more complex approach. At the very least, proper strategies on IP require an adequate level of understanding and knowledge of the underlying technology. For example, the rights to the data, the model, the platform and the infrastructure may all have separate strategies to ensure the appropriate government ownership or use rights. When it comes to the model itself, the government should work to accept commercial terms and customary rights to leverage more industry participants. The infrastructure and platforms, however, may need more government control and ownership as any vendor-lock on large platforms may inhibit the government's ability to compete among various models that may connect to the platform via API. Finally, data rights offer unique considerations and will typically be case by case. The IP Cadre, set up by OUSD A&S, provides very good guidance for acquisition professionals, particularly on setting up innovative IP strategies for the procurement of emerging technologies.

Agile Contracting for Value-Added End User Adoption

While the term "agile" is more commonly associated with the development and deployment of software, there is a need for contracting teams to engage in flexible contracting that is iterative and nimble. It is also worth mentioning that agile contracting has always been a method that the FAR has recommended for large IT procurements. FAR 39.103 describes the process of modular contracting, which provide the necessary agility to

pivot as circumstances change throughout the development and delivery phases of AI acquisition.

Incentives for the AI Ecosystem—Carrots and Sticks for Successful AI Development and Adoption

In this final section, this paper will address the very relevant and oft neglected area of meaningful incentives that bring about true AI adoption success. The importance of this subject area warrants a special attention and therefore a separate section. The incentive principles discussed in this section should be applied in all stages of the AI life cycle discussed above. Incentives need to be designed to help understand the dynamic and uncertain goals and objectives of the Government use cases that embed AI technologies as part of the solution. The real explainability challenging the Government may be less about the internal technology and more about of inherent unsureness in the problem statement, challenge, or statement of objectives. Solving successfully for the wrong objective is a major challenge for AI systems which must be addressed by iterative feedback, testing, and evaluation to ensure fidelity with the intended objectives. If pivots or adjustments are identified then changing the system to address the new direction should be encouraged, not discouraged. This, again, emphasizes a common theme exhibited throughout this paper: the core challenges of explainability, trust, and successful AI adoption are primarily centered on the way the Government executes an AI acquisition through diligent contracting and product management.

As always, the behavior of people as they respond to their incentives is critical to achieving these acquisition improvements. The non-financial and financial incentives that may be applied in a contract setting will be important determinants of overall project success. Furthermore, thoughtful incentives for AI projects will help bridge the much publicized "gap" between Government missions and private industry by providing real and meaningful benefits for high tech companies. Financial and non-financial rewards for the contractor project team can be tied to model and mission outcomes. Reliability, robustness, efficiency, and value for money metrics in best value procurements can be factored in the project team incentives. Criteria should be designed to be easy to measure and be one of the outputs of doing the work, not requiring separate effort.

Type	Organization	Individual
Financial	Incentive fee, SLA based, quantitative award fee or withhold based on subjective evaluation Award term additional period of performance	Bonus Promotion
Non- financial	Team Awards & Recognition CPARS, Past performance Intellectual property Customer Experience Mission impact & Reputation	Teamwork & Partnership Recognition & Reputation Learning new skills Management Excellence Flexible work hours & location

Figure 2. Traditional Performance Incentives for AI Solution Contractors

The performance incentives should include the following: a) clear and measurable performance objectives; b) financial and non-financial performance incentives and penalties; c) regular monitoring and evaluation of contractor performance; and d) flexibility for contractors to adapt their approach based on performance feedback. In addition, rewards for individual functional specialists should be based in part on the team's overall performance to ensure alignment of effort.

As in any third-party effort or contract, the design of incentives should consider the Principal-Agent relationship between the Government, the Principal, and the performing contractor, the Agent. A divergence between their goals and objectives creates the potential for inefficiency and rent-seeking behaviors.

Non-Monetary Incentives

Managers and employees of solution delivery firms are expected by their company to behave in ways that increase the value of the enterprise. Significant determinants of company value include top-line revenue growth and bottom-line earnings. The successful companies in this market are the IT and professional services firms that have invested in the knowledge of how to work with the Government to overcome the many barriers to entry and business model requirements to be successful. However, critical mission success is dependent on diversifying the market participants, and therefore it is necessary to discuss how the Government can first incentivize new technology players in the acquisition processes before describing incentives for successful contract performance.

To incentivize new entrants, the Government should engage in selective cost-sharing. This will assist in building industry capability to meet government specific requirements like FedRamp. Defraying the costs of the Government-specific regulations would encourage more businesses to be willing to forego other business opportunities as the Government accommodates the business needs of industry for things like timely cash flow and reduced costs surrounding regulation compliance. For example, Other Transaction Agreements (OTAs) are, by design, a contractual mechanism to reduce barriers to entry because they are not subject to specific regulations of the FAR. The Government's new Office of Strategic Capital should make this their primary mission and goal, but any acquisition professionals can execute these incentives through thoughtful structuring of the contracts and agreements.

Traditional Government suppliers and their employees often share many of the attributes of their government clients, most importantly a dedication to the mission and its success. Such firms would be particularly sensitive to reputation-based incentives as follows:

- Public recognition: The Government publicly acknowledges the achievements of high-performing contractors, increasing their visibility and credibility in the market.
- Preferred contractor status: The Government grants preferred status to contractors that have demonstrated exceptional performance, making it easier for them to secure future contracts.

A company's excellent reputation can lead to more business opportunities and higher financial returns.

Government Incentives

Since the purpose of examining incentives in the Government AI acquisition is to ensure mission success, it is necessary to consider the incentives of all team members within an AI project, to include government professionals. To neglect the personal motivations of such key members is to ignore the impact that a high-morale government

workforce has on an office's successful technology adoption. Government program managers are the first line of accountability for the performance of AI systems. They respond to the objectives, constraints, and incentives of their organization and in their careers. Financial incentives such as bonuses for government officials are limited by law, and not usually major incentives, either in value or influence.

Instead, the non-financial incentive of government contracts mission success is often the dominant incentive. Government managers have the same kind of devotion to the mission. Government buying teams have the desire to excel, but only if management fosters this desire and provides the rewarding environment. Team members should be measured on common program goals, not the goals of their specialist silos. They all need to transparently see measurable improvements in mission performance, new mission capabilities, and improved service levels, and to be recognized for those improvements. Program performance dashboards should be made easily accessible to all members of the Acquisition team. Team cohesion ensures mission focus, which will consequently transfer to cohesion with the contractor, who will merely be brought into the well-functioning team.

Monetary Incentives

Having addressed various non-monetary incentives necessary for overall program success, it is important to examine the alignment of contract incentives with meaningful performance metrics based on actual end user value. This will, in turn, ensure that monetary awards are properly aligned with progressive delivery and adoption of the AI technology.

Milestone-Based Payments

The Government releases funds to the contractor based on the successful completion of predefined milestones. This approach ensures that the contractor is financially incentivized to meet key project objectives and deliverables. It reduces the Government's exposure to financial risk while accounting for the cash flow needs of suppliers. As long as the Government buyers are satisfied with the performance of the contractor and has an associated exit strategy, milestone payments create powerful alignment between the goals of the Government buying organization and the contractor. Such milestone payments tie to performance with an exit strategy also serve to prevent lock-in effects where the Government develops such an integrated relationship with a provider that the switching costs effectively turn competition into a negotiation in which the service provider wields significant market power. When highly integrated teams engage in iterative development and delivery, these milestones may take on a different nature. Instead of large looming milestones, for example, it is possible for mature organizations to create "foot pebbles"—or shorter rewards based on successful completion of "done" for each sprint. It will still be essential to keep the team focused on the larger goals within a project by marking large distances with landmarks (a.k.a., a "milestone"), but perhaps it is equally as important to mark and reward progress based on some notable pebbles within the path itself. This form of iterative incentive awards could be another step closer to providing a realistic means for the Government teams to shift their thinking to agile development.

Technical Metrics for Performance-Based Contracts

While incentives should be tied to performance metrics, metrics for successful performance are not synonymous with incentives. Performance metrics are based on deployment, adoption, and refinement are essential measuring tools for successful performance regardless of whether an AI contract utilizes monetary incentives.

Technical implementation metrics are vital for tracking deployment, adoption, and refinement of AI systems:

Deployment Metrics (Guest Contributor, 2019)
Deployment Frequency—How long the AI takes to deploy
Failed Deployments—Deployments that cause issues or outages
Code Committed—The number of Commits the team makes to the software before it can be deployed into production
Lead Time—Measuring the amount of time passing between inception and the actual production and deployment, the team's ability to adapt to change
Error Rate—A function of the transactions that result in an error during a particular time window
Mean Time to Detection (MTTD)—The amount of time that passes between the beginning of the issue and the time when the issue gets detected and some remedial action is taken
Mean Time to Recovery (MTTR)—The average time taken by the team to repair a failure in the system

Adoption Metrics (Dilmegani, 2023)
Total Users—The total number of people using the AI tool
Active Users—The number of people who access the AL tool
Engaged Users—The number of people who interact with AI tool
New Users—The increase in users after initial implementation
Retention Rate—The percentage of users that continue use of the AI tool

Refinement (Dilnegani, 2023)
Goal Completion Rate—Captures the percentage of a successful engagement
Goal Completion Time—The time to complete a successful engagement
Failure Rate—Captures the percentage of a non-successful engagement
User Satisfaction—Defined through exit surveys
Change Fail Percentage—Captures the percentage of non-successful changes/updates

       IPTs should work closely enough to know how things are going, thereby negating the need for a monthly report. Performance metrics centered on deployment are more fitting for metrics to measure acceptable contract performance. Performance metrics centered on adoption and refinement are more challenging as such goals may not be attained despite acceptable contractor performance. As such, these metrics could be more fitting as milestone incentives. Foot-pebble incentives can be based on individual sprint success, tracked iteratively and simply, and calculated and paid by the month/quarter/etc.

Conclusion and Final Thoughts

       Artificial intelligence, machine learning, and deep learning represent the future of disruptive information technology innovations. While the promise of AI solutions is great, they are not one-size-fits-all and require varied types of domain knowledge, flexible processes, and capable technologies to effectively develop and implement. In some cases, the value proposition of a proposed AI solution may not outweigh the resource costs of its ultimate development and implementation because of the sheer complexity of the AI development pipeline. Acquisition of a successful AI capability thus requires a combination of cultural changes, technical capacity, contracting flexibility, and effective incentives. Because of these myriad factors, the Government must develop and execute a highly efficient yet carefully curated process that is tailored for the specific needs of the AI acquisition life cycle.

References

Agile Government Contracting. (n.d.). https://www.agilegovernmentcontracting.com/

Blake, A. (2023, March 29). Tech leaders call for development pause for GPT-4.5, GPT-5. *Digital Trends*. https://www.digitaltrends.com/computing/pause-ai-development-open-letter-warning/

Dilmegani, C. (2023, January 19). Top 22 metrics for chatbot analytics in 2023. *AIMultiple*. https://blog.aimultiple.com/chatbot-analytics/

Engstrom, D. F, Ho, D. E., Sharkey, C. M., & Mariano-Florentino, C. (2020, February). *Government by algorithm: Artificial intelligence in federal administrative agencies*. Stanford Law School, Thomson Reuters. https://law.stanford.edu/education/only-at-sls/law-policy-lab/practicums-2018-2019/administering-by-algorithm-artificial-intelligence-in-the-regulatory-state/acus-report-for-administering-by-algorithm-artificial-intelligence-in-the-regulatory-state/

Fountaine, T., McCarthy, B., & Saleh, T. (2020, June 1). Building the AI-powered organization. *Harvard Business Review*,. https://hbr.org/2019/07/building-the-ai-powered-organization

Guest Contributor. (2019, August 20). 7 crucial DevOps metrics that you need to track. *Packt Hub*. https://hub.packtpub.com/7-crucial-devops-metrics-that-you-need-to-track/

Hosanagar, K. (2017, April 13). The democratization of machine learning: What it means for tech innovation. *Knowledge at Wharton*. https://knowledge.wharton.upenn.edu/article/democratization-ai-means-tech-innovation/

Hutson, M. (2018, May 24). How researchers are teaching AI to learn like a child. *Science*. https://www.science.org/content/article/how-researchers-are-teaching-ai-learn-child.

Petrocelli, T. (2018, September 20). Democratization of AI development is beginning. *CMSWire.com.* https://www.cmswire.com/digital-workplace/democratization-of-ai-development-is-beginning/

Pisano, E. D. (2020, January 1). AI shows promise for breast cancer screening. *Nature News*, Nature Publishing Group. https://www.nature.com/articles/d41586-019-03822-8

Rahulbot et al.(2017, December 9). You don't need a data scientist, you need a data culture. *Data Therapy*. https://datatherapy.org/2017/12/06/building-a-data-culture/

Schmidt, E. (2023, March 14). Innovation power. *Foreign Affairs*. https://www.foreignaffairs.com/united-states/eric-schmidt-innovation-power-technology-geopolitics.

SITNFlash. (2020, October 26). Racial discrimination in face recognition technology. *Science in the News*. https://sitn.hms.harvard.edu/flash/2020/racial-discrimination-in-face-recognition-technology/